Aus dem Zentrum für Innere Medizin der Universität Heidelberg
(Zentrumssprecher: Prof. Dr. med. Dirk Jäger)
Klinik für Hämatologie, Onkologie, Reumatologie
(Ärztlicher Direktor: Prof. Dr. med. Carsten Müller-Tidow)

Biological and clinical impact of the immunoglobulin light chains sequence diversity in patients with dominant kidney AL amyloidosis

Inauguraldissertation
Zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
an der
Medizinischen Fakultät Heidelberg
der
Ruprecht-Karls-Universität

vorgelegt von
Sarah Ursula Schreiner

aus
Speyer

2024

Dekan: Herr Prof. Dr. Michael Boutros

Doktormutter: Frau Apl.-Prof. Dr. med. Ute Hegenbart

Table of contents

Abbreviations

AA	Amino acid
AGG	Beta sheet aggregation tendency
AL	AL amyloidosis
AL_D	AL patients with diverse organ involvement
AL_H	AL patients with dominant heart involvement
AL_K	AL patients with dominant kidney involvement
ANS	Autonomic nerve system
AP	Alkaline phosphatase
ASCT	Autologous stem cell transplantation
ATTR	Wild-type transthyretin
BMPC	Bone marrow plasma cell
cDNA	Complementary deoxyribonucleic acid
CDR	Complementary determining regions
CR	Complete remission
CRAB	hypercalcemia, renal insufficiency, anemia and bone lesions
cTnT	Cardiac troponin T
dFLC	difference between disease-associated and uninvolved circulation free light chains
DNA	Deoxyribonucleic acid
DPBS	Dulbecco´s phosphate buffered saline
EDTA	Ethylenediaminetetra acetic acid
eGFR	estimated glomerular filtration rate
EM	Electron microscopy
Fab	Antigen-binding structures
FBS	Fetal bovine serum
Fc	Receptor-binding domain
FR	Framework region
GI	Gastrointestinal tract
GMMG	German-Speaking Myeloma Multicentre Group
GRAVY	Grand average of hydropathicity
GTC	Guanidinium thiocyanate
HC	Heavy chain

iFISH	Interphase fluorescence in situ hybridization
Ig	Immunoglobulins
IGKC	Immunoglobulin kappa constant segment
IGKJ	Immunoglobulin kappa joining segment
IGKV	Immunoglobulin kappa variable segment
IGLC	Immunoglobulin lambda constant segment
IGLJ	Immunoglobulin lambda joining segment
IGLV	Immunoglobulin lambda variable segment
LC	Light chain
MGUS	Monoclonal gammopathy of undetermined significance
miRNA	Micro ribonucleic acid
MM	Multiple myeloma
M-protein	Monoclonal protein
Mw	Molecular weight
n/a	Not available
NGS	Next-generation sequencing
NMR	Nuclear magnetic resonance
pI	Isoelectric point
PI	Proteasome inhibitor
PNS	Peripheral nervous system
RNA	Ribonucleic acid
rr	Relative risk
SMM	Smoldering multiple myeloma
ST	Soft tissue
TAE	Tris acetate-EDTA

List of figures

List of tables

1 Introduction

1.1 Immunoglobulins

In 1890, von Behring and Kitasato first reported the identification of an agent capable of neutralising the diphtheria toxin in the blood. The ability of this agent to discriminate between two immunological compounds was described as “Antikörper”, or antibodies in studies the following year. The substance that triggers the formation of an antibody was thereafter referred to as the "Antisomatogen+Immunkörperbildner". This term was shortened to antigen (Schroeder and Cavacini 2010; Von Behring and Kitasato 1890). The structure and function of antibodies, also called immunoglobulins (Igs), have been studied for more than a century, highlighting the complexity of this protein (Schroeder and Cavacini 2010).

1.1.1 Pathway of B-cell development

Igs are expressed as part of the B cell receptor in form of membrane-bound Igs on the surface of B lymphocytes (mature B cells and memory B cells) or in form of soluble Igs secreted by plasma cells. The B cell differentiation from the haematopoietic stem cell to the memory B cell or plasma cell can be divided into different phases (Figure 1), distinguishing between two main phases, the antigen independent phase occurring in the bone marrow and the antigen dependent phase occurring in the germinal centres of the secondary lymphoid organs (Lefranc and Lefranc 2020). In most higher vertebrates, five different classes of antibodies are formed: IgM, IgG, IgA, IgE and IgD (Feige et al. 2010). The mature B cells express IgM and IgD and the other different classes or subclasses of Igs are expressed by the memory cells or plasma cells and generally require cooperation between B and T cells (Lefranc and Lefranc 2020).

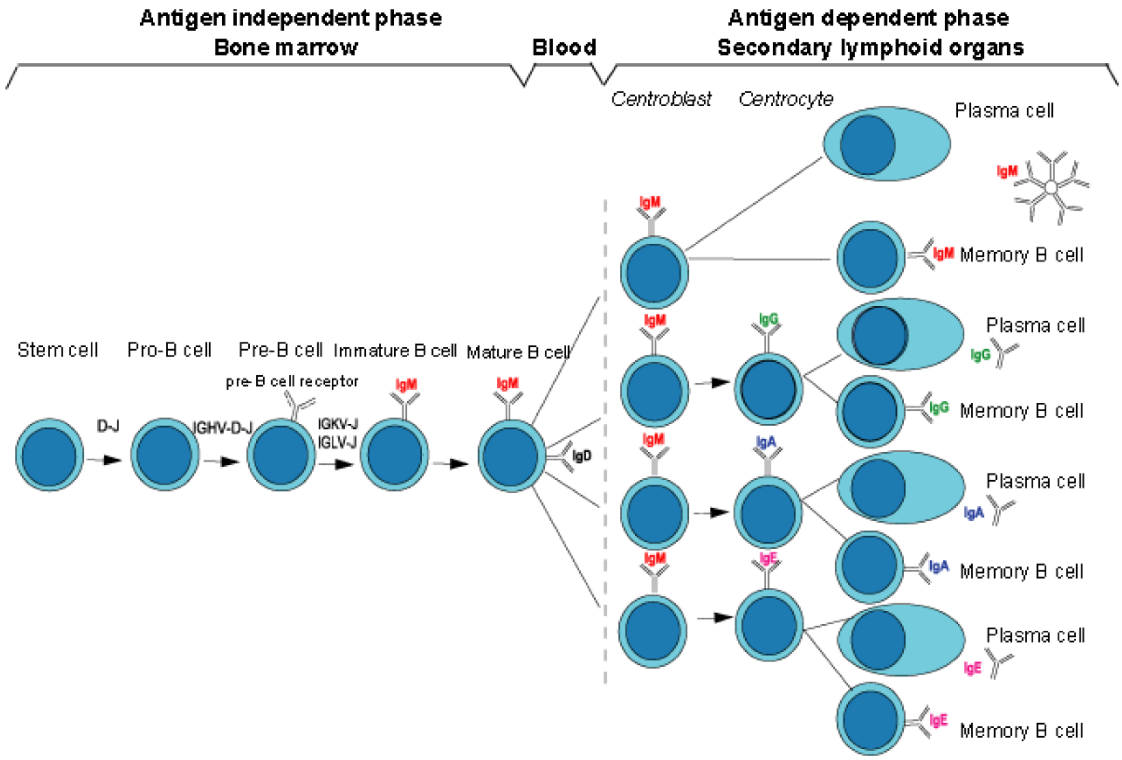

Figure 1. B cell differentiation. (Lefranc and Lefranc 2020). This figure is published under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

1.1.2 Structure of Immunoglobulins

Igs are formed from two heavy chains (HCs) and two light chains (LCs) (Figure 2). The LC can be either a kappa or a lambda chain. Each chain comprises one NH_2-terminal "variable" (V) domain and one (Ig LCs) or more (Ig HCs) COOH-terminal "constant" (C) domains, with consisting of two sandwich β-sheets "held together" by disulfide bridge between two conserved cysteine residues. HCs with three C domains typically have a spacer hinge region across the first (C_H1) and the second (C_H2) domains. In previous studies of Ig structure, enzymes such as papain were used to fragment IgG molecules, resulting in two antigen-binding structures (Fab) and one receptor-binding domain (Fc). The Fab contains an entire LC as well as the V- and C_H1-regions of one HC (Schroeder and Cavacini 2010). The different classes of Igs differ in the HC C-regions used and both types of LCs (kappa or lambda) can assemble with all HC classes (Feige et al. 2010). The V domain primary sequence is functional separated into three hypervariable intervals called complementarity determining regions (CDRs) located between four regions of stable sequence called frameworks (FRs) (Schroeder and Cavacini 2010). In healthy

individuals the kappa to lambda ratio in the serum is approximately 2:1 (Lefranc and Lefranc 2020).

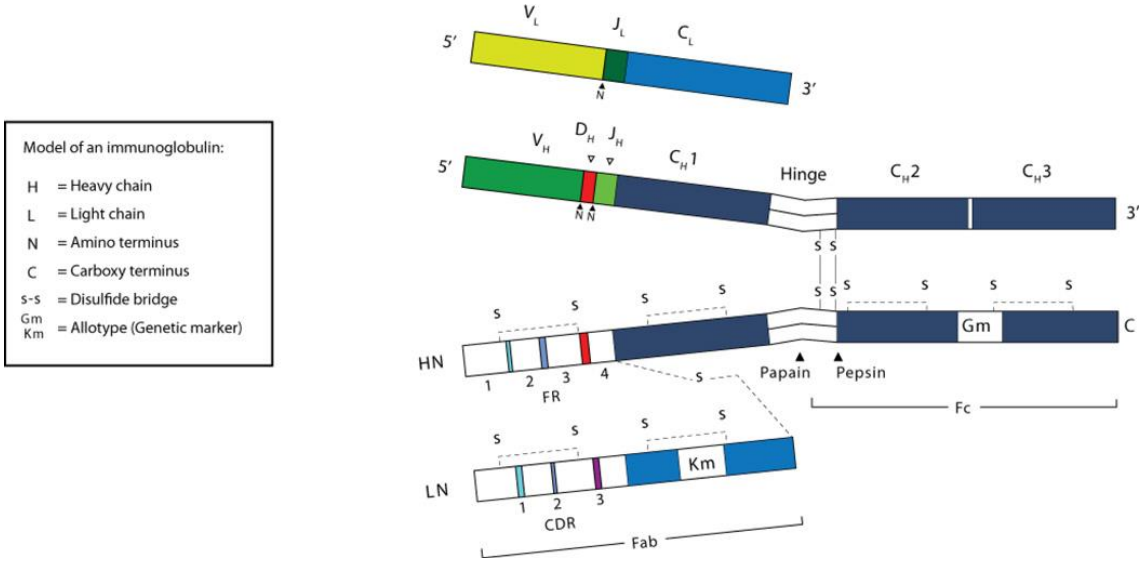

Figure 2. Two-dimensional model of an IgG molecule. (Schroeder and Cavacini 2010) Reproduced with permission from Elsevier (https://www.jacionline.org/article/S0091-6749(09)01465-1/fulltext).

Each of the five different Igs (IgM, IgD, IgG, IgA and IgE) is associated with a specific HC: H-mu, H-delta, H-gamma, H-alpha and H-epsilon, respectively with different biological functions, physicochemical properties, antigenic determinants, and effector properties. Further, the IgG and IgA are separated into four subclasses for IgG (IgG1, IgG2, IgG3 and IgG4) and two IgA subclasses (IgA1 and IgA2). Accordingly, there are nine different HC isotypes in human, namely H-mu, H-delta, H-gamma1, H-gamma2, H-gamma3, H-gamma4, H-alpha1, H-alpha2, H-epsilon. The HC constant region, which is encoded by one of the nine functional immunoglobulin (IG) HC genes (IGHC) (IGHM, IGHD, IGHG3, IGHG1, IGHA1, IGHG2, IGHG4, IGHE and IGHA2), distinguishes the nine isotypes. And Kappa and lambda LCs can be associated with any HC isotype (Lefranc and Lefranc 2020).

1.1.3 The immunoglobulin kappa and lambda loci in homo sapiens

The Ig kappa (*IGK*) and Ig lambda (*IGL*) loci consists of *V*-, *J*- and *C*-genes. The *IGK* locus is located on the short arm of chromosome 2p11.2 and spans 1820 kb. The locus consists of 76 *IGKV*, five *IGKJ* and one *IGKC* gene. The 76 *IGKV* genes can be divided

into six subgroups (Lefranc and Lefranc 2020). And one third of the V kappa gene segments have stop codons or frameshift mutations, resulting in approximately 30 functional V kappa gene segments. The kappa V domains are the joint product of the V kappa and J kappa gene segments while the kappa C domains are encoded by a single C kappa exon (Figure 3). Every active V kappa gene segment can rearrange with any of the five J kappa elements, resulting in a potential combinatorial repertoire of over 140 VJ combinations. The V kappa gene segment comprises FR1, FR2 and FR3, CDR1 and CDR2 and the amino terminus of CDR3, while the J kappa possesses the full carboxy terminus of CDR3 and FR4 (Schroeder and Cavacini 2010). The *IGL* locus, located on the long arm of chromosome 22q11.2 spans 1050 kb and consists of 73-74 *IGLV* genes with 30-36 potentially functional V lambda genes and seven to 11 *IGLJ* and seven to 11 *IGLC* genes, with four functional C lambda exons, associated with its own J lambda gene (Lefranc and Lefranc 2020; Schroeder and Cavacini 2010).

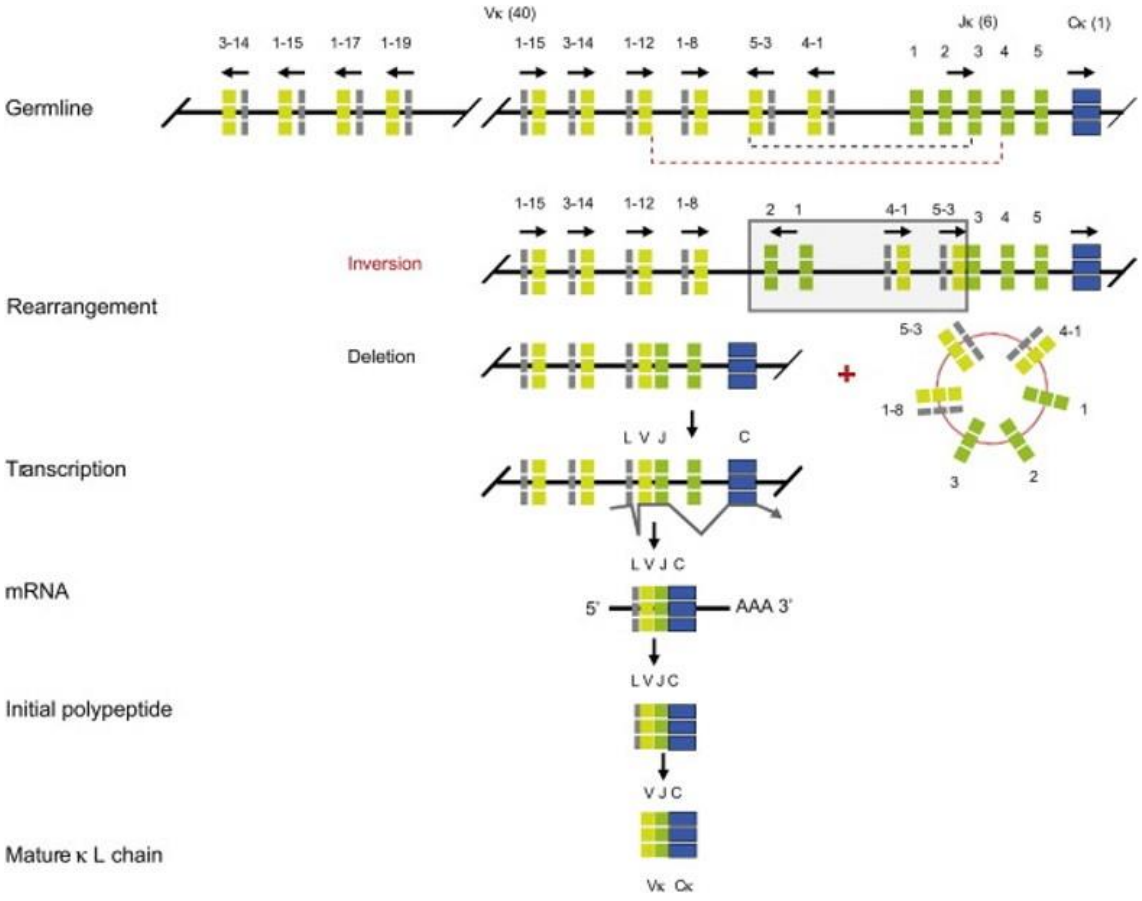

Figure 3. Rearrangement in human kappa locus. (Schroeder and Cavacini 2010). Reproduced with permission from Elsevier (https://www.jacionline.org/article/S0091-6749(09)01465-1/fulltext).

1.1.4 V(D)J- recombination

Immunoglobulin chain variable domain diversity is mainly due to combinatorial diversity of V-J and V-D-J junction, junctional diversity as well as somatic hypermutations. Furthermore, the binding of the variable domains of the HC and the LC, together to build the antigen binding site, generates an additional degree of variation (Lefranc and Lefranc 2020).

To form the V domain a recombination signal sequence is needed. This signal sequence contain highly conserved base pair sequences that contain seven nucleotides (heptamers) and are separated by a base pair spacer from fewer conserved nine base pair sequences (nonamer), thereby avoiding unproductive V-V or J-J rearrangements (Schroeder and Cavacini 2010).

The activating genes *RAG-1* and *RAG-2*, which are expressed in developing lymphocytes, are required to initiate the V(D)J-recombination (Dudley et al. 2005). It leads to DNA double-strand breaks between the ends of the rearranging gene segments and neighbouring recombination signal sequences. These double-strand breaks are repaired by DNA repair processes. Terminal deoxyribonucleotidyltransferases can variably incorporate non-germline-encoded nucleotides (N nucleotides) into the coding ends of the recombination product. After the paired complex formed, RAG proteins truncate the DNA at the heptamer sequences. A hairpin loop is created by the ligation of the 3´-OH with the 5´-phosphate. By cutting the DNA again, a 3´overhang is formed, which enables further modification. The cut ends are then fixed by various proteins and ligases (Schroeder and Cavacini 2010).

1.2 Plasma cell diseases

Plasma cell dyscrasias are a group of heterogeneous diseases caused by the expansion of monoclonal plasma cells that produce a monoclonal Ig or Ig fragments (M-protein) detectable in the serum and/or urine. In the benign form it remains asymptomatic but in the malignant form, plasma cell expansion is uncontrolled and responsible for a wide range of symptoms (Boccadoro and Pileri 1995). Monoclonal gammopathy of undetermined significance (MGUS) is a clinically asymptomatic, premalignant form of

plasma cell disease and usually discovered by chance. 1 - 2 % of people in the United States and Europe were diagnosed with MGUS and the incidence increases with age and is higher in men and in people of African descent. Approximately half of those who were diagnosed with MGUS at age 70 had a monoclonal paraprotein for at least ten years. IgG associated MGUS accounts for 70 % of cases, followed by IgM with 15 % and IgA with 12 %. Also, MGUS with IgE, IgD and only LCs have been described. MGUS is diagnosed by the detection of less than 3 g/dl of monoclonal paraprotein and less than 10 % clonal bone marrow plasma cells (BMPCs) in asymptomatic patients (Castillo 2016; Rajkumar 2022; Rajkumar et al. 2015). Smoldering multiple myeloma (SMM), which is another type of plasma cell disease, can develop from an MGUS and is typically characterized by the presence of monoclonal paraprotein ≥ 3 g/dl and/or 10 % to 60 % BMPCs without signs of end-organ damage, like hypercalcemia, renal insufficiency, anemia and bone lesions (CRAB)-criteria, attributed to the disease. Compared to MGUS, which often occurs in the general population over 50 years of age, SMM is a comparatively rare clinical entity (Rajkumar 2022; Rajkumar et al. 2015). In a study of Kristinsson et al. (2013) it was shown that 14 % of patients with newly diagnosed multiple myeloma (MM) had also SMM. As a multi-step process that begins with precursor disease states, like MGUS and SMM, MM can develop (Figure 4). Even though, MGUS, SMM and MM have distinct clinical characteristics and there are numerous biological similarities across these disease states. MM is a biologically and clinically heterogeneous disease with multiple genetic alterations postulated as driving events in myeloma genesis. Primary genetic events, such as chromosomal translocations and aneuploidy, associated with the development of the precursor states can potentially be linked to the development of MM. While no precise genetic event signals the switch from MGUS and SMM to MM, patients with certain genetic and epigenetic abnormalities, such as DNA methylation and microRNA (miRNA) expression, are at a higher risk for getting MM (Kumar et al. 2017). In 2020, MM accounted for an estimated incidence of 1.78 per 100 000 people globally and 14 % of all haematological malignancies (Huang et al. 2022). The incidence varies between the individual countries, although it is highest in more industrialized countries such as the United States, Western Europe and Australia. This higher incidence in wealthy countries is most likely due to improved diagnosis and increased clinical awareness. MM is 2-3 times more common in black than in white people, while it is less common in Asian and Hispanic people. And MM can further progress to non-bone marrow disease such as extramedullary myeloma and plasma cell leukaemia (Kumar et al. 2017). In addition to

the presence of 10 % BMPCs or biopsy-proven plasmacytoma, the diagnosis of MM requires the identification of one or more myeloma-defining events. Myeloma-defining events were described by the CRAB criteria and three specific biomarkers, bone marrow clonal plasmacytosis of 60 %, serum free light chain (FLC) ratio of 100 and the presence of focal lesions on magnetic resonance imaging (Rajkumar 2022).

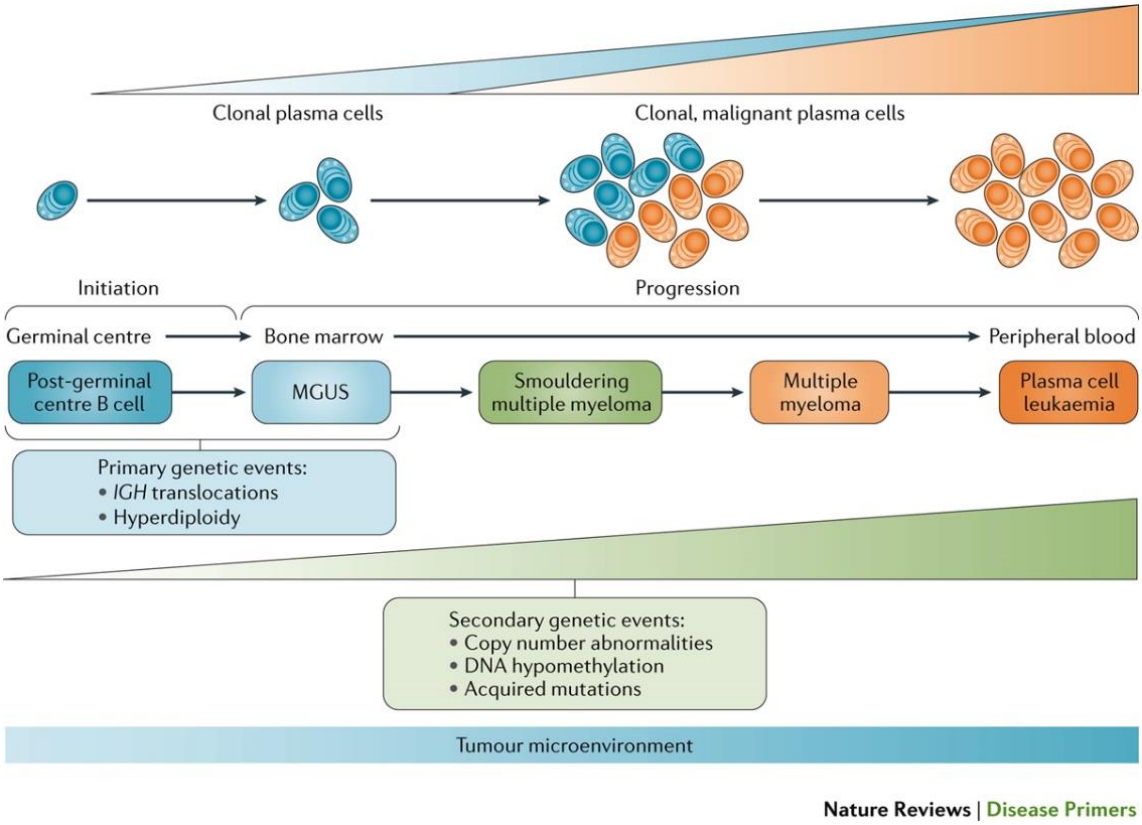

Figure 4. Development of monoclonal gammopathies. (Kumar et al. 2017). Reproduced with permission from Springer Nature (https://www.nature.com/articles/nrdp201746).

1.3 Amyloidosis

Molecular research of amyloidosis began in the middle of the 19th century with pathological classification of human tissues obtained at autopsy by staining with metachromatic dyes such as crystal violet or with Congo red, especially under polarized light, and tioflavins, identifying areas in human tissues belonging to a range of disease as amyloidosis (Buxbaum and Linke 2012). Amyloidosis is caused by misfolding and aggregation of proteins into highly organised amyloid fibrils, which are deposited in the tissues and lead to persistent organ damage. Protein aggregates or preceding intermediates

can lead to cell dysfunction and cell death, a process known as proteotoxicity. Furthermore, the deformation of tissue structure due to amyloid deposition contributes to organ dysfunction (Merlini et al. 2018).

So far, 36 proteins have been identified that can form extracellular amyloid fibrils in humans, including some that form localised deposits, such as β-amyloid in Alzheimer's disease, resulting in localised amyloidosis, while others accumulate throughout body tissues, known as systemic amyloidosis. Over 17 proteins can cause systemic amyloidosis; immunoglobulin heavy or LCs can form both systemic amyloid deposits and localised amyloid deposits. The most common forms of systemic amyloidosis include wild-type transthyretin (ATTR) amyloidosis and monoclonal immunoglobulin LC amyloidosis (known as AL amyloidosis), Wild-type ATTR amyloidosis is characterized by transthyretin aggregation, is age-dependent and predominantly observed in men over 70 years of age, whereas AL amyloidosis (AL) is usually preceded by monoclonal gammopathy. Although the biochemical and aetiological diversity of systemic amyloidosis, clinical manifestations of the various types mostly overlap and are primarily determined by organs affected (Merlini et al. 2018).

1.3.1 Clinical presentation and diagnosis of AL amyloidosis

AL is most typically diagnosed in patients with less than 10 % plasma cells in the bone marrow, but it can also be associated with MM, or other disease like non-Hodgkin’s lymphoma, Waldenström’s macroglobulinemia, chronic lymphocytic leukaemia (Baker 2022). AL is mostly acquired, only in one case a mutation in the constant region has been reported, which is due to a hereditary cause (Benson et al. 2015; Merlini et al. 2018). It’s a rare disease with an incidence of 5.1 to 12.8 cases per million people worldwide. By age, the incidence of AL is increasing, with a median age at diagnosis of 64 years. In fact, only about 5 % of patients have not reached the age of 40. Like MM, more men than woman are affected (3:2) (Baker 2022). The diagnosis of AL depends on the identification of the monoclonal Ig. In most cases, the diagnosis is achieved with serum and urine protein electrophoresis (PEL), serum and urine protein immunofixation (IFE) and/or quantification of serum free light chain (FLC). Immunofixation is often required to detect the paraprotein in AL patients due to the low amount of circulating monoclonal protein typically associated with AL compared to MM (Feitosa et al. 2022; Katzmann et al. 2009).

Organ involvement

AL shows a variety of disease manifestations, whereby different tissues and organs can be affected. Especially, the heart and kidneys are the most frequently affected organs. But other organs, such as the liver or also the soft tissue, the peripheral nervous system or gastrointestinal tract can be affected by the deposits (Figure 5) (Merlini et al. 2018). Thereby the clinical manifestation does not depend on the different amyloidosis subtype (Feitosa et al. 2022). Involvement of the heart in particular is a prognostically unfavourable factor or the involvement of three or more organs (Nienhuis et al. 2016). In addition, a correlation between LCs of certain IG families and preferential deposition in different organs has been found. For example, lambda LCs of the *IGLV1-44* family preferentially deposit in the heart and *IGLV6-57* LCs in the kidneys (Kourelis et al. 2017).

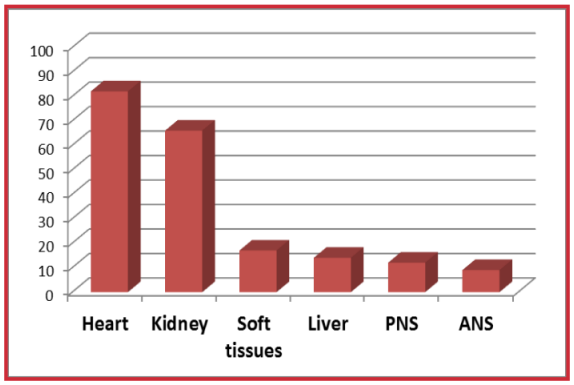

Figure 5. Organ involvement in AL amyloidosis. (Milani et al. 2018). (https://creativecommons.org/licenses/by-nc/4.0/). PNS= peripheral nervous system, ANS= autonomic nervous system.

Clinical staging systems in the context of AL amyloidosis

Within diagnostics, clinical prognostic markers for heart and kidney failures are routinely determined in AL amyloidosis and used for classification into different risk groups. N-terminal pro-brain natriuretic peptide (NT-proBNP), cardiac troponin T (cTnT) and the difference of involved and uninvolved free LC (dFLC) are used for classification according to Mayo for cardiac involvement (Dispenzieri et al. 2004; Kumar et al. 2012; Wechalekar et al. 2013) and proteinuria and the estimated glomerular filtration rate (eGFR) for renal involvement (Palladini et al. 2014) (Table 1).

Table 1. Clinical staging systems in AL amyloidosis. Mayo 2004 by (Dispenzieri et al. 2004), Mayo 2004 European by (Wechalekar et al. 2013), Mayo 2012 by (Kumar et al. 2012), renal staging by (Palladini et al. 2014).

Model	**clinical parameter and cut-offs**	**Stages**
Mayo 2004	NT-proBNP= 332 ng/L cTnT= 0.035 ng/mL	I: no clinical parameter >cut-offs II: one clinical parameter >cut-offs III: both clinical parameters >cut-offs
Mayo 2004 European	The Mayo 2004 stage III was differentiated by: NT-proBNP= 8500 ng/L	IIIa: Mayo 2004 III and NT-proBNP <8500 ng/L IIIb: Mayo 2004 III and NT-proBNP >8500 ng/L
Mayo 2012	NT-proBNP= 1800 ng/L cTnT= 0.025 ng/mL dFLC= 180 mg/L	I: no clinical parameter >cut-offs II: one clinical parameter >cut-offs III: two clinical parameters >cut-offs IV: all clinical parameters >cut-offs
Renal Staging	Proteinuria= 5 g/24h eGFR= 50 mL/min per 1.73 m^2	I: proteinuria ≤cut-off and eGFR ≥cut-off II: either proteinuria >cut-off or eGFR <cut-off III: proteinuria >cut-off and eGFR <cut-off

NT-proBNP= N-terminal pro-B-type natriuretic peptide, cTnT= troponin-T, dFLC= difference between involved and uninvolved serum free light chain concentration, eGFR= estimated glomerular filtration rate

Studies based on 1065 patients newly diagnosed with AL from the Pavia Amyloidosis Research and treatment centre showed the correlation between the different stages of cardiac AL and the median survival or the progression to dialysis in relation to renal involvement. Accordingly, to the results by the Pavia research group, patients with renal stage I have a 1 % risk of requiring dialysis after two years. However, the risk was much higher for stage II patients with 12 % and for stage III patients with 48 % (Milani et al. 2018).

AL amyloidosis and kidney involvement

Clinically, 50-80 % of AL cases affect patients kidneys, making it one of the most common forms of renal amyloidosis globally (Dember 2006). The diagnosis is frequently established following the development of proteinuria ≥ 5 g/24 h, but renal biopsy is also used as a basis for diagnosis. The gold standard for the diagnosis of AL, regardless of the organ affected but not specific for the amyloid subtype, is the detection of positive apple-green birefringence in tissue samples stained with Congo red in polarized light (Figure 6 C). Histologic examination of the affected organ provides the most sensitivity for diagnosis, although other sides such as abdominal fat pads, salivary glands or rectal

mucosa can also be examined. To identify amyloid precursors renal biopsy is used. The deposition of amyloid mainly occurs in the glomeruli. Under the light microscope, amyloid deposits in the glomerulus are visible as amorphous material in the mesangium and capillary loops. The mesangial deposits may be intense and give an amorphous, hyaline appearance with a nodular configuration on haematoxylin-eosin staining (Figure 6 A) (Feitosa et al. 2022). With the use of electron microscopy, the diagnosis can be made even if the Congo red staining is negative, which can occur in approximately 35 % of AL cases (Novak et al. 2004).

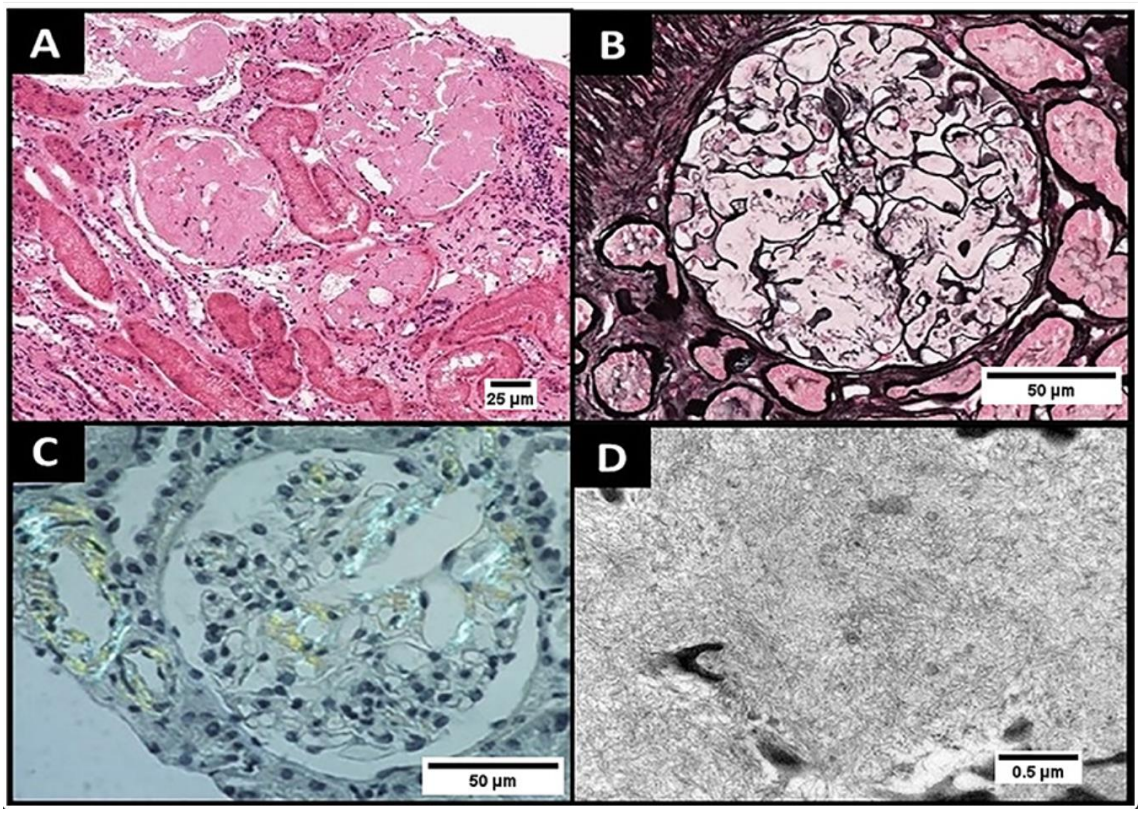

Figure 6. Kidney biopsy from a patient with AL amyloidosis. **A** haematoxylin-eosin staining (original magnification 200x) **B** silver methenamine (400x), **C** Congo red staining (400x) **D** Ultrastructural examination (60,000x) (Feitosa et al. 2022). This figure was published under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

1.4 Fibril formation

Amyloidogenic LCs can exist both, in the native state and in the form of amyloid fibrils. To change between the two states, a high energy barrier must be overcome, by unfolding the native state to enter fibril formation (Hartl and Hayer-Hartl 2009). In general, the *in-*

vitro reaction to form amyloid fibrils is defined as a nucleated self-assembly process, which is favoured under destabilising conditions. This process comprises three characteristic phases, the "lag" phase (Figure 7), which is characterised by a slow oligomerisation of non-native protein conformation. Followed by the "elongation" phase (Figure 7), after reaching a critical concentration of critical fibrillar nucleus forms. It is assumed that the fibril core can interact with the native proteins and integrate them or grow with sequential incorporation of further non-native protein conformations. This critical concentration varies depending on the stability of LCs, it can be low for very unstable LCs or higher for those with higher stability. Once an equilibrium state or "plateau" phase is reached, further polymerisation is suspended (Blancas-Mejía and Ramirez-Alvarado 2013; Merlini et al. 2018). Investigations of the "lag" phase show structural transitions that precede fibril formation (Kazman et al. 2021). It was shown that the process begins with a partial unfolding of the V domain and the formation of small dimers. The hydrophobic core of the LC domain rearranges during oligomerisation and structural changes from an antiparallel to a parallel beta-sheet secondary structure occur before amyloid formation (Kazman et al. 2021).

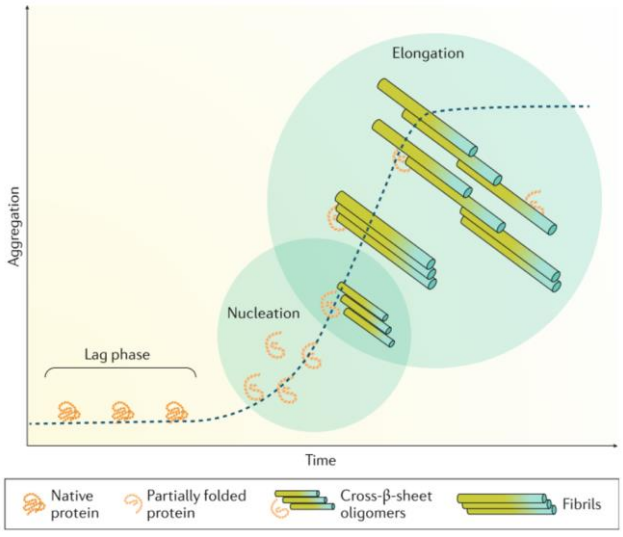

Figure 7. Kinetics of fibril formation in vitro. (Merlini et al. 2018) Reproduced with permission from Springer Nature (https://www.nature.com/articles/s41572-018-0034-3).

1.5 Published structures of amyloid fibrils

Amyloid fibrils are highly ordered aggregates that have a cross beta-sheet structure composed of a highly repetitive intermolecular beta-sheet motif which is unique among protein folds (Greenwald and Riek 2010). Investigations of the molecular structure using cryo-electron microscopy (EM), or nuclear magnetic resonance (NMR) spectroscopy analysis can contribute to the understanding of fibril formation (Iadanza et al. 2018; Makin and Serpell 2005; Sunde and Blake 1997). Previous cryo-EM studies on lambda fibrils revealed the structure of four different lambda subfamilies, which included *IGLV1-51*, *IGLV1-44*, *IGLV3-19* and *IGLV6-57* (Radamaker et al. 2021a; Radamaker et al. 2021b; Swuec et al. 2019). NMR analysis of kappa fibrils have only been carried out in limited cases (Hora et al. 2017; Piehl et al. 2017a; Piehl et al. 2017b). In addition to somatic hypermutations, posttranscriptional modifications such as disulfide bonds and N-glycosylation, particularly in kappa, can favour amyloid formation (Rajkumar et al. 2014; Stevens 2000). Beyond this characteristic, key conformational changes, such as a rotational switch around the molecular disulfide bond, may be responsible for driving fibril formation (Radamaker et al. 2019). The fibrils exhibited the typical cross beta-sheet structure, supported by hydrogen bonds in the backbone and interactions between side chains along the fibril axis (Radamaker et al. 2021a; Swuec et al. 2019). The FOR005 fibrils showed two conformations within a fibril, contributing to structural breaks within the patient amyloid fibrils (Radamaker et al. 2021a).

1.6 Treatment of AL amyloidosis

Therapy of AL focuses on the B-cell clone responsible for the production of the aberrant clonal immunoglobulin protein. The type and extent of treatment rely on a risk assessment based on the patient´s characteristics and clone biology (Merlini et al. 2018). While the current therapies are mainly plasma cell-directed and aim to supress LC synthesis, the investigation of new therapeutic approaches that are able to remove amyloid deposits from involved organs is becoming more important (Absmeier et al. 2022; Chakraborty and Lentzsch 2020; Merlini et al. 2018). For the treatment of low-risk patients, chemotherapy with high-dose melphalan followed by autologous stem cell transplantation (ASCT) is recommended (Comenzo and Gertz 2002; Comenzo et al.

1996; Merlini et al. 2018), which aims to eliminate the amyloidogenic plasma cell clone (Palladini et al. 2004). However, due to the high toxicity, the feasibility of a stem cell transplantation is limited to a minority of patients. Patients who are not eligible for transplantation are usually managed with high-dose dexamethasone in combination with melphalan (Palladini et al. 2004). Proteasome inhibitors (PIs) like bortezomib are also used in combination with cyclophosphamide and dexamethasone (Palladini et al. 2015). Novel treatment approaches aim to remove the deposited fibrils with the help of monoclonal antibodies (Huart 2023; Khwaja et al. 2023; Wechalekar et al. 2021).

1.7 German-speaking Myeloma Multicentre Group - HD6 clinical trial

Treatment of the related disease MM – which was used as control group without amyloid deposits – has become highly variable and complex, targeting the aberrant, monoclonal B-cells and modulating the patient´s immune system. Treatment regimens can incorporate various combinations of immune modulatory drugs, PIs, corticoids, monoclonal antibodies, chemotherapy with high-dose melphalan followed by ASCT and recently CAR-T cells (Rajkumar 2022). The treatment of the MM patient cohort analysed in this study followed the protocol of the clinical trial of the German-speaking Myeloma Multicentre Group (GMMG) randomised, open, multicentre phase III trial. This trial recruited 564 patients with newly diagnosed MM. For the first time, the role of elotuzumab in combination with bortezomib, lenalidomide and dexamethasone (VRD) was investigated as induction/consolidation in combination with lenalidomide maintenance in a high dose setting. Patients received four cycles of VDR as induction therapy and underwent mobilisation and harvesting of peripheral blood stem cells. Subsequently, they were treated with high dose melphalan and ASCT and followed with two cycles of VRD consolidation and lenalidomide maintenance. In arm B1 + B2 patients were additionally treated with elotuzumab in the induction phase, while in A2 + B2 patients received elotuzumab in addition to consolidation and maintenance. Median progression-free survival was the primary endpoint of the trial. Secondary objectives were overall survival, complete remission (CR) after induction, CR after consolidation, time to progression, best treatment response, duration of response, minimal residual disease-negativity, toxicity and quality of life (Salwender et al. 2019).

1.8 Aim of the project

It is still not known why the LC of AL patients form amyloid deposits and why free LCs in other plasma cell dyscrasias, like MM do not. There is indication that the primary structure of the LCs may be a reason for the development of AL, with mutations playing a key role, leading to changes in biochemical properties and ultimately destabilisation. However, previous studies have mainly focussed on the variable part of the LC, rather than the LC in its entirety and science has also concentrated on the more common lambda type, preferably with cardiac involvement. Only a few studies on kappa AL exist. Direct comparisons with other plasma cell diseases, which could shed light on the causes of destabilisation in AL, are also limited.

Consequently, the aim of this project was to investigate the following research questions:

- Are there differences in the LC sequence between kappa AL and MM?
- Are typical sequences features within kappa AL based on a specific organ tropism in particular renal AL?
- How does this compare with lambda AL?
- Do some LCs have a higher amyloidogenic potential than other LCs?

To investigate these questions the following aims were defined:

Aim 1: Development of a molecular genetic assay to identify the full-length LC sequences of rarer kappa AL patients.

Aim 2: Sequence analysis of kappa patient LCs regarding the different organ involvement and comparison with non-amyloidogenic kappa MM sequences to identify characteristics associated with amyloidogenicity.

Aim 3: Sequence analysis and comparison with lambda AL LCs focusing on rarer kidney involvement.

Aim 4: Implementation of a next-generation sequence workflow to verify complex sequences and to identify specific characteristics.

2 Materials and Methods

2.1 Materials

The material used in this study is listed in different categories. With indication of the respective manufacturer, the material is structured as follows: technical equipment (Table 2), consumables (Table 3), chemicals and enzymes (Table 4) and kits (Table 5). Oligonucleotides (Table 6 and Table 7) as well as software and web-based tools (Table 8) are listed separately. The respective company headquarters was specified as the supplier location. Buffers and their formulation are listed in chapter 2.1.6. Patient cohorts are described in chapter 2.1.7.

2.1.1 Technical equipment and consumables

Table 2. Technical equipment

Description	Manufacturer
4200 TapeStation System	Agilent, Santa Clara, CA, USA
Automated Cell Counter Countess™ II	Thermo Fisher Scientific, Waltham, MA, USA
Automated Cell Separator RoboSep™-S	Stemcell™ Technologies, Vancouver, Canada
Bench Mounted Fume Hood Type 1200	WALDNER Holding SE & Co. KG, Wangen, Germany
Centrifuge 5810	Eppendorf, Hamburg, Germany
Centrifuge 5810 R	Eppendorf, Hamburg, Germany
Centrifuge PerfectSpin 24 Plus	Peqlab™ by VWR International, Radnor, PA, USA
Centrifuge Sprout® Plus	Heathrow Scientific®, Vernon Hills, IL, USA
Centrifuge Thermo Scientific™ PICO™ 17 Microcentrifuge	Thermo Fisher Scientific, Waltham, MA, USA
Cuvette Färbekasten nach Hellendahl mit Erweiterung	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Cytocentrifuge Cytospin 4	Thermo Fisher Scientific, Waltham, MA, USA
Electrophoresis Power Supply Lightning Volt OSP-250L	Owl by Thermo Fisher Scientific, Waltham, MA, USA
Eppendorf µCuvette® G1.0	Eppendorf, Hamburg, Germany

Description	**Manufacturer**
Gel chamber multiSUB® MSChoice	Cleaver Scientific, Rugby, UK
HLC Heating-ThermoMixer HTMR-133	DITABIS - Digital Biomedical Imaging Systems AG, Pforzheim, Germany
IKA Vortexer MS3 Basic PCR	Agilent, Santa Clara, CA, USA
Media Storage Bottle CLS-1172 100 mL	SCHOTT AG, Mainz, Germany
Microcentrifuge MC-24 Touch	Benchmark Scientific, Sayreville, NJ, USA
Microscope B3 professional series	Motic®, MoticEurope, S.L.U., Barcelona, Spain
Microwave MICROMAT 135	AEG, Frankfurt am Main, Germany
Molecular Imager® Gel DOC™ XR+ with Image Lab™Software	Bio-Rad Laboratories, Inc., Hercules, CA, USA
NextSeq 550 System	Ilumina, San Diego, CA, USA
Pipette “Discovery comfort” (0.5-10 µL, 2-20 µL, 20-200 µL, 100-1000 µL)	HTL lab solutions, Corning HTL SA, Warsaw, Poland
Pipette “Reference” (0.5-10 µL, 10 - 100 µL)	Eppendorf, Hamburg, Germany
Pipette “Research” (0.5-10 µL, 2 - 200 µL)	Eppendorf, Hamburg, Germany
Pipette controller accu-jet® pro	Brand GmbH + Co. KG, Wertheim, Germany
Precision scale Kern PLJ 3500-2NM	Kern & Sohn GmbH, Balingen, Germany
Quantus™ Flourometer	Promega, Madison, WIS, USA
Spectrometer BioSpectrometer® basic	Eppendorf, Hamburg, Germany
Thermocycler Biometra® TProfessional	Analytik Jena GmbH, Jena, Germany
Vortex Mixer 7-2020	neoLab Migge GmbH, Heidelberg, Germany
Vortex mixer REAX 2000	Heidolph Instruments, Schwabach, Germany
Vortex Mixer Vortex-Genie® 2 - G560E	Scientific Industries SI™ by Thermo Fisher Scientific, Waltham, MA, USA

Table 3. Consumables

Description	Manufacturer
Aqua Spüllösung	B. Braun SE, Melsungen, Germany
Centrifuge Tube PP (50 mL)	Greiner Bio-One International GmbH, Kremsmünster, Austria
Corning®Costar® Stripette®, serological pipettes (5, 10, 25, 50 mL)	Sigma-Aldrich® by Merck, Darmstadt, Germany
Countess™ Cell Counting chamber slides	Invitrogen™ by Thermo Fisher Scientific, Waltham, MA, USA
Countess™ Test Beads	Invitrogen™ by Thermo Fisher Scientific, Waltham, MA, USA
Disposable Hemocytometer C-Chip DHC-N01	NanoEnTek, Seoul, Korea
Double Cytoslide™ Microscope Slides	Epredia™ by Thermo Fisher Scientific, Waltham, MA, USA
Eppendorf LoBind® Tube	Eppendorf, Hamburg, Germany
Eppendorf Safe-Lock® Tubes (1.5, 2 mL)	Eppendorf, Hamburg, Germany
Eppendorf Tubes® (5 mL)	Eppendorf, Hamburg, Germany
Erlenmeyer flask	neoLab Migge GmbH, Heidelberg, Germany
EZ Double Cytofunnel™	Epredia™ by Thermo Fisher Scientific, Waltham, MA, USA
Falcon® Round Bottom Polystyrene Tubes (14 mL)	Corning, Inc., Corning, NY, USA
H_2O Ampuwa Spüllösung	Fresenius Kabi Deutschland GmbH, Bad Homburg, Germany
Kimtech Science™ Precision Wipes™ Tissue Wipers	Kimberly-Clark Professional™ by Thermo Fisher Scientific, Waltham, MA, USA
Parafilm® M	Heathrow Scientific®, Vernon Hills, IL, USA
PCR SingleCap 8er-SoftStrips 0.2 mL	Biozym Scientific GmbH, Hessisch Oldendorf, Germany
pluriStrainer® (100 µm)	pluriSelect Life Science UG & Co.KG, Leipzig, Germany
RoboSep™ Filter Tips	Stemcell™ Technologies, Vancouver, Canada
SafeSeal Tips Professional (10, 20, 100, 200, 1250 µL)	Biozym Scientific GmbH, Hessisch Oldendorf, Germany
Sterile Pasteur pipettes	LP ITALIANA SPA, Milano, Italy
Vernichtungsbeutel	neoLab Migge GmbH, Heidelberg, Germany

Description	Manufacturer
Water, DNase, RNase-free	MP Biomedicals, LLC, Irvine, CA, USA

2.1.2 Chemicals and enzymes

Table 4. Chemicals and enzymes

Description	Manufacturer
2- Mercaptoethanol	Gibco™ by Thermo Fisher Scientific, Waltham, MA, USA
Acetic acid 100%	Carl Roth™ by Thermo Fisher Scientific, Waltham, MA, USA
Agarose basic for biochemistry	neoFroxx, Einhausen, Germany
Ammonium chloride solution	Stemcell™ Technologies, Vancouver, Canada
AmpliTaq Gold™ DNA polymerase with buffer II and $MgCl_2$	Applied Biosystems™ by Thermo Fisher Scientific, Waltham, MA, USA
Buffer RLT	Qiagen, Hilden, Germany
D1000 Reagents	Agilent Technologies, Inc., Waldbronn, Germany
D1000 Screen Tape	Agilent Technologies, Inc., Waldbronn, Germany
dNTP-Set, 100 mM-solution	Thermo Scientific™ by Thermo Fisher Scientific, Waltham, MA, USA
Dulbecco's phosphate buffered saline (DPBS), 10x	Gibco™ by Thermo Fisher Scientific, Waltham, MA, USA
Ethylenediaminetetraacetic acid (EDTA) - Solution pH 8.0 (0.5 M) for molecular biology	PanReac AppliChem ITW Reagents, Darmstadt, Germany
Ethanol absolut ≥99,8%	VWR International, Radnor, PA, USA
Fetal bovine serum (FBS), Qualified, HI	Gibco™ by Thermo Fisher Scientific, Waltham, MA, USA
Histopaque® 1077	Sigma-Aldrich® by Merck, Darmstadt, Germany
Methanol 99.9 %, p.a.	Carl Roth™ by Thermo Fisher Scientific, Waltham, MA, USA
NaOH 10M BioUltra	Sigma-Aldrich® by Merck, Darmstadt, Germany
Nuclease free water, Ambion®	Invitrogen™ by Thermo Fisher Scientific, Waltham, MA, USA

Description	Manufacturer
Oligo(dT)18 Primer	Thermo Scientific™ by Thermo Fisher Scientific, Waltham, MA, USA
RoboSep™ Buffer	Stemcell™ Technologies, Vancouver, Canada
SYBR® safe DNA gel stain	Invitrogen™ by Thermo Fisher Scientific, Waltham, MA, USA
Take5™ 50 bp DNA ladder	highQu GmbH, Kraichtal, Germany
Take5™6x loading dye solution	highQu GmbH, Kraichtal, Germany
Tris acetate-EDTA (TAE) buffer, 10x	Sigma-Aldrich® by Merck, Darmstadt, Germany
Tris-hydrochlorid, 1 M-solution (pH 7.0/Mol. Biol.)	Fisher BioReagents by Thermo Fisher Scientific, Waltham, MA, USA
Trypan Blue solution, 0.4 %	Gibco™ by Thermo Fisher Scientific, Waltham, MA, USA
UltraPure™ 0,5 M EDTA, pH 8,0	Invitrogen™ by Thermo Fisher Scientific, Waltham, MA, USA

2.1.3 Kits

Table 5. Kits

Description	Manufacturer
AllPrep DNA/RNA/Protein Mini Kit	Qiagen, Hilden, Germany
EasySep™ Human CD138 Positive Selection Kit II	Stemcell™ Technologies, Vancouver, Canada
High capacity cDNA reverse transcriptase Kit	Applied Biosystems™ by Thermo Fisher Scientific, Waltham, MA, USA
High Pure PCR-Product Purification Kit	Roche, Basel, Switzerland
IDT® for Illumina® DNA/RNA UD Indexes Set B, (96 Indexes)	Illumina, San Diego, CA, USA
Illumina® DNA Prep, (M) tagmentation (24 samples)	Illumina, San Diego, CA, USA
NextSeq 500/550 Mid Output Kit v2.5 (150 cycles)	Illumina, San Diego, CA, USA
NextSeq PhiX Control Kit	Illumina, San Diego, CA, USA
QIAshredder Kit	Qiagen, Hilden, Germany

Description	Manufacturer
QuantiFluor® ONE dsDNA System	Promega, Fitchburg, WI, USA

2.1.4 Oligonucleotides

The oligonucleotides used in this study were produced by Eurofins. T_m calculation was done by the manufacturer.

Table 6. Oligonucleotides for kappa approach

Description	Sequence (5' → 3')	T_m [°C]
VKKL_A_fw_SaS	CCAGATGACCCAGTCTCC	58.2
VKKL5_A_fw_SaS	CATGTCAGCGACTCCAGG	63,8
VKKL2a_fw_Huhn[1]	CTCCATCTCCTGCAGGTCTAG	57.1
VKKL3a_fw_Huhn[1]	CTGCAGGGCCAGTCAGAG	60.5
CKKL_A_rv_SaS	CACTCTCCCCTGTTGAAGC	57.9

[1]adapted from (Huhn 2018)

Table 7. Oligonucleotides for lambda approach

Description	Sequence (5' → 3')	T_m [°C]
VLKL12a_Huhn[1]	GGTCCTGGGCTCAGTCTG	60.5
VLKL3c_Huhn[1]	TGGTACCAGCAGAAGCCAGG	61.4
VLKL4a_Huhn[1]	CCAGCCTGTGCTGACTCA	58.2
VLKL7a_Huhn[1]	CAGACTGTGGTGACTCAGGAG	61.8
VLKL3_A_fw_NB[2]	CCTATGAGCTGACACAGCC	58.2
VLKL6_A_fw_NB[2]	CAGCCCCACTCTGTGTCG	60.5
VLKL3_H_fw_NB[2]	GCTGACTCAGGACCCTGC	60.5
VLKL3_I_fw_NB[2]	CCCTCAGTGTCCGTGTCC	65.0
CLKL_A_rv_NB[2]	CACTGTCTTCTCCACGGTG	58.8

Description	Sequence (5' → 3')	T_m [°C]
CLKL_B_rv_NB[2] (discarded in the run of the project)	CCCTTCATGCGTGACCTG	58.2

[1]adapted from (Huhn 2018)

[2]designed by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and published in parts (Berghaus et al. 2022)

2.1.5 Software and web-based tools

Table 8. Software and web-based tools

Description	Application	Source
2200 TapeStation Controller and Analysis Software 3.2.037131	Next-generation sequencing quality control	Agilent, Santa Clara, CA, USA
abYsis	Sequence analysis	http://www.abysis.org/abysis/index.html [27.11.2022] (Swindells et al. 2017)
AL-Base	Database	https://wwwapp.bumc.bu.edu/BEDAC_ALBase [27.11.2022] (Bodi et al. 2009)
Clustal Omega	Sequence alignment	https://www.ebi.ac.uk/Tools/msa/clustalo/ [27.11.2022] (Madeira et al. 2022)
Ensembl	Database	https://www.ensembl.org/index.html [27.11.2022] (Howe et al. 2021)
Ensembl Blast	Sequence BLAST	https://www.ensembl.org/Multi/Tools/Blast?db=core [27.11.2022] (Howe et al. 2021)
ExPasy - Compute pI/MW	pI calculation	https://web.expasy.org/compute_pi/ [27.11.2022] (Bjellqvist et al. 1994; Bjellqvist et al. 1993; Gasteiger et al. 2005)
ExPasy - ProtParam	Amino acid composition and grand average of hydropathicity score calculation	https://web.expasy.org/protparam/ [10.01.2023] (Gasteiger et al. 2005)
ExPasy - Translate	cDNA translation	https://web.expasy.org/translate/ [27.11.2022] (Duvaud et al. 2021)
IBM SPSS ® statistics, version 29.0.0.0	Statistical analysis	https://www.ibm.com/de-de/products/spss-statistics [06.11.2022]
IMGT	Database	https://www.imgt.org/ [27.11.2022] (Lefranc et al. 2015)
Local Run Manager 2.2.1.1645	Firmware for next-generation sequencing	https://support.illumina.com/sequencing/sequencing_software/local-run-manager/downloads.html [26.08.2021]
Local Run Manager	Generating FASTQ files,	https://support.illumina.com/downloads/local-run-manager-generate-fastq-module-v2.html [26.08.2021]

Description	Application	Source
Generate FastQ Analysis Module 2.0.1	next-generation sequencing	
MEGA - Molecular Evolutionary Genetic Analysis MEGA-X Version 10.1.8	Sequence analysis	https://www.megasoftware.net/ [27.11.2022] (Kumar et al. 2018)
MiXCR	Analysis FASTQ-files	https://github.com/milaboratory/mixcr/ [03.01.2023] http://mixcr.milaboratory.com/ [21.09.2021] (Bolotin et al. 2015)
Multiple Oligonucleotide Analyser	Oligonucleotide design	https://www.thermofisher.com/de/de/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyser.html [27.11.2022]
NetNGlyc 1.0	N-glycosylation	https://services.healthtech.dtu.dk/services/NetCGlyc-1.0/ [29.12.2023] (Julenius 2007)
UniProt	Database	https://www.uniprot.org/ [02.02.2023] (Consortium 2023)
R x64 4.1.3	Data visualization and statistical analysis	https://www.r-project.org/ [13.03.2022]
R-studio 2022.02.0+443	User interface	https://www.rstudio.com/ [13.03.2022]
Relative risk calculator	MedCalc Software Ltd. Relative risk calculator	https://www.medcalc.org/calc/relative_risk.php Version 20.218 [05.03.2023]
Sequence massager	Sequence analysis	https://biomodel.uah.es/en/lab/cybertory/analysis/massager.htm [27.11.2022]
TANGO	Calculation Aggregation propensities	http://tango.crg.es/ [28.03.2023] (Fernandez-Escamilla et al. 2004; Linding et al. 2004; Rousseau et al. 2006)
Vbase2	Sequence BLAST	http://www.vbase2.org/ [27.11.2022] (Retter et al. 2005)
V_LAmY-Pred	Calculation of aggregation potential	https://web.iitm.ac.in/bioinfo2/vlamy-pred/ [16.03.2023] (Rawat et al. 2021)

2.1.6 Buffers and their formulation

Glacial acetic acid: 20 mL Acetic acid 100% + 60 mL Methanol 99.9 %, p.a

Guanidinium thiocyanate (GTC) -buffer: 250 µL 2-Mercaptoethanol + 25 mL RLT-buffer

NH_4Cl -cell-lysis buffer: 500 mL Ammonium chloride solution + 100 mL H_20 distilled

PBS- buffer (1x): 100 mL 10x DPBS + 900 mL H_20 distilled

PBS-EDTA-FBS-buffer: 100 mL 10x DPBS + 1 mL FBS + 4 mL 0.5 M EDTA - Solution (pH 8.0) + 900 mL H_20 distilled

TAE-buffer (1x): 500 mL 10x TAE + 4500 mL H_20 distilled

2.1.7 Patient cohorts

2.1.7.1 AL Amyloidosis patients

This project included 41 kappa and 52 lambda AL patients from the Amyloidosis Centre of the University Hospital Heidelberg recruited from January 2019 to May 2022 and was approved by the Ethics Committee of the University Heidelberg (S-123/2006, last time renewed 07.12.2021) following the Helsinki guidelines for research of human subjects. All includes patients provided informed consent for utilization of their biomaterial and clinical data for research purpose. As part of their clinical diagnostic workup, bone marrow was aspirated among others for iFISH diagnostics and therefor processed in the GMMG central laboratory and biobank multiple myeloma. Clinical data was obtained and analysed from the respective physician´s letters and diagnostic laboratory results. Patients were categorized based on their organ involvement in accordance to the established criteria (Gertz et al. 2005) and stratified into the following groups: AL patients with “dominant kidney involvement” (AL_K) or “dominant heart involvement” (AL_H) in the absence of other clinically relevant organ involvement. AL patients who had other clinically relevant organ involvement were defined as “diverse” (AL_D) with one AL_D patient having heart and kidney involvement in equal parts in the kappa cohort. Consequently, the kappa cohort consisted of 14 AL_K and 8 AL_H, and 19 AL_D patients. The lambda cohort was composed of 34 AL_K and 18 AL_D patients. AL lambda patients with dominant cardiac involvement were studied as part of another PhD thesis (Natalie Berghaus, co-doctoral student, University Hospital Heidelberg).

2.1.7.2 Multiple myeloma patients

In order to identify AL specific sequence characteristics, besides alignment with reference sequences from databases, LC sequences from patients with MM were included in this study. A total of 83 newly diagnosed kappa MM and 52 lambda MM patients fulfilling the International Myeloma Working Group criteria and treated in the German-Speaking Myeloma Multicentre Group (GMMG) HD6 clinical trial (EudraCT No.: 2014-003079-40) were studied. The clinical trial was conducted in compliance with the ethical guidelines of the Helsinki Declaration and the European Clinical Trial Directive (2005). Approval for the study was obtained from all local ethics committees of the participating institutions and written informed consent was obtained from all patients. Within this trial, simultaneous presentation of AL at time of study inclusion as well as intermediate presentation (mean observation period 49.8 month) were set as exclusion criteria (Goldschmidt et al. 2021; Salwender et al. 2019). Therefore, two patients were excluded at inclusion timepoint and one patient was dropped out during the study period due to co-occurring AL. LC sequences of the MM patients were identified using a bulk RNA sequencing approach. The initial bioinformatic raw data processing for obtaining the LC sequences out of the bulk data was provided by Dr. rer. nat. Alexandra M. Poos. Detailed LC analysis for this project was performed as described in chapter 2.2.9.2-2.2.9.4. The GMMG study group consented to the use of the data and provided the associated clinical data. The verification as well as the extraction of the clinical lambda MM data were part of another PhD thesis (Natalie Berghaus, co-doctoral student, University Hospital Heidelberg) and used in this project for comparison purpose.

2.2 Methods

Parts of these methods have already been published in (Feurstein et al. 2022), (Berghaus et al. 2022) and (Schreiner et al. 2024).

2.2.1 Processing of bone marrow aspirates

In the course of diagnosis, 22.5 - 92.5 mL of bone marrow was collected from patients and processed according to the standardized procedures of the GMMG Central Laboratory and Biobank Multiple Myeloma (SOP für Dichtegradienten-Zentrifugation

von PB und KM, SOP zur Roboterunterstützte CD138$^+$ Sortierung von Probenmaterial nach dem StemCell Protokoll, SOP zum Zählen der Zellzahl von Proben mittels Countess Cell-Counter, SOP zur Erstellung von Präparaten für die Zytogenetische Analyse von Myelomzellen).

2.2.1.1 Isolation of mononuclear cells

To isolate the mononuclear cells, a density gradient centrifugation was performed. For this purpose, up to 4x 20 mL bone marrow aspirate was filtered (pluriStrainer® (100 µm), pluriSelect Life Science UG & Co.KG, Leipzig, Germany) and diluted to a final volume of 50 mL with 1x PBS+EDTA+FBS-buffer (chapter 2.1.6). In order to obtain the maximum quantity of cells, the remaining cells in the filter were washed out. The diluted bone marrow was layered onto 15 mL Histopaque®1077 (Sigma-Aldrich® by Merck, Darmstadt, Germany) without mixing and centrifuged at 2800 rpm for 30 min without brake. The layer of mononuclear cells was transferred via a filter, washed to a final volume of 50 mL with 1x PBS+EDTA+FBS-buffer and centrifuged at 1800 rpm for 10 min. Next, erythrocytes were lysed. The cell pellet was resuspended in 4-6 mL NH_4Cl - cell-lysis buffer (chapter 2.1.6, Stemcell™ Technologies, Vancouver, Canada) and incubated for 10 min, followed diluted to a final volume of 50 mL with 1x PBS+EDTA+FBS-buffer and centrifuged at 1500 rpm for 5 min. The cell pellet was resuspended in 10 mL 1x PBS+EDTA+FBS-buffer and 500 µL of the cell suspension were harvested for biobanking and stored as dry pellet.

2.2.1.2 CD138$^+$ automated bead-based cell separation

Basis for the molecular characterization of tumour cells is the enrichment of malignant cells in the sample. In this context, CD138$^+$ automated bead-based cell separation was performed. CD138 is a specific antigen, which is highly expressed on the surface of plasma cells and by separating the antigen carrying cells a high purity of the samples can be ensured (Roccatello et al. 2020).

For CD138$^+$ cell enrichment, the cell quantity (chapter 2.2.1.1) was first calculated (1:10 dilution Trypan Blue Solution, 0.4 %) with Automated Cell Counter Countess™ II (Thermo Fisher Scientific, Waltham, MA, USA), centrifuged at 1500 rpm for 5 min and adjusted to an optimum of $1x10^8$ cells/mL with 1x PBS/EDTA/FBS-buffer. Cell sorting was performed according to the manufacturer´s protocol “Human CD138 WB and BM

Positive Selection II 17887" (Stemcell™ Technologies, Vancouver, Canada) with "Automated Cell Separator RoboSep™-S" (StemcellTM Technologies, Vancouver, Canada) and the required reagents "EasySep™ Human CD138 Positive Selection Kit II with RoboSep™ Buffer" (StemcellTM Technologies, Vancouver, Canada). After the run, the $CD138^+$ cell fraction was dissolved in 5 mL 1x PBS/EDTA/FBS-buffer, washed through a filter and the cell count of a 1:2 dilution (Trypan Blue Solution, 0.4 %) was quantified microscopically. As part of the clinical diagnostics, samples were made for cytogenetic analysis. This required 5.0 x 10^4 cells per slide of the positive fraction (maximum 6 slides). Preparation of the slides is described in detail in chapter 2.2.1.3. The remaining $CD138^+$ cells were used for sequence analysis of the present work. Therefore, cell pellets with an optimal cell concentration of 2.0 x 10^6 cells per aliquot in 350 µL GTC-buffer (chapter 2.1.6) were prepared and stored at -80 °C for further analysis. The cell count of the negative fraction was also measured (1:10 dilution Trypan Blue Solution, 0.4 %) and the cell suspension was centrifuged at 1500 rpm for 10 min to prepare pellets with a maximum cell count of 2.0 x 10^7 cells for biobanking.

2.2.1.3 Cytospin preparation

For the detection of chromosomal numerical abnormalities in malignant plasma cells samples for interphase fluorescence in situ hybridization (iFISH) were prepared. The required number of cells (5.0 x 10^4 cells per slide) was centrifuged at 1500 g for 5 min and resuspended in 200 µL cold 1x PBS-buffer (chapter 2.1.6). 100 µL of cell suspension were added per chamber and centrifuged at 600 rpm for 6 min. After centrifugation the slides were dried and cells were fixed with glacial acid (chapter 2.1.6) for 10 min followed by microscopic quality control. The slides were stored at -20 °C for further processing. IFISH diagnostics were performed in the Laboratory of Molecular Cytogenetics (Prof. Dr. sc. hum. Anna Jauch) and were not part of this project.

2.2.2 Purification of total RNA

To purify total RNA from samples derived from $CD138^+$ cell fraction, the "AllPrep DNA/RNA/Protein Mini Kit" (Qiagen, Hilden, Germany) was used according to the manufacturer´s instructions ("Simultaneous Purification of Genomic DNA, Total RNA and Total Protein from Animal and Human Cells"). Due to the sample storage in GTC-buffer (chapter 2.1.6), the first step of adding RLT-buffer with 2-Mercaptoethanol

was not required and the cell lysate was homogenized directly using a "QIAshredder" (Qiagen, Hilden, Germany). The optional centrifugation step of the RNeasy spin column before RNA elution was applied as described in the manufacturer´s protocol. RNA elution was done in two steps by adding 15 µL RNase-free water directly onto the spin column membrane and centrifuged for 1 min at 13000 g two times to achieve a final volume of 30 µL total RNA. The RNA concentration was checked by measuring 2 µL of the sample after initializing the system with 2 µL RNase-free water as blank.

2.2.3 cDNA synthesis

For reverse transcription of total RNA, the "High capacity cDNA reverse transcriptase Kit" (Applied Biosystems™ by Thermo Fisher Scientific, Waltham, MA, USA) with RNase inhibitor was used according to the manufacturer´s protocol. Oligo(dT)18 Primer (Thermo Scientific™, by Thermo Fisher Scientific, Waltham, MA, USA) were used instead of 10 x RT Random Primers (Kit component).

2.2.4 Polymerase chain reaction of the kappa cohort

To amplify the LCs of the kappa patients a primer set was established. First, two primer, a forward (VKKL_A_fw_SaS, Table 6) primer which bind equally well to the main *IGKV* LC family sequences and one reverse primer (CKKL_A_rv_SaS, Table 6) to represent the complete *IGKC*-segment were designed and checked for their melting point, GC content and possible dimers with the "Multiple Oligonucleotide Analyser" (chapter 2.1.5). In a few cases, the sequences could not be clearly captured with the initially primer set. For these cases, a multiplex primer set with two additional forward primer (VKKL2a_fw_Huhn, VKKL3a_fw_Huhn, Table 6) adapted from a previously published method (Huhn 2018) was used. In case of further presence of multiple sequence signals, a PCR with the *IGKV5* family specific singleplex primer VKKL5_A_fw_SAS (Table 6) was done.

The PCR reaction mix was prepared according to the following protocol (Table 9):

Table 9. PCR reaction mix

Reagents per sample	**Volume [µL]**
10x buffer	2.5
$MgCl_2$ (25 mM)	1.5
dNTPs (10 mM each)	0.5
Primer forward (10 mM)	0.5
Primer reverse (10 mM)	0.5
Polymerase[1]	0.125
cDNA from $CD138^+$ cell fraction	1
H_2O	ad 25

[1]AmpliTaq Gold 250 Units, 5 U/µL

Under the following PCR conditions (Table 10):

Table 10. PCR conditions

Step[1]	**Temperature [°C]**	**Time**
1	95	5 min
2	95	30 sec
3	60	45 sec
4	72	1 min
5	72	10 min

[1] step 2 to 4 was repeated 40x

The PCR performance was monitored by gel electrophoresis and for visualization the "Image Lab™Software" (Molecular Imager® Gel DOCTM XR+, Bio-Rad Laboratories, Inc., Hercules, CA, USA) was used.

2.2.5 Polymerase chain reaction of the lambda cohort

The primer used for the lambda LC sequencing are listed in Table 7. For this approach, adapted primer were used (Huhn 2018) and the primer design as indicated in Table 7 was performed by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg). In a first approach a multiplex primer set of four forward primer (VLKL12a_Huhn, VLKL3c_Huhn, VLKL4a_Huhn, VLKL7a_Huhn) located in *IGLV* and one reverse primer (CLKL_A_rv_NB), located in *IGLC* was applied. After the first sequencing showed overlapping signals, a second PCR was performed using an *IGLV* family specific singleplex primer. In this case, VLKL6_A_fw_NB primer was used to examine possible *IGLV6-57* usage, which was not covered by the first primer set. The sequences assigned to *IGLV3* were re-sequenced with the following two primer VLKL3_H_fw_NB and VLKL3_I_fw_NB to improve sequence quality. The PCR was performed under the same conditions as the kappa PCR (Table 9,Table 10).

2.2.6 Purification of polymerase chain reaction product for sanger sequencing

Purification of the PCR products was done using the "High Pure PCR-Product Purification Kit" (Roche, Basel, Switzerland) following the manufacturer´s protocol "Purification of PCR Products in Solution after Amplification", except for the elution step. Elution was performed in two steps with 2x 15 µL elution buffer.

2.2.7 DNA Sanger sequencing

Sanger sequencing of the purified PCR products was performed by the Eurofins GATC service. Sequencing was carried out using cycle sequencing technology, a modification of the traditional Sanger sequencing method (SupremeRun Tube). For this approach, ABI3730XL sequencing machine was used (Genomics 2023). The required concentration as well as amount of forward and reverse primer (10 mM) were sent to Eurofins and both forward and reverse sequencing were performed.

2.2.8 Next-generation sequencing of lambda AL Amyloidosis patients

In order to investigate the AL lambda *IGLV3-21*-subfamily in more detail, a next-generation sequencing (NGS) analysis was performed. For this purpose, an evaluation pipeline based on MiXCR (Table 8) was established in this project and the sequence characteristics of two lambda AL_K and two AL_D patients were analysed.

2.2.8.1 Experimental setup

NGS was done on the NextSeq 550 System (Ilumina, San Diego, CA, USA) according to the Illumina "DNA Prep Reference Guide" (document # 1000000025416 v09 June 2020), "NextSeq System Denature and Dilute Libraries Guide" (document # 15048776 v09 December 2018), "NextSeq 550 System Guide" (document # 15069765 v06 June 2019) and "Local Run Manager Generate FASTQ Analysis Module Workflow Guide" (document # 1000000003344 v02 July 2018). The following kits were used: IDT® for Illumina® DNA/RNA UD Indexes Set B, (96 Indexes), Illumina® DNA Prep, (M) tagmentation (24 samples), NextSeq 500/550 Mid Output Kit v2.5 (150 cycles) and NextSeq PhiX Control Kit (Illumina, San Diego, CA, USA).

Figure 8 shows the NGS workflow:

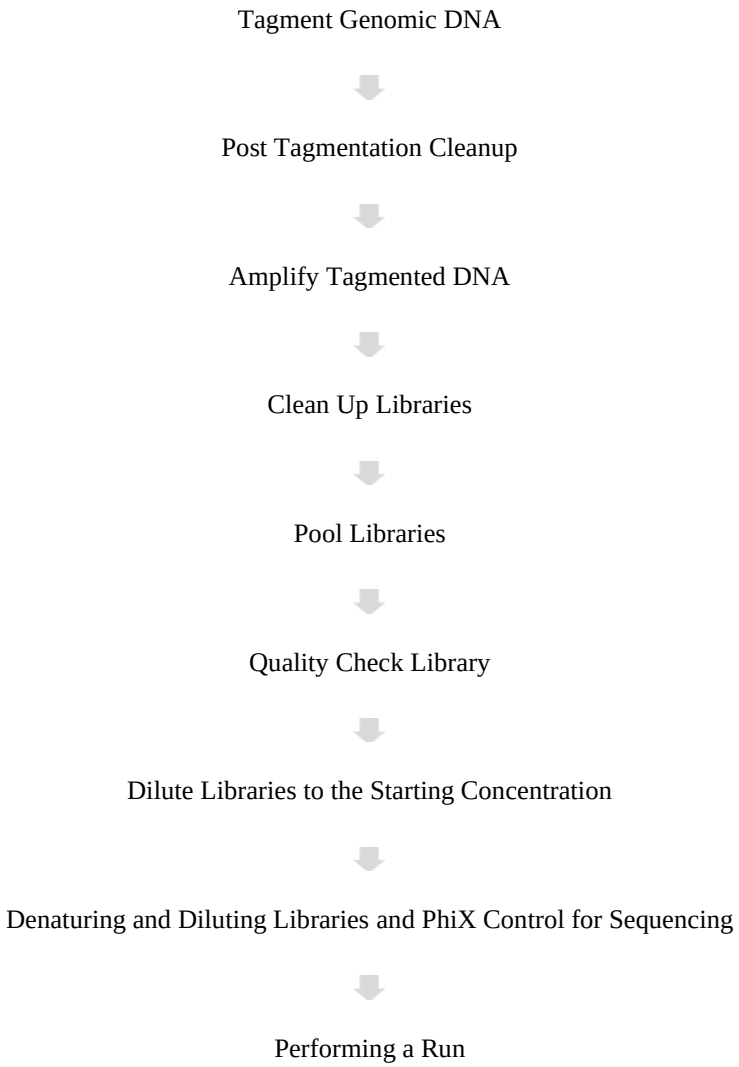

Figure 8. Next-generation sequencing workflow.

Quality control was done with a 4200 TapeStation System (Agilent, Santa Clara, CA, USA) following the manufacture´s protocol "Agilent D1000 ScreenTape System Quick Guide". The measurement was done with D1000 ScreenTape with D1000 Reagents (Agilent, Santa Clara, CA, USA) as well as the Agilent Software packages (2200 TapeStation Controller Software, and TapeStation Analysis Software). As an additional method, the Quantus™ Flourometer with QuantiFluor® ONE dsDNA System (Promega, Madison, WIS, USA) was applied for measurement of the quality and product concentration according to the manufacturer's instructions. Library molarity was calculated using the formula provided in the manufacturer's protocol and libraries were diluted to a concentration of 2 nM in 15 µL RBS buffer. The libraries and PhiX was diluted to a starting concentration of 1.4 pM.

2.2.8.2 Establishment of a next-generation sequencing evaluation pipeline

During this project, Sanger sequencing of the *IGLV3-21* sequences repeatedly showed difficulties in resolving sequences specifically assigned to the *IGKV3-21* germline. To address this problem, the sequences in question were analysed in more detail using NGS. The NGS application, in combination with an appropriate bioinformatics algorithm, offers the possibility of resolving individual partial sequences. Within the framework of this thesis, an evaluation pipeline was established to extract the raw data output of NGS. For this purpose, MiXCR (Table 8) was used. MiXCR is described as a useful tool to process large immune data from raw sequence data. Thereby, it considers sequence quality and performs PCR error correction and follows the shown pipeline (Figure 9) (Bolotin et al. 2015).

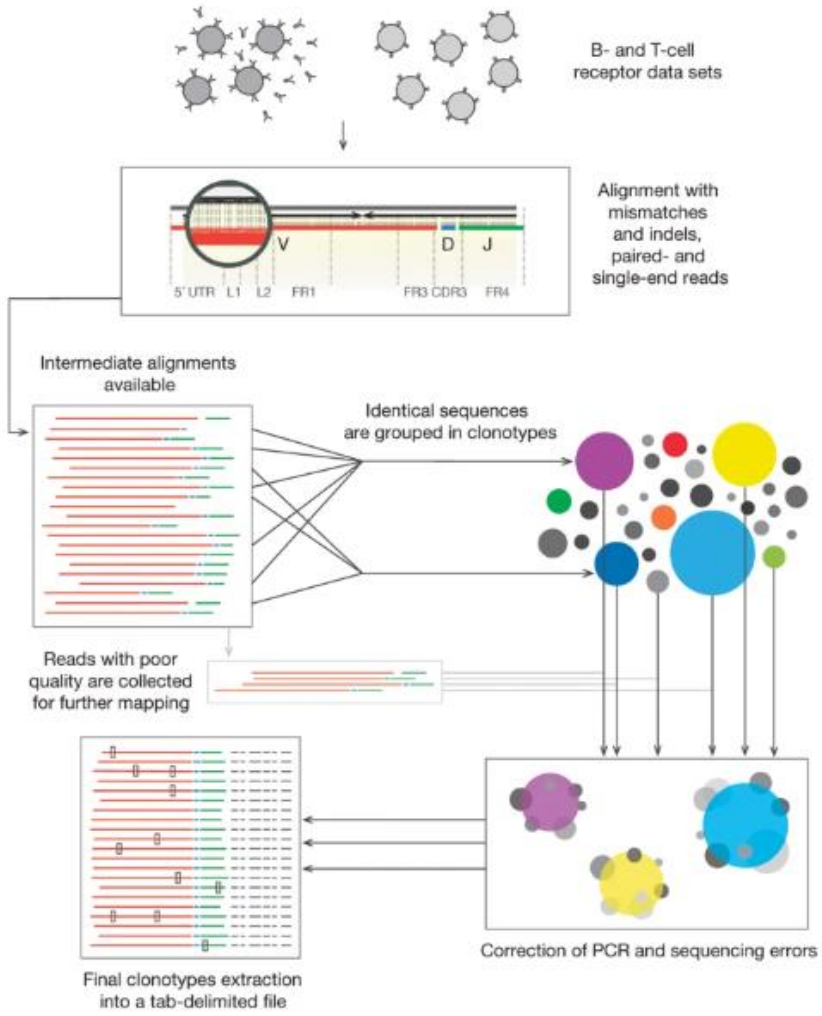

Figure 9. MiXCR pipeline. (Bolotin et al. 2015) Reproduced with permission from Springer Nature (https://www.nature.com/articles/nmeth.3364).

MiXCR version 3.0.13 was installed following the manual installation instructions (Bolotin et al. 2015) and the fastq.gz files from all lines were extracted with the following command:

```
$ gunzip *.gz
```

As the evaluation programme only allows the use of one file, all contents have been combined into one file using the following command:

```
$ cat *.fastq > sum.fastq
```

The following commands were used to execute MiXCR:

```
$ java -jar mixcr.jar analyse shotgun --verbose --species
hs --starting-material dna --receptor-type bcr --contig-
assembly sum.fastq analysis
```

2.2.9 Bioinformatic analysis

2.2.9.1 Generation of a final sequence from Sanger sequencing

The Ab1-files from forward and reverse Sanger sequencing were analysed with MEGA - Molecular Evolutionary Genetic Analysis- software (Table 8) and initially checked for inconsistencies and sequence ambiguities. Heterozygous signals without a clear dominant major signal were defined using the IUPAC-IUB code (Figure 10) (NC-IUB 1985).

Symbol	Base
W	A or T
S	C or G
M	A or C
K	G or T
R	A or G
Y	C or T
B	C, G or T
D	A, G or T
H	A, C or T
V	A, C or G
N	any base

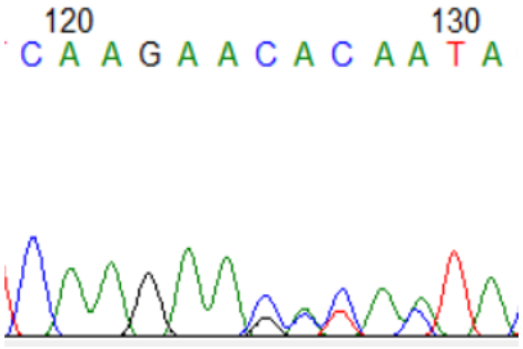

Figure 10. Management of heterozygous signals. Heterozygous signals were defined according to the IUPAC-IUB code (NC-IUB 1985).

A final sequence was generated based on forward and reverse sequencing. All final DNA sequences of the LCs are attached in the appendix (see page 152 ff.).

2.2.9.2 *IGK/L* family assignment and amino acid translation

To assign the *IGK/LV*- and *IGK/LJ*-gene families Vbase2 was used (Table 8, (Retter et al. 2005)). Translation into an amino acid (AA) sequence was performed with Expasy

(Table 8, (Duvaud et al. 2021)) and Ensembl BLAST (Table 8, (Howe et al. 2021)) was used for *IGK/LC* assignment. The *IGK/LV*-assignment was confirmed with Ensembl Blast and in case of discrepancies or if no determination was possible because of too many heterozygous signals, the sequence was defined as not evaluable. All final translated AA sequences are attached in the appendix (see page 152 ff.).

2.2.9.3 Mutation analysis

For identification of patient-specific variants, each sequence was aligned with the identified *IGK/LV*- and *IGK/LJ* specified reference sequence using Clustal Omega (Table 8). For *IGK/LV*, *IGK/LJ* and *IGK/LC* the Vbase2 reference, IMGT reference (Table 8) and the Ensembl reference (Table 8) were aligned, respectively. In addition to the Vbase2 reference, the Ensembl reference was also used for the *IGK/LV* regions. If these differed at one position and the patient sequence had one of these AAs, it was not considered as a mutation. Due to the unspecified AA at the first position of the *IGK/LC* reference, the UniProt reference (Table 8) was used at this point. In the case of lambda analysis, *IGLJ2*01* (M15641) and *IGLJ3*01* (M15642) shared the same reference sequence in IMGT, consequently they were defined as *IGLJ2* and *IGLJ3*02* (D87023) as *IGLJ3*.

2.2.9.4 Analysis of the amino acid composition and biophysical parameters

To achieve comparability despite the different sequence lengths resulting from the different sequencing methods, the AL and MM sequences were shortened to the same length at the N- and C-terminus. For ambiguous positions, the AA of the reference was assumed. If the Vbase2 reference and the Ensembl reference of the *IGK/LV* regions showed two different AA at the relevant position, the AA of the Vbase2 reference was taken. Mutation hotspots were defined as positions that were mutated to or more than 50 % as compared to the respective reference sequence. To calculate the median mutation count, only the mutated regions and segments were considered. The Expasy Tool: Compute pI/MW was used to evaluate the isoelectric point (pI) (Table 8, (Bjellqvist et al. 1994; Bjellqvist et al. 1993; Gasteiger et al. 2005)). For calculation of the AA composition and grand average of hydropathicity (GRAVY)-score the ProtParam-Tool from Expasy was used (Table 8, (Gasteiger et al. 2005)). The theoretical beta sheet aggregation tendency (AGG) score of the individual AA sequences was calculated using the prediction tool TANGO (Table 8, (Fernandez-Escamilla et al. 2004; Linding et al. 2004;

Rousseau et al. 2006)). The analysis of a potential N-glycosylation site was done by using the bioinformatic tool NetNGlyc 1.0 (Julenius 2007) with default settings N-Xaa-S/T. The analysis was performed on the respective AL or MM representatives of the different *IGK/LV*-subfamilies, furthermore the effect of the presence of a heavy chain (HC) of the antibody in the serum was studied.

2.2.10 Analysis of the MM cohort

The MM data were not generated as part of this project. The data were based on a bulk RNA-sequencing approach. Preparation of sequencing libraries were done with Illumina TruSeq stranded mRNA kit (San Diego, CA, USA) on the Illumina NovaSeq 6000 PE 100 S1 platform. Raw data processing was done by Dr. rer. nat Alexandra Poos using MIXCR (option: conting-assembly) and provided with the permission of the GMMG study group. $CD138^+$ cells were also enriched (Milteny) and RNA extraction was performed using the same methods as in this project (chapter 2.2.1).

2.2.10.1 Verification of the bulk RNA sequencing data

Given that the sequences from the AL and MM cohort were produced by different methods, randomly chosen kappa MM samples from the bulk RNA sequencing set representing all five major *IGKV*- family members were analysed with the same protocol used for the AL sequences for validation purposes (see appendix SI Figure1).

2.2.11 Statistics and data visualisation

Data analysis and visualization was performed using Excel as well as R x64 4.1.3/R-studio 2022.02.0+443 with ggplot packages. P-values ≤ 0.05 were considered to be significant and indicated at appropriate position. P-value correction was performed to exclude false positives results in case of subgroup-analysis. Statistical analysis was carried out using IBM SPSS Statistics, version 29.0.0.0. First, the data were tested for normal distribution with Shapiro-Wilk test. In case of normal distribution, the T-test was applied, if not, the non-parametric Mann-Whitney U-test was used. Nominal data were tested with Fisher´s exact test. Two-sided testing was applied in all cases. For calculation of the relative risk (rr) the MedCalc Software Ltd. Relative risk calculator was used.

3 Results

This chapter is divided into two parts. The analysis of the kappa cohort is the main focus of this thesis. Here, the entire AL patient collective visiting the Amyloidosis Centre Heidelberg between January 2019 and May 2022 was analysed and compared with MM patients of the GMMG HD6 clinical trial (Goldschmidt et al. 2021; Salwender et al. 2019). Initially, the patients of both cohorts were clinically characterized. Subsequently, the LC sequences of the AL and MM patients were analysed for their germline and family usage. In this context, the *IGKV*-, *IGKJ*- and *IGKC*- family assignments were determined. To identify patient-specific variants a detailed mutation analysis was performed. Next, the AA composition of the LC, categorized according to their different *IGK*-families, as well as the analysis of various biophysical parameters in order to draw potential conclusions about the aggregation behaviour was done. Thereby, the pI, as an indicator for the protein charge, the GRAVY- score, as an indicator for the protein hydrophobicity and the AGG-score, as indicator for β-sheet aggregation was analysed. Considering glycosylation is a potential cause of LC aggregation, bioinformatical analysis of the protein sequence for glycosylation was performed. In addition to the comparison of the AL with the MM LC sequences, a subgroup analysis was performed for the presence of a HC in the blood. Further, the novel V_LAmY-Pred-tool, which is designed to predict the aggregation of amyloidogenic IG sequences based on machine learning, was tested with the LC sequences investigated in this study. In the second part of this chapter, the LC sequences of lambda amyloidosis patients with kidney involvement (AL_K) or diverse organ involvement (AL_D) was analysed with respect to their family usage. This was followed by a detailed mutation analysis and an analysis of the biophysical properties of the most common AL subfamily in comparison with the most common MM subfamily. Further an NGS-analysis was performed to study specific sequence characteristics.

3.1 Kappa patient cohort

Parts of these results have already been published in Schreiner et al. (2024).

3.1.1 Clinical characteristics of the kappa patient cohort

In this project 41 kappa AL patients and 83 kappa MM patients were included. The AL patients were classified in terms of their organ involvement, resulting in 14 AL_K, eight AL_H and 19 AL_D patients. Clinical data of the AL patients were obtained from physician report and laboratory results. Parameters such as age, sex, the classification of the affected organs were evaluated by the attending physician. The laboratory parameters as AP, quantity of free LCs, or M-gradient were determined by the central laboratory (University Hospital Heidelberg). Plasma cell infiltration was extracted from the bone marrow report ("KM-Labor MED V", University Hospital Heidelberg). AL patients were classified using established variables and cut-offs according to the standardized staging models (Dispenzieri et al. 2004; Palladini et al. 2014; Wechalekar et al. 2013). The clinical data of the MM patients were provided by the GMMG study group as documented in the electronical patients case report forms (eCRF) of the trial. Table 11 summarizes the clinical data of the AL patients categorized by organ involvement, as well as of the MM patients. Both patient cohorts had a similar range of age at diagnosis. The median age ranged from 61 to 63 years. More males than females were affected in AL_K (n= 10 vs. 4), AL_D (n= 13 vs. 6) and MM (n= 53 vs. 30). The AL_H subgroup was represented by equal ratios of men and women (n= 4 vs. 4). Bone marrow were obtained at the time of initial diagnosis in all AL_K and MM patients, in two cases in AL_H and in three cases in AL_D at time of relapse. Besides the most frequently affected organs in AL, the heart and kidneys, also the liver was affected in 11 cases in the AL cohort of 41 patients, including nine in the AL_D patient group. As expected (Dittrich et al. 2017), the AL_K patients showed a lower dFLC than the AL_H and AL_D patients (median dFLC: AL_K= 97 mg/L vs. AL_H= 543 mg/L, $p< 0.001$; AL_K= 97 mg/L vs. AL_D= 472 mg/L, $p< 0.001$). Also, the dFLC of the AL_K patients was significantly lower compared to the MM patients (median dFLC: AL_K= 97 mg/L vs. 349 mg/L, $p= 0.009$). Within the AL group, AL_K patients had the lowest AP (AP: AL_K= 74 U/L, AL_H= 125 U/L, AL_D= 120 U/L) and compared with MM, the AP of AL_D was significantly higher than of the MM patients (AP: MM= 70 U/L, $p< 0.001$). A clonal heavy chain in serum was detected in more MM patients than AL_K and AL_D patients (HC: MM= 73/83 vs. AL_K= 8/14, $p= 0.033$; MM= 73/83 vs. AL_D= 9/19, $p<0.001$) More precisely, AL_K patients showed IgG exclusively (8/14), whereas AL_D patients had in one case IgA (1/19) and in eight cases IgG (8/19). For MM patients, IgM and IgD were detected in one

case (each 1/83) and in 12 cases IgA (12/83) and in 59 cases IgG (59/83). AL_H displayed IgG and IgA in equal parts (each 2/8). A t(11;14) known to be associated with a loss of a heavy chain (Bochtler et al. 2008) was detected in 4/8 AL_H, 5/13 AL_K and 12/15 AL_D patients. In the MM cohort, 16/82 had a t(11;14). The iFISH analysis further indicated a 1q21 in 2/8 AL_H, 3/13 AL_K, 3/15 AL_D and 8/82 MM patients.

Table 11. Clinical characteristics of 41 kappa AL amyloidosis and 83 kappa multiple myeloma patients.(Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

	AL_H (n= 8)	**AL_K (n= 14)**	**AL_D (n= 19)**	**MM (n= 83)**
Age [y], median (range)	62 (52-80)	61 (34-69)	63 (39-78)	61 (42-70)
Sex female/male [n]	4/4	4/10	6/13	30/53
Newly diagnosed [n]	6/8	14/14	16/19	83/83
Relapse/progress [n]	2/8	0/14	3/19	0/83
AL and concomitant MM	3/8	0/14	5/19	
No. of organs involved, no. of patients				
1	2	11	2	
2	5	1	5	
3	1	2	7	
4	0	0	2	
More than 4	0	0	3	
different organs involved [n]				
Heart	8/8	0/14	16/19	
Kidney	0/8	14/14	5/19	
Liver	0/8	2/14	9/19	
ANS	1/8	0/14	5/19	
PNS	0/8	0/14	4/19	
GI	2/8	2/14	7/19	
ST	4/8	1/14	9/19	
dFLC [mg/L], median (range)[I]	543 (207-1891)	97 (8-649)	472 (54-3784)	349 (0-30124)
FLC ratio >100	3/8	0/8	4/19	
Cardiac stage[1]				
I	0/8	7/13	2/18	
II	1/8	3/13	6/18	
IIIa	3/8	2/13	7/18	
IIIb	4/8	1/13	3/18	
Renal stage				
I	5/7	2/14	15/18	
II	2/7	8/14	3/18	
III	0/7	4/14	0/18	
AP [U/L] median (range)[2II]	125 (52-199)	74 (41-240)	120 (45-2137)	70 (28-349)[3]
M-gradient [n][III]	5/8	8/14	8/19	74/83
Plasma cell infiltration [%] (range)[IV]	12 (8-43)	10 (5-20)	13 (2-67)	42 (2-100)
Monoclonal protein in the serum [n][V]	4/8	8/14	9/19	73/83
IgG	2/8	8/14	8/19	59/83

	AL_H (n= 8)	AL_K (n= 14)	AL_D (n= 19)	MM (n= 83)
IgA	2/8	0/14	1/19	12/83
IgM	0/8	0/14	0/19	1/83
IgD	0/8	0/14	0/19	1/83
Light chain only	4/8	6/14	10/19	10/83
t(11;14) [n]	4/8	5/13	12/15	16/82
1q21 [n]	2/8	3/13	3/15	8/82

AL= AL amyloidosis, MM= multiple myeloma, AL_H= AL patients with dominant heart involvement, AL_K= AL patients with dominant kidney involvement, AL_D= AL patients with diverse organ involvement, ANS= autonomic nervous system, PNS= peripheral nervous system, GI= gastrointestinal tract, ST= soft tissue, dFLC = difference between disease-associated und uninvolved circulation free light chains, AP = alkaline phosphatase, cardiac and renal staging was done according to (Dispenzieri et al. 2004; Palladini et al. 2014; Wechalekar et al. 2013), n/a= not available

[1] EURO score,[2] Reference range obtained from the laboratory results: AP 40-130 U/L, [3]4 n/a

[I] AL_H vs. AL_K p< 0.001, AL_K vs. AL_D p< 0.001, AL_K vs. MM p< 0.001

[II] AL_D vs. MM p< 0.001

[III] AL_K vs. MM p= 0.021, AL_D vs. MM p< 0.001

[IV] AL_K vs. MM p< 0.001, AL_H vs. AL_D p< 0.001, AL_H vs. MM p< 0.001, AL_D vs. MM p< 0.001

[V] AL_K vs. MM p= 0.033, AL_D vs. MM p< 0.001

3.1.2 Analysis of the kappa multiple myeloma sequence composition

MM sequences were generated using a bulk RNA-sequencing approach. Due to the bioinformatics evaluation pipeline, different sequence segments can be attributed to sequence variations from individual positions. In the following, identical sequences were grouped together and their percentage occurrence in relation to the total sum of the different sequence variations was summarised, which allows indications of possible sub clonal sequences. The most frequent cohort was referred to as sequence fraction_1 and the second most frequent cohort was referred to as sequence fraction_2. Only sequences with a sequence fraction size >1 % were analysed (Table 12). The highest *IGKV* representative, *IGKV1* (n= 44) with a median sequence fraction_1 of 99.52 % had seven sub-sequences_2 with a sequence fraction >1 %. Including one sequence, assigned to *IGKV1/D-39* and *IGKJ2*, with a sequence fraction_2 of 42.11 % (vs. 57.47 % sequence fraction_1). Detailed sequence analysis revealed that the two sequences differed only in the length. Sequence_2 was N-terminally four and C-terminally six nucleotides longer than sequence_1. *IGKV2* (n= 8) with a median sequence fraction_1 of 99.01 % showed two sub-sequences_2 with a sequence fraction >1 %. However, with a sequence fraction_2 of 2.07 % and 5.65 %, the occurrence was comparatively low. One further major *IGKV* representative for MM, *IGKV3* (n= 25) (median sequence fraction_1= 99.31 %) displayed one sequence, assigned to *IGKV3/D-11* and *IGKJ1* with a sequence fraction_1 of 64.92 %. Overall, this group had only one sequence with a fraction_2 about

1.05 %. *IGKV4* (n= 5) and *IGKV5* (n= 1) showed no sub-sequences_2 >1 %. For further analysis, only sequences of sequence fraction_1 were used.

Table 12. Multiple myeloma sequence composition of bulk RNA-sequencing. Analysis of the sequences from kappa multiple myeloma cohort (n=83) with a sequence fraction >1 %.

	n sequences	Median sequence fraction_1 [%] (range)	n sub-sequences_2 >1 %	Median sequence fraction_2 [%] (range)
IGKV1	44	99.52 (57.47 - 99.95)	7	4.3 (1.11 - 42.11)
IGKV2	8	99.01 (91.47 - 99.69)	2	3.86 (2.07 - 5.65)
IGKV3	25	99.31 (64.92 - 99.96)	1	1.05
IGKV4	5	99.49 (97.67 - 99.92)	0	n/a
IGKV5	1	99.93	0	n/a

n/a= not available

3.1.3 *IGK* family usage and organ tropism of the kappa cohort

3.1.3.1 *IGKV* region in kappa AL amyloidosis and multiple myeloma

AL and MM sequences were assigned to the respective *IGKV*- family using Vbase2 based on the cDNA level and family assignment was confirmed using the AA sequences and Ensembl Blast. Figure 11 shows the family assignment of the AL and MM kappa patient cohort.

Representatives of the *IGKV1* family were highly prevalent in AL as well as MM, with higher frequency in AL (AL= 80 %, MM= 53 %, p= 0.002) (Figure 11A). *IGKV2* was only detectable in MM (10 %). *IGKV3* was underrepresented in AL (AL= 10 %, MM= 30 %, p= 0.014). 5 % of the sequences from patients with AL and 6 % from MM were assigned to *IGKV4*. *IGKV5* was identified in 2 % of AL and in 1 % of MM cases. In one (2 %) of the 41 AL patients, the *IGKV* gene family could not be clearly assigned and was defined as not evaluable.

Regarding the *IGKV*-subgroups in AL and MM (Figure 11B), significantly differences could be found for *IGKV1/D-33*. *IGKV1/D-33* was more common in AL (32 % vs. 13 %, p= 0.016). Also, differences in the *IGKV1-16* expression could be identified with detection of 10 % in the AL cohort but not in the MM (p= 0.010). Within the *IGKV1*-subgroups, *IGKV1-5* (AL= 12 %, MM= 17 %) and *IGKV1/D-39* (AL= 20 %, MM= 17 %) were highly represented in both cohorts. *IGKV3/D-11* was higher in MM (14 %) than in

AL (2 %). Analysing the two cohorts for further presence of *IGKV*-subgroups exclusive to one of the two diseases, *IGKV1/D-13* (2 %) was detected only in the AL cohort. In contrast, *IGKV1-6* (1%), *IGKV2/D-28* (5 %), *IGKV2/D-30* (4 %), *IGKV2/D-40* (1 %), *IGKV3/D-20* (11 %, p= 0.030) were detected exclusively in MM. *IGKV1-9* (AL: 2 %, MM: 2 %), *IGKV1-12* (AL: 2 %, MM: 2 %), *IGKV3-15* (AL: 7 %, MM: 5 %), *IGKV4-1* (AL: 5 %, MM: 6 %) and *IGKV5-2* (AL: 2 %, MM: 1 %) was less prominent in both cohorts and did not differ in abundance. Based on the frequency, the relative risk (rr), or disease risk, was also calculated. This calculation resulted in a rr for AL of 2.452 (p= 0.013) for the *IGKV1/D-33*-subfamily and a rr of 18.439 (p= 0.049) for *IGKV1-16*. *IGKV3/D-11* had a rr of 0.173, but without statistical significance. Further data on the rr can be obtained from SI Table 1 (see appendix).

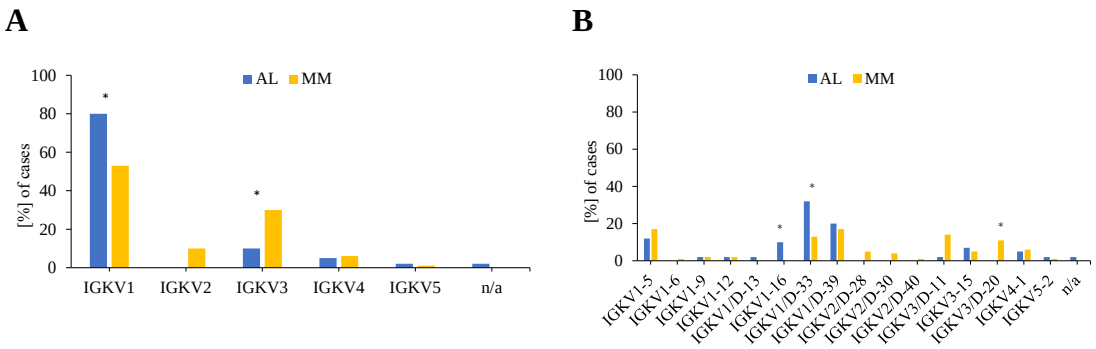

Figure 11. Variable region gene usage in kappa AL amyloidosis and multiple myeloma patients. AL= AL amyloidosis, MM= multiple myeloma, AL n= 41, MM n= 83, *p-values ≤ 0.05 were considered to be significant, n/a= not available **A** *IGKV*-family usage in AL and MM, **B** *IGKV*-subgroups in AL and MM. (Schreiner et al. 2024) This figure was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

Next, the influence of the organ tropism and the *IGKV*-family assignment was studied (Figure 12).

IGKV1 was highly represented in all three groups of AL patients, AL_K (64 %), AL_H (88 %) and AL_D (89 %) (Figure 12A). IGKV3 was detected in 21 % of AL_K patients and in 13 % of AL_H patients; no AL_D patient was assigned to *IGKV3*. *IGKV4* was only detected in AL patients with diverse organ involvement (11 %) and *IGKV5* only in AL_K patients (7 %). One AL_K patient (7 %) could not be assigned to any *IGKV* family.

Concerning organ tropism in AL and *IGKV*-subgroups, differences in *IGKV1*-subgroups were observed (Figure 12B). *IGKV1/D-33* was predominantly detected in patients with heart involvement (6/8, 75 %, p= 0.024), while *IGKV1/D-39* was more frequently detected in patients with dominant kidney involvement (5/14, 36 %). When analysing AL_D patients, no clear predominant *IGKV1*-subgroup could be identified. *IGKV1-5* was detected in 21 %, *IGKV1-12* in 5 %, *IGKV1/D-13* in 5 %, *IGKV1-16* in 16 %, *IGKV1/D-33* in 26 % and *IGKV1/D-39* in 16 %, respectively. Regarding *IGKV3*, *IGKV3/D-11* was identified in 7 % of AL_K patients, whereas *IGKV3-15* was identified in 14 % of AL_K patients and 13 % of AL_H patients. 11 % of AL_D had *IGKV4-1*. The 7 % of AL_K patients with *IGKV5* were assigned to the *IGKV5-2*-subgroup.

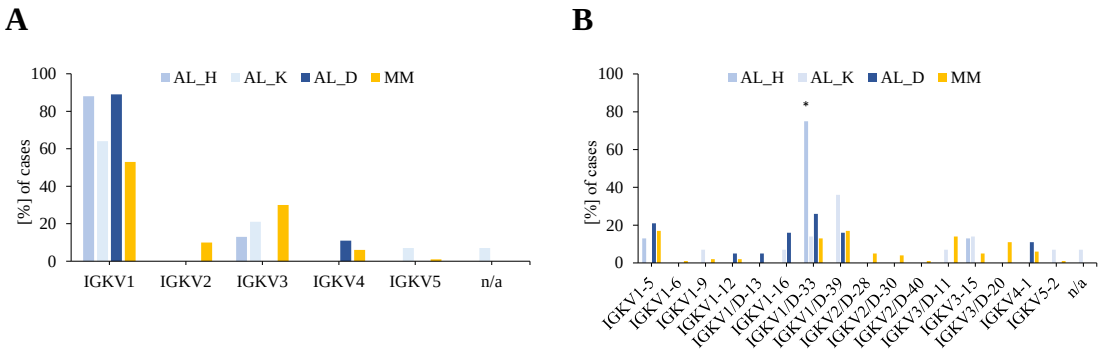

Figure 12. Variable region gene usage in kappa AL amyloidosis with different clinical presentation and multiple myeloma patients. AL= AL amyloidosis, MM= multiple myeloma, AL n= 41, MM n= 83, AL_H= AL patients with dominant heart involvement, AL_K= AL patients with dominant kidney involvement, AL_D= AL patients with diverse organ involvement, AL_H n= 8, AL_K n= 14, AL_D n= 19, *p-values ≤ 0.05 were considered to be significant, n/a= not available **A** *IGKV*-family usage in AL (organ tropism) and MM, **B** *IGKV*-subgroups in AL (organ tropism) and MM. (Schreiner et al. 2024) This figure was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

The effect of the presence or absence of a Ig HC in serum on the *IGKV* usage of the most prevalent subgroups *IGKV1-5*, *IGKV1/D-33* and *IGKV1/D-39* in AL as well as MM was also analysed (Figure 13). Twenty of the kappa AL patients (vs. AL_HC n= 21) and ten of the MM patients (vs. MM_HC n= 73) had no HC. In both AL and MM, a trend towards *IGKV1-5* was observed (Figure 13A). In kappa AL, patients without a detectable HC in serum were four times more frequent than patients with a detectable HC (AL_noHC= 20 %, AL_HC= 5 %). A threefold difference was found in kappa MM (MM_noHC= 40 %, MM_HC= 14 %). When analysing the organ tropism in AL, the

general differences between HC and no_HC was driven by the AL_H and AL_D patients (Figure 13B).

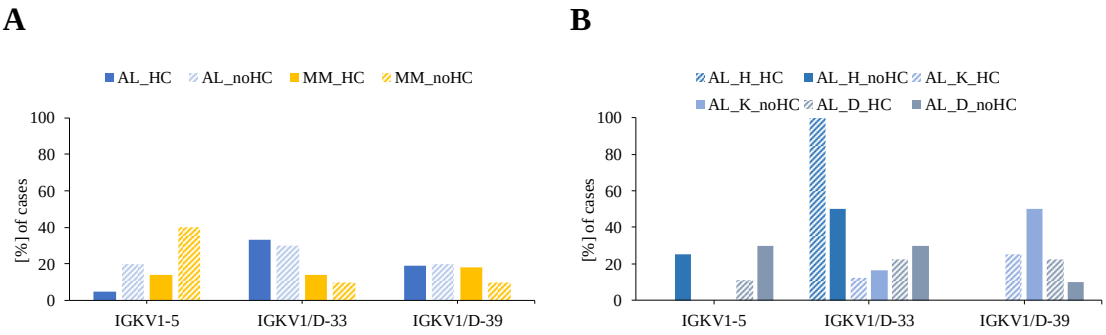

Figure 13. Influence of the presence of a heavy chain in serum on the most frequent *IGKV*-subgroups. AL= AL amyloidosis, MM= multiple myeloma, AL n= 41, MM n= 83, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 21, AL_noHC n=20, MM_HC n= 73, MM_noHC n= 10, AL_H= AL patients with dominant heart involvement, AL_K= AL patients with dominant kidney involvement, AL_D= diverse AL patients, AL_H n= 8, AL_K n= 14, AL_D n= 19, **A** Comparison of the most abundant *IGKV*-subgroups with respect to a heavy chain in serum in AL and MM, **B** Influence of organ tropism on the expression of the most abundant *IGKV*-subgroups with respect to a heavy chain in serum. (Schreiner et al. 2024) This figure was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

Kappa AL, unlike lambda, has been reported to be associated with liver involvement (Comenzo et al. 2001). Hence, AL patients and liver involvement was analysed with respect to their *IGKV* assignment separately (Table 13). In the kappa AL cohort 11 of 41 patients had liver involvement and nine of those were in the AL diverse group. Analysis of their *IGKV*-subfamily usage showed the highest difference in *IGKV1-16*. 27 % of the AL with liver involvement were assigned to *IGKV1-16*, whereas only 3 % of patients without liver involvement did. *IGKV1-12* and *IGKV1/D-13* were only be detected in AL patients with liver involvement (9 %). *IGKV1-5* was only detected in AL patients without liver involvement (17 %). No differences were found for patients with and without liver involvement assigned to *IGKV1/D-33* and *IGKV1/D-39* (*IGKV1/D-33*= 27 % vs. 33 %, *IGKV1/D-39*= 18 % vs. 20 %).

Table 13. Correlation of *IGKV* gene usage and liver involvement in kappa AL amyloidosis patients. (Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

IGKV [%]	AL_liver involvement (n= 11)	AL_without liver involvement (n= 30)
V1-5	0	17
V1-9	0	3
V1-12	9	0
V1/D-13	9	0
V1-16	27	3
V1/D-33	27	33
V1/D-39	18	20
V3/D-11	0	3
V3-15	0	10
V4-1	9	3
V5-2	0	3
n/a	0	3

AL= AL amyloidosis, n/a= not available

3.1.3.2 *IGKJ* region in kappa AL amyloidosis and kappa multiple myeloma

Once the *IGKV* gene usage was identified, the corresponding joining regions of the AL, including organ tropism, and of the MM cohort were examined (Table 14). All in all, *IGKJ2* was most commonly observed in the AL cohort (39 %). In MM, both *IGKJ1* (30 %) and *IGKJ2* (29 %) were highly frequent. Furthermore, the expression of *IGKJ3* and *IGKJ5* differed between the two diseases. The expression of *IGKJ3* was higher in MM (13 %) than in the AL cohort (5 %). *IGKJ5* was twice as frequent in AL (10 %) than in MM (5 %). *IGKJ1* was associated with 30 % of MM and 17 % of AL patients.

In terms of organ involvement in AL (Table 14), cardiac involvement was primarily related to *IGKJ2* (63 %), while in AL_K patients *IGKJ1*, *IGKJ2* and *IGKJ3* were equally presented (29 %). AL_D patients were mainly associated with *IGKJ2* and *IGKJ4* (37 %).

Table 14. Analysis of the *IGKJ* usage in kappa AL amyloidosis and multiple myeloma. (Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

	IGKJ1	*IGKJ2*	*IGKJ3*	*IGKJ4*	*IGKJ5*	n/a
AL_H [%] (n= 8)	13	63	13	0	13	0
AL_K [%] (n= 14)	29	29	0	29	7	7
AL_D [%] (n= 19)	11	37	5	37	11	0
AL, whole cohort [%] (n= 41)	17	39	5	27	10	2
MM [%] (n= 83)	30	29	13	23	5	0

AL= AL amyloidosis, MM= multiple myeloma, AL_H= AL patients with dominant heart involvement, AL_K= AL patients with dominant kidney involvement, AL_D= diverse AL patients, n/a= not available

3.1.3.3 Linkage between the most frequent *IGKV*-subfamilies - and *IGKJ* in kappa AL amyloidosis and multiple myeloma

Analysis of the association of the most frequent *IGKV*-subgroups with *IGKJ* (Table 15) showed that *IGKV1-5* was most frequently associated with *IGKJ2* in both AL (10 %) and MM (8 %). 6 % of the MM patients were assigned to *IGKV1-5* connected with *IGKJ1*, whereas none of the AL patients were. The two patient cohorts did not differ regarding *IGKV1-5* and *IGKJ4* (AL: 2 %, MM: 1 %) and *IGKJ5* (AL: 0 %, MM: 1 %) usage. In no case *IGKV1-5* was linked to *IGKJ3*. The highest representative in AL, *IGKV1/D-33*, was also mainly related to *IGKJ2* (17 %). In MM, *IGKV1/D-33* was equally linked to *IGKJ2* (5 %) and *IGKJ4* (5 %) and only 1 % were associated with *IGKJ1* and 2 % with *IGKJ3*. No association of *IGKV1/D-33* and *IGKJ5* was shown in MM. In AL, 7 % of patients assigned to *IGKV1/D-33* were associated with *IGKJ4*. *IGKJ1*, *IGKJ3* and *IGKJ5* were equally associated (2 %). For *IGKV1/D-39*, no major linkage to *IGKJ* was found in AL patients. 7 % of AL patients assigned to *IGKV1/D-39* were associated with *IGKJ1* and 5 % with *IGKJ2* and *IGKJ4*, respectively. 2 % were assigned to *IGKJ5* and none to *IGKJ3*. 6 % of MM patients assigned to *IGKV1/D-39* were linked to *IGKJ2* and 4 % to *IGKJ1*. *IGKJ3*, *IGKJ4* and *IGKJ5* were associated with *IGKV1/D-39* in 2 % of the MM cohort.

With regard to organ involvement in AL (Table 15), the group of *IGKV1-5* associated with *IGKJ2*, included patients with dominant heart involvement (13 %) and patients with diverse organ involvement (16 %). 5 % of AL_D patients were assigned to *IGKV1-5* and *IGKJ4*. *IGKV1/D-33* associated with *IGKJ2* was mainly related to patients with heart involvement (50 %), while patients with kidney involvement and patients with diverse organ involvement showed a comparable association with *IGKJ2* and *IGKJ4* (AL_K= 7 %, AL_D= 11 %). All AL patients assigned to *IGKV1/D-39* and linked to *IGKJ2* had kidney involvement (14 %). Whereas 14 % of the remaining AL_K and 5 % of AL_D patients were associated to *IGKJ1* and 7 % of AL_K and 5 % of AL_D patients were linked to *IGKJ4*.

Table 15. Analysis of the IGK*V*/*J* usage of the most common *IGKV*-subfamilies in kappa AL and MM. (Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

[%]	*IGKJ1*	*IGKJ2*	*IGKJ3*	*IGKJ4*	*IGKJ5*	n/a
IGKV1-5						
AL_H (n= 8)	0	13	0	0	0	0
AL_K (n= 14)	0	0	0	0	0	7
AL_D (n= 19)	0	16	0	5	0	0
AL, whole cohort (n= 41)	0	10	0	2	0	2
MM (n= 83)	6	8	0	1	1	0
IGKV1/D-33						
AL_H (n= 8)	13	50	13	0	0	0
AL_K (n= 14)	0	7	0	7	0	7
AL_D (n= 19)	0	11	0	11	5	0
AL, whole cohort (n= 41)	2	17	2	7	2	2
MM (n= 83)	1	5	2	5	0	0
IGKV1/D-39						
AL_H (n= 8)	0	0	0	0	0	0
AL_K (n= 14)	14	14	0	7	0	7
AL_D (n= 19)	5	0	0	5	5	0

[%]	*IGKJ1*	*IGKJ2*	*IGKJ3*	*IGKJ4*	*IGKJ5*	n/a
AL, whole cohort (n= 41)	7	5	0	5	2	2
MM (n= 83)	4	6	2	2	2	0

AL= AL amyloidosis, MM= multiple myeloma, AL_H= AL patients with dominant heart involvement, AL_K= AL patients with dominant kidney involvement, AL_D= diverse AL patients, n/a= not available

3.1.4 Detailed analysis of most common *IGKV*-subfamilies in AL

The following chapter presents a detailed sequence analysis of *IGKV1/D-33* (AL= 32 %), *IGKV1/D-39* (AL= 20 %) and *IGKV1-5* (AL= 12 %), as the three most common *IGKV* representatives in AL. This includes the analysis of the mutation distribution with mutation count and frequency as well as the AA composition of the complete LC. In order to gain information about the theoretical stability of the LC, the theoretical pI, as indication for solubility, the GRAVY-score, which is defined by the hydrophobicity of the protein sequence and the prediction for a β-sheet aggregation (AGG-score) were calculated. In this context, the patient derived LC were studied as well as the difference to the specific reference sequence (delta pI, delta GRAVY, delta AGG). Moreover, a detailed mutation analysis was performed, identifying mutation hotspots, which means position mutated to or more than 50 % compared to the reference sequence and the influence of variations in the charge of AA was studied. Further, mutations that only affected one of the two diseases were also specified. Despite from that, a sub-analysis of the influence of the presence of a potential binding partner of the LC (HC in serum) was performed.

3.1.4.1 Overview about patients with *IGKV1/D-33*

With about 32 % *IGKV1/D-33* is the most abundant *IGKV* family representative in AL. In MM, *IGKV1/D-33* was detected in 13 % of patients. Table 16 shows an overview of the AL and MM patients. Overall, sequencing of the patients assigned to *IGKV1/D-33* resulted clear sequences, without heterogeneous sequencing signals. Except FOR1013, which showed heterogeneous signals at two positions. In the AL cohort, six patients had dominant heart involvement and two had dominant kidney involvement. Five patients had other clinically relevant organ involvement and were defined as diverse. Regarding the *IGKJ* usage, all *IGKJ* variants were represented, with *IGKJ2* being the most common in

AL (54 %), MM showed no preferred representative. Seven out of 13 AL patients had a HC in the serum, with either IgG or IgA. In MM, seven patients presented IgG and three had IgA. Only one MM patient had no HC.

Table 16. Overview about the *IGKV1/D-33* AL amyloidosis and multiple myeloma patients.

	Disease	X AA [n]	Organ involvement	*IGKJ*	HC in serum
FOR1001	AL	0	Diverse	*IGKJ2*	IgA
FOR1006	AL	0	Heart	*IGKJ3*	IgG
FOR1007	AL	0	Heart	*IGKJ2*	-
FOR1010	AL	0	Heart	*IGKJ2*	-
FOR1013	AL	2	Kidney	*IGKJ4*	IgG
FOR1018	AL	0	Heart	*IGKJ2*	IgA
FOR1020	AL	0	Kidney	*IGKJ2*	-
FOR1021	AL	0	Heart	*IGKJ2*	IgG
FOR1023	AL	0	Heart	*IGKJ1*	IgA
FOR1024	AL	0	Diverse	*IGKJ4*	-
FOR1028	AL	0	Diverse	*IGKJ2*	-
FOR1036	AL	0	Diverse	*IGKJ5*	IgG
FOR1039	AL	0	Diverse	*IGKJ4*	-
MM1009	MM	0	-	*IGKJ2*	IgG
MM1021	MM	0	-	*IGKJ1*	IgG
MM1024	MM	0	-	*IGKJ2*	IgG
MM1043	MM	0	-	*IGKJ4*	IgG
MM1048	MM	0	-	*IGKJ2*	IgA
MM1050	MM	0	-	*IGKJ4*	-
MM1055	MM	0	-	*IGKJ3*	IgA
MM1056	MM	0	-	*IGKJ3*	IgG
MM1062	MM	0	-	*IGKJ4*	IgG
MM1068	MM	0	-	*IGKJ4*	IgG
MM1074	MM	0	-	*IGKJ2*	IgA

AL= AL amyloidosis, MM= multiple myeloma, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal.

3.1.4.2 Median mutation count and frequency of *IGKV1/D-33* in AL amyloidosis and multiple myeloma and the presence of a heavy chain in serum

Sequences assigned to *IGKV1/D-33* were analysed with respect to the distribution of the mutations in the corresponding regions and segments and the mutation frequency (Table 17). Regarding the *IGKV*-regions of the two cohorts, differences in the median mutation count could be observed without reaching statistical significance (median mutation count: AL= 12.0 vs. MM= 10.0, mutation frequency AL and MM= 100 %). 77 % of AL patients and 64 % of MM patients displayed mutations in the *IGKJ*-region (mutation frequency), although the median mutation count of the mutated MM sequences seemed to be higher (median mutation count: AL= 1.0 vs. MM= 2.0). In the sequence regions covered by both cohorts, no mutation could be observed in the *IGKC*-region. Considering the median mutation frequency from the level of the different *IGK*-segments, the most pronounced differences were found in FR1. In FR1 69 % of AL patients had at least one mutation (median mutation count: AL and MM= 1.0), whereas only 36 % of MM patients did. The highest median mutation count could be observed in FR3 for both diseases (median mutation count: AL= 4.0, MM= 3.5) with a mutation frequency of approximately 90 % for both.

In addition to the comparison of the two diseases, the influence of a HC was also investigated (Table 17). With a median mutation count of 12.0 the AL sequences with a detectable HC were higher mutated in the *IGKV*-segment than the AL sequences without a HC (mutation count= 10.5) with the same mutation frequency for both (AL_HC and AL_noHC= 100 %). In contrast, the mutation frequency of the *IGKJ*-segment of the AL_noHC was higher (AL_noHC= 83 % vs. AL_HC= 71 %) with the same median mutation count of 1.0.

Table 17. Median mutation count and mutation frequency of *IGKV1/D-33* sequences. (Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

			IGKV 89 AA							***IGKJ*** 12 AA
	n			FR1 18 AA	CDR1 11 AA	FR2 15 AA	CDR2 7 AA	FR3 32 AA	CDR3	
AL	13	[%]	100	69	77	85	77	92	77	77
		[n]	12.0	1.0	1.0	1.0	1.0	4.0	2.5	1.0
MM	11	[%]	100	36	73	82	64	91	100	64
		[n]	10.0	1.0	2.0	1.0	2.0	3.5	2.0	2.0

			IGKV 89 AA							***IGKJ*** 12 AA
	n			FR1 18 AA	CDR1 11 AA	FR2 15 AA	CDR2 7 AA	FR3 32 AA	CDR3	
AL_HC	7	[%]	100	71	71	100	71	86	86	71
		[n]	12.0	1.0	1.0	1.0	1.0	4.0	2.5	1.0
AL_noHC	6	[%]	100	67	100	67	83	100	67	83
		[n]	10.5	1.0	1.5	1.0	1.0	3.5	2.0	1.0
MM_HC	10	[%]	100	30	70	80	60	90	100	60
		[n]	9.5	1.0	2.0	1.0	2.0	3.0	2.0	1.5
MM_noHC	1	[%]	-	-	-	-	-	-	-	-
		[n]								

[%]= percentage of mutated segments, [n]= median mutation count. For the CDR3-region, the patient-specific linker and the first two AA of the J-segment were included and the AA length was between 8 and 14 AA. AL= AL amyloidosis, MM= multiple myeloma, HC= heavy chain in serum, no_HC= no heavy chain in serum. All values are given as medians.

3.1.4.3 Amino acid composition of *IGKV1/D-33* in AL amyloidosis and multiple myeloma and the presence of a heavy chain in serum

In order to study the potential effect on the complete LC, the total composition of the AA was calculated. Table 18 summarizes all median percentages of the different AA of the AL and MM cohort assigned to *IGKV1/D-33* and the stratification into HC and noHC. Differences in the median proportion of the AA composition were detected between AL and MM (Table 18). The AL sequences showed a significantly higher median percentage of asparagine (median: AL= 4.6 % vs. MM= 4.0 %, p= 0.007, Figure 14). The following AA also tended to be higher in AL compared to MM: glutamine (median: AL= 8.0 % vs. MM= 7.4 %), glutamine acid (median: AL=3.5 % vs. MM= 3.4 %), isoleucine (median: AL= 4.0 % vs. MM= 3.4 %) and leucine (median: AL= 8.0 % vs. MM= 7.4 %). Conversely, MM sequences had an increased median percentage of phenylalanine (median: MM= 5.1 % vs. AL= 4.6 %) and serine (median: MM= 15.3 % vs. AL= 14.3 %) (Table 18).

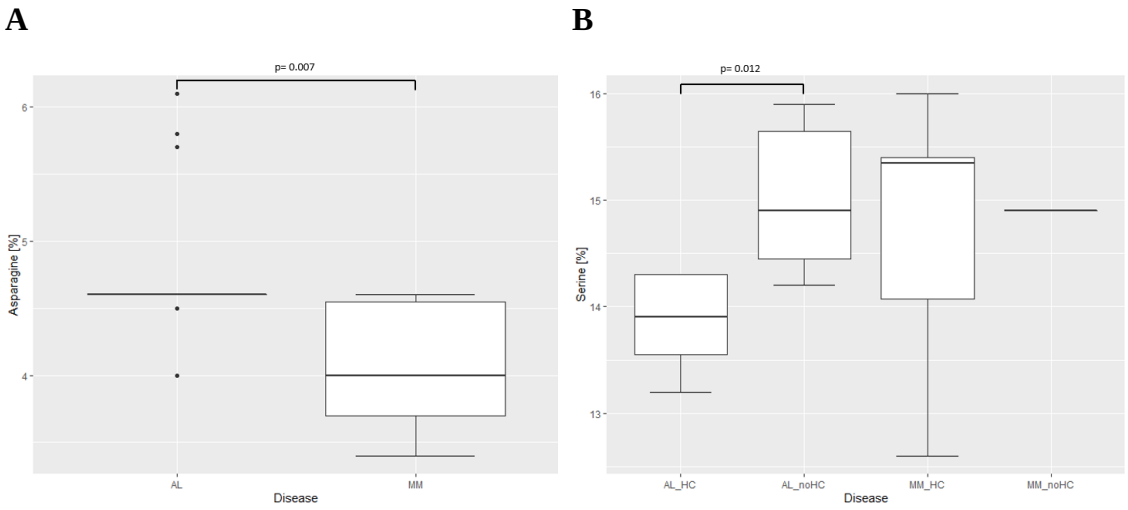

Figure 14. Comparison of the amino acid amount between AL amyloidosis and multiple myeloma sequences assigned to *IGKV1/D-33*. AL= AL amyloidosis, MM= multiple myeloma, HC= heavy chain in serum, noHC= no heavy chain in serum. All p-values presented are 2-sided and p-values ≤ 0.05 were considered to be significant. **A** Comparison of AL and MM assigned to *IGKV1/D-33*, AL n= 13, MM n= 11, **B** Stratification in the presence or absence of a heavy chain in serum, AL_HC n= 7, AL_noHC= 6, MM_HC= 10, MM_noHC= 1. (Schreiner et al. 2024) This figure was published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

The stratification of the sequences according to a possible binding partner for the LC revealed another significant difference in the AA composition. By comparing the AL sequences for the presence or absence of a HC, the amount of serine was significantly lower in AL with HC in the serum (median: AL_HC= 13.9 % vs. AL_noHC= 14.9 %, p= 0.012, Figure 14). The following AA also tended to be lower in AL with detectable HC in serum compared to AL without a HC: arginine (median: AL_HC= 2.3 % vs. AL_noHC= 2.8 %), glutamine acid (median: AL_HC= 3.4 % vs. AL_noHC= 4.0 %) and leucine (median: AL_HC= 7.4 % vs. AL_noHC= 8.0 %). Asparagine (median: AL_noHC= 5.7 % vs. AL_HC= 5.4 %), glutamine (median: AL_noHC= 3.4 % vs. AL_HC= 4.0 %), glycine (median: AL_noHC= 6.3 % vs. AL_HC= 6.2 %), isoleucine (median: AL_noHC= 4.0 % vs. AL_HC= 3.7 %), phenylalanine (median: AL_noHC= 5.1 % vs. AL_HC= 4.6 %), proline (median: AL_noHC= 5.7 % vs. AL_HC= 5.4 %), valine (median: AL_noHC= 6.3 % vs. AL_HC= 6.0%) had an increased median percentage in AL without a HC (Table 18).

Table 18. Proportion of the different amino acids in AL amyloidosis and multiple myeloma with *IGKV1/D-33*.

	AL [%]	MM [%]	AL_HC [%]	AL_noHC [%]	MM_HC [%]	MM_noHC [%]
Ala	5.7	5.7	5.7	5.7	5.7	5.1
Arg	2.3	2.3	2.3	2.8	2.3	2.9
Asn	4.6*	4.0*	4.6	4.6	4.0	4.0
Asp	5.7	5.7	5.7	5.4	5.7	5.7
Cys	1.7	1.7	1.7	1.7	1.7	1.7
Gln	8.0	7.4	8.0	7.7	7.4	6.9
Glu	3.5	3.4	3.4	4.0	3.7	3.4
Gly	6.3	6.3	6.3	6.2	6.3	6.3
His	0.0	0.0	0.0	0.0	0.0	0.1
Ile	4.0	3.4	4.0	3.7	3.7	2.9
Leu	8.0	7.4	7.4	8.0	7.7	6.9
Lys	5.1	5.1	5.1	5.1	5.1	5.1
Met	0.0	0.0	0.0	0.0	0.0	0.6
Phe	4.6	5.1	5.1	4.6	4.9	5.1
Pro	5.7	5.7	5.7	5.4	5.7	6.3
Ser	14.3	15.3	13.9*	14.9*	15.4	14.9
Thr	8.6	8.6	8.6	8.6	8.6	9.1
Trp	1.1	1.1	1.1	1.1	1.1	1.1
Tyr	4.6	4.6	4.6	4.6	4.3	4.6
Val	6.3	6.3	6.3	6.0	6.3	6.3

AL= AL amyloidosis, MM= multiple myeloma, AL n= 13, MM n= 11, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 7, AL_noHC= 6, MM_HC= 10, MM_noHC= 1, Ala= alanine, Arg= arginine, Asn= asparagine, Asp= aspartic acid, Cys= cysteine, Gln= glutamine, Glu= glutamic acid, Gly= glycine, His= histidine, Ile= isoleucine, leu= leucine, Lys= lysine, Met= methionine, Phe= phenylalanine, pro= proline, Ser= serine, Thr= threonine, Trp= tryptophan, Tyr= tyrosine, Val= valine, All values are given as medians, *p-values ≤ 0.05 were considered to be significant.

3.1.4.4 Biochemical properties of *IGKV1/D-33*

To gain insights into the theoretical stability of the LC, the biochemical properties of the sequences associated with *IGKV1/D-33* were studied. In this context, the theoretical pI of the full-length LC, which is an indication for solubility, was calculated as well as the GRAVY- (protein hydrophobicity) and AGG-score (aggregation propensity). For this

purpose, the absolute value of the patient specific LC and the difference to the respective references was calculated as well as the influence of a potential binding partner for the LC (HC) was observed (Table 19).

The analysis of the pI of both cohorts showed no differences (median pI: AL= 4.90, MM= 4.87; median delta pI: AL= 0.14, MM= 0.11) (Table 19). Likewise, when the AL sequences were differentiated according to the presence of a HC (median pI: AL_HC= 4.80, AL_noHC= 4.99; median delta pI: AL= 0.04, MM= 0.23). The calculation of the GRAVY-score showed only minor differences between the two patient groups (median GRAVY: AL= -0.418, MM= -0.434; median delta GRAVY: AL= -0.418, MM= -0.434), even though the GRAVY-score of the AL_HC was slightly higher (median GRAVY: AL_HC= -0.416, AL_noHC= -0.462; median delta GRAVY: AL_HC= 0.027, AL_noHC= -0.462). Considering the AGG-score, the AL cohort had a higher AGG-score than the MM (median AGG: AL= 499.374, MM= 458.062), which was also reflected in the delta AGG (median delta AGG: AL= 39.899, MM= -0.204), but without being statistically significant. When comparing the AL_HC and AL_noHC, a higher AGG-score was detected in the AL_HC group (median AGG: AL_HC= 560.343, AL_noHC= 477.208; median delta AGG: AL_HC= 101.748, AL_noHC= 14.733).

Table 19. Biochemical properties of *IGKV1/D-33* AL amyloidosis and multiple myeloma sequences. (Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

	pI *IGKVJC*	**delta pI *IGKVJC***	**GRAVY**	**delta GRAVY**	**AGG**	**delta AGG**
AL	4.90	0.14	-0.418	0.021	499.374	39.899
MM	4.87	0.11	-0.434	0.011	458.062	-0.204
AL_HC	4.80	0.04	-0.416	0.027	560.343	101.748
AL_noHC	4.99	0.23	-0.462	-0.005	474.208	14.733
MM_HC	4.87	0.11	-0.409	0.024	456.931	-1.940
MM_noHC	5.63	0.87	-0.501	-0.085	571.726	112.251

AL= AL amyloidosis, MM= multiple myeloma, AL n= 13, MM n= 11, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 7, AL_noHC= 6, MM_HC= 10, MM_noHC= 1, pI= isoelectric point, GRAVY= grand average of hydropathy, AGG= β-sheet aggregation, The corresponding parameter was calculated as well as the difference to the respective difference. All values are given as medians.

3.1.4.5 Mutation analysis of *IGKV1/D-33*

A detailed sequence analysis of *IGKV1/D-33* sequences from AL and MM patients was performed to investigate mutation hotspots (mutated ≥ 50 %), variations in AA charge and mutations affecting only one of the two diseases (Figure 15). This analysis revealed mutation hotspots in the FR3 and CDR3 for AL. More precisely: In FR3, two positions were mutated more than 50 % (65S, 83I). In CDR3, there was only one position (93N). The MM sequences showed mutation hotspots only in the CDR3. This also affected position 93N, but also position 101. In terms of the influence of mutations on the charge of AA, the following positions stood out. At position 55E, located in the CDR2, mutations in five of the 13 AL sequences caused a reversal of charge from negative to positive or a loss of charge. More precisely, in two cases a mutation was detected towards lysine (charge exchange) and in the case of a loss in charge a mutation towards glutamine (1x) or leucine (2x). For MM, sequences were mutated only in one case towards glutamine. The same pattern was detected for AL sequences at position 70D, located in the FR3 and next to an AL specific mutation hotspot (65S). But in this case, a mutation towards histidine (change of charge) was found in two cases and a mutation towards asparagine (loss of charge) in three cases. In MM, at position 70D three mutations resulted in a loss or reversal of charge (D70H, D70G, D70N). Position 92D, also affected by changes in charge and next to the mutation hotspot 93N (AL) showed a loss or reversal of charge from negative to positive in four AL sequences (D92H, D92L, D92S, D92Q). Only in one MM sequence a reversal of charge could be detected at this position (D92K). Mutations affecting the MM mutation hotspot 93N inserted negative charges in two of the MM sequences (2x N93D) and in one of the sequences a positive charge (N93H). Only in one of the AL sequences a mutation resulted on the insertion of a positive charge (N93K).

In 13 positions, only the AL sequences were mutated (10S, 13A, 15V, 17D, 19V, 20T, 26S, 36Y, 39K, 78L, 79H, 86Y, 105T) and mutations in 12 positions could only be observed in MM cases (14S, 18R, 24Q, 27Q, 37Q, 56T, 66G, 74T, 75I, 87Y, 89Q, 90Q) (Figure 15).

Considering *IGKV1/D-33* linked to *IGKJ2*, which is the most common *IGKV-IGKJ* linkage, mutation hotspot 83I is of particular interest. This affects the following AL sequences: FOR1001, FOR1007, FOR1010, FOR1018, FOR1020, FOR1021 and FOR1028. Interestingly, mutations at this position (83I) were associated with dominant

cardiac involvement in all cases (4/4) and diverse organ involvement in one case. This AL patient (FOR1001) has also heart in addition to kidney involvement. Furthermore, this patient was the only case showing a two AA insertion in the *IGKJ*-region (108insVE), resulting in an additional negative charge.

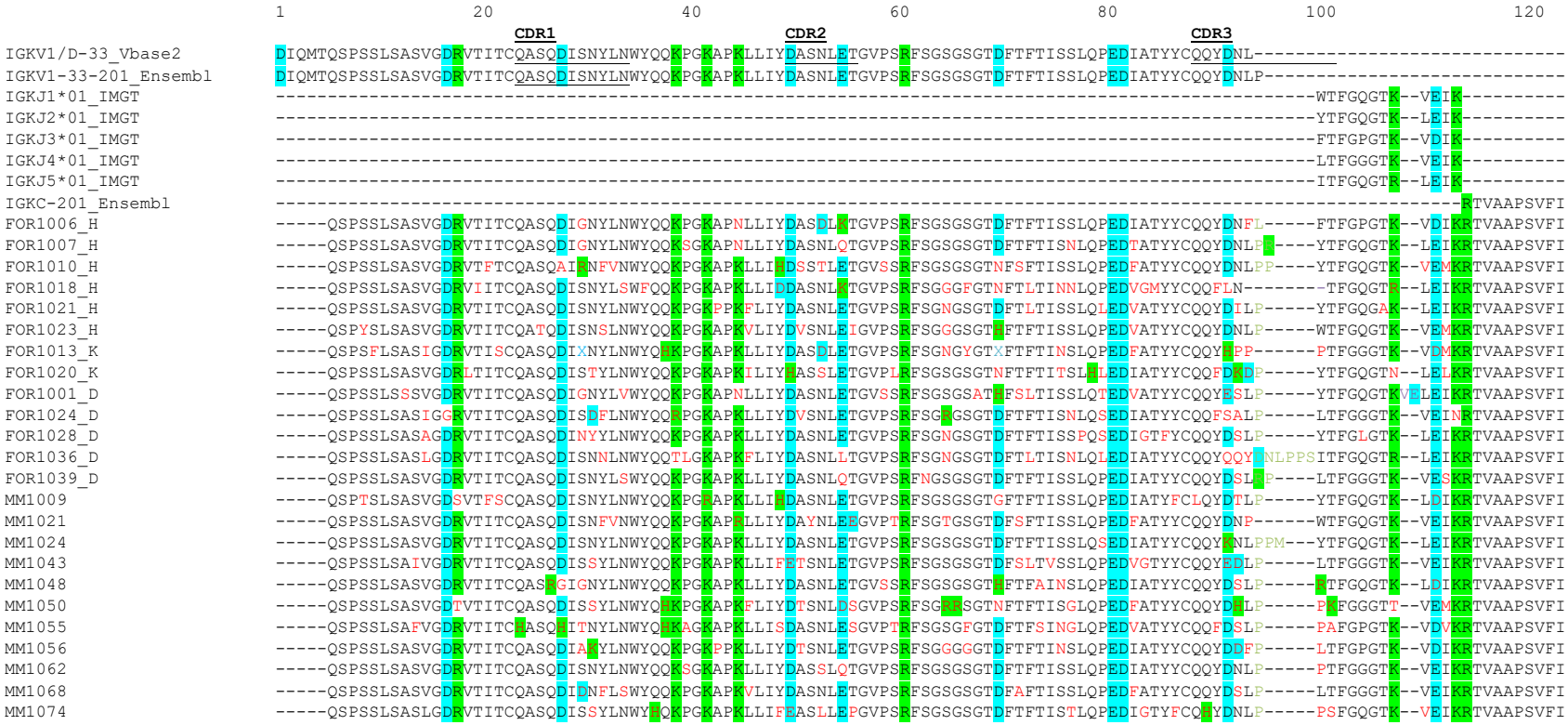

Figure 15. Sequence Alignment of *IGKV1/D-33* AL amyloidosis and multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. H= dominant heart involvement, K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis. (Schreiner et al. 2024) This figure was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

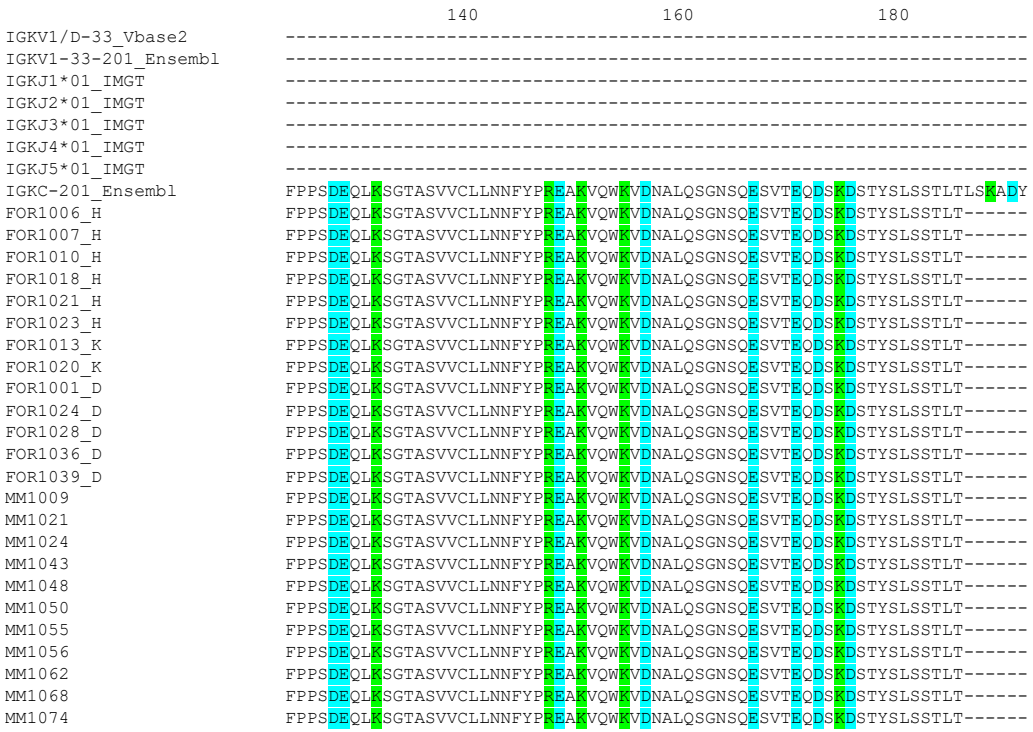

Figure 15. Sequence Alignment of *IGKV1/D-33* AL amyloidosis and multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. H= dominant heart involvement, K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis. (Schreiner et al. 2024) This table was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

3.1.4.6 Overview about patients with *IGKV1/D-39*

As the second most common representative in AL, *IGKV1/D-39* was analysed in detail (AL= 20 %, MM= 17 %). Table 20 summarises the associated patients. Sequencing of patients assigned to *IGKV1/D-39* showed clear, unambiguously assignable signals and correspondingly clear sequences. Exceptions were FOR1026 with 13 positions of unclear signals and FOR1004 with seven positions, where translation into a unique AA was not possible. The AL cohort comprises five patients with dominant kidney involvement and three AL_D patients. No clear *IGKJ* representative was identified in AL, in MM *IGKJ2* was most frequently detected (36%). Four of the eight AL patients had a HC in the serum, all with IgG. In MM, the number of patients with a HC was higher; 13 of the 14 MM patients had a HC, including 1x IgA and 12x IgG.

Table 20. Overview about the *IGKV1/D-39* AL amyloidosis and multiple myeloma patients.

	Disease	X AA [n]	Organ involvement	*IGKJ*	HC in serum
FOR1002	AL	0	Kidney	*IGKJ2*	-
FOR1004	AL	7	Kidney	*IGKJ2*	IgG
FOR1011	AL	0	Kidney	*IGKJ1*	IgG
FOR1012	AL	0	Kidney	*IGKJ4*	-
FOR1019	AL	0	Kidney	*IGKJ1*	-
FOR1026	AL	13	Diverse	*IGKJ4*	-
FOR1032	AL	0	Diverse	*IGKJ5*	IgG
FOR1034	AL	0	Diverse	*IGKJ1*	IgG
MM1003	MM	0	-	*IGKJ2*	IgG
MM1004	MM	0	-	*IGKJ2*	IgG
MM1005	MM	0	-	*IGKJ2*	IgG
MM1006	MM	0	-	*IGKJ1*	IgG
MM1020	MM	0	-	*IGKJ2*	IgG
MM1031	MM	0	-	*IGKJ3*	IgG
MM1034	MM	0	-	*IGKJ4*	IgG
MM1036	MM	0	-	*IGKJ4*	IgA
MM1052	MM	0	-	*IGKJ2*	IgG
MM1057	MM	0	-	*IGKJ5*	IgG
MM1061	MM	0	-	*IGKJ5*	IgG

	Disease	X AA [n]	Organ involvement	*IGKJ*	HC in serum
MM1069	MM	0	-	*IGKJ3*	-
MM1079	MM	0	-	*IGKJ1*	IgG
MM1084	MM	0	-	*IGKJ1*	IgG

AL= AL amyloidosis, MM= multiple myeloma, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal.

3.1.4.7 Median mutation count and frequency of *IGKV1/D-39* in AL amyloidosis and multiple myeloma and the presence of a heavy chain in serum

Analysis of the mutation distribution of *IGKV1/D-39* sequences (Table 21) revealed a higher median mutation count of the *IGKV*-regions for AL compared to MM (median mutation count: AL= 14.0, MM= 9.5, mutation frequency: AL and MM= 100 %), without reaching statistically significant. 75 % of the *IGKJ*-regions of AL sequences and 71 % of MM were mutated (mutation frequency) with a median mutation count of 1.5 for AL and 1.0 for MM. No mutations could be observed in the *IGKC*-regions. Examining the mutations at the level of different segments, it can be noted that, with exception of the CDR2 and CDR3, the mutation frequency of all segments of AL sequences was higher. The median mutation count of FR1 (AL= 2.0, MM= 1.0, p= 0.020) and FR3 (AL= 3.5, MM= 2.0, p= 0.017) was significantly higher in AL. For AL the highest mutation count could be observed in CDR1 (median mutation count: AL= 3.0) and FR3 (median mutation count: AL= 3.5). In MM, besides CDR1 (median mutation count: MM= 2.5), CDR3 (median mutation count: MM= 3.0) was the highest mutated segment.

When the AL sequences were stratified according to the presence of a HC in the serum (Table 21), the sequences without a detectable HC had a higher median mutation count in the *IGKV*-region (AL_noHC= 14.0 vs. AL_HC= 12.0) with the highest in CDR1 (AL_noHC= 4.0) and FR3 (AL_noHC= 4.0). In contrast, the AL sequences with a detectable HC in the serum displayed the highest median mutation count in the CDR3 (AL_HC= 4.0). In the *IGKJ*-region the AL_HC sequences displayed a higher mutation frequency than the AL_noHC (AL_HC= 100 % vs. AL_noHC= 50 %). Concerning the level of the different segments the sequences differed only in the CDR1 (AL_HC= 100 % vs. AL_noHC= 75 %) and in the CDR2 (AL_HC= 50 % vs. AL_noHC= 75 %).

Table 21. Median mutation count and mutation frequency of *IGKV1/D-39* sequences.

			IGKV 88 AA							***IGKJ*** 12 AA
	n			FR1 17 AA	CDR1 11 AA	FR2 15 AA	CDR2 7 AA	FR3 31 AA	CDR3	
AL	8	[%]	100	50	88	75	63	75	88	75
		[n]	14.0	2.0*	3.0	2.0	1.0	3.5*	2.0	1.5
MM	14	[%]	100	64	100	86	71	93	79	71
		[n]	9.5	1.0*	2.5	1.5	2.0	2.0*	3.0	1.0
AL_HC	4	[%]	100	50	100	75	50	75	75	100
		[n]	12.0	2.0	2.5	2.0	1.5	3.0	4.0	1.5
AL_noHC	4	[%]	100	50	75	75	75	75	75	100
		[n]	14.0	2.5	4.0	2.0	1.0	4.0	1.5	1.5
MM_HC	13	[%]	100	69	100	85	77	92	77	69
		[n]	10.0	1.0	2.0	2.0	2.0	2.0	2.5	1.0
MM_noHC	1	[%]	-	-	-	-	-	-	-	-
		[n]								

[%]= percentage of mutated segments, [n]= median mutation count. For the CDR3-region, the patient-specific linker and the first two AA of the J-segment were included and the AA length was between 9 and 11 AA. All values are given as medians, *a p-value ≤ 0.05 was considered to be significant.

3.1.4.8 Amino acid composition of *IGKV1/D-39* in AL amyloidosis and multiple myeloma and the presence of a heavy chain in serum

In a next step, the total AA composition of the *IGKV1/D-39* LC sequences of both cohorts were calculated, as well as the potential influence of a HC in serum was studied. All data are shown in Table 22. Only minor differences in the amino acid composition were found between the two cohorts, with the largest difference being in the positively charged AA arginine (median: AL= 2.8 % vs. MM= 3.4 %). The following AA differed less: alanine (median: AL= 6.7 % vs. MM= 6.5 %), phenylalanine (median: AL= 4.5 % vs. MM= 4.4 %), serine (median: AL= 17.1 % vs. MM= 17.0 %) and histidine (median: AL= 0.3% vs. MM= 0.6 %).

Stratification into HC and no_HC also revealed only minor differences in the AA composition (Table 22). The following AA levels were higher in AL with HC in serum: alanine (median: AL_HC= 7.3 % vs. AL_noHC= 6.7 %), asparagine acid (median: AL_HC= 4.8 % vs. AL_noHC= 4.5 %), glutamine (median: AL_HC= 7.0 % vs. AL_noHC= 6.7 %), glycine (median: AL_HC= 6.2 % vs. AL_noHC= 6.1 %), lysine (median AL_HC= 5.4 % vs. AL_noHC= 5.3 %), methionine (median AL_HC= 0.3 % vs. AL_noHC= 0.0 %), proline (median: AL_HC= 5.6 % vs. AL_noHC= 5.3 %), serine

(median: AL_HC= 17.4 % vs. AL_noHC= 16.8 %), tyrosine (median: AL_HC= 4.0 % vs. AL_noHC= 3.9 %). AL_HC sequences had lower levels of arginine (median: AL_HC= 2.8 % vs. AL_noHC= 3.1 %), asparagine (median: AL_HC= 3.4 % vs. AL_noHC= 5.1 %), leucine (median: AL_HC= 8.1 % vs. AL_noHC= 8.4 %), threonine (median: AL_HC= 8.2 % vs. AL_noHC= 8.7 %) and valine (median: AL_HC= 5.9 % vs. AL_noHC= 6.1 %).

Table 22. Proportion of the different amino acids in AL amyloidosis and multiple myeloma with *IGKV1/D-39*.

	AL [%]	MM [%]	AL_HC [%]	AL_noHC [%]	MM_HC [%]	MM_noHC [%]
Ala	6.7	6.5	7.3	6.7	6.2	6.7
Arg	2.8	3.4	2.8	3.1	3.4	2.8
Asn	3.4	3.4	3.4	5.1	3.4	3.9
Asp	4.5	4.5	4.8	4.5	4.5	5.0
Cys	1.7	1.7	1.7	1.7	1.7	2.2
Gln	6.7	6.7	7.0	6.7	6.7	6.7
Glu	3.4	3.4	3.4	3.4	3.4	2.8
Gly	6.1	6.1	6.2	6.1	6.1	6.7
His	0.3	0.6	0.3	0.3	0.6	0.0
Ile	3.4	3.4	3.4	3.4	3.4	3.4
Leu	8.4	8.4	8.1	8.4	8.4	7.8
Lys	5.4	5.4	5.4	5.3	5.1	5.6
Met	0.0	0.0	0.3	0.0	0.0	0.0
Phe	4.5	4.4	4.5	4.5	4.4	3.9
Pro	5.6	5.6	5.6	5.3	5.6	5.6
Ser	17.1	17.0	17.4	16.8	17.2	15.6
Thr	8.4	8.4	8.2	8.7	8.4	8.9
Trp	1.1	1.1	1.1	1.1	1.1	1.1
Tyr	3.9	3.9	4.0	3.9	3.9	4.5
Val	6.1	6.1	5.9	6.1	6.1	6.7

AL= AL amyloidosis, MM= multiple myeloma, AL n= 8, MM n= 14, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 4, AL_noHC= 4, MM_HC= 13, MM_noHC= 1, Ala= alanine, Arg= arginine, Asn= asparagine, Asp= aspartic acid, Cys= cysteine, Gln= glutamine, Glu= glutamic acid, Gly= glycine, His= histidine, Ile= isoleucine, leu= leucine, Lys= lysine, Met= methionine, Phe= phenylalanine, pro= proline, Ser= serine, Thr= threonine, Trp= tryptophan, Tyr= tyrosine, Val= valine. All values are given as medians.

3.1.4.9 Biochemical properties of *IGKV1/D-39*

Next, the theoretical stability of the *IGKV1/D-39* sequences were investigated using the theoretical pI, GRAVY- and AGG-score, examining the values of the LC itself and the differences from the corresponding reference sequence (Table 23). Calculation of the pI revealed only minor differences between the two cohorts (median pI: AL= 7.06, MM= 7.62, median delta pI: AL= -0.54, MM= 0.02), but with a higher difference when AL sequences were analysed for the presence of a HC in serum. In this case the pI values differed about +1 pH unit for AL_noHC (median pI= 7.61; median delta pI= 0.01) sequences compared to AL_HC (median pI= 6.37; median delta pI= -1.23), which might indicate changes in solubility. Only small differences in the GRAVY-score between the two cohorts could be found (median GRAVY: AL= -0.387, MM= -0.346, median delta GRAVY: AL= -0.019, MM= -0.024). Also, the stratification in AL_HC and AL_noHC did not resulted in a potential change in the GRAVY-score (median GRAVY: AL_HC= -0.373, AL_noHC= -0.387; median delta GRAVY: AL_HC= 0.004, AL_noHC= -0.034). The AL sequences presented a lower AGG score than the MM sequences (median AGG: AL= 691.521, MM= 721.383), which was also reflected in the delta AGG (median delta AGG: AL= 66.415, MM= 95.570). AL_HC showed a lower AGG with a lower delta AGG compared to AL_noHC (median AGG: AL_HC= 687.988, AL_noHC= 708.733; median delta AGG: AL_HC= 64.310, AL_noHC= 87.271).

Table 23. Biochemical properties of *IGKV1/D-39* AL amyloidosis and multiple myeloma sequences.

	pI *IGKVJC*	delta pI *IGKVJC*	GRAVY	delta GRAVY	AGG	delta AGG
AL	7.06	-0.54	-0.387	-0.019	691.521	66.415
MM	7.62	0.02	-0.346	-0.024	721.383	95.570
AL_HC	6.37	-1.23	-0.373	0.004	687.988	64.310
AL_noHC	7.61	0.01	-0.387	-0.034	708.733	87.271
MM_HC	7.64	0.05	-0.346	-0.037	747.400	127.365
MM_noHC	7.55	-0.04	-0.345	-0.001	695.366	63.774

AL= AL amyloidosis, MM= multiple myeloma, AL n= 8, MM n= 14, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 4, AL_noHC= 4, MM_HC= 13, MM_noHC= 1, pI= isoelectric point, GRAVY= grand average of hydropathy, AGG= β-sheet aggregation, The corresponding parameter was calculated as well as the difference to the respective reference. All values are given as medians.

3.1.4.10 Mutation analysis of *IGKV1/D-39*

Detailed sequence analysis of *IGKV1/D-39* sequences from AL and MM patients (Figure 16) revealed mutation hotspots (mutated ≥ 50 %) in the CDR1, CDR2 and CDR3 for both cohorts. In AL three positions in CDR1 (28S, 30S and 31S), one position in CDR2 (53S) and two positions in CDR3 (93S, 94T) were identified as mutation hotspots. In MM, position 28S and 31S (CDR1), 50A (CDR2) and position 97 (CDR3) were also mutated in at least 50 % of the sequences. Due to the mutations located near the mutation hotspots, the accumulation of additional charges was observed, with more insertion of positive charges. In AL, accumulation of positive charges occurred in four out of five cases in region 26 to 34 of the CDR1. In three of these cases it was a mutation towards histidine, more precisely Q27H, Y32H and N34H. And in one case a mutation towards arginine could be identified (S31R). Only one mutation resulted in the insertion of a negative charge (S28D). In MM, positive charges were inserted in six cases (S26H, Q27R, S30K, 2x S31R, Y32H) and negative charges in three cases (2x Q27E, S28D). In CDR2, one position stood out. At position 55, positive charges were inserted in four of the MM sequences (Q55H) and a negative charge (Q55E) in only one case. In one AL sequence, a mutation towards histidine (Q55H) could also be detected. Between position 87 to 101, including the CDR3 region, positive charges were also inserted at position Q89H, Y97R (2x), Y87H as well as one negative charge on position Q90E. Seven of the 14 MM sequences displayed an insertion of at least one positive charge within the described range (2x Q89H, T94R, ins95R, 2x Y97H, W97R, Q101R). Besides this, one MM sequence showed the addition of a negative charge (Q90E).

In AL, 11 positions were mutated exclusively, whereas in MM there were 15 positions (Figure 16). It concerns 13A, 16G, 18R, 40P, 42K, 43A, 60S, 61R, 63S, 68G and 69T in AL and 19V, 22T, 25A, 38Q, 39K, 51A, 56S, 65S, 67S, 74T, 75I, 81E, 98T, 101Q and 103T in MM.

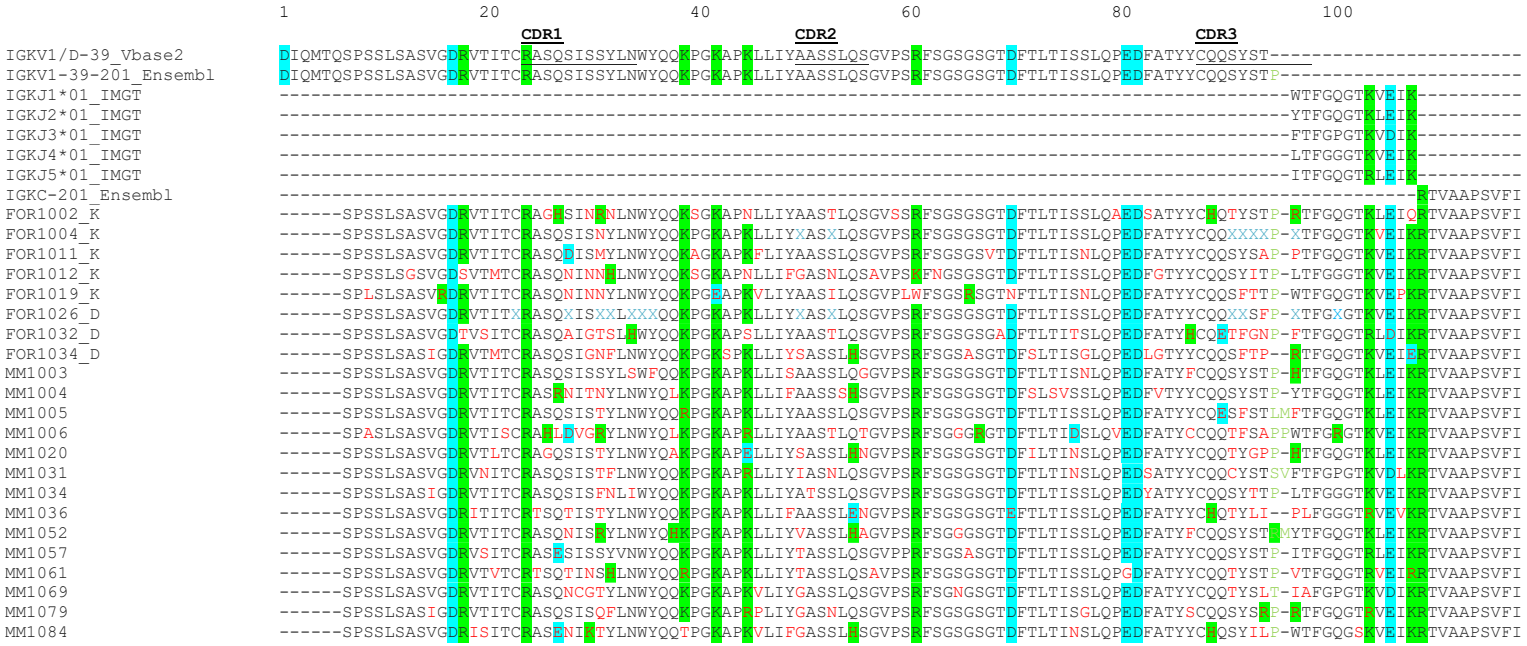

Figure 16. Sequence Alignment of *IGKV1/D-39* AL amyloidosis and multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids., X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis.

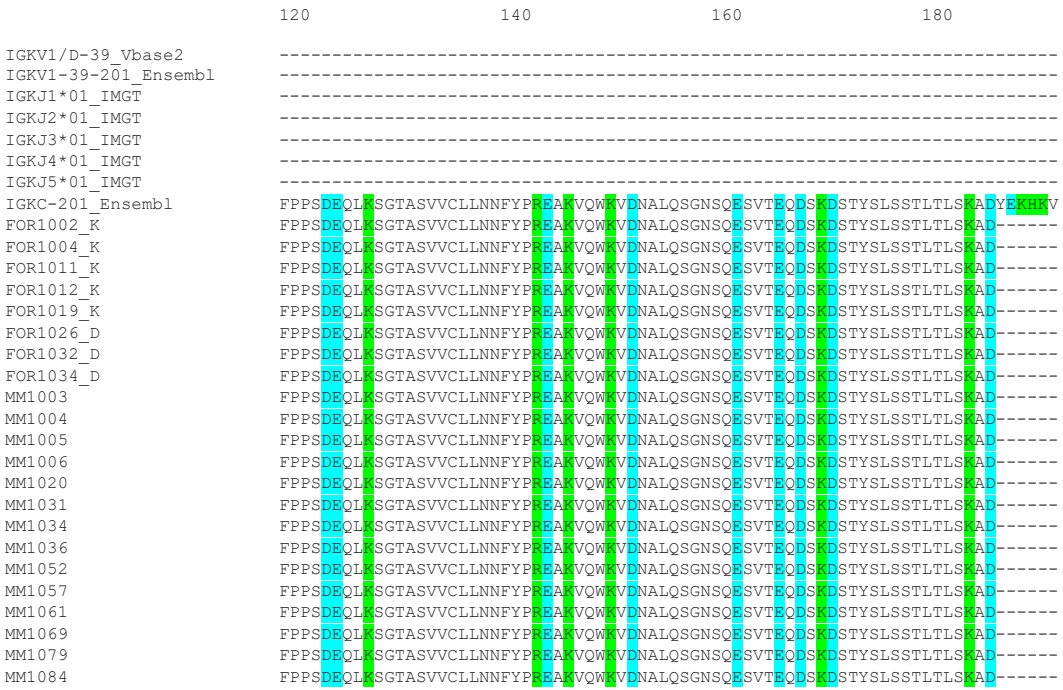

Figure 16. Sequence Alignment of *IGKV1/D-39* AL amyloidosis and multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis.

3.1.4.11 Overview about patients with *IGKV1-5*

Table 24 provides an overview of the patients assigned to *IGKV1-5*. Five (12 %) of the AL patients, including four AL_D patients and one AL_H patient, could be assigned to *IGKV1-5* as well as 14 MM patients (17 %). One AL sequence (FOR1009) showed heterogeneous signals at four positions, whereas the other sequences showed clear signals at all positions. The majority of AL patients and MM patients were linked to *IGKJ2* (AL= 80 %, MM= 50 %). One AL patient had a HC in the serum, the remaining four did not. Ten of the 14 MM patients had a HC, with 2x IgA and 8x IgG.

Table 24. Overview about the *IGKV1-5* AL amyloidosis and multiple myeloma patients.

	Disease	X AA [n]	Organ involvement	*IGKJ*	HC in serum
FOR1009	AL	4	Diverse	*IGKJ2*	IgG
FOR1014	AL	0	Diverse	*IGKJ2*	-
FOR1027	AL	0	Diverse	*IGKJ4*	-
FOR1040	AL	0	Heart	*IGKJ2*	-
FOR1041	AL	0	Diverse	*IGKJ2*	-
MM1012	MM	0	-	*IGKJ1*	IgG
MM1016	MM	0	-	*IGKJ2*	IgA
MM1017	MM	0	-	*IGKJ1*	IgG
MM1032	MM	0	-	*IGKJ1*	IgG
MM1035	MM	0	-	*IGKJ2*	-
MM1040	MM	0	-	*IGKJ1*	IgG
MM1041	MM	0	-	*IGKJ1*	-
MM1042	MM	0	-	*IGKJ2*	IgG
MM1045	MM	0	-	*IGKJ2*	-
MM1054	MM	0	-	*IGKJ2*	-
MM1058	MM	0	-	*IGKJ5*	IgA
MM1060	MM	0	-	*IGKJ4*	IgG
MM1065	MM	0	-	*IGKJ2*	IgG
MM1078	MM	0	-	*IGKJ2*	IgG

AL= AL amyloidosis, MM= multiple myeloma, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal.

3.1.4.12 Median mutation count and frequency of *IGKV1-5* in AL amyloidosis and multiple myeloma and the presence of a heavy chain in serum

Next, sequences belonging to *IGKV1-5* were studied for their mutation distribution in the respective regions and segments as well as mutation frequency (Table 25). All *IGKV1-5* sequences had mutations in the *IGKV*-region with a median mutation count of 6.0 for AL and 7.5 for MM. In AL, 60 % of the sequences showed mutations in the *IGKJ*-region, whereas only 36 % of the MM did (median mutation count: AL= 1.0, MM= 1.0). Concerning the different segments, the highest differences were detected in FR1 and CDR2. In FR1, the AL sequences presented a higher mutation frequency (80 %) than the MM sequences (43 %) (median mutation count: AL= 1.0, MM= 1.0). In contrast, the AL sequences showed a lower mutation frequency in the CDR2 (40 %), with a median mutation count of 2.0 (vs. MM median mutation frequency= 71 %, median mutation count= 2.5).

As only one AL patient had a HC in serum, no comparison was possible regarding the presence of a HC in AL patients assigned to *IGKV1-5*. Ten of the MM patients associated with *IGKV1-5* had a HC in serum and four did not (Table 25). The MM patients with a HC showed a higher median mutation count in the *IGKV*-region (MM_HC= 8.5, MM_noHC= 6.0), with the highest in the CDR2 (MM_HC= 3.0 vs. MM_noHC= 1.0). Regarding the mutation frequency in the individual segments, in FR1 and CDR3 the mutation frequency of the MM_HC sequences was higher than the MM_noHC (FR1: 50 % vs. 25 %, CDR3: 100 % vs. 50%), with the similar mutation count (FR1: 1.0 for both, CDR3: 1.5 vs. 2.0).

Table 25. Median mutation count and mutation frequency of *IGKV1-5* sequences.

			IGKV 88 AA							***IGKJ*** 12 AA
	n			FR1 17 AA	CDR1 11 AA	FR2 15 AA	CDR2 7 AA	FR3 32 AA	CDR3	
AL	5	[%]	100	80	80	80	40	80	80	60
		[n]	6.0	1.0	1.5	2.0	2.0	2.0	1.5	1.0
MM	14	[%]	100	43	100	79	71	79	86	36
		[n]	7.5	1.0	2.0	1.0	2.5	2.0	2.0	1.0
AL_HC	1	[%]	-	-	-	-	-	-	-	-
		[n]								
AL_noHC	4	[%]	100	75	75	100	50	75	75	50
		[n]	8.0	1.0	2.0	2.0	2.0	2.0	1.0	1.0
MM_HC	10	[%]	100	50	100	80	60	80	100	40

			IGKV 88 AA							***IGKJ*** 12 AA
	n			FR1 17 AA	CDR1 11 AA	FR2 15 AA	CDR2 7 AA	FR3 32 AA	CDR3	
		[n]	8.5	1.0	2.0	1.0	3.0	2.0	1.5	1.5
MM_noHC	4	[%]	100	25	100	75	100	75	50	25
		[n]	6.0	1.0	2.0	1.0	1.0	1.0	2.0	1.0

[%]= percentage of mutated segments, [n]= median mutation count. For the CDR3-region, the patient-specific linker and the first two AA of the J-segment were included and the AA length was between 8 and 11 AA. All values are given as medians.

3.1.4.13 Amino acid composition of *IGKV1-5* in AL amyloidosis and multiple myeloma and the presence of a heavy chain in serum

Analysis of the AA composition of *IGKV1-5* AL and MM LCs (Table 26) revealed a higher proportion of alanine (median: AL= 7.2 % vs. MM= 6.7 %), leucine (median: AL= 8.3 % vs. MM= 7.8 %) and proline (median: AL= 5.6 % vs. MM= 5.3 %) in AL. Glycine (median: AL= 6.1 % vs. MM= 5.9 %) and serine (median: AL= 16.1 % vs. MM= 16.0 %) was less increased in AL. And threonine (median: AL= 8.3 % vs. MM= 8.9 %) and valine (median: AL= 5.6 % vs. MM= 6.1 %) was lower in AL.

Given that only one of the four AL patients had a HC the analysis of the influence of a HC in serum for AL was limited. This was not the case for MM, where stratification was possible (Table 26). By stratifying MM into HC and noHC, significantly higher amount of tyrosine was observed in the MM patients without a HC (median: MM_HC= 4.7 % vs. MM_noHC= 5.3 %, p= 0.011). Other small differences were detected for arginine (median: MM_HC= 3.1 % vs. MM_noHC= 3.3 %), asparagine (median: MM_HC= 3.3 % vs. MM_noHC= 3.1 %), asparagine acid (median: MM_HC= 4.5 % vs. MM_noHC= 4.7 %), leucine (median: MM_HC= 7.8 % vs. MM= 8.1 %) and proline (median: MM_HC= 5.0 % vs. MM_noHC= 5.6 %), alanine (median: MM_HC= 7.0 % vs. MM_noHC= 6.7 %), isoleucine (median: MM_HC= 3.4 % vs. MM_noHC= 3.3 %) and leucine (median: MM_HC= 7.8 % vs. MM_noHC= 8.1 %).

Table 26. Proportion of the different amino acids in AL amyloidosis and multiple myeloma with *IGKV1-5*.

	AL [%]	MM [%]	AL_HC [%]	AL_noHC [%]	MM_HC [%]	MM_noHC [%]
Ala	7.2	6.7	6.6	7.2	7.0	6.7
Arg	3.3	3.3	3.3	3.1	3.1	3.3
Asn	3.3	3.3	2.2	3.6	3.3	3.6
Asp	5.6	4.5	6.1	5.3	4.5	4.7
Cys	1.7	1.7	1.7	1.7	1.7	1.7
Gln	6.6	6.7	6.6	6.4	6.7	6.7
Glu	3.9	3.9	3.9	3.6	3.9	3.9
Gly	6.1	5.9	6.1	5.9	5.9	5.9
His	0.0	0.0	0.0	0.0	0.0	0.0
Ile	2.8	3.3	3.9	2.8	3.4	3.3
Leu	8.3	7.8	8.3	8.6	7.8	8.1
Lys	5.6	5.6	5.5	5.6	5.6	5.6
Met	0.0	0.0	0.0	0.0	0.0	0.0
Phe	3.9	3.9	3.9	3.9	3.9	3.9
Pro	5.6	5.6	5.5	5.6	5.0	5.6
Ser	16.1	16.1	16.6	15.9	16.0	15.6
Thr	8.3	8.3	7.7	8.3	8.9	8.6
Trp	1.7	1.7	1.7	1.7	1.7	1.7
Tyr	5.0	5.0	5.0	5.3	4.7*	5.3*
Val	5.6	5.6	5.5	5.6	6.1	6.1

AL= AL amyloidosis, MM= multiple myeloma, AL n= 5, MM n= 14, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 1, AL_noHC= 4, MM_HC= 10, MM_noHC= 4, Ala= alanine, Arg= arginine, Asn= asparagine, Asp= aspartic acid, Cys= cysteine, Gln= glutamine, Glu= glutamic acid, Gly= glycine, His= histidine, Ile= isoleucine, leu= leucine, Lys= lysine, Met= methionine, Phe= phenylalanine, pro= proline, Ser= serine, Thr= threonine, Trp= tryptophan, Tyr= tyrosine, Val= valine, all values are given as medians, *a p-value ≤ 0.05 was considered to be significant.

3.1.4.14 Biophysical properties of *IGKV1-5*

Analysis of the biophysical properties of the *IGKV1-5* sequences (Table 27) revealed a lower median pI for AL sequences than for MM (median pI: AL=5.15 vs. MM= 7.07; median delta pI: AL= -0.30 vs. MM= -0.53). No differences could be noted in the pI when stratifying the MM sequences for the presence of a HC in serum (median pI: MM_HC=

7.07, MM_noHC= 7.06; median delta pI: MM_HC= -0.52, MM_noHC= -0.53). When considering the GRAVY-scores for both cohorts, only marginal differences could be found (median GRAVY: AL= -0.412, MM= -0.407; median delta GRAVY: AL= -0.011, MM= -0.006). Stratification of the MM sequences by the presence of a HC in serum resulted in a higher GRAVY-score for the MM_HC sequences compared to the MM_noHC sequences (median GRAVY: MM_HC= -0.377, MM_noHC= -0.429; median delta AGG: MM_HC= 0.023, MM_noHC= -0.030). Concerning the AGG-score, the AL sequences presented a lower value than the MM sequences (555.625 vs. 650.341) with a negative delta AGG for AL (-53.453) and a positive for MM (43.041). Analysis of MM sequences for the presence of a HC in serum revealed a higher AGG-score for MM_HC sequences (708.846; median delta AGG: 93.722) than for MM_noHC (585.262; median delta AGG: 3.368).

Table 27. Biochemical properties of *IGKV1-5* AL amyloidosis and multiple myeloma sequences.

	pI *IGKVJC*	delta pI *IGKVJC*	GRAVY	delta GRAVY	AGG	delta AGG
AL	5.15	-0.30	-0.412	-0.011	555.625	-53.453
MM	7.07	-0.53	-0.407	-0.006	650.341	43.041
AL_HC	5.15	-0.28	-0.411	-0.010	519.469	-89.609
AL_noHC	5.68	-0.33	-0.430	-0.029	559.981	-49.097
MM_HC	7.07	-0.52	-0.377	0.023	708.846	93.722
MM_noHC	7.06	-0.53	-0.429	-0.030	585.262	3.368

AL= AL amyloidosis, MM= multiple myeloma, AL n= 5, MM n= 14, HC= heavy chain in serum, no_HC= no heavy chain in serum, AL_HC n= 1, AL_noHC= 4, MM_HC= 10, MM_noHC= 4, pI= isoelectric point, GRAVY= grand average of hydropathy, AGG= β-sheet aggregation, The corresponding parameter was calculated as well as the difference to the respective difference. All values are given as medians.

3.1.4.15 Mutation analysis of *IGKV1-5*

Detailed analysis of *IGKV1-5* sequences (Figure 17) revealed mutation hotspots (mutations ≥ 50 %) for AL in FR1, CDR1, FR2 and CDR3. This affected one position in each of the mentioned regions. More precisely in FR1 position 21I, in CDR1 position 31S, in FR2 position 45K and in CDR3 92N. In MM, four mutation hotspots were detected. Two in CDR1 (30S, 31S), one in the FR2 (45K) and CDR2 (53S). Both cohorts

shared two mutation hotspots (31S and 45K). Analysing additional charges in the regions surrounding the mutation hotspots the addition of positive charges was detected in two of the five AL sequences at position 31S located in the CDR1. In four of the 14 MM sequences, positive charges were inserted at position Q26H, S31R (2x), W32R. Within the CDR3 region, mutations resulted in the insertion of a negative charge in three of the five AL sequences (N92D (2x), Y94D, Y98D,) and in five of the 14 MM sequences (N92D (3x), S93D, Y94D, ins95E). Furthermore, three additional positive charges were detected at position Q90H, W98R and S93H. Moreover, at mutation hotspot 45K, a loss of positive charges was observed in AL (3/5) and in MM (6/14).

In eight positions, only the AL sequences were mutated (20T, 40P, 42K, 65S, 66G, 71F, 72T, 107E) and mutations in 24 positions could only be observed in MM cases (7S, 14S, 22T, 27H, 32L, 33A, 36Y, 38H, 39K, 61R, 63S, 68G, 70E, 74T, 76S, 78L, 80P, 83F, 85T, 87Y, 89Q, 90Q, 93S, 104T) (Figure 17).

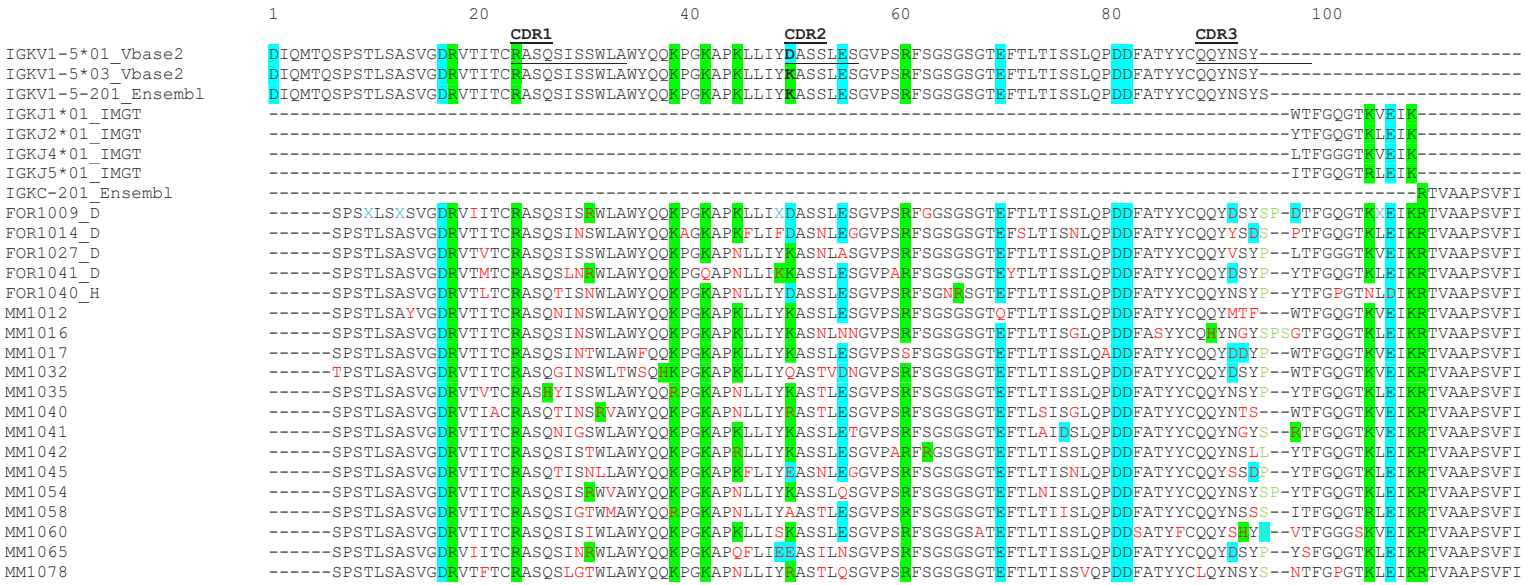

Figure 17. Sequence Alignment of *IGKV1-5* AL amyloidosis and multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. D= diverse organ involvement, H= dominant heart involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis.

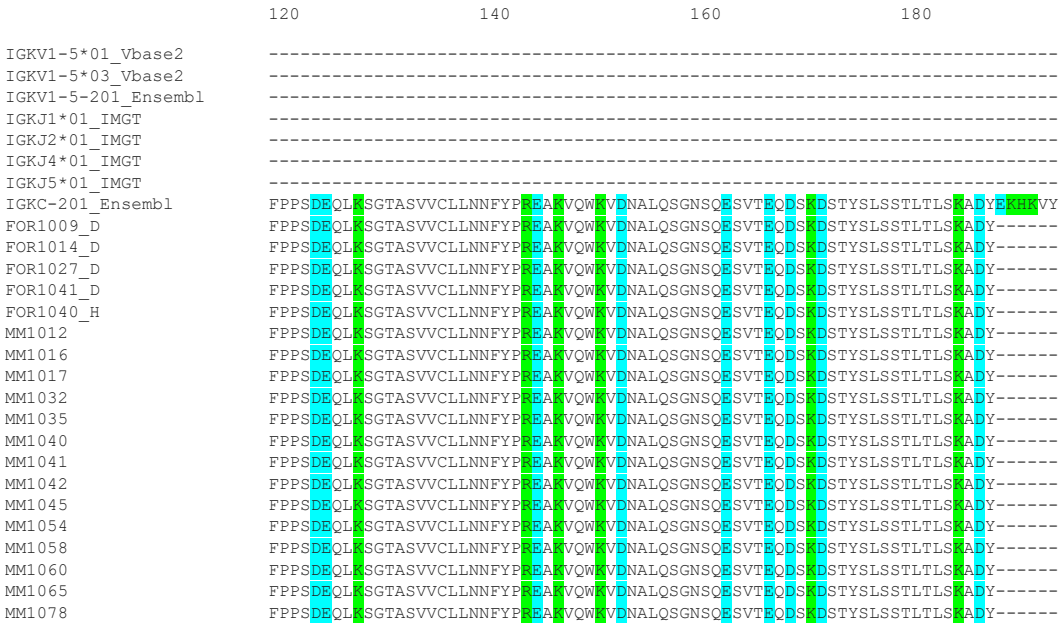

Figure 17. Sequence Alignment of *IGKV1-5* AL amyloidosis and multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. D= diverse organ involvement, H= dominant heart involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis.

3.1.5 Comparative analysis of one AL and MM representative - *IGKV1-16* and *IGKV3/D-11*

3.1.5.1 Overview about AL amyloidosis patients with *IGKV1-16* and multiple myeloma patients with *IGKV3/D-11*

As a next step, a comparative analysis of *IGKV1-16* and *IGKV3/D-11* was performed. *IGKV3/D-11* was detected as one of the most abundant subfamilies in MM (MM= 14 % vs. AL= 2 %) and *IGKV1-16* was detected exclusively in AL (AL= 10 %). Based on this divergent behaviour, which could be an indication of a different amyloidogenic potential, the following analysis was performed. Four *IGKV1-16* AL patients were compared with 12 *IGKV3/D-11* MM patients. To ensure comparability, only the matching regions of the *IGKV1-16* and *IGKV3/D-11* sequences were analysed. Table 28 summarises the associated *IGKV1-16* patients. Sequencing revealed clear, unambiguous signals in all cases. One of the four AL patients had dominant kidney involvement and the remaining patients were classified as AL_D. Two of the patients were assigned to *IGKJ4*, one to *IGKJ1* and one to *IGKJ2*. About half of the patients had a HC (IgG). Table 29 presents an overview about the MM patients assigned to *IGKV3/D-11*. All possible *IGKJ* linkages were represented, with *IGKJ1* (4x) the most common. *IGKJ3* and *IGKJ4* was assigned in three cases and *IGKJ2* and *IGKJ5* in one case, respectively. All patients had a HC with 8x IgG and 4x IgA.

Table 28. Overview about *IGKV1-16* AL amyloidosis patients.

	Disease	X AA [n]	Organ involvement	*IGKJ*	HC in serum
FOR1025	AL	0	Diverse	*IGKJ2*	0
FOR1031	AL	0	Diverse	*IGKJ4*	IgG
FOR1033	AL	0	Diverse	*IGKJ1*	0
FOR1035	AL	0	Kidney	*IGKJ4*	IgG

AL= AL amyloidosis, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal.

Table 29. Overview about *IGKV3/D-11* MM patients.

	Disease	X AA [n]	Organ involvement	*IGKJ*	HC in serum
MM1007	MM	0	-	*IGKJ4*	IgG
MM1008	MM	0	-	*IGKJ3*	IgG
MM1011	MM	0	-	*IGKJ1*	IgG
MM1014	MM	0	-	*IGKJ3*	IgG
MM1015	MM	0	-	*IGKJ4*	IgG
MM1018	MM	0	-	*IGKJ1*	IgG
MM1022	MM	0	-	*IGKJ1*	IgA
MM1023	MM	0	-	*IGKJ1*	IgG
MM1046	MM	0	-	*IGKJ4*	IgA
MM1071	MM	0	-	*IGKJ2*	IgA
MM1077	MM	0	-	*IGKJ5*	IgG
MM1080	MM	0	-	*IGKJ3*	IgA

MM= Multiple myeloma, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal.

3.1.5.2 Median mutation count and frequency of *IGKV1-16* in AL amyloidosis and *IGKV3/D-11* in multiple myeloma

In the following, a comparative analysis regarding the overall mutation distribution and frequency of *IGKV1-16* AL sequences and *IGKV3/D-11* MM sequences was performed. (Table 30). Both cohorts showed significant differences in the median mutation count and frequency. More precisely, *IGKV1-16* AL sequences showed significantly more mutations in the *IGKV*-segment than the *IGKV3/D-11* MM sequences (median mutation count: *IGKV*: AL (*IGKV1-16*)= 9.0, MM (*IGKV3/D-11*)= 5.5, p= 0.033). Interestingly, none of the *IGKV3/D-11* sequences had mutations in the FR1-region, whereas 50 % of the *IGKV1-16* AL sequences were mutated (mutation frequency: FR1: MM (*IGKV3/D-11*)= 0 %, AL (*IGKV1-16*)= 50 %, p= 0.050). Overall, the mutation frequency of the different FR- and CDR-regions of all *IGKV1-16* AL sequences was higher compared to *IGKV3/D-11* MM sequences with the highest mutation count in the FR3 for the *IGKV1-16* AL sequences (median mutation count: FR3: AL= 3.0). The *IGKJ*-regions of both cohorts had a similar mutation frequency (AL (*IGKV1-16*)= 75 %, MM(*IGKV3/D-11*)= 67 %) with a twofold higher mutation count of the *IGKV1-16* AL

sequences (AL (*IGKV1-16*)= 3.0 vs. MM (*IGKV3/D-11*)= 1.5). No mutations could be detected in the *IGKC*-regions.

The stratification of *IGKV1-16* sequences according to the presence of a HC in the serum (Table 30) shows that all patients with a HC had a higher mutation frequency with a lower mutation count in all segments with the exception of *IGKJ*. In this case the mutation frequency of the patients without a HC was higher and the mutation count was lower (mutation frequency: *IGKJ*: AL *IGKV1-16*_HC= 50 % vs. AL *IGKV1-16*_noHC= 100 %; mutation count: IGKJ: AL *IGKV1-16*_HC= 4.0 vs. AL *IGKV1-16*_noHC= 2.0).

Table 30. Median mutation count and mutation frequency of *IGKV1-16* and *IGKV3/D-11* sequences.

			IGKV 89 AA							***IGKJ*** 12 AA
	n			FR1 18 AA	CDR1 11 AA	FR2 15 AA	CDR2 7 AA	FR3 32 AA	CDR3	
AL *IGKV1-16*	4	[%]	100	50*	100	75	100	100	100	75
		[n]	9.0*	2.0	1.0	2.0	1.0	3.0	1.0	3.0
AL *IGKV1-16*_HC	2	[%]	100	100	100	100	100	100	100	50
		[n]	9.0	2.0	1.0	1.5	1.0	3.0	1.0	4.0
AL *IGKV1-16*_noHC	2	[%]	100	0	100	50	100	100	100	100
		[n]	9.5	0.0	2.0	3.0	1.5	3.0	1.5	2.0
MM *IGKV3/D-11*	12	[%]	100	0*	75	50	75	75	75	67
		[n]	5.5*	-	2.0	1.0	1.0	1.0	2.0	1.5

[%]= percentage of mutated segments, [n]= median mutation count. For the CDR3-region, the patient-specific linker and the first two AA of the J-segment were included and the AA length was between 8 and 10 AA. All values are given as medians. *a p-value ≤ 0.05 was considered to be significant.

3.1.5.3 Biophysical properties of *IGKV1-16* AL amyloidosis and *IGKV3/D-11* multiple myeloma sequences

Given that two different subfamilies were studied in this comparative analysis, the overall AA composition could not be directly compared. However, consideration of the biochemical properties of the LCs in question with the respective delta values (reference values), allows a comparative analysis (Table 31). Interestingly, the mutations of the *IGKV3/D-11* MM sequences had no influence on the median pI (median pI: MM (*IGKV3/D-11*)= 6.49; median delta pI: MM (*IGKV3/D-11*)= 0.00). No differences could be found for the median delta GRAVY for the *IGKV1-16* AL and

IGKV3/D-11 MM sequences (median delta GRAVY: AL *(IGKV1-16*)= -0.389, MM (*IGKV3/D-11*)= -0.450). The *IGKV1-16* AL and *IGKV3/D-11* MM sequences had a comparable median delta AGG (median delta AGG: AL *IGKV1-16*= -62.938, MM *IGKV3/D-11*= -29.932).

When analysing the influence of the presence of a HC upon the *IGKV1-16* AL sequences (Table 31), it was shown that the sequences without a HC in the serum had a lower median pI than the sequences with a detectable HC (median pI: AL (*IGKV1-16*_HC)= 8.13 vs. AL (*IGKV1-16*_noHC)= 7.10; median delta pI: AL (*IGKV1-16*_HC)= 0.30, AL (*IGKV1-16*_noHC)= -0.73). With regard to the GRAVY-score, the AL *IGKV1-16* sequences with and without a HC did not differ (median GRAVY: AL (*IGKV1-16*_HC)= -0.389, AL (*IGKV1-16*_noHC)= -0.373; median delta GRAVY: AL (*IGKV1-16*_HC)= -0.032, AL (*IGKV1-16*_noHC)= 0.030). Notably, when considering the AGG-score, the *IGKV1-16* AL sequences without a HC showed a much higher score than the *IGKV1-16*_HC sequences (median AGG: AL (*IGKV1-16*_noHC)= 942.376 vs. AL (*IGKV1-16*_HC)= 584.495; median delta AGG: AL (*IGKV1-16*_noHC)= 197.193, AL (*IGKV1-16*_HC)= -158.068).

Table 31. Biochemical properties of *IGKV1-16* AL amyloidosis and *IGKV3/D-11* multiple myeloma sequences.

	pI *IGKVJC*	**delta pI *IGKVJC***	**GRAVY**	**delta GRAVY**	**AGG**	**delta AGG**
AL *IGKV1-16*	8.13	0.30	-0.389	-0.017	684.822	-62.938
MM *IGKV3/D-11*	6.49	0.00	-0.450	-0.019	740.871	-29.932
AL *IGKV1-16*_HC	8.13	0.30	-0.389	-0.032	584.495	-158.068
AL *IGKV1-16*_noHC	7.10	-0.73	-0.373	0.030	942.376	197.193

AL= AL amyloidosis, MM= multiple myeloma, AL n= 4, MM n= 12, AL_HC n= 2, AL_noHC n= 2, pI= isoelectric point, GRAVY= grand average of hydropathy, AGG= β-sheet aggregation, the corresponding parameter was calculated as well as the difference to the respective reference. All values are given as medians.

3.1.5.4 Mutation analysis of *IGKV1-16* AL amyloidosis and *IGKV3/D-11* multiple myeloma sequences

Comparative mutation analysis of the *IGKV1-16* AL and *IGKV3/D-11* MM sequences indicated 10 mutation hotspots (mutations ≥ 50 % of sequences) for *IGKV1-16* and only two mutation hotspots for *IGKV3/D-11* (Figure 18). For *IGKV1-16* sequences this affected the CDR1 (28G, 32Y), CDR2 (57Q) and CDR3 (97S), the FR2 (37F, 43K) and the FR3 (72D, 80S, 83P). One mutation hotspot was also detected in the *IGKJ*-segment (109E). At mutation hotspot 28G (CDR1), the mutation towards aspartic acid resulted in the insertion of a negative charge in two of the three mutated sequences and mutations towards histidine on position 32Y resulted in the insertion of positive charge in one sequence. In CDR2, a mutation towards histidine was also detected (57Q). In one of the hotspots localised in FR3 (72D), the mutation caused a loss of charges. In MM the two mutation hotspots were in the CDR1 at position 30S and 31S. Regarding charge changes in the hotspots for MM, mutation hotspot 30S showed the insertion of a positive charge (arginine) in three of the seven mutated sequences and the insertion of a negative charge (asparagine acid) in one case.

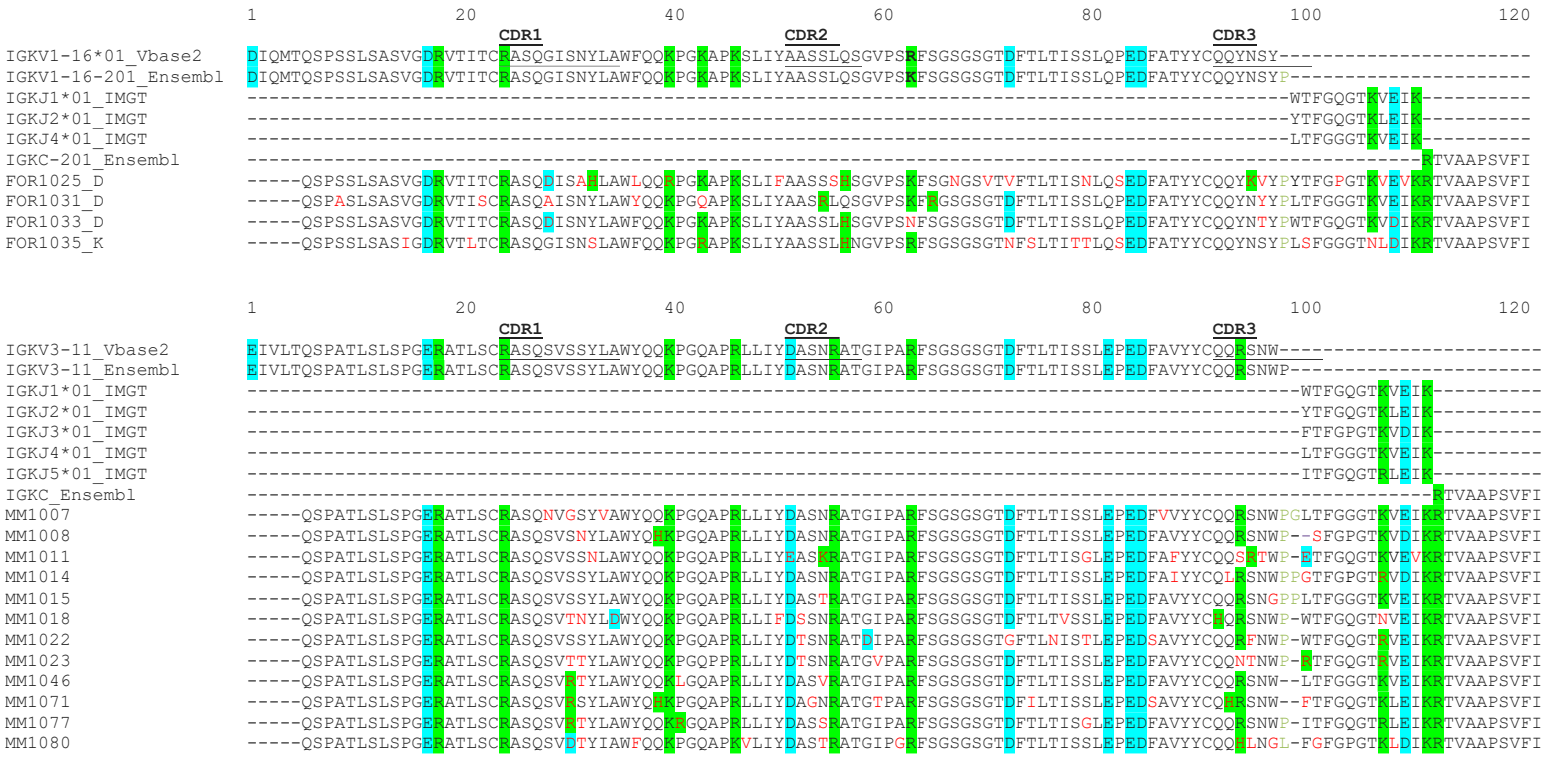

Figure 18. Sequence Alignment of *IGKV1-16* AL amyloidosis and *IGKV3/D-11* multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis.

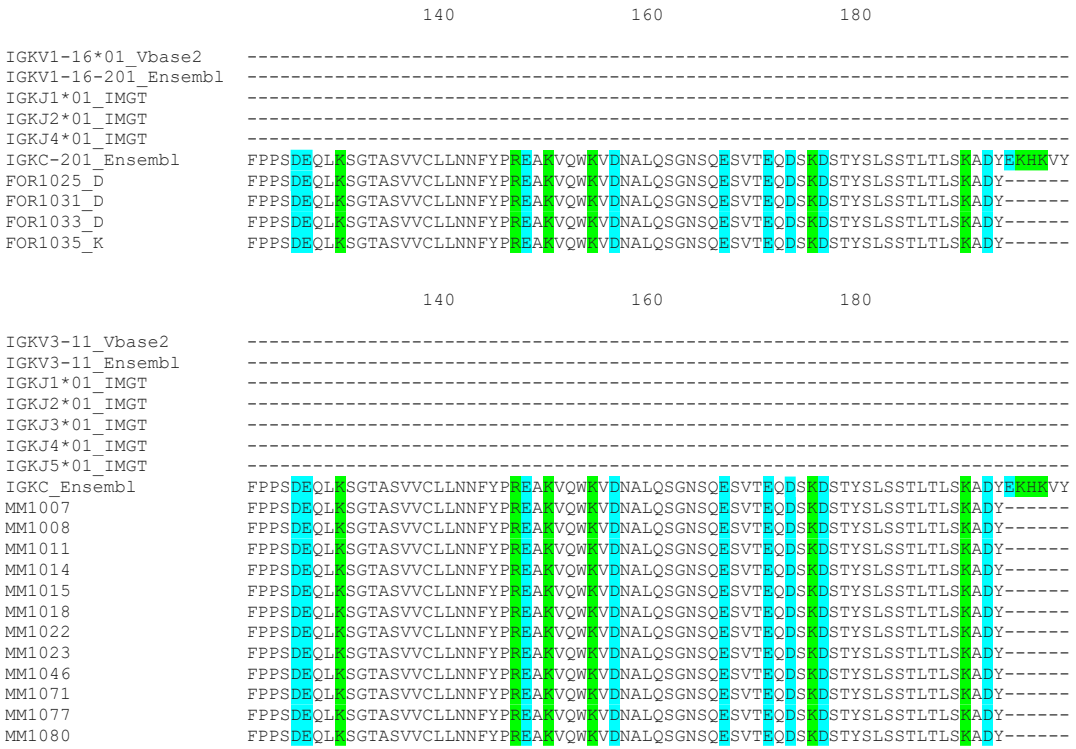

Figure 18. Sequence Alignment of *IGKV1-16* AL amyloidosis and *IGKV3/D-11* multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P01834). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis.

3.1.6 Analysis of a potential N-glycosylation site of the most abundant kappa light chains

As glycosylation has been implicated as a potential cause of LC aggregation (Dispenzieri et al. 2020), the AA sequences of kappa LC was investigated using the bioinformatics tool NetNGlyc 1.0 (Julenius 2007). As in the previous analysis, the focus was on the major AL *IGKV* representatives, *IGKV1/D-33*, *IGKV1/D-39* and *IGKV1-5*, a comparison was also made between *IGKV1-16* as a representative for AL and *IGKV3/D-11* for MM (Table 32). The evaluation of the entire cohort can be found in the appendix (SI Table 2). In this analysis, differences were found in the *IGKV1/D-33*-subgroup, where 61.5 % of the AL sequences showed potential N-glycosylation sites, but only 9.0 % of the MM sequences. Also, 75 % of the *IGKV1-16* sequences representing AL showed potential glycosylation, whereas only 8.3 % of the *IGKV3/D-11* MM representatives did. The association between a potential glycosylation site with respect to one of the two entities was also reflected in the whole cohorts, with significantly more AL sequences having a potential N-glycosylation site than MM sequences (42.5 % vs. 16.9 %, p= 0.004).

Table 32. Analysis of a potential N-glycosylation site of the most abundant kappa light chains.

	potential N-glycosylation site [n]	no potential N-glycosylation site [n]	Percentage of sequences with potential N-glycosylation site [%]
IGKV1-5			
AL [n= 5]	1	4	20.0
MM [n= 14]	3	11	21.4
IGKV1/D-33			
AL [n= 13]	8	5	61.5
MM [n= 11]	1	10	9.0
IGKV1/D-39			
AL [n= 8]	2	6	25.0
MM [n= 14]	4	10	28.6
IGKV1-16			
AL [n= 4]	3	1	75.0
IGKV3/D-11			
AL [n= 1]	0	1	0.0
MM [n= 12]	1	11	8.3

Analysis according to (Julenius 2007), AL= AL amyloidosis, MM= multiple myeloma

3.1.7 Analysis of the prediction of kappa light chain aggregation using V_LAmY-Pred- algorithm

Over the past few years, machine learning has become increasingly important in life sciences and is also gaining acceptance in the medical field. Two new bioinformatics tools based on machine learning have been developed to calculate the aggregation potential of LCs and thus to infer their amyloidogenic potential (Garofalo et al. 2021; Rawat et al. 2021). Whereby with Lictor only lambda LC sequences can be analysed for their aggregation potential (Garofalo et al. 2021), V_LAmY-Pred allows the classification of both, lambda and kappa LCs (Rawat et al. 2021). In the following, the prediction of the LC aggregation was investigated using the V_LAmY-Pred algorithm by analysing the three major AL *IGKV* representatives, *IGKV1/D-33*, *IGKV1D-39* and *IGKV1-5*. In addition, the prediction probability of *IGKV1-16* as a representative for AL and *IGKV3/D-11* as a representative for MM was calculated (SI Table 3) and summarized in Table 33. For *IGKV1-5*, five AL and 14 MM sequences were analysed. 40 % of the AL sequences were classified as amyloidogenic and 64 % of the MM sequences as non-amyloidogenic. A comparable assessment was given for the *IGKV1/D-39* sequences. 50 % of the AL sequences were considered to be amyloidogenic and 79 % of the MM sequences as non-amyloidogenic. Different for *IGKV1/D-33*, where all AL sequences were assessed as amyloidogenic, but also all MM sequences. About 50 % of the *IGKV1-16* AL sequences were classified as amyloidogenic. All MM sequences assigned to *IGKV3/D-11* were classified as non-amyloidogenic (12/12 MM sequences).

Table 33. Analysis of prediction of amyloidogenicity with V_LAmY-Pred- algorithm.

	classified as amyloidogenic [n]	classified as non-amyloidogenic [n]	correctness [%]
IGKV1-5			
AL [n= 5]	2	3	40
MM [n= 14]	5	9	64
IGKV1/D-33			
AL [n= 13]	13	0	100
MM [n= 11]	11	0	0
IGKV1/D-39			
AL [n= 8]	4	4	50
MM [n= 14]	3	11	79
IGKV1-16			

	classified as amyloidogenic [n]	classified as non-amyloidogenic [n]	correctness [%]
AL [n= 4]	2	2	50
IGKV3/D-11			
MM [n= 12]	0	12	100

Analysis according to (Rawat et al. 2021), AL= AL amyloidosis, MM= multiple myeloma

3.1.8 Summary kappa cohort

3.1.8.1 Summary patient cohorts and clinical data

In this project, 41 kappa AL patients, visiting the Amyloidosis Centre Heidelberg between January 2019 to May 2022, were consecutively enrolled and compared to 83 MM patients of the GMMG-HD6 clinical trial. The AL patients were classified according to their clinical organ presentation. This resulted in eight patients with dominant heart involvement (AL_H), in 14 patients with dominant kidney involvement (AL_K) and the remaining 19 patients were classified as diverse (AL_D) (Table 11). The AL and MM cohorts did not differ in their median age at diagnosis (median age between 61 and 63 years) and in AL_H males and females were equally represented (n= 4 vs. 4), while there were more males than females in the other groups (AL_K (n=10 vs. 4), AL_D (n=13 vs. 6), MM (n=53 vs. 30)). Bone marrow was collected in all AL_K and MM patients at the time of new diagnosis and in two AL_H and three AL_D patients at time of relapse. In addition to the most commonly affected organs in AL, the heart and kidneys, the liver was involved in 11 cases in the AL cohort. Nine of these cases were in the AL_D group. The AL_K patients had a lower median dFLC than the other AL patient groups (median dFLC: AL_K= 97 mg/L vs. AL_H= 543 mg/L, p< 0.001; AL_K= 97 mg/L vs. AL_D= 472 mg/L, p< 0.001). Moreover, the dFLC of AL_K patients was significantly lower compared to MM patients (median dFLC: AL_K= 97 mg/L vs. 349 mg/L, p= 0.009). In the AL group, the AL_K patients had the lowest AP (AP: AL_K= 74 U/L, AL_H= 125 U/L, AL_D= 120 U/L) and compared to MM, the AP of the AL_D patients was significantly higher than of the MM patients (AP: MM= 70 U/L, p< 0.001). An M-gradient was more common in MM patients compared to AL_K and AL_D patients (M-gradient: MM= 74/83 vs. AL_K= 8/14, p= 0.021; MM= 74/83 vs. AL_D, p< 0.001). A clonal heavy chain in serum was detected in more MM patients than AL_K and AL_D patients (HC: MM= 73/83 vs. AL_K= 8/14, p= 0.033; MM= 73/83 vs. AL_D= 9/19,

p< 0.001). In particular, AL_K patients had only IgG (8/14), whereas AL_D patients had in one case IgA (1/19) and in eight cases IgG (8/19). In MM patients, IgM and IgD were detected in one case (1/83 each), IgA in 12 cases (12/83) and IgG in 59 cases (59/83). AL_H presented IgG and IgA in equal parts (each 2/8). The iFISH- analysis showed a translocation t(11;14) in 4/8 AL_H, 5/13 AL_K, 12/15 AL_D and 16/82 MM patients, whereas a deletion 1q21 was detected in 2/8 AL_H, 3/13 AL_K, 3/15 AL_D and 8/82 MM patients.

3.1.8.2 Summary multiple myeloma sequence composition

Because the LC sequences of the AL and MM patients were generated by two different methods the MM sequences were first successfully verified by the same protocol used for the AL sequences (SI Figure 1). Then the composition of the MM bulk RNA data was analysed. Analysis for the presence of sub-sequences (sequence fraction_2) revealed a highly amount of median sequence fraction_1 between 99.01 % to 99.93 % for all *IGKV* representatives (Table 12). *IGKV1*, as the main *IGKV* representative (n= 44), had 7 sub-sequences (> 1 %), whereby one of the *IGKV1/D-39/J2* sequences with a median sequence fraction_2 of 42.11 % (vs. 57.47 % sequence fraction_1) stood out. However, detailed analysis revealed that the two sequences differed only in their length and were of the same origin. *IGKV2* (n= 8) had two sub-sequences with a low median sequence fraction_2 of 2.07 % and 5.65 %, respectively. One of the *IGKV3/D-11/IGKJ1* sequences had a median sequence fraction_1 of only 64.92 %, but this case had no other sub-sequences > 1 %. *IGKV4* and *IGKV5* had no sub-sequences > 1%. Due to the homogeneous high sequence fraction size of the sequences_1 across all *IGKV* families, sequences from sequence fraction_1 were used for further analysis.

3.1.8.3 Summary *IGKV/IGKJ* assignment

Next, the sequences of the AL and MM cohort were analysed according to their *IGK* family usage. *IGKV1* was highly expressed in AL and MM, with a higher frequency in AL (AL= 80 %, MM= 53 %, p= 0.002) (Figure 11A). 10 % of the MM patients were assigned to *IGKV2*, whereas no AL patient did. *IGKV3* was higher in MM (MM= 30 %, AL= 10 %, p= 0.014) and both cohorts did not differ in *IGKV4* (AL= 5 %, MM= 6 %) and *IGKV5* (AL= 2 %, MM= 1 %) expression. Among the *IGKV*-subgroups in AL and MM (Figure 11B), *IGKV1/D-33* was significantly more represented in AL (32 % vs.

13 %, p= 0.016), with a rr of 2.452 (p= 0.013). *IGKV1-16* was only identified in the AL cohort (10 %, p= 0.010, rr= 18.439, p= 0.049). *IGKV1-5* and *IGKV1/D-39* was highly represented in AL and MM (*IGKV1-5*: AL= 12 %, MM= 17 %; *IGKV1/D-39*: AL= 20 %, MM= 17 %). IGKV3/D-20 was only detected in MM (11 %, p= 0.030) and IGKV3/D-11 was detected as one of the most abundant subfamilies in MM (MM= 14 % vs. AL= 2 %, rr= 0.173). Analysis of AL patients with respect to their clinical organ presentation in relation to their *IGKV* usage (Figure 12B) revealed an association of *IGKV1/D-33* and heart involvement (75 %, p= 0.024) and *IGKV1/D-39* was more frequently in patients with dominant kidney involvement (36 %), but without reaching statistical significance. Because of the higher frequency of liver involvement in kappa AL (Comenzo et al. 2001) liver involvement in combination with *IGKV* usage was investigated separately (Table 13) and a trend towards *IGKV1-16* was observed. 27 % of AL with liver involvement were assigned to *IGKV1-16*, while this was only the case for 3 % of patients without liver involvement. In addition to organ tropism in AL, the influence of a HC in serum on the *IGKV* usage of the three most abundant *IGKV* representatives in AL, *IGKV1-5*, *IGKV1/D-33* and *IGKV1/D-39* was investigated (Figure 13). The presence of a HC had an impact on the *IGKV1-5* usage in both, AL and MM. *IGKV1-5* was four times higher in AL patients without a HC (AL_noHC= 20 %, AL_HC= 5 %) and a threefold difference was found in MM (MM_noHC= 40 %, MM_HC= 14 %). Besides the analysis of the *IGKV* assignment, the associated *IGKJ* usage was investigated to draw conclusions about a potentially specific *IGKV-IGKJ* assembly. Overall, *IGKJ2* was most observed in the AL cohort (39 %) and *IGKJ1* (30 %) as well as *IGKJ2* (29 %) in MM (Table 14). *IGKJ3* was higher in MM (13 % vs. 5 %) and *IGKJ5* was higher in AL (10 % vs. 5 %). 30 % of MM and 17 % of AL patients were assigned to *IGKJ1*. Considering AL patients for their clinical presentation (Table 14), *IGKJ2* was primarily associated with dominant cardiac involvement (63 %) and *IGKJ1*, *IGKJ2* and *IGKJ3* was equally presented (29 %) in AL_K patients. *IGKJ2* and *IGKJ4* were predominantly detected in the AL_D group (37 %). Analysis of the linkage of the most frequent *IGKV*-subfamilies to *IGKJ* (Table 15) revealed *IGKV1-5* was most frequently linked to *IGKJ2* in AL (16 %) and MM (8 %). And *IGKV1/D-33* was most frequently linked to *IGKJ2* (17 %) and no dominant linkage was present in MM. No clear linkage was found for *IGKV1/D-39*. Regarding organ involvement in this context, *IGKV1-5* linked to *IGKJ2* was represented by AL patients with dominant cardiac involvement (13 %) and AL_D patients (16 %). None of the patients with dominant kidney involvement showed this linkage.

IGKV1/D-33 linked to *IGKJ2* was the major association for patients with dominant heart involvement (50 %). And *IGKV1/D-39* was equally linked with *IGKJ1* (14 %) and *IGKJ2* (14 %) in patients with dominant kidney involvement.

3.1.9 Summary detailed analysis of most common *IGKV*-subfamilies in AL and comparative analysis of one AL and MM representative

In a further step, the three most common representatives of the AL *IGKV*-subfamilies, *IGKV1/D-33* (32 %), *IGKV1/D-39* (20 %) and *IGKV1-5* (12 %) were studied in detail. For this purpose, the median mutation count, mutation frequency and AA composition were investigated. In order to gain knowledge about the potential stability of the LCs, the theoretical pI, the GRAVY-score and AGG-score were calculated. Furthermore, a detailed analysis of the individual variants was performed to analyse mutations hotspots as well as the effect of mutations on the AA charges. Beyond this, a comparative analysis of *IGKV1-16*, as an exclusive representative of AL (10 %) and *IGKV3/D-11*, as one of the most frequent MM *IGKV*-subfamily (14 %) but detected in only 2 % of AL patients, was performed to gain insight into a potentially divergent amyloidogenic behaviour.

Table 34. Summary of the detailed analysis of the most common *IGKV*-subfamilies in AL.

	IGKV1/D-33		***IGKV1/D-39***		***IGKV1-5***		***IGKV1-16***	***IGKV3/D-11***
	AL n= 13	**MM n= 11**	**AL n= 8**	**MM n= 14**	**AL n= 5**	**MM n= 14**	**AL n= 4**	**MM n= 12**
median mutation count *IGKV/J*	IGKV: 12.0 ↑ FR3: 4.0 ↑ HC (IGKV): 12.0 vs. 10.5	IGKV: 10.0 FR3: 3.5 ↑	IGKV: 14.0 ↑ FR1: 2.0 ↑* FR3: 3.5 ↑* noHC (IGKV): 14.0 ↑ vs. 12.0	IGKV: 9.5 FR1: 1.0 FR3: 2.0	IGKV: 6.0	IGKV: 7.5	IGKV: 9.0 ↑* FR3: 3.0 ↑	IGKV: 5.5
mutation frequency *IGKV/J*	FR1: 69 % ↑	FR1: 36 %	HC (IGKJ): 100 % ↑ noHC: 50 %	higher, exception of CDR3	FR1: 80 % ↑ CDR2: 40 %	FR1: 43 % CDR2: 71 % ↑	FR1: 50 % ↑* higher in all other FRs and CDRs	FR1: 0 %
AA composition	Asn ↑* HC: Ser ↓*	-	-	-	-	noHC: Tyr ↑*	-	-

	IGKV1/D-33		***IGKV1/D-39***		***IGKV1-5***		***IGKV1-16***	***IGKV3/D-11***
	AL n= 13	**MM n= 11**	**AL n= 8**	**MM n= 14**	**AL n= 5**	**MM n= 14**	**AL n= 4**	**MM n= 12**
median pI	4.90	4.87	7.06 HC: 6.37 vs. 7.61	7.65	5.15	7.07	delta= 0.30 HC: 8.13 vs. 7.10	delta= 0.00
median GRAVY	-0.418	-0.434	-0.387	-0.346	-0.412	-0.407	delta= -0.017	delta= -0.019
median AGG	499.374	458.062	691.521	721.383	555.625	650.341	delta= -62.938 noHC: 942.376 ↑ vs. 584.495	delta= -29.932
mutation hotspots	2x FR3 (65S, 83I) 1x CDR3 (93N)	2x CDR3 (93N, 101)	3x CDR1 (28S, 30S, 31S) 1x CDR2 (53S) 2x CDR3 (93S, 94T)	2x CDR1 (28S, 31S) 1x CDR2 (50A) 1x CDR3 (97)	1x FR1 (21I) 1x CDR1 (31S) 1x FR2 (45K) 1x CDR3 (92N)	2x CDR1 (30S, 31S) 1x FR2 (45K) 1x CDR2 (53S)	1x CDR1 (28G, 32Y) 1x CDR2 (57Q) 1x CDR3 (97S) 2x FR2 (37F, 43K) 3x FR3 (72D, 80S, 83P) 1x IGKJ (109E)	2x CDR1 (30S, 31S)
additional charge	loss or reversal of charge	loss or reversal of charge	positive charges ↑	positive charges ↑	CDR1: positive charges ↑ FR2: loss of charges CDR3: negative charges ↑	CDR1: positive charges ↑ FR2: loss of charges CDR3: positive and negative charges	CDR1: negative and positive charges ↑ CDR2: positive charges ↑ FR3: loss of charges	CDR1: negative and positive charges ↑
Potential N-glycosylation site [%]	61.5	9.0	25.0	28.6	20.0	21.4	75.0	8.3
Relative risk	2.452*		1.186		0.741		18.440*	0.173

AL= AL amyloidosis, MM= multiple myeloma, pI= isoelectric point, GRAVY= grand average of hydropathy, AGG= β-sheet aggregation, mutation hotspots were defined as positions which were mutated in ≥50 % of cases, *a p-value ≤ 0.05 was considered to be significant.

The main findings are summarized in Table 34. Sequences assigned to *IGKV1/D-33*, representing the highest AL *IGKV1*-subfamily (AL= 32 %, MM= 13 %, p= 0.016), showed a higher median mutation count of the *IGKV*-region for AL (median mutation count: AL= 12.0 vs. MM= 10.0) with the highest in FR3 for both cohorts (median mutation count: AL= 4.0, MM= 3.5). If stratifying the AL sequences for the presence of a HC, the sequences without a HC in serum has a lower median mutation count (AL_noHC= 10.5) as the sequences with a HC (AL_HC= 12.0). AL sequences had a higher mutation frequency in the FR1 than the MM (mutation frequency: AL= 69 % vs. MM= 36 %). Calculation of the AA composition revealed a significantly higher median proportion of asparagine (median: AL= 4.6 vs. MM= 4.0 %, p= 0.007) and considering AL sequences according to the presence of a HC in serum, serine was significantly lower in AL with detectable HC (median: AL_HC= 13.9 % vs. AL_noHC= 14.9 %, p= 0.012). The AL and MM cohort did not differ in the pI (AL= 4.90, MM= 4.87) as well as the GRAVY- (AL= -0.418, MM= -0.434) and AGG-score (AL= 499.374, MM= 458.062). AL sequences had more mutation hotspots than MM (3 vs. 2), localized in FR3 and CDR3 for AL and only in CDR3 for MM with sharing one mutation hotspot in the CDR3 (93N). In both cohorts mutations resulted in reversal or loss of charge. For example, at position 70D, next to the mutation hotspot 65S, AA exchanges led to a reversal of charge in two cases (2x D70H) and a loss of charge in three cases (3x D70N). In MM, three mutations resulted in a loss or reversal of charge (D70H, D70G, D70N). 61.5 % of AL sequences belonging to *IGKV1/D-33* showed potential N-glycosylation sites and only 9.0 % of the MM did.

Analysis of *IGKV1/D-39* (Table 34), which is highly represented in AL and MM (AL= 20 %, MM= 17 %) also revealed a higher median mutation count of the *IGKV*-region in AL compared to MM (median mutation count: AL= 14.0, MM= 9.5), in particular FR1 (median mutation count: AL= 2.0, MM= 1.0, p= 0.020) and FR3 (AL= 3.5, MM= 2.0, p= 0.017) were significantly more mutated in AL. But MM, with exception of CDR3, had a higher mutation frequency than AL. In AL, patients without a HC had a higher median mutation count than patients with a HC (median mutation count: AL_noHC= 14.0 vs. AL_H= 12.0). Both cohorts had a neutral pI (median pI: AL= 7.06, MM= 7.65) and when considering AL sequences with respect to the presence of a HC, a pI difference of +1,0 was detected for sequences without a HC (median pI: AL_HC= 6.37, AL_noHC= 7.61), which might indicate changes in solubility.

No differences were found for the GRAVY- (AL= -0.387, MM= -0.346) and AGG-score (AL= 691.521, MM= 721.383). AL and MM showed mutation hotspots in the CDR1, CDR2 and CDR3 with higher frequency for AL (6 vs. 4). And mutation hotspot 28S and 31S was shared by both. Mutations close to mutation hotspots resulted in additional charges, with positive charges being dominant. 25 % of the *IGKV1/D-39* AL sequences showed potential N-glycosylation sites and 28.6 % of the MM did.

IGKV1-5 (Table 34), detected in 12 % of the AL and in 17 % of the MM cohort had a median mutation count of 6.0 for AL and 7.5 for MM. Regarding the mutation frequency in both cohorts differences were found for FR1 and CDR2. In AL 80 % of the sequences were mutated and in contrast only 43 % of the MM sequences did. But vice versa MM sequences showed higher mutation frequency in CDR2 (mutation frequency AL= 40 % vs. MM= 71 %). Given that only one of the four AL sequences had a HC, the analysis regarding the presence of a HC for AL was limited. In contrast, in MM ten patients had a HC and four did not. Analysing the AA composition of these two groups, significant differences in tyrosine expression were found. MM patients without a HC had a significantly higher amount than patients with a HC (MM_HC= 4.7 %, MM_noHC= 5.3 %, p= 0.011). Differences were also found in the pI and AGG-scores of AL and MM patients, but not in the GRAVY-score. AL patients had a lower pI (AL= 5.15, MM= 7.07) and also a lower AGG-score (AL= 555.625, MM= 650.341), but the GRAVY-score was similar (AL= -0.412, MM= -0.407). AL sequences showed four mutation hotspots with one mutation hotspots in FR1, CDR1, FR2 and CDR3, in each case. The MM sequences also showed four mutation hotspots, but with two in CDR1, one in CDR1, one in FR2 and one in CDR2. Mutation hotspot 31S and 45 K was detected in both cohorts. In CDR1, next to the mutation hotspot 30S and 31S, mutations resulted in more positive charges in AL and MM. In contrast, at position 45K, a loss of charge was observed in three of the five AL and in six of the 14 MM sequences. Interestingly, in CDR3, mutations only caused additional negative charges in AL, in contrast to MM, where both negative and positive charges were inserted. The amount of potential N-Glycosylation sites between AL and MM was approximately the same (AL= 20.0 %, MM= 21.4 %).

In a next step, a comparative analysis of *IGKV3/D-11*, representing one of the most common MM subfamilies (MM= 14 % vs. AL= 2 %) and *IGKV1-16*, only found in AL

(AL= 10 %), was performed (Table 34). The *IGKV*-segment of the *IGKV1-16* AL sequences showed a significantly higher median mutation count than *IGKV3/D-11* MM sequences (median mutation count: AL= 9.0, MM= 5.5, p= 0.033). The *IGKV3/D-11* MM sequences showed no mutations in the FR1, but 50 % of the *IGKV1-16* AL sequences did (mutation frequency: AL= 50 %, MM= 0 %, p= 0.050). Overall, the *IGKV1-16* AL sequences showed higher mutation frequencies in all FR- and CDR-regions compared to the *IGKV3/D-11* MM sequences with the highest median mutation count in FR3 for AL. Analysis of the pI showed that mutations detected in *IGKV3/D-11* MM sequences had no influences on the delta pI (median delta pI: MM (*IGKV3/D-11*)= 0.00 vs. AL (*IGKV1-16*)= 0.30). *IGKV1-16* AL sequences without a HC had a lower pI than the sequences with a HC (median pI: AL (*IGKV1-16*_HC)= 8.13 vs. AL (*IGKV1-16*_noHC)= 7.10). And when stratifying the AL sequences again for a HC, the *IGKV1-16* AL sequences without a HC showed a much higher AGG-score than the *IGKV1-16*_HC sequences (median AGG: AL (*IGKV1-16*_noHC)= 942.376 vs. AL (*IGKV1-16*_HC)= 584.495). *IGKV1-16* AL sequences showed nine mutation hotspots in comparison to *IGKV3/D-11* MM sequences with showing only two. Mutation hotspots detected in the *IGKV1-16* AL sequences covered CDR1, CDR2, CDR3, FR2 and FR3 and one position in the *IGKJ*-region. *IGKV3/D-11* MM sequences had only mutation hotspots in the CDR1. Regarding the effect of mutations on the charge of the AAs, in both AL and MM sequences, mutations in CDR1 resulted in additional negative and positive charges. In CDR2 (position 57Q), the mutation towards histidine introduced positive charges in three of the four *IGKV1-16* AL sequences. And at mutation hotspot 72D, AA exchanges towards valine and asparagine resulted in a loss of charge. 75 % of the *IGKV1-16* sequences revealed potential N-glycosylation sites whereas only 8.3 of the *IGKV3/D-11* MM did.

3.1.10 Summary analysis of the prediction of light chain aggregation using V_LAmY-Pred- algorithm

Nowadays, the application of machine learning-based tools is becoming increasingly important in life sciences and in the medical contexts. V_LAmY-Pred is a tool developed to predict the aggregation potential of individual LC sequences and allows the analysis of both lambda and kappa sequences (Rawat et al. 2021). The algorithm was tested on the LC sequences studied in detail (Table 33). 40 % of *IGKV1-5* AL sequences were

classified as amyloidogenic and 64 % of MM patients as non-amyloidogenic. Comparable for *IGKV1/D-39*, where 50 % of the AL were classified as amyloidogenic and 79 % of the MM sequences as non-amyloidogenic. This was different for *IGKV1/D-33*, where all AL and MM sequences were classified to be amyloidogenic. 50 % of the *IGKV1-16* AL sequences were categorised as amyloidogenic and all of the *IGKV3/D-11* MM sequences as non-amyloidogenic.

3.2 Lambda cohort

3.2.1 Clinical data of the lambda patient cohort

As part of this project, 34 lambda AL patients with dominant kidney involvement and 18 lambda patients classified as diverse patients were studied. The diverse group of patients included all lambda patients who did not have dominant heart involvement, dominant kidney involvement or equal heart and kidney involvement. For comparison purpose 52 lambda MM patients were also included in this project. The extraction of the clinical data of the 52 lambda MM patients as well as the verification of the LC sequences was not part of this project, data were provided by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and used in this project for comparison purpose. In the framework of this thesis clinical data of the 34 lambda AL patients were obtained from the physician reports and laboratory findings. Table 35 shows the clinical data of the lambda AL and MM patients. Characteristics including age, sex and organ involvement were assessed by the treating physician. Laboratory findings such as AP, quantity of free LCs or M-gradient were evaluated by the central laboratory of the University Hospital Heidelberg and plasma cell infiltration was taken from the bone marrow laboratory report ("KM-Labor MED V", University Hospital Heidelberg). AL patients were staged according to established variables and cut-offs (Dispenzieri et al. 2004; Palladini et al. 2014; Wechalekar et al. 2013). All three patient groups had a comparable age at diagnosis (median age between 59 and 61 years). In AL_K there were 15 woman and 19 men, in AL_D 12 woman and six men and in MM 22 woman and 30 men. Bone marrow was collected from 30 of the AL_K patients at the time of new diagnosis and from four patients at time of relapse. For AL_D patients, bone marrow was aspirated in 16 patients at the time of new diagnosis and at the time of relapse in two cases. All MM patients provided their bone marrow at the time of new diagnosis. Apart from the organs most frequently affected in AL, the heart and kidney, soft tissue (n= 10), followed by the liver (n=9), were highly involved. AL_K patients had a lower median dFLC than the AL_D (median dFLC: AL_K= 58 mg/L vs. AL_D= 239 mg/L, $p< 0.001$) and MM patients (median dFLC: AL_K= 58 mg/L vs. MM= 411 mg/L, $p< 0.001$). The AL_K and MM patients did not differ in the median AP (median AP: AL_K= 71 U/L, MM= 70 U/L), but the AL_D patients had a higher (median AP: AL_D= 122 U/L vs. AL_K= 71 U/L, $p< 0.001$, AL_D= 122 U/L vs. MM= 70 U/L, $p< 0.001$). A M-gradient was detected more

frequently in MM (M-gradient: MM= 45/52 vs. AL_K= 15/33, p< 0.001, MM= 45/52 vs. AL_D= 9/17, p= 0.014). MM had a higher median plasma cell infiltration than AL (median plasma cell infiltration: MM= 60 % vs. AL_K= 7 %, p< 0.001, MM= 60 % vs. AL_D= 11 %, p< 0.001) and AL_D had a higher than AL_K (median plasma cell infiltration: AL_D= 11 % vs. AL_K= 7 %, p= 0.004). More MM patients had a clonal heavy chain in the serum (MM= 41/52 vs. AL_K= 17/34, p= 0.018 or AL_D= 11/18) and IgG was the most common in all lambda patients (IgG: MM= 33/52, AL_K= 10/34, AL_D= 6/18). A t(11;14) was detected in more AL than in MM patients (t(11;14): AL_K= 21/31 vs. MM= 10/52, p< 0.001, AL_D= 7/12 vs. MM= 10/52, p= 0.022). And a 1q21 was detected in 6/31 AL_K, 3/12 AL_D and 15/52 MM patients.

Table 35. Clinical characteristics of the 52 lambda AL amyloidosis and 52 lambda multiple myeloma patients.

	AL_K (n= 34)	**AL_D (n= 18)**	**MM (n= 52)**
Age [y], median (range)	59 (38-77)	62 (49-76)	61 (37-70)
Sex female/male [n]	15/19	12/6	22/30
Newly diagnosed [n]	30/34	16/18	52/52
Relapse/progress [n]	4/34	2/18	0/52
No. of organs involved, no. of patients			
1	24	2	
2	7	4	
3	2	6	
4	1	3	
More than 4	0	2	
different organs involved [n]			
Heart	6/34	15/18	
Kidney	34/34	9/18	
Liver	1/34	8/18	
ANS	0/34	3/18	
PNS	2/34	4/18	
GI	3/34	8/18	
ST	2/34	8/18	
dFLC [mg/L], median (range)[I]	58 (3-743)	239 (7-1276)	411 (2-16344)[3]
Cardiac stage[1]			
I	15/33	2/17	
II	9/33	1/17	
IIIa	8/33	9/17	
IIIb	1/33	5/17	
Renal stage			
I	12/34	10/14	
II	16/34	3/14	
III	6/34	1/14	
AP [U/L] median (range)[2 II]	71 (34-201)	122 (47-1887)[3]	70 (32-247)[4]
M-gradient [n][III]	15/33	9/17	45/52
Plasma cell infiltration, median [%] (range)[IV]	7 (2-16)[3]	11 (6-30)[5]	60 (5-100)[4]

	AL_K (n= 34)	AL_D (n= 18)	MM (n= 52)
Monoclonal protein in the serum [n][V]	17/34	11/18	41/52
IgG	10/34	6/18	33/52
IgA	5/34	2/18	7/52
IgM	1/34	2/18	1/52
IgD	1/34	1/18	0/52
Light chain only	17/34	7/18	11/52
t(11;14) [n][VI]	21/31	7/12	10/52
1q21 [n]	6/31	3/12	15/52

AL= AL amyloidosis, MM= multiple myeloma, AL_H= AL patients with dominant heart involvement, AL_K= AL patients with dominant kidney involvement, AL_D= AL patients with diverse organ involvement, ANS= autonomic nervous system, PNS= peripheral nervous system, GI= gastrointestinal tract, ST= soft tissue, dFLC = difference between disease-associated und uninvolved circulation free light chains, AP = alkaline phosphatase, n/a= not available, cardiac and renal staging was done according to (Dispenzieri et al. 2004; Palladini et al. 2014; Wechalekar et al. 2013). The extraction of the clinical data of the MM patients were provided by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and used in this thesis for comparison purpose.
[1] EURO score, [2]Reference range obtained from the laboratory results: AP 40-130 U/L, [3] 1 n/a, [4] 2 n/a, [5] 3 n/a
[I] AL_K vs. AL_D p< 0.001, AL_K vs. MM p< 0.001
[II] AL_K vs. AL_D p< 0.001, AL_D vs. MM p<0.001
[III] AL_K vs. MM p< 0.001, AL_D vs. MM p= 0.014
[IV] AL_K vs. AL_D p= 0.004, AL_K vs. MM p< 0.001, AL_D vs. MM p< 0.001
[V] AL_K vs. MM p= 0.018
[VI] AL_K vs. MM p<0.001, AL_D vs. MM p= 0.022

3.2.2 *IGL*-family usage of the lambda cohort

3.2.2.1 *IGLV* region in lambda AL amyloidosis and multiple myeloma

Lambda AL and MM patients were assigned to their corresponding *IGLV* families using Vbase2 (cDNA level) and Ensembl Blast (AA level). Figure 19 presents the family usage of the AL and MM lambda patient cohort.

IGLV1 was comparatively represented in AL and MM (AL= 29 %, MM= 27 %) (Figure 19A). 19 % of AL sequences showed *IGLV2* and 33 % of MM sequences. *IGLV3* was detectable in 25 % of AL patients and in 33 % of MM. Significant differences were observed in the expression of *IGLV6*, with 19 % of AL patients having *IGLV6* and only 2 % of MM (p= 0.004). *IGLV7* was presented only in MM (2 %). Four AL (8 %) and 2 MM patients (4 %) could not be classified and were defined as non-evaluable.

In terms of *IGLV*-subgroup usage (Figure 19B), significant differences were found in the expression of *IGLV2-23* and *IGLV6-57*. *IGLV2-23* could only be detected in MM (15 %, p= 0.006) and *IGLV6-57* was detected in 19 % of AL and only 2 % of MM patients

(p= 0.004). Differences were also observed for *IGLV3-21* (AL= 12 %, MM= 21 %), but without reaching statistical significance. The following subfamilies were detected only in AL: *IGLV2-8* (AL= 4 %), *IGLV3-9* (AL= 2 %), whereas *IGLV2-11* (MM= 2 %) and *IGLV7-46* (MM= 2 %) were detected only in MM. No differences were found for *IGLV1-40* (AL= 4 %, MM= 4 %), *IGLV1-44* (AL= 17 %, MM= 17 %), *IGLV1-47* (AL= 2 %, MM= 2 %), *IGLV2-14* (AL= 15 %, MM= 15 %), *IGLV3-25* (AL= 2 %, MM= 2 %). *IGLV1-51* was detected in 6 % of AL and in 4 % of MM patients, *IGLV3-1* in 8 % of AL and 6 % of MM patients, *IGLV3-19* in 2 % of AL and in 4 % of MM patients.

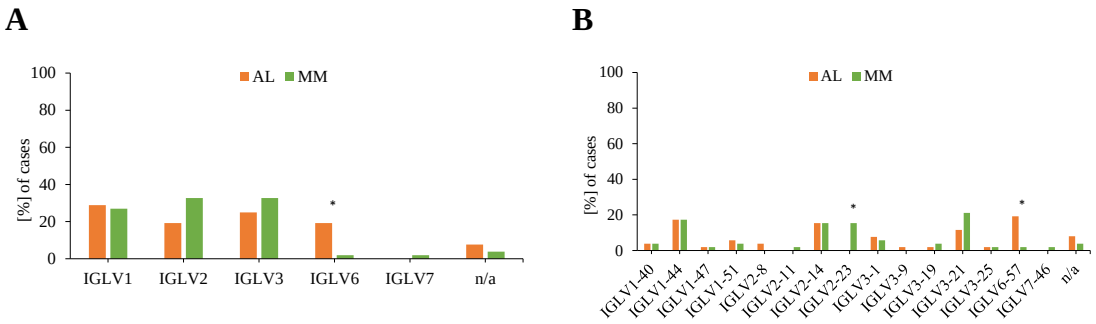

Figure 19. Variable region gene usage in lambda AL amyloidosis and multiple myeloma patients. AL= AL amyloidosis, MM= multiple myeloma, AL n= 52, MM n= 52, *p-values ≤ 0.05 were considered to be significant, n/a= not available **A** *IGLV*-family usage in AL and MM, **B** *IGLV*-subgroups in AL and MM. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

Afterwards, the influence of organ tropism and the assignment to the *IGLV*-family was investigated (Figure 20).

IGLV1 and *IGLV6* was higher represented in AL_K than in AL_D patients, with 35 % of *IGLV1* in AL_K and 17 % in AL_D and 24 % of *IGLV6* in AL_K and 11 % in AL_D (Figure 20A). *IGLV2* was detected in 18 % of AL_K and in 22 % of AL_D patients and *IGLV3* was detected in 24 % of AL_K and in 28 % of AL_D patients. 4 AL_D patients (22 %) could not be assigned to any *IGLV* family.

Analysis of AL_K and AL_D patients for *IGLV*-subfamilies (Figure 20B) showed that *IGLV1-44* was the highest *IGLV1* representative for AL_K (21 %) and AL_D (11 %). *IGLV1-47* and *IGLV1-51* were only detected in AL_K patients (*IGLV1-47*: AL_K= 3 %, *IGLV1-51*: AL_K= 9 %) and *IGLV1-40* was detected in 3 % of AL_K and in 6 % of

AL_D patients. *IGLV2-14* was detected in 15 % of AL_K and in 17 % of AL_D patients. And *IGLV2-8* was detected in 3 % of AL_K and in 6 % of AL_D patients. *IGLV3-21* was the most frequent *IGLV3*-subgroup in AL_K and AL_D, with 12 % in AL_K and 11 % in AL_D. *IGLV3-9* and *IGLV3-19* were only present in AL_D (6 % each) and *IGLV3-25* only in AL_K (3 %). 9 % of AL_K and in 6 % of AL_D patients had *IGLV3-1*. *IGLV6-57* was identified in 24 % of AL_K and in 11 % of AL_D patients. Accordingly, *IGLV6-57* was the most common subfamily of all AL patients with dominant kidney involvement (24 %) and *IGLV2-14* of all patients classified as diverse (17 %).

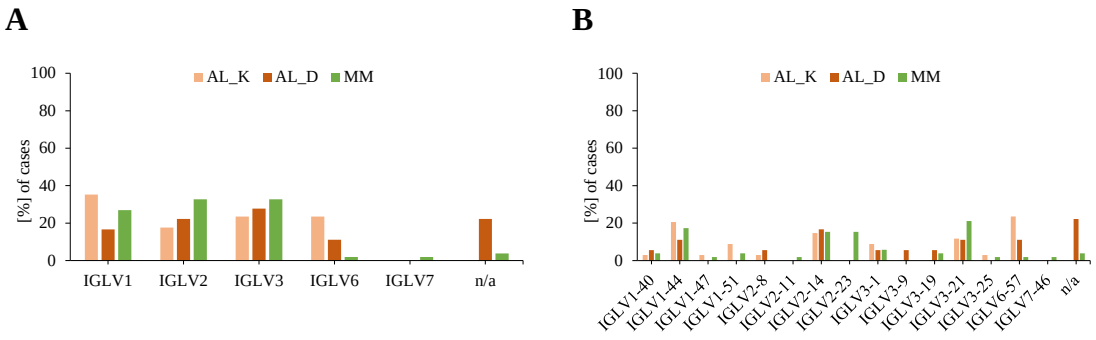

Figure 20. Variable region gene usage in lambda AL amyloidosis with different clinical presentation and multiple myeloma patients. AL= AL amyloidosis, MM= multiple myeloma, AL n= 52, MM n= 52, AL_K= AL patients with dominant kidney involvement, AL_D= AL patients with diverse organ involvement, AL_K n= 34, AL_D n= 18, **A** *IGLV*-family usage in AL (organ tropism) and MM, n/a= not available **B** *IGLV*-subgroups in AL (organ tropism) and MM. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.2.2 *IGLJ/C* usage in lambda AL amyloidosis and multiple myeloma

In addition to the analysis of *IGLV*, the respective *IGLJ* and *IGLC* usage was also studied (Table 36). Overall, the association between *IGLJ2* and *IGLC2* was most frequent in AL (27 %) and in MM (42 %). *IGLJ1* with *IGLC1* were more frequently linked in AL (15 %) than in MM (10 %), similarly for *IGLJ3* with *IGLC2* (AL= 8 %, MM= 2 %) and *IGLJ3* with *IGLC3* (AL= 19 %, MM= 13 %). *IGLJ2* associated with *IGLC3* was only detected in AL (2 %) and *IGLJ7* associated with *IGLC7* was only detected in 2 % of MM.

Regarding the different organ presentation in AL (Table 36), the sequences of AL_K and AL_D patients did not differ substantially in the presence of *IGLJ1* and *IGLC1* (AL_K= 15 %, AL_D= 17 %) and *IGLJ2* and *IGLC2* (AL_K= 26 %, AL_D= 28 %)

linkage. *IGLJ2* linked to *IGLC3* was only detected in AL_D (6 %) and *IGLJ3* linked to *IGLC2* was only detected in AL_K (12 %). 21 % of AL_K and 17 % of AL_D patients showed *IGLJ3/C3* linkage.

Table 36. Analysis of the *IGLJ/C* usage in AL amyloidosis and multiple myeloma.

	IGLJ1		*IGLJ2*			*IGLJ2/3*			*IGLJ3*			*IGLJ7*	n/a
	IGLC1	*IGLC1/2*	*IGLC2*	*IGLC2/3*	*IGLC3*	*IGLC2*	*IGLC2/3*	*IGLC3*	*IGLC2*	*IGLC2/3*	*IGLC3*	*IGLC7*	
AL_K [%] (n= 34)	15	3	26	6	0	9	0	0	12	9	21	0	0
AL_D [%] (n= 18)	17	0	28	0	6	0	0	0	0	11	17	0	22
AL, whole cohort [%] (n= 52)	15	2	27	4	2	6	0	0	8	10	19	0	8
MM [%] (n= 52)	10	0	42	8	0	6	2	4	2	8	13	2	4

AL= AL amyloidosis, MM= multiple myeloma, AL=K= AL patients with dominant kidney involvement, AL_D= diverse AL patients, n/a= not available. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.2.3 Linkage between the most frequent *IGLV*-subfamilies - and *IGLJ/C* in lambda AL amyloidosis and multiple myeloma

Analysis of the association of the most common *IGLV*-subfamilies with *IGLJ/C* (Table 37) revealed that *IGLV1-44* was equally associated with *IGLJ2/C2* (4 %), *IGLJ3/C2* (4 %) and *IGLJ3/C3* (4 %) in AL. *IGLV2-14* was most frequently linked to *IGLJ1/C1* (6 %) and to *IGLJ2/C2* (6 %) in AL and to *IGLJ2/C2* (10 %) in MM. In AL, *IGLV6-57* was most frequently associated with *IGLJ2/C2* (6 %) and *IGLJ3/C3* (6 %). Only one of the MM sequences showed *IGLV6-57*, which was associated with *IGLJ3/C3* (2 %).

Regarding organ involvement in AL (Table 37), all patients with *IGLV1-44* and *IGLJ2/C2* had dominant kidney involvement (AL_K= 6 %). 3 % of AL patients with dominant kidney involvement presented *IGLV1-44* linked to *IGLJ3/C2*, *IGLJ3/C2/C3* and *IGLJ3/C3*, respectively. Whereas 6 % of diverse patients had *IGLV1-44* associated with *IGLJ3/C2* or *IGLJ3/C3*. AL patients with *IGLV2-14* and *IGLJ1/C1* were more likely to

have diverse organ involvement (AL_D= 11 %) than dominant kidney involvement (AL_K= 3 %). Patients with *IGLV2-14* and *IGLJ2/C2* did not differ in their organ presentation (AL_K= 6 %, AL_D= 6 %). *IGLV6-57* associated with *IGLJ2/C2* was equally present in AL_K (6 %) and in AL_D (6 %) patients. In contrast, all patients with *IGLV6-57* and *IGLJ3/C3* had dominant kidney involvement (AL_K= 9 %).

Table 37. Analysis of the *IGLJ/C* linkage of the most common *IGLV*-subfamilies in lambda AL and MM.

[%]	*IGLJ1*		*IGLJ2*			*IGLJ2/3*		*IGLJ3*			n/a
	IGL C1	*IGL C1/C2*	*IGL C2*	*IGL C2/3*	*IGL C3*	*IGL C2*	*IGL C2/C3*	*IGL C2*	*IGL C2/C3*	*IGL C3*	
IGLV1-44											
AL_K (n= 34)	0	0	6	0	0	3	0	3	3	3	0
AL_D (n= 18)	0	0	0	0	0	0	0	6	6	6	22
AL, whole cohort (n= 52)	0	0	4	0	0	2	0	4	4	4	8
MM (n= 52)	2	0	2	4	0	2	2	2	2	2	4
IGLV2-14											
AL_K (n= 34)	3	3	6	0	0	0	0	0	0	3	0
AL_D (n= 18)	11	0	6	0	0	0	0	0	0	0	22
AL, whole cohort (n= 52)	6	2	6	0	0	0	0	0	0	2	8
MM (n= 52)	2	0	10	0	0	0	0	0	4	0	4
IGLV6-57											
AL_K (n= 34)	0	0	6	0	0	0	0	3	6	9	0
AL_D (n= 18)	0	0	6	0	6	0	0	0	0	0	22
AL, whole cohort (n= 52)	0	0	6	0	2	0	0	2	4	6	8
MM (n=52)	0	0	0	0	0	0	0	0	0	2	4

AL= AL amyloidosis, MM= multiple myeloma, AL_K= AL patients with dominant kidney involvement, AL_D= diverse AL patients, n/a= not available. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.2.4 Comparison of *IGLV6-57* AL- and *IGLV2-23* MM-sequences

In the following, the AL sequences of the *IGLV6-57*-subfamily were compared with the MM sequences of the *IGLV2-23*-subfamily. *IGLV6-57* was with about 19 % the most frequent subfamily in AL (vs. 2 % MM, p= 0.004). And *IGLV2-23* were only represented in MM (15 %, p= 0.006). Based on this divergent behaviour, both subgroups were analysed in detail to draw conclusions about a possible different amyloidogenic potential. Ten *IGLV6-57* AL patients (Table 38) were compared with eight *IGLV2-23* MM patients (Table 39). Only the matching regions of the *IGLV6-57* AL and *IGLV2-23* MM sequences were compared to ensure comparability. Sequencing of *IGLV6-57* (Table 38) revealed clear sequences, with unambiguous signals in seven cases. FOR207 showed heterogeneous signals in four positions, FOR221 in seven positions and FOR151 in 11 positions. Eight of the patients had dominant kidney involvement and two of the patients were classified as diverse. Three of the *IGLV6-57* AL patients were assigned to *IGLJ2/C2* and one to *IGLJ2/C3*. *IGLJ3/C2* was assigned to one and *IGLJ3/C3* to three patients. No clear distinction of *IGLC2* and *IGLC3* was possible in two of the patients and both options were considered in the following analysis. Six of the patients had a HC in serum, with 3x IgG and 1x IgM, IgD or IgA. Two of the *IGLV2-23* MM patients (Table 39) were assigned to *IGLJ2/C2* and three to *IGLJ3/C3*. In three of the patients, no clear differentiation between *IGLJ2* and *IGLJ3* was possible and like for *IGLV6-57* both options were considered in the analysis. Seven of the patients had a HC in serum, with 5x IgG and 2x IgA.

Table 38. Overview about *IGLV6-57* AL amyloidosis patients.

	Disease	X AA [n]	Organ involvement	*IGLJ*	*IGLC*	HC in serum
FOR125	AL	0	Kidney	*IGLJ3*	*IGLC3*	IgM
FOR143	AL	0	Kidney	*IGLJ2*	*IGLC2*	0
FOR158	AL	0	Kidney	*IGLJ3*	*IGLC3*	IgG
FOR199	AL	0	Kidney	*IGLJ3*	*IGLC3*	0
FOR207	AL	4	Kidney	*IGLJ3*	*IGLC2/3*	IgD
FOR209	AL	0	Kidney	*IGLJ2*	*IGLC2*	IgG
FOR212	AL	0	Kidney	*IGLJ3*	*IGLC2*	IgA
FOR221	AL	7	Kidney	*IGLJ3*	*IGLC2/3*	0
FOR151	AL	11	Diverse	*IGLJ2*	*IGLC3*	IgG

	Disease	X AA [n]	Organ involvement	*IGLJ*	*IGLC*	HC in serum
FOR183	AL	0	Diverse	*IGLJ2*	*IGLC2*	0

AL= AL amyloidosis, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal.

Table 39. Overview about *IGLV2-23* MM patients.

	Disease	X AA [n]	Organ involvement	IGLJ	IGLC	HC in serum
MM107	MM	0	-	IGLJ2/3	IGLC2	IgA
MM114	MM	0	-	IGLJ3	IGLC3	IgG
MM117	MM	0	-	IGLJ2	IGLC2	0
MM125	MM	0	-	IGLJ3	IGLC3	IgG
MM132	MM	0	-	IGLJ2/3	IGLC2	IgG
MM133	MM	0	-	IGLJ3	IGLC3	IgA
MM135	MM	0	-	IGLJ2	IGLC2	IgG
MM149	MM	0	-	IGLJ2/3	IGLC3	IgG

MM= Multiple myeloma, X AA means the number of undefined amino acids, resulting from heterozygous sequencing signal without a clear dominant major signal. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.2.5 Median mutation count and frequency of *IGLV6-57* in AL amyloidosis and *IGLV2-23* in multiple myeloma

Next, the overall mutation distribution and frequency of *IGLV6-57* AL and *IGLV2-23* MM sequences were comparatively analysed (Table 40). For comparison purpose, only the C- and N-terminal overlapping regions were considered, resulting in a missing FR1- and beginning of CDR1-region (5 AA). Furthermore, it must be noted that the corresponding references differ in the length by one AA in the CDR1-region and in two AA in the FR3-region. The *IGLV2-23* MM sequences showed a significantly higher median mutation count of the *IGLV*-segment (median mutation count: *IGLV*: MM (*IGLV2-23*)= 10.0, AL (*IGLV6-57*)= 4.5, $p < 0.001$) with a higher mutation frequency than the AL sequences (mutation frequency: *IGLV*: MM (*IGLV2-23*)= 100 %, AL (*IGLV6-57*)= 80 %). Especially the CDR1- and FR2-regions were mutated more frequently in MM than in AL (mutation frequency: CDR1: MM (*IGLV2-23*)= 75 %,

AL (*IGLV6-57*)= 20 %, FR2: MM (*IGLV2-23*)= 88 %, AL (*IGLV6-57*)= 30 %, p= 0.025). Interestingly, one of the MM sequences showed a mutation in the *IGLC*-segment (median mutation count: *IGLC*: MM (*IGLV2-23*)= 13 %).

Stratification of AL *IGLV6-57* sequences by the presence of a HC in the serum (Table 40) revealed that the *IGLV*-segment was mutated in all cases in patients with a HC, but only in 75 % of patients with a loss of a HC. However, the *IGLJ*-segment of patients without a HC was mutated more frequently but with the same median mutation count (mutation frequency: *IGLJ*: AL *IGLV6-57*_noHC= 100 %, AL *IGLV6-57*_HC= 50 %; median mutation count: *IGLJ*: AL *IGLV6-57*_noHC= 1.0, AL *IGLV6-57*_HC= 1.0). Focusing on the level of the different regions, the most notable differences were found in the median mutation count of the CDR1-region (median mutation count: CDR1: AL *IGLV6-57*_HC= 1.0, AL *IGLV6-57*_noHC= 3.0).

Table 40. Median mutation count and mutation frequency of *IGLV6-57* and *IGLV2-23*.

			IGLV 71/70 AA							***IGLJ*** 12/12 AA	***IGLC*** 74/74 AA
	n			FR1	CDR1 8/9 AA	FR2 15/15 AA	CDR2 7/7 AA	FR3 34/32 AA	CDR3		
AL *IGLV6-57*	10	[%]	80	n/a	20	30*	80	80	70	60	0
		[n]	4.5*	n/a	2.0	1.0	1.5	1.0	1.5	1.0	-
AL *IGLV6-57*_HC	6	[%]	100	n/a	17	33	83	83	83	50	0
		[n]	5.0	n/a	1.0	1.5	2.0	1.0	1.0	1.0	-
AL *IGLV6-57*_noHC	4	[%]	75	n/a	25	25	75	75	50	100	0
		[n]	3.0	n/a	3.0	1.0	1.0	1.0	2.0	1.0	-
MM *IGLV2-23*	8	[%]	100	n/a	75	88*	88	100	100	75	13
		[n]	10.0*	n/a	2.5	1.0	2.0	2.5	2.5	1.0	1.0
MM *IGLV2-23*_HC	7	[%]	100	n/a	86	86	86	100	100	71	14
		[n]	10.0	n/a	2.5	1.0	2.0	2.0	3.0	1.0	1.0
MM *IGLV2-23*_noHC	1	[%]	-	n/a	-	-	-	-	-	-	-
		[n]									

[%]= percentage of mutated segments, [n]= median mutation count. For the CDR3-region, the patient-specific linker and the first two AA of the J-segment were included and the AA length was between 9 and 11 AA. All values are given as medians. *a p-value ≤ 0.05 was considered to be significant. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.2.6 Biophysical properties of *IGLV6-57* AL amyloidosis and *IGLV2-23* multiple myeloma sequences

The biophysical properties of the two subfamilies *IGLV6-57*, representative of lambda AL, and *IGLV2-23*, representative of lambda MM, were investigated below (Table 41). As the comparison is between two different subfamilies, the following analysis focuses on the differences to the respective delta values (reference values). If the analysis of the assignment to the specific *IGLJ/C* segment resulted in two variants with the same probability, both reference variants were included in the subsequent analysis and the mean value was used. The *IGLV2-23* MM sequences showed a higher pI than the *IGLV6-57* AL sequences (8.12 vs. 5.0) and considering the delta pI, a significant difference from the respective reference was observed, with a larger one for the *IGLV2-23* MM sequences (delta pI *IGLVJC*: MM *IGLV2-23*= -0.30, AL *IGLV6-57*= 0.18, p= 0.043). This was also reflected in the GRAVY- and AGG-scores, where the impact of the mutations and therefore the delta values were also higher (delta GRAVY: MM *IGLV2-23*= 0.035, AL *IGLV6-57*= 0.006; delta AGG: MM *IGLV2-23*= 207.998, AL *IGLV6-57*= 9.406).

No differences in the pI of the *IGLV6-57* AL sequences was found when they were differentiated according to the presence of a HC (pI *IGLVJC*: AL *IGLV6-57*_HC= 5.15, AL *IGLV6-57*_noHC= 4.92), although the delta pI of the *IGLV6-57* sequences with a detectable HC was higher (delta pI *IGLVJ*: AL *IGLV6-57*_HC= 0.33, AL *IGLV6-57*_noHC= 0.006) (Table 41). The same was shown for the GRAVY- and AGG-score (GRAVY: AL *IGLV6-57*_HC= -0.391, AL *IGLV6-57*_no HC= -0.390; delta GRAVY: AL *IGLV6-57*_HC= 0.013, AL *IGLV6-57*_noHC= 0.006; AGG: AL *IGLV6-57*_HC= 663.629, AL *IGLV6-57*_no HC= 635.528; delta AGG: AL *IGLV6-57*_HC= 26.856, AL *IGLV6-57*_noHC= -1.234).

Table 41. Biochemical properties of *IGLV6-57* AL amyloidosis and *IGLV2-23* multiple myeloma sequences.

	pI *IGLVJC*	**delta pI *IGLVJC***	**GRAVY**	**delta GRAVY**	**AGG**	**delta AGG**
AL *IGLV6-57*	5.00	0.18*	-0.391	0.006	648.883	9.406

	pI *IGLVJC*	delta pI *IGLVJC*	GRAVY	delta GRAVY	AGG	delta AGG
MM *IGLV2-23*	8.12	-0.30*	-0.253	0.035	978.737	207.998
AL *IGLV6-57*_HC	5.15	0.33	-0.391	0.013	663.629	26.856
AL *IGLV6-57*_noHC	4.92	0.09	-0.390	0.006	635.528	-1.234

AL= AL amyloidosis, MM= multiple myeloma, AL n= 10, MM n= 8, AL_HC n= 6, AL_noHC n= 4, pI= isoelectric point, GRAVY= grand average of hydropathy, AGG= β-sheet aggregation, the corresponding parameter was calculated as well as the difference to the respective reference. All values are given as medians. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.2.7 Mutation analysis of *IGLV6-57* AL amyloidosis and *IGLV2-23* multiple myeloma sequences

In the following, a detailed comparative mutation analysis of the *IGLV6-57* AL and *IGLV2-23* MM sequences was performed (Figure 21). In particular, mutation hotspots (mutation ≥ 50 % of sequences) and individual charge exchanges were investigated. This analysis revealed only one position (53N) for AL, which was mutated in at least 50 % of the sequences. For MM, a total of eight positions were found. This affected three positions in CDR1 (29V, 31S, 32Y), one in FR2 (49M), two in CDR2 (53G, 54S) and CDR3 (96S, 97S). Considering the mutations in terms of an associated change in charge, mutation hotspot 53N stood out. In five of the six mutated positions, an additional charge was inserted. To be more precise, in four cases a negative charge was inserted and in one case a positive charge. Furthermore, the addition of positive charges was also observed in four cases involving the linker region and in one case at the first position of the *IGLJ*-segment.

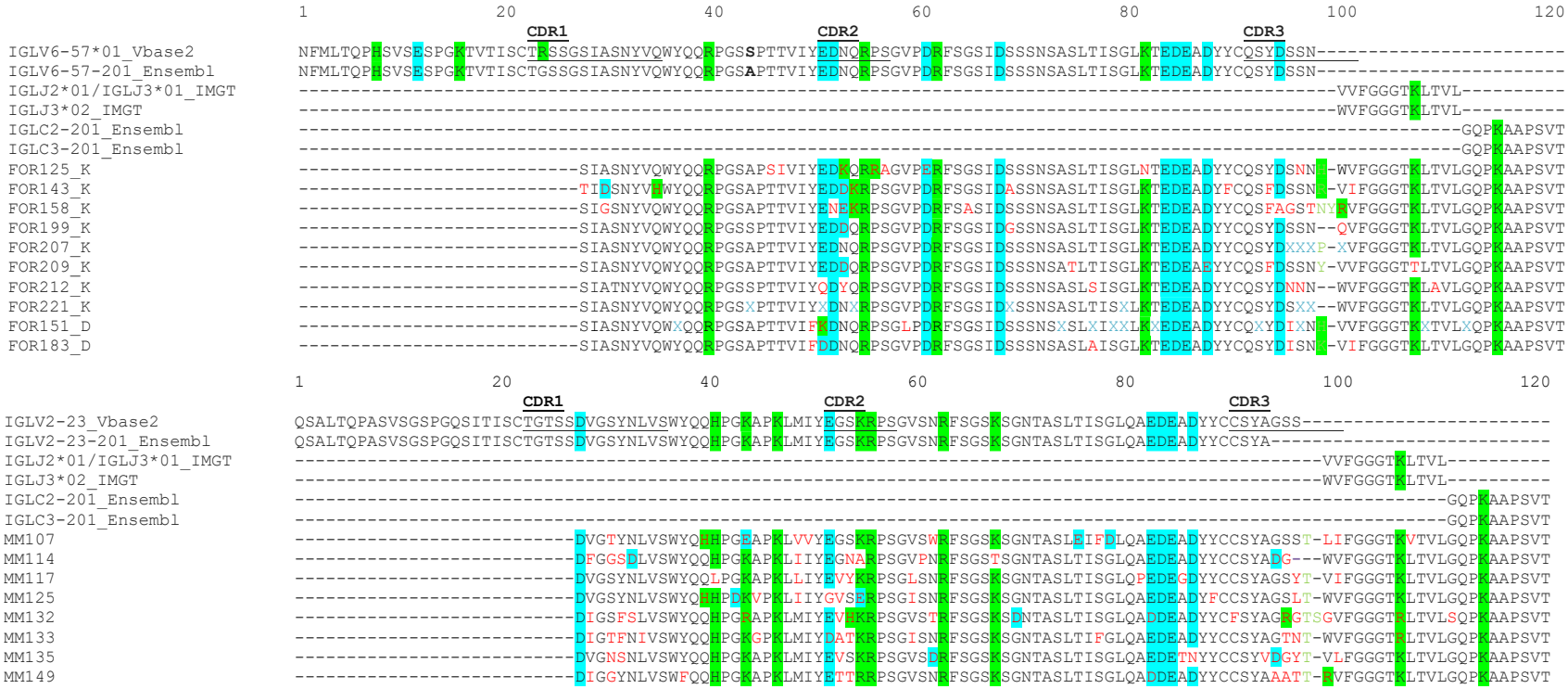

Figure 21. Sequence Alignment of *IGLV6-57* AL amyloidosis and *IGLV2-23* multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P0DOY2/3). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

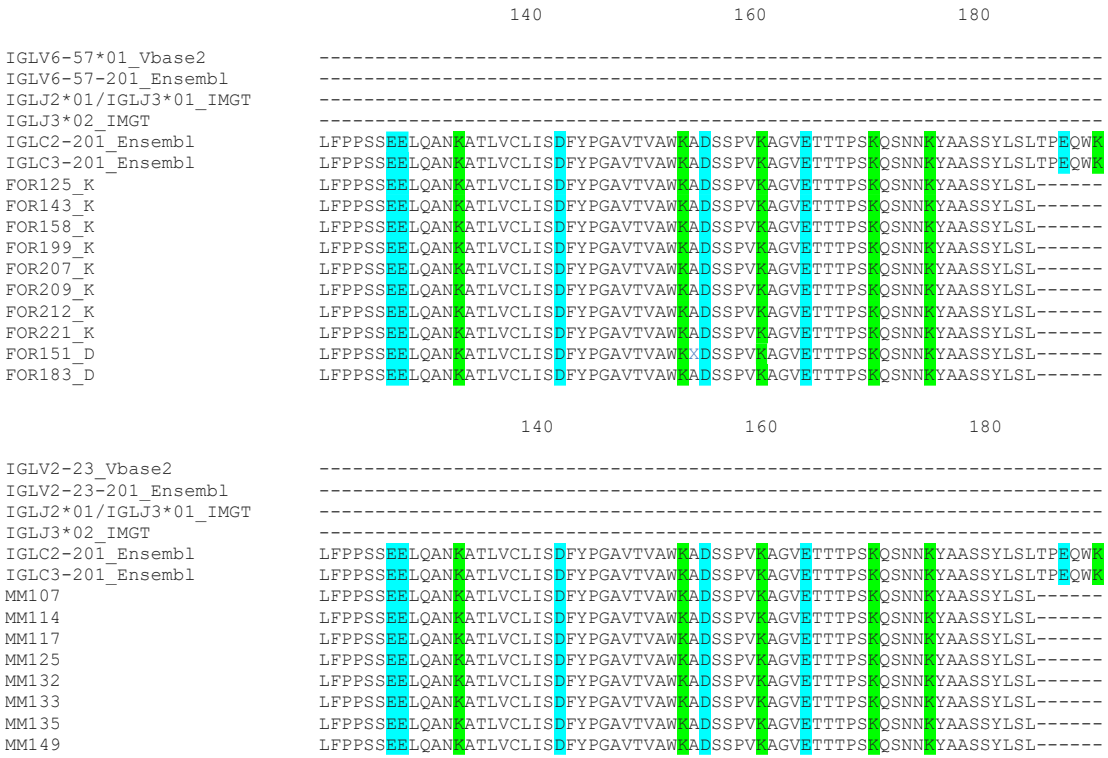

Figure 21. Sequence Alignment of *IGLV6-57* AL amyloidosis and *IGLV2-23* multiple myeloma sequences. FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients. K= dominant kidney involvement, D= diverse organ involvement, red= nonsynonymous substitutions, green= linker-region, purple= insertions and deletions, highlighted in green= positive charged amino acids, highlighted in blue= negative charged amino acids, X= undefined amino acid. The first position of the *IGKC*- Ensembl reference was completed by the Uni Prot reference (P0DOY2/3). The amino acids were numbered according to the Vbase2 reference. Complementary-determining regions (CDRs) were aligned according to Kabat using abYsis. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.3 Analysis of a potential N-glycosylation site of lambda *IGLV6-57* AL and *IGLV2-23* MM sequences

As before, the *IGLV6-57* sequences as representative for AL and the *IGLV2-23* sequences as representatives for MM were analysed. The analysis of the sequences for a potential N-glycosylation site did not reveal in any of the AL sequences or in any of the MM sequences a theoretical glycosylation site (Table 42).

Table 42. Analysis of a potential N-glycosylation site of *IGLV6-57* and *IGLV2-23* light chains.

	potential N-glycosylation site [n]	no potential N-glycosylation site [n]	Percentage of sequences with potential N-glycosylation site [%]
IGLV6-57			
AL [n= 10]	0	10	0
IGLV2-23			
AL [n= 8]	0	8	0

Analysis according to (Julenius 2007). AL= AL amyloidosis, MM= multiple myeloma

3.2.4 Analysis of the prediction of lambda light chain aggregation using V_LAmY-Pred- algorithm

As for the studied kappa cohort, the theoretical aggregation behaviour and associated amyloidogenic potential of the lambda *IGLV6-57* AL and *IGLV2-23* MM sequences were investigated using the bioinformatical tool V_LAmY-Pred. The aggregation probability was calculated (SI Table 3) and summarized in Table 43. Eight of the *IGLV6-57* AL sequences were classified as amyloidogenic and two as non-amyloidogenic. All of the *IGLV2-23* MM sequences (8/8) were classified as non-amyloidogenic.

Table 43. Analysis of prediction of amyloidogenicity with V_LAmY-Pred- algorithm.

	classified as amyloidogenic [n]	classified as non-amyloidogenic [n]	correctness [%]
IGLV6-57			
AL [n= 10]	8	2	80
IGLV2-23			
MM [n= 8]	0	8	100

Analysis according to (Rawat et al. 2021). AL= AL amyloidosis, MM= multiple myeloma. Verification of MM data was done by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and data were used in this thesis for comparison purpose.

3.2.5 Analysis of *IGLV3-21*

In the course of this project, it was observed that the Sanger sequencing of *IGLV3-21* sequences was repeatedly hampered by overlapping signals, predominantly in the N-terminal FR1 and CDR1 regions. To solve this problem, an *IGLV3*-specific primer (VLKL3_A_fw_NB) was applied in a first step. But in most cases, this did not lead to an improvement in sequencing quality. Which could lead to the assumption that further sub-sequences, particularly in this subfamily, may be present. To analyse this in more detail, NGS technology was used. For this purpose, a suitable evaluation pipeline had to be established (chapter 2.2.8.2) that allowed the analysis of individual sub-signals. In this context, two samples from patients with dominant kidney involvement and two samples from patients classified as diverse were examined. Analysis was performed using a multiplex primer similar to the first sequencing approach of Sanger sequencing and all sub-signals/sub-sequences with a fraction size > 1 % were considered. An overview of the examined samples with analysis of the sequence composition is given in Table 44 and all results are shown in the appendix (SI Table 4). Sub-sequences per sample varied between four and 11 detected sequences (fraction size of > 1 %). FOR129 and FOR170 presented a total of four sub-sequences with a median sequence fraction size of 1.98 % (range: 2.84 - 1.09 %) for FOR129 and 1.20 % (range: 1.09 - 1.98 %) for FOR170. FOR146 had seven sub-sequences with a median fraction size of 1.15 % (range: 1.05 - 1.44) and FOR178 had the most sub-sequences (n= 11) associated with the highest variance in fraction size (median sequence fraction: 1.42 %, range: 1.02 - 12.04 %). *IGLV3-21* was detected in all samples, although not as the most frequent *IGLV* subfamily. In both FOR129 and FOR146, *IGLV3-19* was the most abundant subfamily, although with comparable frequency to *IGLV3-21* (FOR129: *IGLV3-21*= 1.53 %, *IGLV3-19*= 2.84 %; FOR146: *IGLV3-21*= 1.06 %, *IGLV3-19*= 1.44 %). FOR170 indicated *IGLV7-43* as most common *IGL*-subfamily with 1.98 % sequence fraction size and *IGLV3-21* as the second most common with 1.26 %. FOR178 had the highest deviation with 12.04 % for *IGLV3-25* compared to 5.90 % for *IGLV3-21*. Overall, it was found that, with exception of FOR178, all detected sub-sequences including *IGLV3-21* had approximately the same, but low frequency.

Table 44. Next- generation sequencing analysis of sub-sequences of two kidney and two diverse *IGLV3-21* patients.

	Organ involvement	n sequences >1 %	Median sequence fraction [%] (range)	Sequence fraction *IGLV3-21* [%]	Most abundant *IGLV* / [%]
FOR129	Kidney	4	1.98 (1.09 - 2.84)	1.53	*IGLV3-19* / 2.84
FOR146	Diverse	7	1.15 (1.05 - 1.44)	1.06	*IGLV3-19* / 1.44
FOR170	Diverse	4	1.20 (1.09 - 1.98)	1.26	*IGLV7-43* / 1.98
FOR178	Kidney	11	1.42 (1.02 - 12.04)	5.90	*IGLV3-25* / 12.04

Sequencing was performed using a multiplex primer and sub-sequences were extracted using MiXCR. Sub-sequences with a fraction size > 1 % were analysed.

3.2.6 Summary lambda cohort

3.2.6.1 Summary patient cohorts and clinical data

In the context of this project, 34 lambda AL patients with dominant kidney involvement and 18 patients with diverse organ involvement were enrolled. For comparison purpose, a cohort consisting of 52 lambda MM patients was also included. The extraction of the clinical data of the MM patients as well as the verification of the sequence data was not performed as part of this work. The data were provided by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and used in this study for comparison purpose. The AL and MM patients did not differ in age of initial diagnosis (median age 59 - 61 years) (Table 35). In the MM cohort, there were more males than females (n= 30 vs. 22), while in the AL cohort, both sexes were equally affected (n= 27 vs. 25), with a contrasting trend within the two AL groups (AL_K: female= 15, male= 19; AL_D: female= 12, male= 6). Bone marrow was aspirated at the time of new diagnosis, with exception of four cases in AL_K and two cases in AL_D. Patients differed in their median dFLC, with the lowest in AL_K (median dFLC: AL_K= 58 mg/L vs. AL_D= 239 mg/L, p< 0.001, vs. MM= 411 mg/L, p< 0.001). The AL_D had a higher median AP than the AL_K (median AP: AL_D= 122 U/L vs. AL_K= 71 U/L, p< 0.001, vs. MM= 70 U/L, p< 0.001). A M-gradient was detected more frequently in MM patients (M-gradient: MM= 45/52 vs. AL_K= 15/33, p< 0.001, vs. AL_D= 9/17, p= 0.014) and MM patients differs in their median plasma cell infiltration rate (median plasma cell infiltration: MM= 60 % vs. AL_K= 7 %, p< 0.001, vs. AL_D= 11 %, p< 0.001). Considering the AL cohort, AL_D showed a higher median plasma cell infiltration rate than AL_K (median plasma cell

infiltration: AL_D= 11 % vs. AL_K= 7 %, p= 0.004). More MM than AL_K patients had a heavy chain in the blood (MM= 41/52 vs. AL_K= 17/34, p= 0.018) and a t(11;14) was detected in more AL than MM patients (t(11;14): MM= 10/52 vs. AL_K= 21/31, p< 0.001, vs. AL_D= 7/12, p= 0.022).

3.2.6.2 Summary *IGLV-, J/C* assignment

Assignment of lambda AL and MM sequences revealed an equal abundance for *IGLV1* (AL= 29 %, MM= 27 %) (Figure 19A). *IGLV2* was detected in 19 % of AL and in 33 % of MM sequences. 25 % of AL and 33 % of MM sequences were assigned to *IGLV3*. Significant differences were observed in *IGLV6*, with 19 % in AL and only 2 % in MM (p=0.004). *IGLV7* was represented only in MM (2 %).

Analysis of *IGLV*-subgroup usage (Figure 19B) showed significant differences for *IGLV2-23* and *IGLV6-57*. *IGLV2-23* was only represented in MM (15 %, p= 0.006) and *IGLV6-57* was significantly more represented in AL (AL= 19 % vs. MM= 2 %, p= 0.004). *IGLV3-21* had a higher abundance in MM (AL= 12 %, MM= 21 %), but without reaching statistical significance. Moreover, *IGLV2-8* (AL= 4 %) and *IGLV3-9* (AL= 2 %) were only detected in AL, in contrast to *IGLV2-11* (MM= 2 %) and *IGLV7-46* (MM= 2 %), which were only detected in MM.

Examination of organ tropism and *IGLV* usage (Figure 20A) revealed a higher occurrence of *IGLV1* and *IGLV6* in AL_K (*IGLV1*: AL_K= 35 %, AL_D= 17 %, *IGLV6*: AL_K= 24 %, AL_D= 11 %). As for *IGLV*-subfamilies (Figure 20B), *IGLV6-57* was the highest representative for AL_K (AL_K= 24 %, AL_D= 11 %). *IGLV1-44* was detected in 21 % of AL_K and in 11 % of AL_D patients and *IGLV2-14* was detected in 15 % of AL_K and in 17 % of AL_D patients, respectively.

In a next step, the *IGLJ/C* linkage of lambda AL patients was studied and compared with MM (Table 36). In both, the most common linkage was *IGLJ2/C2* (AL= 27 %, MM= 42 %). 15 % of AL patients had *IGLJ1* with *IGLC1*, whereas this was the case in 10 % of MM patients. 8 % of AL and 2 % of MM patients were associated with *IGLJ3/C2*, compared with 19 % and 13 % with *IGLJ3/C3*. *IGLJ2/C3* was only detected in AL (2 %) and *IGLJ7/C7* in MM (2 %). Regarding the clinical organ presentation in AL, AL_K and AL_D differed in the linkage of *IGLJ2/C3* (AL_K= 0 %, AL_D= 6 %) and *IGLJ3/C2* (AL_K= 12 %, AL_D= 0 %). In both groups, *IGLJ1/C1*, *IGLJ2/C2* and *IGLJ3/C3* was

highly represented (*IGLJ1/C1*: AL_K= 15 %, AL_D= 17 %; *IGLJ2/C2*: AL_K= 26 %, AL_D= 28 %, *IGLJ3/C3*: AL_K= 21 %, AL_D= 17 %).

Comparison of the most frequent *IGLV*-subfamilies with their linkage to *IGLJ/C* (Table 37) showed an equivalent association of *IGLV1-44* to *IGLJ2/C2*, *IGLJ3/C2* and *IGLJ3/C3* (4 %) in AL. *IGLV2-14* was most frequently associated with *IGLJ1/C1* and *IGLJ2/C2* (6 %) in AL and with *IGLJ2/C2* (10 %) in MM. *IGLV6-57* was most frequently associated with *IGLJ2/C2* (6 %) and *IGLJ3/C3* (6 %) in AL. Analysing the clinical organ presentation in AL (Table 37) showed, that all patients with *IGLV1-44* and *IGLJ2/C2* had dominant kidney involvement (AL_K= 6 %). Patients with diverse organ involvement had an equal association of *IGLV1-44* with *IGLJ3/C2* and *IGLJ3/C3* (6 %). *IGLV2-14* and *IGLJ1/C1* was more frequent in patients with diverse organ involvement (AL_D= 11 %, AL_K= 3 %). *IGLV6-57* linked to *IGLJ2/C2* was equally present in AL_K and AL_D patients (6 %). Whereas all patients with *IGLV6-57* and *IGLJ3/C3* had dominant kidney involvement (AL_K= 9 %).

3.2.6.3 Summary comparative analysis of *IGLV6-57* AL and *IGLV2-23* MM sequences

Due to the significantly different abundance of *IGLV6-57* and *IGLLV2-23* in both diseases (*IGLV6-57*: AL= 19 %, MM= 2 %, p= 0.004; *IGLV2-23*: AL= 0 %, MM= 15 %, p= 0.006) (Figure 19), leading to a potentially different amyloidogenic character, sequences from both subfamilies (*IGLV6-57* n= 10, *IGLV2-23* n= 8) were analysed comparatively. Eight of the corresponding AL patients had dominant kidney involvement and two belonged to the diverse group. The analysis of the individual mutations (Table 40) showed a significantly higher mutation count of the *IGLV* segment of the *IGLV2-23* MM sequences (median mutation count: *IGLV*: MM (*IGLV2-23*)= 10.0, AL (*IGLV6-57*)= 4.5, p< 0.001) combined with a higher mutation frequency (mutation frequency: *IGLV*: MM (*IGLV2-23*)= 100 %, AL (*IGLV6-57*)= 80 %). This was mainly reflected in the CDR1 and FR2 region (mutation frequency: CDR1: MM (*IGLV2-23*)= 75 %, AL (*IGLV6-57*)= 75 %, FR2: MM (*IGLV2-23*)= 88 %, AL (*IGLV6-57*)= 30 %, p= 0.025). One mutation was detected in the *C*-segment of one MM sequences (median mutation count: *IGLC*: MM (*IGLV2-23*)= 13 %). Stratification of the AL sequences (Table 40) showed differences in the median mutation count in the CDR1 region, where the AL *IGLV6-57*_HC sequences had more mutations (median mutation count: CDR1: AL *IGLV6-57*_HC= 1.0,

AL *IGLV6-57*_noHC= 3.0), but without reaching statistical significance. The analysis of the biophysical properties of the sequences belonging to the *IGLV6-57*- and *IGLV2-23*-subfamily (Table 41) showed a higher pI of the *IGLV2-23* MM sequences (pI: *IGLV2-23*: 8.12, *IGLV6-57*: 5.0), with a significantly differences in the delta pI (delta pI *IGLVJC*: MM *IGLV2-23*= -0.30, AL *IGLV6-57*= 0.18, p= 0.043). This was confirmed by the GRAVY and AGG scores, where the effect of the mutations and thus the delta values were also higher for the *IGLV2-23* sequences (delta GRAVY: MM *IGLV2-23*= 0.035, AL *IGLV6-57*= 0.006; delta AGG: MM *IGLV2-23*= 207.998, AL *IGLV6-57*= 9.406). Investigation of the presence of a HC in the serum (Table 41) revealed a higher delta pI of the *IGLV6-57* sequences with a HC (delta pI *IGLVJ*: AL *IGLV6-57*_HC= 0.33, AL *IGLV6-57*_noHC= 0.006). The same pattern was shown for the GRAVY- and AGG-scores (GRAVY: AL *IGLV6-57*_HC= -0.391, AL *IGLV6-57*_no HC= -0.390; delta GRAVY: AL *IGLV6-57*_HC= 0.013, AL *IGLV6-57*_noHC= 0.006; AGG: AL *IGLV6-57*_HC= 663.629, AL *IGLV6-57*_no HC= 635.528; delta AGG: AL *IGLV6-57*_HC= 26.856, AL *IGLV6-57*_noHC= -1.234). Analysis of individual mutated sequence positions (Figure 21) revealed only one mutation hotspot (mutations ≥ 50% of the sequences) in AL (CDR2: 53N) compared to eight in MM (CDR1: 29V, 31S, 32Y; FR2: 49M; CDR2: 53G, 54S; CDR3: 96S, 97S). Moreover, mutation hotspot 53N was associated with the insertion of an additional charge in five of the six mutated sequences. None of the analysed lambda sequences showed potential N-glycosylation sites Furthermore, the sequences were studied for their theoretical aggregation behaviour and amyloidogenic potential using the machine learning based tool V_LAmY-Pred (Table 43). Of the *IGLV6-57* AL sequences, eight were categorised as amyloidogenic and two as non-amyloidogenic. All *IGLV2-23* MM sequences were categorised as non-amyloidogenic.

3.2.6.4 Summary IGLV3-21

In the course of this project, the *IGLV3-21* sequences showed accumulated sub-signals in the Sanger sequencing. Therefore, sequences from *IGLV3-21* were analysed in more detail using NGS technology in order to determine the sequence composition and probably draw conclusions about the possible presence of sub-sequences. For this purpose, two sequences from patients with dominant kidney involvement and two sequences from patients with diverse organ involvement were examined. All sequences with a fraction size > 1 % were considered. Sub-sequences per sample differed between four and 11 detected sequences and an occurrence of the median sequence fractions

between 1.15 % and 1.98 % (Table 44). *IGLV3-21* was detected in all sequences, but other subfamilies were also detected with mostly comparable high abundance. For example, in FOR146 *IGLV3-19* was detected as the most frequent subfamily with an occurrence of 1.44 % and *IGLV3-21* with 1.06 %. Overall, it could be observed that further subfamilies despite from *IGLV3-21* were present, but almost all of them were in a similar range in terms of occurrence.

4 Discussion

In recent years, lambda LCs have been extensively examined in the context of AL, while knowledge about kappa sequences remains scarce. In addition, most of the previous studies have focussed on the variable part of the LC and have not considered the LC as its whole. This project offers a comprehensive analysis of the full-length LC sequence of kappa LC in AL compared to MM. The analysis was based on two precisely defined clinical cohorts including forty-one patients with kappa AL. AL patients were consecutively enrolled in this study and sequence analysis was not possible for only one patient. Patients with MM were thoroughly characterized according to the GMMG-HD6 study guidelines to rule out concurrent AL with a median follow-up of 48.9 month (Goldschmidt et al. 2021; Salwender et al. 2019). As a result, three patients diagnosed with both kappa MM and AL were omitted from the project, thereby rendering the remaining total of 83 MM patients particularly qualified as control group. An initial step in this project was to analyse the LC sequences based on their *IGKV/J/C* classification. For this purpose, a PCR protocol was developed, which allowed the analysis of the entire LC, from its *IGKV* to its *IGKC* part. The classification of the specific *IGKV/J/C* segments formed the basis for all subsequent research. To identify an AL specific pattern, the *IGKV/J/C* classifications of the AL sequences were compared with MM LC sequences. The next step involved determining whether AL specific characteristics were associated with one subgroup within the AL cohort, achieved by stratifying the AL sequences according to their clinical organ presentation.

Beyond the *IGKV/J/C* assignment, the LC sequences were evaluated further focusing on mutation patterns, their amino acid composition, as well as various biochemical properties. This included investigation of a potential interaction with a LC binding partner (the HC). A bioinformatic analysis was performed to identify the effect of mutations on the theoretical isoelectric point (pI), protein hydrophobicity (GRAVY-score), aggregation capacity (AGG-score) and the presence of a potential N-glycosylation site.

Furthermore, the aggregation tendencies of the kappa LCs were evaluated using a bioinformatics tool based on machine learning technology. While the study predominantly focussed on kappa LCs, it also includes an analysis of 52 lambda AL LC sequences for comprehensive understanding and comparative analysis.

4.1 Is there a dominant light chain type in kappa AL amyloidosis patients compared to multiple myeloma?

In an initial step, the obtained LC sequences were matched with the *IGKV/J/C* segments with the highest similarity. This served not only as a preliminary categorization but also laid the basis for further analysis, like studying mutations, as the assigned germline sequences served as references for the unmutated state. In general, the analysis of AL kappa sequences with respect to *IGKV* usage has been studied in prior research. However, in the most cases, these studies included a smaller patient cohort (Abraham et al. 2003; Comenzo et al. 2001; Connors et al. 2007). Furthermore, there has been no direct comparative analysis regarding the usage of *IGKV* in MM patients. This comparative study offers the potential not only to find the predominance *IGKV* usage within AL but also provide indications about the amyloidogenic propensity of particular LC group.

As shown in this study, *IGKV1* was the most common LC type in both AL and MM, however with significant higher frequency in AL (AL= 80 %, MM= 53 %, p= 0.002) (Figure 11). This corresponds to the result of Abraham et al., who also identified the *IGKV1* family as the most frequently gene usage in their study cohort (Abraham et al. 2003). In their investigation, 28 kappa AL patients were analysed, whereby a dominant *IGKV* LC was detected in 26 AL patients. Among these 26 AL patients, 77 % were linked to the *IGKV1* family, while the remaining 19 % were associated with *IGKV4*. In the present project, *IGKV4* was with 5 % the third most common *IGKV* assignment in AL and *IGKV3* had a higher incidence with about 10 % (Figure 11). Nonetheless, if this is compared with the incidence in MM (30 %, p= 0.014), it is evident that *IGKV3* is markedly underrepresented in AL, indicating that it cannot be considered characteristic of AL (Figure 11). Further, *IGKV2* was only detectable in MM (10 %, Figure 11) in this study. Focussing on the level of *IGKV*-subgroups, it was shown that *IGKV1/D-33* is linked to AL in 32 % of cases, while in MM, only 13 % of cases exhibited this subtype (p= 0.016, Figure 11). These findings align also with the research by Abraham et al. (2003), where *IGKV1/D-33* was also identified as the most prevalent subgroup within their study. In addition, significant differences could be found in the expression of *IGKV1-16*, with 10 % of the AL patients showing *IGKV1-16*, but none of the MM patients (p= 0.010, Figure 11). Also, in the calculation of the relative risk (rr), these two subgroups stood out. With a rr of 2.452 (p= 0.013) for *IGKV1/D-33* and 18.44 (p= 0.049) for

IGKV1-16, patients showing one of these LC subtypes indicate an increased risk of developing AL (SI Table 1). In alignment with the study of Abraham et al. (2003) *IGKV1/D-39* was detected as the second most frequent *IGKV*-subgroup (Abraham et al. 2003). However, it must be noted that *IGKV1-39* was with 17 % also highly represented in the MM cohort in this study and the calculation of the rr (rr= 1.186, (SI Table 1)) also revealed no statistical significance. Focussing on *IGKV2*, which only occurred in MM, the subgroups *IGKV2/D-28* (5 %, rr= 0.228), *IGKV2/D-30* (4 %, rr= 0.023), *IGKV2/D-40* (1 %, rr= 0.683) could be detected. In addition to the representatives of *IGKV2* family, a further *IGKV*-subgroup, being *IGKV3-/D-20* (11 %, p= 0.030, Figure 11), was only detectable in MM with a rr of 0.111, but without statistical significance. Furthermore *IGKV3/D-11* was detected as one of the abundant subfamilies in MM (MM= 14 % vs. AL= 2 %, Figure 11, rr= 0.173, (SI Table 1)).

This study also examined the impact of the primary clinical presentation on the expression of the *IGKV*-subtypes (Figure 12). For this purpose, three patient groups with AL were defined based on their predominant organ involvement: those with predominant renal or cardiac involvement and the third category summarises patients with diverse organ manifestations. Consequently, the cohort was divided as follows: 34 % of the patients displayed dominant renal symptoms (AL_K), 20 % had primarily cardiac symptoms (AL_H) and the remaining 46 % of patients were defined as diverse (AL_D), encompassing a range of different clinical presentations. In addition to multi-organ involvement, patients with an isolated organ presentation that was not a kidney or heart involvement were also assigned to this group, as their number was too small to consider them as a separate group. This concerned one patient with isolated liver involvement and one patient with symptoms of polyneuropathy. In the analysis of the different organ presentations, *IGKV1* was highly represented in all three groups of AL patients (AL_K (64 %) AL_H (88 %) and AL_D (89 %), Figure 12). In 21 % of AL_K patients and 13 % of AL_H patients, *IGKV3* was found and no AL_D patient showed *IGKV3*. *IGKV4* was only detected in AL_D (11 %) and *IGKV5* only in AL_K patients (7 %). Focussing again on the *IGKV*-subgroups *IGKV1/D-33* stood out, as *IGKV1/D-33* was primarily detected in patients heart involvement (6/8, 75 %, p= 0.024, Figure 12), while *IGKV1/D-39* was more frequently detected in patients with dominant kidney involvement (5/14, 36 %, Figure 12), but without statistical significance. When analysing the patients with a diverse organ involvement no clear predominant *IGKV1*-subgroup could be identified.

Comparing these results with the results of a study by Kourelis et al. (2017), who also investigated the influence of organ tropism on the *IGKV* usage, differences were found for *IGKV1/D-33* and *IGKV1-05*. In the study by Kourelis *IGKV1/D-33* was associated with liver involvement (27 % vs. 7 %, p= 0.0002) and patients showing *IGKV1-05* had a trend towards kidney involvement (53 % vs. 18 %, p= 0.02) (Kourelis et al. 2017). In the present project, the influence of liver involvement on the *IGKV* usage was also investigated. In this analysis the greatest differences were found for *IGKV1-16*, with 27 % of AL patients had liver involvement and only 3 % had no liver involvement. For *IGKV1/D-33* no differences were found between patients with and without liver involvement (27 % vs. 33 %). *IGKV1-5* was only detected in AL_H patients (13 %) and AL_D patients (21 %).

Previous studies on the variability of the kappa LC sequences mainly concentrated on the *IGKV*-segments and neglected the *IGKJ*-segments (Connors et al. 2007). In this project, in addition to *IGKV* diversity, the combination of the different *IGKV/J*-segments was also analysed. This was possible by the approach of sequencing the complete patient-specific LC. Thus, this study analysed the combination of *IGKV* and *IGKJ* for the first time and shed light on the complex process of *IGKV/J* recombination in AL in comparison to MM, although the sample size is relatively small and further analyses would be of great interest. In AL, *IGKJ2* (39 %) was the most predominant and in MM *IGKJ1* (30 %) and *IGKJ2* (29 %) were similarly present (Table 14). The two diseases also differ in the expression of *IGKJ3* and *IGKJ5*. *IGKJ3* was more abundant in MM patients (13 % vs. 5 %) and in AL the expression of *IGKJ5* was higher (10 % vs. 5 %) (Table 14). In the context of organ tropism, a notable preference in the association of *IGKJ*, beyond the formerly characterized *IGKV* usage patterns, was shown. Specifically, *IGKV1/D-33* associated with heart failure, exhibits a predominant recombination with *IGKJ2*, observed in 50 % of cases (Table 15). Conversely, association with *IGKJ1* and *IGKJ3* were identified in a minority of cases, accounting for 13 %, while recombination involving *IGKJ4* or *IGKJ5* were absent in this cohort. IGKV1/D-39 which is predominantly recognized in patients with dominant kidney involvement, was equally combined to *IGKJ1* and *IGKJ2* (14 %) (Table 15). The presence of *IGKJ4* was detected in 7 %, however, no recombination with *IGKJ3* or *IGKJ5* was detected among patients with *IGKV1/D-39* and dominant kidney involvement.

This investigation of the association between the different *IGKV* and *IGKJ* patterns in AL in contrast to MM offers new insight into the molecular mechanisms, on the one hand into the association with different clinical entities and on the other hand into an organ-specific tendency in the deposition of a specific *IGKV/J* LC in AL. This could serve as a basis for preventive therapeutic interventions and diagnostic approach in AL.

4.1.1 Lambda versus kappa – What is the dominant light chain type in lambda AL amyloidosis?

In addition to the previously discussed findings regarding the predominant kappa LC type, lambda AL LC sequences were also analysed in this project. In contrast to the kappa studies, the complete patient cohort was not included. In this case, the cohort consisted of AL patients with a dominant kidney involvement or a diverse organ involvement. Patients with dominant heart involvement were part of another PhD project. Also in this context, the AL data were compared with MM data.

Major differences were observed in the frequencies of *IGLV2* and *IGLV6*. Specifically, *IGLV2* exhibited a frequency of 19 % in AL compared to 33 % in MM (Figure 19). Conversely, *IGLV6* had a higher frequency in AL (19 %) than in MM (2 %), with reaching statistical significance (p= 0.004) (Figure 19). Focussing on *IGLV*-subgroups, significant differences were detected for *IGLV2-23* and *IGLV6-57*. *IGLV2-23* was only detected in MM (15 %, p= 0.006) whereas *IGLV6-57* was detected in 19 % of AL and only in 2 % of MM patients (p= 0.004) (Figure 19). Analysing the AL patients showing *IGLV6-57* with regards to their organ involvement revealed primarily patients with a dominant renal involvement. Among all identified *IGLV* subgroups in AL, *IGLV6-57* (24 %) was the most common subgroup in patients with dominant kidney involvement, besides *IGLV1-44* (21 %) (Figure 20). The association between *IGLV1-44* and dominant kidney involvement has already been shown in a previously published subgroup of this data (Berghaus et al. 2022). Here, 25 patients (compared to 34 patients in the present work) with dominant kidney involvement were analysed and *IGLV1-44* was the most frequently detected subgroup (20 %). *IGLV6-57* was detected in fewer patients with dominant kidney involvement (12 %), which can be explained by the smaller sample size. Berghaus et al. (2022) also included patients with concomitant cardiac and renal involvement and detected *IGLV6-57* as the most prevalent subgroup of these patients. Considering that by increasing the sample size in the present work, *IGLV6-57* emerged as the most abundant

IGLV subgroup, it can be speculated that the high frequency of *IGLV6-57* in patients with concomitant cardiac and renal involvement is due to the proportion of renal involvement. In further studies *IGLV6-57* has also been identified as the most frequently detected *IGLV*-subgroup in patients with kidney involvement (Comenzo et al. 2001; Kourelis et al. 2017; Perfetti et al. 2002).

In addition, the *IGLJ* and *IGLC* usage was analysed in the different cohorts of patients. In both lambda AL and MM, the joining of *IGLJ2* and *IGLC2* was the most frequently detected (AL= 27 %; MM= 42 %), without differing between the two AL subgroups AL_K (26 %) and AL_D (28 %) (Table 36). Due to the small number of samples, only tendencies could be identified for the combination of the individual *IGLV*-subgroups with *IGLJ/IGLC*. In line with Berghaus et al. (2022) *IGLV6-57* was preferably associated with *IGLJ2/IGLC2* (6 %) in AL, with representing AL_K and AL_D equally (6 %) (Table 37). *IGLV2-14* was also most frequently combined with *IGLJ2/IGLC2* in MM (10 %), whereas 6 % were combined with *IGLJ1/IGLJ1* and *IGLJ2/IGLC2* in AL. Differentiating AL according to organ tropism, the combination with *IGLJ1/IGLC1* was more frequently detected in AL_D (11 %) than AL_K (3 %) and *IGLJ2/IGLC2* was represented in equal parts (AL_K= 6 %; AL_D= 6 %) For *IGLV1-14*, no favoured combination could be determined (Table 37). In summary, the interpretation of the data regarding to potential triplet combinations is challenging, not least because most of the available studies focussing on the variable lambda domain alone. Nevertheless, it is important to consider the entire LC as its whole, because in addition to the variable domain, the constant domain is also crucial for the investigation of amyloidogenicity (Rottenaicher et al. 2023; Rottenaicher et al. 2021).

4.2 Mutations can have an effect on the amyloidogenic potential

Previous studies have investigated the impact of mutations on the native conformation of LCs. The findings suggest that mutations, depending upon their position and the biochemical characteristics of the substituted AAs, affect the stability of the LCs, thereby influencing their propensity to aggregate. It could be shown that mutations within the FRs of the LCs impair their stability by altering intramolecular interactions or the orientation of the antibody domains (Kazman et al. 2020). In addition, mutations in the CDRs have

also been identified as amyloidogenic drivers (Rottenaicher et al. 2021). However, not only mutations in the variable LC domain, but also in the constant domain can affect the amyloid pathway (Rottenaicher et al. 2023). The mutations identified in the current study were compared with those identified in prior research to infer their amyloidogenic potential. In addition to the comparison with the unmutated state, an extensive comparison with patient-specific non-amyloidogenic LC sequences of MM patients was performed. In this analysis, the focus was not only on individual point mutations, but also on the distribution of the various mutations across the respective regions or segments. The abundance of different AAs was examined, as well as the biochemical properties of the LCs regarding the theoretical stability based on the pI, the GRAVY-score as well as the AGG-score, all considering the unmutated state as well as the comparison between AL and MM. Furthermore, the potential influence of a binding partner, namely a HC, was also considered.

4.2.1 *IGKV1/D-33* – sequence characteristics of the most prevalent subfamily in kappa AL amyloidosis

Initially focussed on *IGKV1/D-33*, identified as the most abundant *IGKV* representative, the analysis disclosed a higher median mutation count within the *IGKV*-segment for the AL sequences relative to the MM (AL= 12.0 vs. MM= 10.0, Table 17). Also, differences in the occurrence of mutations in the *IGKJ*-segment were detected for both diseases, but with higher median mutation count in MM (AL = 1.0 vs. MM = 2.0, Table 17). At the level of individual regions, FR3 and CDR3 were most frequently mutated in both diseases, with FR3 showing the highest median mutation count (AL= 4.0; MM= 3.5, Table 17). Further analysis suggests variances in the AA composition of *IGKV1/D-33* sequences, particularly regarding to asparagine. Compared to MM, AL sequences demonstrated a significant increased concentration of asparagine (AL= 4.6 % vs. MM= 4.0 %, p= 0.007, Table 18). Additionally, the analysis of the influence of the presence versus absence of a HC revealed differences in serine levels within AL (AL_HC= 13.9 % vs. AL= noHC= 14.9 %, p= 0.012, Table 18). Given the polar, uncharged character of these AAs, no difference in the pI of these sequences was observed. Calculating the GRAVY- and the AGG-score also did not provide any clear indication of potential differences in the aggregation behaviour of the two diseases or in stratification regarding the presence of a HC. Consequently, the amyloidogenic potential

of the examined kappa LCs cannot be attributed to disparities in the pI or the calculation of the GRAVY- and AGG-score and might be due to the presence of specific mutations. Poshusta and colleagues postulates that amyloidogenicity cannot be adequately assessed by quantification of mutations, rather the position of mutations is crucial (Poshusta et al. 2009). In their study they highlighted an accumulation of somatic hypermutations within the CDR1, CDR2, CDR3 as well as FR3. In the present project, an accumulation of mutations in FR3 and CDR3 was also detected in AL (Figure 15), which was mainly related a mutation hotspot which was defined as positions that were mutated to or more than 50 % as compared to the respective reference sequence. This concerned 65S, 83I and 93N in this case. Whereas the MM sequences only showed mutation hotspots in CDR3, where position 93N was also affected, as well as position 101 (Figure 15). Analysing mutations in terms of the influence on the charge of AA, different positions stood out. In AL, position 70D, located within the FR3 and next to the AL specific mutation hotspot at 65S showed mutations towards histidine in two cases and leading to a change in charge. Moreover, mutations towards asparagine were noted in three cases, culminating in a loss of charge (Figure 15). In MM mutation at the identical position resulted in either a loss or reversal of charge in three cases (D70H, D70G, D70N). In their study, Randles and colleagues also reported a mutation at position 70 to histidine and revealed that it is not part of the dimer interface, but possibly influence of the overall stability of the protein (Randles et al. 2009).

4.2.2 How does the divergent mutation behaviour affect other *IGKV* kappa subfamilies?

Next to the most dominant *IGKV*-subfamily, further subfamilies widely represented in kappa AL were also analysed. The higher median mutation counts of the kappa sequences, which was already observed in the most frequent *IGKV1/D-33*-subfamily in AL, was also confirmed by analysing the *IGKV1/D-39* sequences. Also in this case, the median mutation count of AL was higher compared to MM (AL= 14.0 vs. MM= 9.5), which was particularly evident in FR1 (AL= 2.0, MM= 1.0, p= 0.020) and FR3 (AL= 3.5, MM= 2.0, p= 0.017) (Table 21). These results were also supported by the comparative analysis of *IGKV1-16* sequences, which only occurred in AL and *IGKV3/D-11* MM sequences, as one of the most frequent representatives in MM. In this case, significantly more mutations were found in the *IGKV*-segment of *IGLV1-16* compared to *IGKV3/D-11* (median

mutation count: *IGLV1-16*= 9.0, *IGKV3/D-11*= 5.5, p= 0.033, Table 30). And again, FR1 and FR3 were primarily affected. Especially the *IGKV3/D-11* MM sequences showed no mutations in the FR1, but 50 % of the *IGKV1-16* AL sequences did (AL= 50 %, MM= 0 %, p= 0.050, Table 30). Interestingly, within *IGKV1/D-39*, AL patients without a detectable HC in the serum showed more mutations than patients with a HC (AL_noHC= 14.0 vs. AL_HC= 12.0, Table 21), which might suggest an effect of the HC. Further studies would be of interest, also in comparison to MM. In general, the comparison of mutations with existing published data is difficult, because most results are based on *IGKV1/D-33*-subfamily. In this study, the *IGKV1/D-39*-subfamily showed mutation hotspots in the CDR-regions in both the AL and MM sequences, which were characterised by the insertion of positive charges (Figure 16). In contrast, the *IGKV1-5* sequences showed mutation hotspots in both FR- and CDR-regions in AL and MM, which led to the insertion of positive charges in the CDR1 and a loss of charge in FR2, whereas negative charges were inserted in AL, but positive and negative charges were inserted in MM (Figure 17). Due to the higher median mutation count, *IGKV1-16* AL sequences showed more mutation hotspots than the *IGKV3/D-11* MM sequences and the changes in charge were correspondingly more frequent (Figure 18).

Looking again at the higher median mutation count of the kappa AL sequences compared to the MM sequences, the lambda sequences analysed in this project are of particular interest. Here, a contrasting behaviour was observed. Once again, one subfamily was elected as a representative for AL (*IGLV6-57*: AL= 19 %, MM= 2 %, p= 0.004) and another as a representative for MM (*IGLV2-23*: AL= 0 %, MM= 15 %, p= 0.006), leading to a potentially different amyloidogenic character. Surprisingly, theses analysis showed a significant higher median mutation count of the studied MM sequences compared to AL (median mutation count: *IGLV*: MM (*IGLV2-23*)= 10.0, AL (*IGLV6-57*)= 4.5, p< 0.001, Table 40) combined with a higher mutation frequency (mutation frequency: *IGLV*: MM (*IGLV2-23*)= 100 %, AL (*IGLV6-57*)= 80 %). This was mainly reflected in the CDR1 and FR2 region (mutation frequency: CDR1: MM (*IGLV2-23*)= 75 %, AL (*IGLV6-57*)= 75 %, FR2: MM (*IGLV2-23*)= 88 %, AL (*IGLV6-57*)= 30 %, p= 0.025, Table 40). Even in the study by Berghaus et al. (2023), who analysed the *IGLV2-14* AL subfamily in detail, a significantly higher median mutation count was found for MM sequences compared to AL. These results suggest that progenitors of kappa and lambda AL undergo differential somatic hypermutation and antigenic selection. The evidence of

differential intraclonal diversification between kappa and lambda AL compared to the respective MM sequences suggest that the development into an amyloidogenic cell population is fundamentally different between kappa and lambda. In addition to the discussed differences in mutation behaviour, an interaction between further effects based on biophysical or biochemical properties is conceivable, which will be discussed in more detail below.

4.3 Influence of different biophysical properties on light chain sequences

Previous studies investigated the effect of various biophysical properties of amyloidogenic LCs such as the pI, hydrophobicity and aggregation properties (Kaplan et al. 2007; Schmittschmitt and Scholtz 2003). In the study by Baur et al. (2022), which also included some of the lambda LC sequences from kidney AL patients of this thesis, it was shown that mutations affect the net charge of the LCs and thus may have an influence on their stability. In another study using four of the lambda LCs from this project, it was shown that AL-associated LCs have a different charge behaviour then MM-associated LCs. It was also shown that the presence of a HC can influence the pI and thus the stability of the LC (Berghaus et al. 2023). When analysing the sequences belonging to the main representatives of the kappa cohort of this project regarding their pI, GRAVY-score and AGG-score, no clear pattern could be identified. But differences were found for the studied lambda sequences. Once again, *IGLV6-57* was used as a representative for AL and *IGLV2-23* as a representative for MM. The analysis of the pI of the patients LCs compared to the pI of the unmutated state showed significantly differences between the two subfamilies (delta pI *IGLVJC*: MM *IGLV2-23*= -0.30, AL *IGLV6-57*= 0.18, p= 0.043, Table 41), whereby the pI of the *IGLV2-23* MM patients sequences (pI: *IGLV2-23*: 8.12) was closer to the physiological pH of the blood (7.4) and thus has a higher stability than the *IGLV6-67* AL LCs (pI: *IGLV6-57*: 5.0) (Table 41). This was confirmed by the GRAVY and AGG scores, where the effect of mutations and thus the delta values were also higher for the *IGLV2-23* MM sequences (delta GRAVY: MM *IGLV2-23*= 0.035, AL *IGLV6-57*= 0.006; delta AGG: MM *IGLV2-23*= 207.998, AL *IGLV6-57*= 9.406).

In summary, these data indicate differences in the biophysical properties of lambda AL sequences compared to their unmutated state and MM. The data also suggest that kappa AL is basically different to lambda AL. Further studies will be necessary to gain more insight into the findings of this research, especially as the findings are based on *in silico* studies, they should be verified using *in vitro* ad finally *in vivo* assays.

4.4 N-glycosylation as a potential driver for amyloidogenicity

It is assumed that glycosylation influences the mechanism of amyloid fibril formation (Bellotti et al. 2000; Merlini and Bellotti 2003; Omtvedt et al. 2000) and kappa in particular have been shown to be associated with it (Stevens 2000). The assumption that glycosylation could be a risk factor for the development of AL initially seems paradoxical, as posttranscriptional modifications leads to stabilisation of natively folded proteins (Solá and Griebenow 2009). In the context of AL, it has been discussed that glycosylation leads to stabilisation of fibril folding and protection against proteolytic degradation (Radamaker et al. 2021b). In the present study, bioinformatic analyses were used to investigate the kappa and lambda sequences about a potential glycosylation site; both the AL and MM sequences were considered. Again, the *IGKV1/D-33* subgroup, which was highly represented in kappa AL, was outstanding. In this case, 61.5 % had a potential glycosylation site but only 9.0 % of MM cases. This was reflected in the comparison of *IGKV1-16*, which was considered typical for kappa AL and *IGKV3/D-11*, mainly found in kappa MM. 75.0 % of the kappa AL sequences with *IGKV1-16* and only 8.3 % of kappa MM sequences with *IGKV3/D-11* had a potential glycosylation site (Table 32). In contrast, none of the analysed lambda sequences showed a potential N-glycosylation site. This emphasises the hypothesis that kappa AL is associated with glycosylation, which was confirmed by considering the whole kappa cohort, with significantly more kappa AL having a potential N-glycosylation site than kappa MM sequences (42.5 % vs. 16.9 %, p= 0.004). In accordance to previous studies by Conners et al., who described the replacement of D70 with asparagine as an introduction of an N-glycosylation site, this specific exchange was also found in three of the 13 *IGKV1/D-33* AL sequences but only in one of the MM sequences in the current study (Connors et al. 2007). Accordingly, these results of the N-glycosylation analysis provide validation of the findings by Nevone et al. describing a N-glycosylation hotspot in FR3 associated with

AL (Nevone et al. 2022). In this study, a predicted N-glycosylation site was detected in 32 – 42 % of AL and 12-14 % of MM cases, mostly located in the FR3 region.

4.5 Is artificial intelligence able to predict the aggregation tendency of light chains?

Computer systems based on artificial intelligence, including deep learning are becoming increasingly popular, also in the medical sector (Basu et al. 2020; Garofalo et al. 2021). It is used in diagnosis and prognosis estimation of patients (Lotter et al. 2021; Saltz et al. 2018; Shen et al. 2019), in the development of new drugs (Olivecrona et al. 2017; Sarkar et al. 2023) or even in drug sensitivity calculation (Costello et al. 2014; Yarahuan et al. 2024). Efforts are also being made to utilise deep learning technology in the field of AL. So far, two bioinformatic tools were developed that aimed at calculating the aggregation potential of LCs and thus allowing conclusions about their toxicity (Garofalo et al. 2021; Rawat et al. 2021). With the tool developed by Garofalo et al. (2021) only lambda sequences can be analysed, whereas the tool by Rawat et al. (2021) allows the analysis of both lambda and kappa sequences. In this project, the aggregation potential of the main representatives of the kappa sequences, as well as representatives for the lambda AL and MM patient sequences were analysed using V_LAmY-Pred (Rawat et al. 2021).

The analysis of the various kappa sequences revealed divergent predictions concerning the amyloidogenic potential across the different subfamilies (Table 33). For the *IGKV1-5*, an amyloidogenic potential was predicted in only 40 % of the analysed AL sequences. In the *IGKV1-16* and *IGKV1/D-39* subfamilies this prediction increased to 50 % of the analysed cases. *IGKV1/D-33* subfamily, which predominates within AL, clearly showed amyloidogenicity in 100 % of the AL sequences. But also, the *IGKV1/D-33* MM sequences were consistently classified as amyloidogenic. The analysis of the *IGKV3/D-11* MM sequences showed a prediction of non-amyloidogenicity in 100 % of cases. Analysing the *IGLV6-57* AL sequences as representative of lambda AL sequences and *IGLV2-23* MM sequences as representative for lambda MM, again appropriate results were obtained. 80 % of the AL sequences were classified as amyloidogenic and 100 % of the MM sequences as non-amyloidogenic.

The development of a tool that predicts the toxicity of LC is still in early development, but the tool performed quite good in the tested cohort. Further studies testing the use of

the V_LAmY-Pred-tool are rare, which makes comparison difficult. However, it should be noted that the performance of an artificial intelligence-based application depends on the data sets used as basis for the development as well as the training of the method (Basu et al. 2020). Assuming that there is a shift in the data sets towards one subgroup, whether for biological causes or due to a selection bias, this would result in certain subgroups being predicted with a higher accuracy than other. This might explain why the MM sequences associated with the *IGKV1/D-33*-subfamily, which is highly represented in AL, were also classified as amyloidogenic. The ability to use a tool that provides access to databases of pathogenic and non-pathogenic LC sequences would greatly improve patient treatment and prognosis by enabling early detection and treatment of amyloidosis. The development is still in its beginnings and further developments and improvements are necessary.

4.6 Limitations and perspectives of this study

This study has some limitations. The low incidence of kappa and/or renal AL resulting in a low sample size hampered this study. This was particularly evident in the subgroup analyses, such as the *IGKV* -in combination with the different *IGKJ*-subfamilies, or the mutations analysis with combined organ stratification. As a result, some analyses only showed tendencies. Analyses based on triple combination of different characteristic, like *IGKVJ* junction in combination with organ involvement should be performed with a larger number of samples to confirm the results. Nonetheless, this study has shown that the multi-dimensional analysis of complex data can indicate causal mechanisms in AL biology.

In terms of lambda analyses, only a selected subset of the entire lambda AL patient cohort was included in this study as this project focussed on patients with dominant kidney involvement and, for comparison, patients with a diverse organ presentation. Further the sequencing of the *IGLV3-21*-subfamily of kidney patients proved to be difficult, because of the increased number of overlapping signals in the Sanger sequencing. To solve this problem NGS targeted sequencing technology was chosen for a detailed analysis, which allows a higher sequencing depth and thus identification of potential sub clonal sequences responsible for overlapping signals in the Sanger sequencing. To analyse the obtained

NGS data, a new analysis pipeline had to be implemented. In addition to the assumed *IGLV3-21*-subfamily, these analyses also revealed other sub-sequences in comparable abundance (Table 44). To clarify this problem, further analyses would be of interest. A potential approach would be, bulk RNA sequencing, which allows the deeper investigation of the sequences as well as their composition. Nevertheless, the inclusion of additional NGS analysis showed results that encourages the investigation of competing sub clones in AL which recently have been shown to play a significant role in progression of MM patients (Poos et al. 2023).

In summary, this study contributes to a broader understanding of AL. Promising results were achieved, for the significantly different subfamily classification in comparison to MM. Differences in the mutation count of kappa versus lambda AL compared to the respective MM were found, showing a differential mutational behaviour of kappa and lambda. Furthermore, the findings on the identification of a significantly increased relative risk of certain patient groups to develop kappa AL could contribute to the development of a personalised treatment strategy for AL.

5 Summary

AL amyloidosis is caused by the misfolding and aggregation of proteins into highly organised amyloid fibrils, which are deposited in the tissues and lead to persistent organ damage. The disease shows a variety of disease manifestations, whereby different organs can be affected. Besides the heart, the kidneys are most frequently affected. And different to healthy individuals and multiple myeloma patients, kappa is less common than lambda and thus data on kappa AL amyloidosis are limited. Aim of this project was the comparative analysis of light chain sequences of kappa AL amyloidosis patients, especially with kidney involvement, in comparison to multiple myeloma patients. 41 kappa AL amyloidosis patients, including 14 patients with dominant kidney involvement, were consecutively enrolled and compared with 83 kappa multiple myeloma patients. Kappa light chain sequences were analysed for specific sequence pattern and their amyloidogenic potential. For this purpose, mutation analyses were performed, but also various biophysical properties such as theoretical isoelectric point, potential N-glycosylation sites, hydrophobicity (GRAVY-score) and the prediction for a β-sheet aggregation (AGG-score) were calculated. The overall aggregation potential was also analysed using an artificial intelligence-based tool. Subsequently, the findings were compared with data from lambda AL amyloidosis patients and for studying specific sequence features the implementation of a next-generation sequences analysis pipeline was another aim of this project. Results revealed a higher frequency of *IGKV1* in AL amyloidosis compared to multiple myeloma patients (AL= 80 %, MM= 53 %, p= 0.002) and *IGKV3* was higher in multiple myeloma (MM= 30 %, AL= 10 %, p= 0.014). *IGKV1/D-33* was significantly more expressed in AL amyloidosis (32 % vs. 13 %, p= 0.016), especially in patients with heart involvement (75 %, p= 0.024). *IGKV1/D-39* was more frequently in patients with kidney involvement (36 %). *IGKV1-16* was only identified in the AL cohort (10 %, p= 0.010) and *IGKV3/D-11* was detected as one of the most abundant subfamilies in multiple myeloma (MM= 14 % vs. AL= 2 %). *IGKJ2* was primarily associated with cardiac involvement (63 %) and *IGKJ1*, *IGKJ2* and *IGKJ3* was equally presented (29 %) in AL amyloidosis patients with kidney involvement. The analysis of the most abundant representatives for AL amyloidosis subfamilies showed a higher median mutation count of the IGKV-regions of kappa AL amyloidosis compared

to multiple myeloma sequences (*IGKV1/D-33 (*AL= 12.0 vs. MM= 10.0); *IGKV1/D-39* (AL= 14.0, MM= 9.5)), which was confirm by the comparative analysis of *IGKV1-16* AL amyloidosis sequences and *IGKV3/D-11* multiple myeloma sequences (AL (*IGKV1-16*)= 9.0, MM (*IGKV3/D-11*)= 5.5, p= 0.033). Interestingly, the lambda AL amyloidosis sequences had a lower median mutation count in comparison to multiple myeloma (IGLV: MM (*IGLV2-23*)= 10.0, AL (*IGLV6-57*)= 4.5, p< 0.001). The analysis of the biophysical properties did not provide a clear pattern, however significantly differences were found for the isoelectric point of *IGLV6-57* AL amyloidosis and *IGLV2-23* multiple myeloma sequences *(*MM (*IGLV2-23*)= 10.0, AL (*IGLV6-57*)= 4.5, p< 0.001), whereby the isoelectric point of the *IGLV2-23* multiple myeloma patients sequences (pI: *IGLV2-23*: 8.12 vs. pI: *IGLV6-57*: 5.0) was closer to the physiological pH and thus indicates greater stability. Patients with kappa AL amyloidosis had significantly more potential N-glycosylation sites than kappa multiple myeloma patients (42.5 % vs. 16.9 %, p= 0.004), but none of the lambda patients showed potential N-glycosylation sites. Systems based on artificial intelligence becoming increasingly popular in the medical sector and the analysis in this project showed also interesting results. 50 % of the *IGKV1-16* AL amyloidosis sequences were classified as amyloidogenic and all the *IGKV3/D-11* multiple myeloma sequences as non-amyloidogenic. In the case of *IGKV1/D-33*, the performance was not consistent, as all AL amyloidosis sequences were classified as amyloidogenic, but also all multiple myeloma sequences. The possibility of the prediction of an increased risk to develop AL amyloidosis would support personalised screening and treatment options. In addition to the increased abundance of *IGKV1/D-33* and *IGKV1-16* in AL amyloidosis, the calculation of the relative risk of these two subgroups stood out. With a rr of 2.452 (p= 0.013) for *IGKV1/D-33* and 18.44 (p= 0.049) for *IGKV1-16*, patients showing an increased risk of developing AL amyloidosis. In conclusion, this study contributes to a better understanding of AL amyloidosis. The results on the identification of an increased risk of certain patient groups for the tendency to develop kappa AL amyloidosis could contribute to the development of an individual treatment strategy.

6 Zusammenfassung

Die AL-Amyloidose wird durch die Fehlfaltung und Aggregation von Proteinen zu hoch organisierten Amyloidfibrillen verursacht, die sich im Gewebe ablagern und zu anhaltenden Organschäden führen. Die Erkrankung zeigt eine Vielzahl von Erscheinungsformen, wobei unterschiedliche Organe betroffen sein können. Neben dem Herzen sind die Nieren am häufigsten betroffen. Anders als in Gesunden und Patient*innen mit Multiplen Myelom ist kappa seltener als lambda und es liegen daher weniger Daten über kappa AL-Amyloidose vor. Ziel dieses Projektes war die vergleichende Analyse der Leichtketten-Sequenzen von Patient*innen mit kappa AL-Amyloidose, insbesondere mit Nierenbeteiligung, im Vergleich zu Patient*innen mit Multiplen Myelom. 41 kappa AL-Amyloidose Patient*innen, darunter 14 Patient*innen mit dominanter Nierenbeteiligung, wurden nacheinander eingeschlossen und mit 83 kappa Multiplen Myelom Patient*innen verglichen. Kappa Leichtkettensequenzen wurden auf spezifische Sequenzmuster und ihr amyloidogenes Potential untersucht. Zu diesem Zweck wurden Mutationsanalysen durchgeführt, aber auch verschiedene biophysikalische Eigenschaften wie der theoretische isoelektrische Punkt, potenziell N-Glykosylierungsstellen, die Hydrophobizität (GRAVY-Score) und die Vorhersage für eine β-Faltblatt Aggregation (AGG-Score) kalkuliert. Das allgemeine Aggregationspotential wurde mit Hilfe eines auf künstliche Intelligenz basierenden Programms analysiert. Anschließend wurden die Ergebnisse mit Daten von lambda AL-Amyloidose Patient*innen verglichen und zur Untersuchung spezifischer Sequenzmerkmale war die Implementierung einer Next-Generation-Sequencing Analysepipeline ein weiteres Ziel dieses Projektes. Die Ergebnisse zeigten eine erhöhte Häufigkeit von *IGKV1* bei AL-Amyloidosen im Vergleich zu Patient*innen mit Multiplen Myelom (AL= 80 %, MM= 53 %, p= 0,002) und *IGKV3* war bei Multiplen Myelom häufiger (MM= 30 %, AL= 10 %, p= 0,014). *IGKV1/D-33* war bei AL-Amyloidose signifikant häufiger exprimiert (32 % vs. 13 %, p= 0,016), insbesondere bei Patient*innen mit Herzbeteiligung (75 %, p= 0,024). *IGKV1/D-39* wurde häufiger bei Patient*innen mit Nierenbeteiligung identifiziert (36 %). *IGKV1-16* wurde nur in der AL-Amyloidose Kohorte nachgewiesen (10 %, p= 0,010) und *IGKV3/D-11* wurde als eine der häufigsten vorkommenden Subgruppe beim Multiplen Myelom detektiert (MM= 14 % vs. AL= 2 %). *IGKJ2* war mit Herzbeteiligung assoziiert (63 %) und *IGKJ1*, *IGKJ2* und *IGKJ3* (29 %) kamen bei AL-Amyloidose Patient*innen mit

Nierenbeteiligung gleichermaßen vor. Die Analyse der häufigsten AL-Amyloidose Subfamilien ergab ein höhere mediane Mutationshäufigkeit der *IGKV*-Regionen der kappa AL-Amyloidose im Vergleich zu Multiplen Myelom Sequenzen (*IGKV1/D-33 (*AL= 12,0 vs. MM= 10,0); *IGKV1/D-39* (AL= 14,0, MM= 9,5)), was durch die vergleichende Analyse der *IGKV1-16* AL-Amyloidose und *IGKV3/D-11* Multiple Myelom Sequenzen bestätigt wurde (AL (*IGKV1-16*)= 9,0, MM (*IGKV3/D-11*)= 5,5, p= 0,033). Interessanterweise wiesen die lambda AL-Amyloidose Sequenzen im Vergleich zu Multiplen Myelom eine geringere mediane Mutationshäufigkeit auf (IGLV: MM (*IGLV2-23*)= 10,0, AL (*IGLV6-57*)= 4,5, p< 0,001). Die Analyse der biophysikalischen Eigenschaften ergab kein klares Muster, jedoch wurden signifikante Unterschiede im isoelektrischen Punkt von *IGLV6-57* AL-Amyloidose und *IGLV2-23* Multiplen Myelom Sequenzen festgestellt (MM (*IGLV2-23*)= 10,0, AL (*IGLV6-57*)= 4,5, p< 0,001), wobei der isoelektrische Punkt der *IGLV2-23* Multiplen Myelom Sequenzen (pI: *IGLV2-23*: 8,12 vs. pI: *IGLV6-57*: 5,0) näher am physiologischen pH-Wert lag, was eine höhere Stabilität andeutet. Patient*innen mit kappa AL-Amyloidose wiesen signifikant mehr potentielle N-Glykosylierungsstellen als Patient*innen mit kappa Multiplen Myelom (42,5 % vs. 16,9 %, p= 0,004), aber keiner der lambda Patient*innen wies potentielle N-Glykosilierungsstellen auf. Systeme, die auf künstlicher Intelligenz basieren werden im medizinischen Bereich immer populärer und die Untersuchung in diesem Projekt zeigte ebenfalls interessante Ergebnisse. 50 % der *IGKV1-16* AL-Amyloidose Sequenzen wurden als amyloidogen eingestuft und alle *IGKV3/D-11* Multiplen Myelom Sequenzen als nicht-amyloidogen. Bei *IGKV1/D-33* war die Leistung nicht befriedigend, da alle AL-Amyloidose Sequenzen als amyloidogen eingestuft wurden, aber auch alle Multiplen Myelom Sequenzen. Die Möglichkeit der Vorhersage eine AL-Amyloidose zu entwickeln, würde personalisierte Screening- und Behandlungsansätze unterstützen. Neben der gesteigerten Häufigkeit von *IGKV1/D-33* and *IGKV1-16* bei AL-Amyloidose, stach die Berechnung des relativen Risikos dieser beiden Untergruppen hervor. Mit einem rr von 2,452 (p= 0,013) für *IGKV1/D-33* und 18,44 (p= 0,049) für *IGKV1-16*, zeigten die Patient*innen ein erhöhtes Risiko der Entwicklung einer kappa AL-Amyloidose. Diese Studie trägt zu einem besseren Verständnis der AL-Amyloidose bei. Die Ergebnisse zur Identifizierung eines erhöhten Risikos bestimmter Patient*innengruppen für die Neigung zur Entwicklung einer kappa AL-Amyloidose könnten die Entwicklung einer individualisierten Behandlungsstrategie vorantreiben.

7 Reference List

Abraham, R. S., Geyer, S. M., Price-Troska, T. L., Allmer, C., Kyle, R. A., Gertz, M. A. and Fonseca, R. (2003). **Immunoglobulin light chain variable (V) region genes influence clinical presentation and outcome in light chain-associated amyloidosis (AL)**. Blood *101 (10)*, 3801-3808, doi: 10.1182/blood-2002-09-2707.

Absmeier, R. M., Rottenaicher, G. J., Svilenov, H. L., Kazman, P. and Buchner, J. (2022). **Antibodies gone bad - the molecular mechanism of light chain amyloidosis**. Febs j *290 (6)*, 1398-1419, doi: 10.1111/febs.16390.

Baker, K. R. (2022). **Light Chain Amyloidosis: Epidemiology, Staging, and Prognostication**. Methodist Debakey Cardiovasc J *18 (2)*, 27-35, doi: 10.14797/mdcvj.1070.

Basu, K., Sinha, R., Ong, A. and Basu, T. (2020). **Artificial Intelligence: How is It Changing Medical Sciences and Its Future?** Indian J Dermatol *65 (5)*, 365-370, doi: 10.4103/ijd.IJD_421_20.

Baur, J., Berghaus, N., Schreiner, S., Hegenbart, U., Schönland, S. O., Wiese, S., Huhn, S. and Haupt, C. (2022). **Identification of AL proteins from 10 λ-AL amyloidosis patients by mass spectrometry extracted from abdominal fat and heart tissue**. Amyloid *30 (1)*, 27-37, doi: 10.1080/13506129.2022.2095618.

Bellotti, V., Mangione, P. and Merlini, G. (2000). **Review: immunoglobulin light chain amyloidosis--the archetype of structural and pathogenic variability**. J Struct Biol *130 (2-3)*, 280-289, doi: 10.1006/jsbi.2000.4248.

Benson, M. D., Liepnieks, J. J. and Kluve-Beckerman, B. (2015). **Hereditary systemic immunoglobulin light-chain amyloidosis**. Blood *125 (21)*, 3281-3286, doi: 10.1182/blood-2014-12-618108.

Berghaus, N., Schreiner, S., Granzow, M., Müller-Tidow, C., Hegenbart, U., Schönland, S. O. and Huhn, S. (2022). **Analysis of the complete lambda light chain germline usage in patients with AL amyloidosis and dominant heart or**

kidney involvement. PLoS One *17 (2)*, e0264407, doi: 10.1371/journal.pone.0264407.

Berghaus, N., Schreiner, S., Poos, A. M., Raab, M. S., Goldschmidt, H., Mai, E. K., Salwender, H. J., Bernhard, H., Thurner, L., Müller-Tidow, C., Weinhold, N., Hegenbart, U., Schönland, S. O. and Huhn, S. (2023). **Comparison of IGLV2-14 light chain sequences of patients with AL amyloidosis or multiple myeloma**. Febs j *290 (17)*, 4256-4267, doi: 10.1111/febs.16805.

Bjellqvist, B., Basse, B., Olsen, E. and Celis, J. E. (1994). **Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions**. Electrophoresis *15 (3-4)*, 529-539, doi: 10.1002/elps.1150150171.

Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J. C., Frutiger, S. and Hochstrasser, D. (1993). **The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences**. Electrophoresis *14 (10)*, 1023-1031, doi: 10.1002/elps.11501401163.

Blancas-Mejía, L. M. and Ramirez-Alvarado, M. (2013). **Systemic amyloidoses**. Annu Rev Biochem *82*, 745-774, doi: 10.1146/annurev-biochem-072611-130030.

Boccadoro, M. and Pileri, A. (1995). **Plasma cell dyscrasias: classification, clinical and laboratory characteristics, and differential diagnosis**. Baillieres Clin Haematol *8 (4)*, 705-719, doi: 10.1016/s0950-3536(05)80255-0.

Bochtler, T., Hegenbart, U., Cremer, F. W., Heiss, C., Benner, A., Hose, D., Moos, M., Bila, J., Bartram, C. R., Ho, A. D., Goldschmidt, H., Jauch, A. and Schonland, S. O. (2008). **Evaluation of the cytogenetic aberration pattern in amyloid light chain amyloidosis as compared with monoclonal gammopathy of undetermined significance reveals common pathways of karyotypic instability**. Blood *111 (9)*, 4700-4705, doi: 10.1182/blood-2007-11-122101.

Bodi, K., Prokaeva, T., Spencer, B., Eberhard, M., Connors, L. H. and Seldin, D. C. (2009). **AL-Base: a visual platform analysis tool for the study of amyloidogenic immunoglobulin light chain sequences**. Amyloid *16 (1)*, 1-8, doi: 10.1080/13506120802676781.

Bolotin, D. A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I. Z., Putintseva, E. V. and Chudakov, D. M. (2015). **MiXCR: software for comprehensive adaptive immunity profiling**. Nat Methods *12 (5)*, 380-381, doi: 10.1038/nmeth.3364.

Buxbaum, J. N. and Linke, R. P. (2012). **A molecular history of the amyloidoses**. J Mol Biol *421 (2-3)*, 142-159, doi: 10.1016/j.jmb.2012.01.024.

Castillo, J. J. (2016). **Plasma Cell Disorders**. Prim Care *43 (4)*, 677-691, doi: 10.1016/j.pop.2016.07.002.

Chakraborty, R. and Lentzsch, S. (2020). **Emerging drugs for the treatment of light chain amyloidosis**. Expert Opin Emerg Drugs *25 (3)*, 299-317, doi: 10.1080/14728214.2020.1803829.

Comenzo, R. L. and Gertz, M. A. (2002). **Autologous stem cell transplantation for primary systemic amyloidosis**. Blood *99 (12)*, 4276-4282, doi: 10.1182/blood.v99.12.4276.

Comenzo, R. L., Vosburgh, E., Simms, R. W., Bergethon, P., Sarnacki, D., Finn, K., Dubrey, S., Faller, D. V., Wright, D. G., Falk, R. H. and Skinner, M. (1996). **Dose-intensive melphalan with blood stem cell support for the treatment of AL amyloidosis: one-year follow-up in five patients**. Blood *88 (7)*, 2801-2806.

Comenzo, R. L., Zhang, Y., Martinez, C., Osman, K. and Herrera, G. A. (2001). **The tropism of organ involvement in primary systemic amyloidosis: contributions of Ig V(L) germ line gene use and clonal plasma cell burden**. Blood *98 (3)*, 714-720, doi: 10.1182/blood.v98.3.714.

Connors, L. H., Jiang, Y., Budnik, M., Théberge, R., Prokaeva, T., Bodi, K. L., Seldin, D. C., Costello, C. E. and Skinner, M. (2007). **Heterogeneity in primary structure, post-translational modifications, and germline gene usage of nine full-length amyloidogenic kappa1 immunoglobulin light chains**. Biochemistry *46 (49)*, 14259-14271, doi: 10.1021/bi7013773.

Consortium, U. (2023). **UniProt: the Universal Protein Knowledgebase in 2023**. Nucleic Acids Res *51 (D1)*, D523-d531, doi: 10.1093/nar/gkac1052.

Costello, J. C., Heiser, L. M., Georgii, E., Gönen, M., Menden, M. P., Wang, N. J., Bansal, M., Ammad-ud-din, M., Hintsanen, P., Khan, S. A., Mpindi, J. P., Kallioniemi, O., Honkela, A., Aittokallio, T., Wennerberg, K., Collins, J. J., Gallahan, D., Singer, D., Saez-Rodriguez, J., Kaski, S., Gray, J. W. and Stolovitzky, G. (2014). **A community effort to assess and improve drug sensitivity prediction algorithms**. Nat Biotechnol *32 (12)*, 1202-1212, doi: 10.1038/nbt.2877.

Dember, L. M. (2006). **Amyloidosis-associated kidney disease**. J Am Soc Nephrol *17 (12)*, 3458-3471, doi: 10.1681/asn.2006050460.

Dispenzieri, A., Gertz, M. A., Kyle, R. A., Lacy, M. Q., Burritt, M. F., Therneau, T. M., Greipp, P. R., Witzig, T. E., Lust, J. A., Rajkumar, S. V., Fonseca, R., Zeldenrust, S. R., McGregor, C. G. and Jaffe, A. S. (2004). **Serum cardiac troponins and N-terminal pro-brain natriuretic peptide: a staging system for primary systemic amyloidosis**. J Clin Oncol *22 (18)*, 3751-3757, doi: 10.1200/jco.2004.03.029.

Dispenzieri, A., Larson, D. R., Rajkumar, S. V., Kyle, R. A., Kumar, S. K., Kourelis, T., Arendt, B., Willrcih, M., Dasari, S. and Murray, D. (2020). **N-glycosylation of monoclonal light chains on routine MASS-FIX testing is a risk factor for MGUS progression**. Leukemia *34 (10)*, 2749-2753, doi: 10.1038/s41375-020-0940-8.

Dittrich, T., Bochtler, T., Kimmich, C., Becker, N., Jauch, A., Goldschmidt, H., Ho, A. D., Hegenbart, U. and Schönland, S. O. (2017). **AL amyloidosis patients with low amyloidogenic free light chain levels at first diagnosis have an excellent prognosis**. Blood *130 (5)*, 632-642, doi: 10.1182/blood-2017-02-767475.

Dudley, D. D., Chaudhuri, J., Bassing, C. H. and Alt, F. W. (2005). **Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences**. Adv Immunol *86*, 43-112, doi: 10.1016/s0065-2776(04)86002-4.

Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V. and Durinx, C. (2021). **Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users**. Nucleic Acids Res *49 (W1)*, W216-w227, doi: 10.1093/nar/gkab225.

Feige, M. J., Hendershot, L. M. and Buchner, J. (2010). **How antibodies fold**. Trends Biochem Sci *35 (4)*, 189-198, doi: 10.1016/j.tibs.2009.11.005.

Feitosa, V. A., Neves, P., Jorge, L. B., Noronha, I. L. and Onuchic, L. F. (2022). **Renal amyloidosis: a new time for a complete diagnosis**. Braz J Med Biol Res *55*, e12284, doi: 10.1590/1414-431X2022e12284.

Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. and Serrano, L. (2004). **Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins**. Nat Biotechnol *22 (10)*, 1302-1306, doi: 10.1038/nbt1012.

Feurstein, S., Zoller, J., Schwab, C., Schreiner, S., Mundt, H., Breitkreutz, I., Schneider, B., Beimler, J., Zeier, M., Waldherr, R., Gröschel, S., Müller-Tidow, C., Schönland, S. O. and Hegenbart, U. (2022). **Concurrent light chain amyloidosis and proximal tubulopathy: Insights into different aggregation behavior-A case report**. EJHaem *3 (4)*, 1377-1380, doi: 10.1002/jha2.555.

Garofalo, M., Piccoli, L., Romeo, M., Barzago, M. M., Ravasio, S., Foglierini, M., Matkovic, M., Sgrignani, J., De Gasparo, R., Prunotto, M., Varani, L., Diomede, L., Michielin, O., Lanzavecchia, A. and Cavalli, A. (2021). **Machine learning analyses of antibody somatic mutations predict immunoglobulin light chain toxicity**. Nat Commun *12 (1)*, 3532-3532, doi: 10.1038/s41467-021-23880-9.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005). **Protein Identification and Analysis Tools on the ExPASy Server**. In: The Proteomics Protocols Handbook, ed. Walker, J. M., Humana Press, Totowa, NJ, pp. 571-607.

Genomics, E. (2023). **Frequently Asked Questions About Our Products And Services**. URL: https://eurofinsgenomics.eu/en/eurofins-genomics/product-faqs/sanger-sequencing/ [as of05.01.2023].

Gertz, M. A., Comenzo, R., Falk, R. H., Fermand, J. P., Hazenberg, B. P., Hawkins, P. N., Merlini, G., Moreau, P., Ronco, P., Sanchorawala, V., Sezer, O., Solomon, A. and Grateau, G. (2005). **Definition of organ involvement and treatment response in immunoglobulin light chain amyloidosis (AL): a consensus opinion from the 10th International Symposium on Amyloid and Amyloidosis, Tours, France, 18-22 April 2004**. Am J Hematol *79 (4)*, 319-328, doi: 10.1002/ajh.20381.

Goldschmidt, H., Mai, E. K., Bertsch, U., Besemer, B., Haenel, M., Miah, K., Fenk, R., Schlenzka, J., Munder, M., Dürig, J., Blau, I. W., Huhn, S., Hose, D., Jauch, A., Kunz, C., Neubauer, A., Scheid, C., Schroers, R., Metzler, I. v., Schieferdecker, A., Thomalla, J., Reimer, P., Mahlberg, R., Graeven, U., Kremers, S., Martens, U. M., Kunz, C., Hensel, M., Seidel-Glaetzer, A., Weisel, K., Raab, M.-S. and Salwender, H. (2021). **Elotuzumab in Combination with Lenalidomide, Bortezomib, Dexamethasone and Autologous Transplantation for Newly-Diagnosed Multiple Myeloma: Results from the Randomized Phase III GMMG-HD6 Trial**. Blood *138 (Supplement 1)*, 486-486, doi: 10.1182/blood-2021-147323.

Greenwald, J. and Riek, R. (2010). **Biology of amyloid: structure, function, and regulation**. Structure *18 (10)*, 1244-1260, doi: 10.1016/j.str.2010.08.009.

Hartl, F. U. and Hayer-Hartl, M. (2009). **Converging concepts of protein folding in vitro and in vivo**. Nat Struct Mol Biol *16 (6)*, 574-581, doi: 10.1038/nsmb.1591.

Hora, M., Sarkar, R., Morris, V., Xue, K., Prade, E., Harding, E., Buchner, J. and Reif, B. (2017). **MAK33 antibody light chain amyloid fibrils are similar to oligomeric precursors**. PLoS One *12 (7)*, e0181799, doi: 10.1371/journal.pone.0181799.

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., Gall, A., Garcia Giron, C., Grego, T., Guijarro-Clarke, C., Haggerty, L., Hemrom, A., Hourlier, T., Izuogu, O. G., Juettemann, T., Kaikala, V., Kay, M., Lavidas, I., Le, T., Lemos, D., Gonzalez Martinez, J., Marugán, J. C., Maurel, T., McMahon, A. C., Mohanan, S., Moore, B., Muffato, M., Oheh, D. N., Paraschas, D., Parker, A., Parton, A., Prosovetskaia, I., Sakthivel, M. P., Salam, A. I. A., Schmitt, B. M., Schuilenburg, H., Sheppard, D., Steed, E., Szpak, M., Szuba, M., Taylor, K., Thormann, A., Threadgold, G., Walts, B., Winterbottom, A., Chakiachvili, M., Chaubal, A., De Silva, N., Flint, B.,

Frankish, A., Hunt, S. E., GR, I. I., Langridge, N., Loveland, J. E., Martin, F. J., Mudge, J. M., Morales, J., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S. J., Cunningham, F., Yates, A. D., Zerbino, D. R. and Flicek, P. (2021). **Ensembl 2021**. Nucleic Acids Res *49 (D1)*, D884-d891, doi: 10.1093/nar/gkaa942.

Huang, J., Chan, S. C., Lok, V., Zhang, L., Lucero-Prisno, D. E., 3rd, Xu, W., Zheng, Z. J., Elcarte, E., Withers, M. and Wong, M. C. S. (2022). **The epidemiological landscape of multiple myeloma: a global cancer registry estimate of disease burden, risk factors, and temporal trends**. Lancet Haematol *9 (9)*, e670-e677, doi: 10.1016/s2352-3026(22)00165-x.

Huart, A. (2023). **Birtamimab: a new amyloidosis treatment?** Blood *142 (14)*, 1178-1180, doi: 10.1182/blood.2023021311.

Huhn, S. (2018). **ELDA qASO-PCR for High Sensitivity Detection of Tumor Cells in Bone Marrow and Peripheral Blood**. Methods Mol Biol *1792*, 1-14, doi: 10.1007/978-1-4939-7865-6_1.

Iadanza, M. G., Jackson, M. P., Hewitt, E. W., Ranson, N. A. and Radford, S. E. (2018). **A new era for understanding amyloid structures and disease**. Nat Rev Mol Cell Biol *19 (12)*, 755-773, doi: 10.1038/s41580-018-0060-8.

Julenius, K. (2007). **NetCGlyc 1.0: prediction of mammalian C-mannosylation sites**. Glycobiology *17 (8)*, 868-876, doi: 10.1093/glycob/cwm050.

Kaplan, B., Livneh, A. and Gallo, G. (2007). **Charge differences between in vivo deposits in immunoglobulin light chain amyloidosis and non-amyloid light chain deposition disease**. Br J Haematol *136 (5)*, 723-728, doi: 10.1111/j.1365-2141.2006.06488.x.

Katzmann, J. A., Kyle, R. A., Benson, J., Larson, D. R., Snyder, M. R., Lust, J. A., Rajkumar, S. V. and Dispenzieri, A. (2009). **Screening panels for detection of monoclonal gammopathies**. Clin Chem *55 (8)*, 1517-1522, doi: 10.1373/clinchem.2009.126664.

Kazman, P., Absmeier, R. M., Engelhardt, H. and Buchner, J. (2021). **Dissection of the amyloid formation pathway in AL amyloidosis**. Nat Commun *12 (1)*, 1-10, doi: 10.1038/s41467-021-26845-0.

Kazman, P., Vielberg, M. T., Pulido Cendales, M. D., Hunziger, L., Weber, B., Hegenbart, U., Zacharias, M., Köhler, R., Schönland, S., Groll, M. and Buchner, J. (2020). **Fatal amyloid formation in a patient's antibody light chain is caused by a single point mutation**. Elife *9*, doi: 10.7554/eLife.52300.

Khwaja, J., Bomsztyk, J., Bygrave, C., Forbes, A., Durairaj, S., Fernandes, S., Taylor, J., Paterson, P., Brearton, G., Crawley, C., Sheehy, O., Brown, R., Soutar, R., Garg, M., Rydzewski, A., Jamroziak, K., Dowling, E., Correia, N., Worthington, S., Mahmood, S. and Wechalekar, A. D. (2023). **High Responses Rates with Single Agent Belantamab Mafodotin in Relapsed Systemic AL Amyloidosis**. Blood *142 (Supplement 1)*, 3408-3408, doi: 10.1182/blood-2023-182042.

Kourelis, T. V., Dasari, S., Theis, J. D., Ramirez-Alvarado, M., Kurtin, P. J., Gertz, M. A., Zeldenrust, S. R., Zenka, R. M., Dogan, A. and Dispenzieri, A. (2017). **Clarifying immunoglobulin gene usage in systemic and localized immunoglobulin light-chain amyloidosis by mass spectrometry**. Blood *129 (3)*, 299-306, doi: 10.1182/blood-2016-10-743997.

Kristinsson, S. Y., Holmberg, E. and Blimark, C. (2013). **Treatment for high-risk smoldering myeloma**. N Engl J Med *369 (18)*, 1762-1763, doi: 10.1056/NEJMc1310911.

Kumar, S., Dispenzieri, A., Lacy, M. Q., Hayman, S. R., Buadi, F. K., Colby, C., Laumann, K., Zeldenrust, S. R., Leung, N., Dingli, D., Greipp, P. R., Lust, J. A., Russell, S. J., Kyle, R. A., Rajkumar, S. V. and Gertz, M. A. (2012). **Revised prognostic staging system for light chain amyloidosis incorporating cardiac biomarkers and serum free light chain measurements**. J Clin Oncol *30 (9)*, 989-995, doi: 10.1200/jco.2011.38.5724.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). **MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms**. Mol Biol Evol *35 (6)*, 1547-1549, doi: 10.1093/molbev/msy096.

Kumar, S. K., Rajkumar, V., Kyle, R. A., van Duin, M., Sonneveld, P., Mateos, M. V., Gay, F. and Anderson, K. C. (2017). **Multiple myeloma**. Nat Rev Dis Primers *3*, 17046, doi: 10.1038/nrdp.2017.46.

Lefranc, M. P., Giudicelli, V., Duroux, P., Jabado-Michaloud, J., Folch, G., Aouinti, S., Carillon, E., Duvergey, H., Houles, A., Paysan-Lafosse, T., Hadi-Saljoqi, S., Sasorith, S., Lefranc, G. and Kossida, S. (2015). **IMGT®, the international ImMunoGeneTics information system® 25 years on**. Nucleic Acids Res *43 (Database issue)*, D413-422, doi: 10.1093/nar/gku1056.

Lefranc, M. P. and Lefranc, G. (2020). **Immunoglobulins or Antibodies: IMGT(®) Bridging Genes, Structures and Functions**. Biomedicines *8 (9)*, doi: 10.3390/biomedicines8090319.

Linding, R., Schymkowitz, J., Rousseau, F., Diella, F. and Serrano, L. (2004). **A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins**. J Mol Biol *342 (1)*, 345-353, doi: 10.1016/j.jmb.2004.06.088.

Lotter, W., Diab, A. R., Haslam, B., Kim, J. G., Grisot, G., Wu, E., Wu, K., Onieva, J. O., Boyer, Y., Boxerman, J. L., Wang, M., Bandler, M., Vijayaraghavan, G. R. and Gregory Sorensen, A. (2021). **Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach**. Nat Med *27 (2)*, 244-249, doi: 10.1038/s41591-020-01174-9.

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A. and Lopez, R. (2022). **Search and sequence analysis tools services from EMBL-EBI in 2022**. Nucleic Acids Res *50 (W1)*, W276-279, doi: 10.1093/nar/gkac240.

Makin, O. S. and Serpell, L. C. (2005). **X-ray diffraction studies of amyloid structure**. Methods Mol Biol *299*, 67-80, doi: 10.1385/1-59259-874-9:067.

Merlini, G. and Bellotti, V. (2003). **Molecular mechanisms of amyloidosis**. N Engl J Med *349 (6)*, 583-596, doi: 10.1056/NEJMra023144.

Merlini, G., Dispenzieri, A., Sanchorawala, V., Schönland, S. O., Palladini, G., Hawkins, P. N. and Gertz, M. A. (2018). **Systemic immunoglobulin light chain amyloidosis**. Nat Rev Dis Primers *4 (1)*, 38, doi: 10.1038/s41572-018-0034-3.

Milani, P., Merlini, G. and Palladini, G. (2018). **Light Chain Amyloidosis**. Mediterr J Hematol Infect Dis *10 (1)*, e2018022, doi: 10.4084/mjhid.2018.022.

NC-IUB, N. C. o. t. I. U. o. B. (1985). **Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984**. Eur J Biochem *150 (1)*, 1-5, doi: 10.1111/j.1432-1033.1985.tb08977.x.

Nevone, A., Girelli, M., Mangiacavalli, S., Paiva, B., Milani, P., Cascino, P., Piscitelli, M., Speranzini, V., Cartia, C. S., Benvenuti, P., Goicoechea, I., Fazio, F., Basset, M., Foli, A., Nanci, M., Mazzini, G., Caminito, S., Sesta, M. A., Casarini, S., Rognoni, P., Lavatelli, F., Petrucci, M. T., Olimpieri, P. P., Ricagno, S., Arcaini, L., Merlini, G., Palladini, G. and Nuvolone, M. (2022). **An N-glycosylation hotspot in immunoglobulin κ light chains is associated with AL amyloidosis**. Leukemia *36 (8)*, 2076-2085, doi: 10.1038/s41375-022-01599-w.

Nienhuis, H. L., Bijzet, J. and Hazenberg, B. P. (2016). **The Prevalence and Management of Systemic Amyloidosis in Western Countries**. Kidney Dis (Basel) *2 (1)*, 10-19, doi: 10.1159/000444206.

Novak, L., Cook, W. J., Herrera, G. A. and Sanders, P. W. (2004). **AL-amyloidosis is underdiagnosed in renal biopsies**. Nephrol Dial Transplant *19 (12)*, 3050-3053, doi: 10.1093/ndt/gfh503.

Olivecrona, M., Blaschke, T., Engkvist, O. and Chen, H. (2017). **Molecular de-novo design through deep reinforcement learning**. J Cheminform *9 (1)*, 48, doi: 10.1186/s13321-017-0235-x.

Omtvedt, L. A., Bailey, D., Renouf, D. V., Davies, M. J., Paramonov, N. A., Haavik, S., Husby, G., Sletten, K. and Hounsell, E. F. (2000). **Glycosylation of immunoglobulin light chains associated with amyloidosis**. Amyloid *7 (4)*, 227-244, doi: 10.3109/13506120009146437.

Palladini, G., Hegenbart, U., Milani, P., Kimmich, C., Foli, A., Ho, A. D., Vidus Rosin, M., Albertini, R., Moratti, R., Merlini, G. and Schönland, S. (2014). **A staging system for renal outcome and early markers of renal response to chemotherapy in AL amyloidosis**. Blood *124 (15)*, 2325-2332, doi: 10.1182/blood-2014-04-570010.

Palladini, G., Perfetti, V., Obici, L., Caccialanza, R., Semino, A., Adami, F., Cavallero, G., Rustichelli, R., Virga, G. and Merlini, G. (2004). **Association of melphalan and high-dose dexamethasone is effective and well tolerated in patients with AL (primary) amyloidosis who are ineligible for stem cell transplantation**. Blood *103 (8)*, 2936-2938, doi: 10.1182/blood-2003-08-2788.

Palladini, G., Sachchithanantham, S., Milani, P., Gillmore, J., Foli, A., Lachmann, H., Basset, M., Hawkins, P., Merlini, G. and Wechalekar, A. D. (2015). **A European collaborative study of cyclophosphamide, bortezomib, and dexamethasone in upfront treatment of systemic AL amyloidosis**. Blood *126 (5)*, 612-615, doi: 10.1182/blood-2015-01-620302.

Perfetti, V., Casarini, S., Palladini, G., Vignarelli, M. C., Klersy, C., Diegoli, M., Ascari, E. and Merlini, G. (2002). **Analysis of V(lambda)-J(lambda) expression in plasma cells from primary (AL) amyloidosis and normal bone marrow identifies 3r (lambdaIII) as a new amyloid-associated germline gene segment**. Blood *100 (3)*, 948-953, doi: 10.1182/blood-2002-01-0114.

Piehl, D. W., Blancas-Mejía, L. M., Ramirez-Alvarado, M. and Rienstra, C. M. (2017a). **Solid-state NMR chemical shift assignments for AL-09 V(L) immunoglobulin light chain fibrils**. Biomol NMR Assign *11 (1)*, 45-50, doi: 10.1007/s12104-016-9718-3.

Piehl, D. W., Blancas-Mejía, L. M., Wall, J. S., Kennel, S. J., Ramirez-Alvarado, M. and Rienstra, C. M. (2017b). **Immunoglobulin Light Chains Form an Extensive and Highly Ordered Fibril Involving the N- and C-Termini**. ACS Omega *2 (2)*, 712-720, doi: 10.1021/acsomega.6b00494.

Poos, A. M., Prokoph, N., Przybilla, M. J., Mallm, J. P., Steiger, S., Seufert, I., John, L., Tirier, S. M., Bauer, K., Baumann, A., Rohleder, J., Munawar, U., Rasche, L., Kortüm, K. M., Giesen, N., Reichert, P., Huhn, S., Müller-Tidow, C., Goldschmidt, H., Stegle, O., Raab, M. S., Rippe, K. and Weinhold, N. (2023). **Resolving therapy resistance mechanisms in multiple myeloma by**

multiomics subclone analysis. Blood *142 (19)*, 1633-1646, doi: 10.1182/blood.2023019758.

Poshusta, T. L., Sikkink, L. A., Leung, N., Clark, R. J., Dispenzieri, A. and Ramirez-Alvarado, M. (2009). **Mutations in specific structural regions of immunoglobulin light chains are associated with free light chain levels in patients with AL amyloidosis**. PLoS One *4 (4)*, e5169, doi: 10.1371/journal.pone.0005169.

Radamaker, L., Baur, J., Huhn, S., Haupt, C., Hegenbart, U., Schönland, S., Bansal, A., Schmidt, M. and Fändrich, M. (2021a). **Cryo-EM reveals structural breaks in a patient-derived amyloid fibril from systemic AL amyloidosis**. Nat Commun *12 (1)*, 1-10, doi: 10.1038/s41467-021-21126-2.

Radamaker, L., Karimi-Farsijani, S., Andreotti, G., Baur, J., Neumann, M., Schreiner, S., Berghaus, N., Motika, R., Haupt, C., Walther, P., Schmidt, V., Huhn, S., Hegenbart, U., Schönland, S. O., Wiese, S., Read, C., Schmidt, M. and Fändrich, M. (2021b). **Role of mutations and post-translational modifications in systemic AL amyloidosis studied by cryo-EM**. Nat Commun *12 (1)*, 1-11, doi: 10.1038/s41467-021-26553-9.

Radamaker, L., Lin, Y. H., Annamalai, K., Huhn, S., Hegenbart, U., Schönland, S. O., Fritz, G., Schmidt, M. and Fändrich, M. (2019). **Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis**. Nat Commun *10 (1)*, 1-8, doi: 10.1038/s41467-019-09032-0.

Rajkumar, S. V. (2022). **Multiple myeloma: 2022 update on diagnosis, risk stratification, and management**. Am J Hematol *97 (8)*, 1086-1107, doi: 10.1002/ajh.26590.

Rajkumar, S. V., Dimopoulos, M. A., Palumbo, A., Blade, J., Merlini, G., Mateos, M. V., Kumar, S., Hillengass, J., Kastritis, E., Richardson, P., Landgren, O., Paiva, B., Dispenzieri, A., Weiss, B., LeLeu, X., Zweegman, S., Lonial, S., Rosinol, L., Zamagni, E., Jagannath, S., Sezer, O., Kristinsson, S. Y., Caers, J., Usmani, S. Z., Lahuerta, J. J., Johnsen, H. E., Beksac, M., Cavo, M., Goldschmidt, H., Terpos, E., Kyle, R. A., Anderson, K. C., Durie, B. G. and Miguel, J. F. (2014). **International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma**. Lancet Oncol *15 (12)*, e538-548, doi: 10.1016/s1470-2045(14)70442-5.

Rajkumar, S. V., Landgren, O. and Mateos, M. V. (2015). **Smoldering multiple myeloma**. Blood *125 (20)*, 3069-3075, doi: 10.1182/blood-2014-09-568899.

Randles, E. G., Thompson, J. R., Martin, D. J. and Ramirez-Alvarado, M. (2009). **Structural alterations within native amyloidogenic immunoglobulin light chains**. Journal of molecular biology *389 (1)*, 199-210, doi: 10.1016/j.jmb.2009.04.010.

Rawat, P., Prabakaran, R., Kumar, S. and Gromiha, M. M. (2021). **Exploring the sequence features determining amyloidosis in human antibody light chains**. Sci Rep *11 (1)*, 1-11, doi: 10.1038/s41598-021-93019-9.

Retter, I., Althaus, H. H., Münch, R. and Müller, W. (2005). **VBASE2, an integrative V gene database**. Nucleic Acids Res *33 (Database issue)*, D671-674, doi: 10.1093/nar/gki088.

Roccatello, D., Fenoglio, R., Sciascia, S., Naretto, C., Rossi, D., Ferro, M., Barreca, A., Malavasi, F. and Baldovino, S. (2020). **CD38 and Anti-CD38 Monoclonal Antibodies in AL Amyloidosis: Targeting Plasma Cells and beyond**. Int J Mol Sci *21 (11)*, 4129, doi: 10.3390/ijms21114129.

Rottenaicher, G. J., Absmeier, R. M., Meier, L., Zacharias, M. and Buchner, J. (2023). **A constant domain mutation in a patient-derived antibody light chain reveals principles of AL amyloidosis**. Commun Biol *6 (1)*, 1-11, doi: 10.1038/s42003-023-04574-y.

Rottenaicher, G. J., Weber, B., Rührnößl, F., Kazman, P., Absmeier, R. M., Hitzenberger, M., Zacharias, M. and Buchner, J. (2021). **Molecular mechanism of amyloidogenic mutations in hypervariable regions of antibody light chains**. J Biol Chem *296*, 100334, doi: 10.1016/j.jbc.2021.100334.

Rousseau, F., Schymkowitz, J. and Serrano, L. (2006). **Protein aggregation and amyloidosis: confusion of the kinds?** Curr Opin Struct Biol *16 (1)*, 118-126, doi: 10.1016/j.sbi.2006.01.011.

Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K. R., Zhao, T., Batiste, R., Van Arnam, J., Shmulevich, I., Rao, A. U. K., Lazar, A. J., Sharma, A. and Thorsson, V. (2018). **Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images**. Cell Rep *23 (1)*, 181-193.e187, doi: 10.1016/j.celrep.2018.03.086.

Salwender, H., Bertsch, U., Weisel, K., Duerig, J., Kunz, C., Benner, A., Blau, I. W., Raab, M. S., Hillengass, J., Hose, D., Huhn, S., Hundemer, M., Andrulis, M., Jauch, A., Seidel-Glaetzer, A., Lindemann, H. W., Hensel, M., Fronhoffs, S., Martens, U., Hansen, T., Wattad, M., Graeven, U., Munder, M., Fenk, R., Haenel, M., Scheid, C. and Goldschmidt, H. (2019). **Rationale and design of the German-speaking myeloma multicenter group (GMMG) trial HD6: a randomized phase III trial on the effect of elotuzumab in VRD induction/consolidation and lenalidomide maintenance in patients with newly diagnosed myeloma**. BMC Cancer *19 (1)*, 504, doi: 10.1186/s12885-019-5600-x.

Sarkar, C., Das, B., Rawat, V. S., Wahlang, J. B., Nongpiur, A., Tiewsoh, I., Lyngdoh, N. M., Das, D., Bidarolli, M. and Sony, H. T. (2023). **Artificial Intelligence and Machine Learning Technology Driven Modern Drug Discovery and Development**. Int J Mol Sci *24 (3)*, 2026, doi: 10.3390/ijms24032026.

Schmittschmitt, J. P. and Scholtz, J. M. (2003). **The role of protein stability, solubility, and net charge in amyloid fibril formation**. Protein Sci *12 (10)*, 2374-2378, doi: 10.1110/ps.03152903.

Schreiner, S., Berghaus, N., Poos, A. M., Raab, M. S., Besemer, B., Fenk, R., Goldschmidt, H., Mai, E. K., Müller-Tidow, C., Weinhold, N., Hegenbart, U., Huhn, S. and Schönland, S. O. (2024). **Sequence diversity of kappa light chains from patients with AL amyloidosis and multiple myeloma**. Amyloid, 1-9, doi: 10.1080/13506129.2023.2295221.

Schroeder, H. W., Jr. and Cavacini, L. (2010). **Structure and function of immunoglobulins**. J Allergy Clin Immunol *125 (2 Suppl 2)*, S41-52, doi: 10.1016/j.jaci.2009.09.046.

Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R. and Sieh, W. (2019). **Deep Learning to Improve Breast Cancer Detection on Screening Mammography**. Sci Rep *9 (1)*, 1-12, doi: 10.1038/s41598-019-48995-4.

Solá, R. J. and Griebenow, K. (2009). **Effects of glycosylation on the stability of protein pharmaceuticals**. J Pharm Sci *98 (4)*, 1223-1245, doi: 10.1002/jps.21504.

Stevens, F. J. (2000). **Four structural risk factors identify most fibril-forming kappa light chains**. Amyloid *7 (3)*, 200-211, doi: 10.3109/13506120009146835.

Sunde, M. and Blake, C. (1997). **The structure of amyloid fibrils by electron microscopy and X-ray diffraction**. Adv Protein Chem *50*, 123-159, doi: 10.1016/s0065-3233(08)60320-4.

Swindells, M. B., Porter, C. T., Couch, M., Hurst, J., Abhinandan, K. R., Nielsen, J. H., Macindoe, G., Hetherington, J. and Martin, A. C. (2017). **abYsis: Integrated Antibody Sequence and Structure-Management, Analysis, and Prediction**. J Mol Biol *429 (3)*, 356-364, doi: 10.1016/j.jmb.2016.08.019.

Swuec, P., Lavatelli, F., Tasaki, M., Paissoni, C., Rognoni, P., Maritan, M., Brambilla, F., Milani, P., Mauri, P., Camilloni, C., Palladini, G., Merlini, G., Ricagno, S. and Bolognesi, M. (2019). **Cryo-EM structure of cardiac amyloid fibrils from an immunoglobulin light chain AL amyloidosis patient**. Nat Commun *10 (1)*, 1-9, doi: 10.1038/s41467-019-09133-w.

Von Behring, E. and Kitasato, S. (1890). **Ueber das Zustandekommen der Diphtherie-Immunität und der Tetanus-Immunität bei Thieren**. Deutschen Medicinischen Wochenschrift *No. 49*.

Wechalekar, A. D., Schonland, S. O., Kastritis, E., Gillmore, J. D., Dimopoulos, M. A., Lane, T., Foli, A., Foard, D., Milani, P., Rannigan, L., Hegenbart, U., Hawkins, P. N., Merlini, G. and Palladini, G. (2013). **A European collaborative study of treatment outcomes in 346 patients with cardiac stage III AL amyloidosis**. Blood *121 (17)*, 3420-3427, doi: 10.1182/blood-2012-12-473066.

Wechalekar, A. D., Silowsky, J., Daniel, E., Harnett, M., Spector, M., Sobolov, S. B., Quarta, C. C., Kurman, M. R., Tulchinskiy, M., Bhattacharyya, A. and Liedtke, M. (2021). **Cardiac Amyloid Reaching for Extended Survival (CARES): Study Design of Two Placebo-Controlled, Double-Blind, Randomized, International Phase 3 Trials Assessing Cael-101 in Patients with Mayo Stage**

Iiia or Stage IIIb AL Amyloidosis. Blood *138 (1)*, 1673-1673, doi: 10.1182/blood-2021-152488.

Yarahuan, J. K. W., Kisvarday, S., Kim, E., Yan, A. P., Nakamura, M. M., Jones, S. B. and Hron, J. D. (2024). **An Algorithm to Assess Guideline Concordance of Antibiotic Choice in Community-Acquired Pneumonia**. Hosp Pediatr *14 (2)*, 137-145, doi: 10.1542/hpeds.2023-007418.

8 Personal Contribution to Data Acquisition / Assessment and Personal Publications

This project was performed within the context of the DFG research group project FOR2969 “Mechanisms of antibody light chain misfolding in systemic AL amyloidosis”. Bone marrow from patients with AL amyloidosis was collected by the physicians as part of a routine clinical screening programme. The clinical organ presentation was defined by Prof. Dr. med. Ute Hegenbart (University Hospital Heidelberg). Bone marrow processing (chapter 2.2.1) was performed in equal parts by Natalie Berghaus (co-doctoral student, University Hospital Heidelberg) and me, irrespective of the clinical presentation (kappa or lambda and organ involvement). From the step of purification of total RNA (chapter 2.2.2) onwards, the entire AL kappa patient cohort and the AL lambda AL_K and AL_D patients were processed by me. 10 kappa and 10 lambda AL patients, who were also included in this study, provided their biomaterial and clinical data prior the start of this project. The PCR method for the analysis of the kappa full-length light chain sequences was established by me, for this purpose an existing method developed by Dr. sc. Hum. Stefanie Huhn (University Hospital Heidelberg) was adapted, extended and verified according to the research question. For the sequencing of the lambda full-length light chain sequences a method developed by Natalie Berghaus was applied. The extraction of the final kappa (AL_H, AL_K and AL_D) and lambda AL_K and AL_D light chain sequences, as well as the interpretation of the data, was done entirely by me. The NGS experiments were performed in a larger run together with Natalie Berghaus and Dr. rer. nat. Philipp Reichert (University Hospital Heidelberg). The NGS raw data processing was performed entirely by me, as well as the interpretation of the lambda AL_K and AL_D sequence data. Dr. rer. Nat. Alexandra Poos provided the bulk RNA data and the clinical data of the MM cohort were provided by the GMMG study group. My personal contribution for the kappa MM data was starting with the translation of the sequences, the verification, analysis and interpretation of the data. For comparison purpose regarding the studied lambda cohort (AL_K and AL_D), the verified lambda MM data were provided by Natalie Berghaus.

Parts of this thesis have already been published in the following articles:

1. Berghaus, N., Schreiner, S., Granzow, M., Müller-Tidow, C., Hegenbart, U., Schönland, S. O. and Huhn, S. (2022). **Analysis of the complete lambda light chain germline usage in patients with AL amyloidosis and dominant heart or kidney involvement.** PLoS One *17 (2)*, e0264407, doi: 10.1371/journal.pone.0264407.

2. Baur, J*., Berghaus, N*., Schreiner, S., Hegenbart, U., Schönland, S. O., Wiese, S., Huhn, S. and Haupt, C. (2022). **Identification of AL proteins from 10 λ-AL amyloidosis patients by mass spectrometry extracted from abdominal fat and heart tissue.** Amyloid *30 (1)*, 27-37, doi: 10.1080/13506129.2022.2095618.

   * These authors contributed equally

3. Berghaus, N., Schreiner, S., Poos, A. M., Raab, M. S., Goldschmidt, H., Mai, E. K., Salwender, HJ., Bernhard, H., Thurner, L., Müller-Tidow, C., Weinhold, N., Hegenbart, U., Schönland S. O. and Huhn, S. (2023). **Comparison of IGLV2-14 light chain sequences of patients with AL amyloidosis or multiple myeloma.** Febs j, *290(17)*, 4256-4267, doi:10.1111/febs.16805.

4. Schreiner, S., Berghaus, N., Poos, A. M., Raab, M. S., Besemer, B., Fenk, R., Goldschmidt, H., Mai, E. K., Müller-Tidow, C., Weinhold, N., Hegenbart, U., Huhn, S. and Schönland S. O. (2024). **Sequence diversity of kappa light chains from patients with AL amyloidosis and multiple myeloma.** Amyloid, 1-9, doi:10.1080/13506129.2023.2295221.

Further publication with personal contribution:

5. Radamaker, L., Karimi-Farsijani, S., Andreotti, G., Baur, J., Neumann, M., Schreiner, S., Berghaus, N., Motika, R., Haupt, C., Walther, P., Schmidt, V., Huhn, S., Hegenbart, U., Schönland, S. O., Wiese, S., Read, C., Schmidt, M. and Fändrich, M. (2021). **Role of mutations and post-translational modifications in systemic AL amyloidosis studied by cryo-EM.** Nature Commun *12(1)*, 1-11, doi:10.1038/s41467-021-26553-9.

6. Feurstein, S., Zoller, J., Schwab, C., Schreiner, S., Mundt, H., Breitkreutz, I., Schneider, B., Beimler, J., Zeier, M., Waldherr, R., Gröschel, S., Müller-Tidow, C., Schönland, S. O. and Hegenbart, U. (2022). **Concurrent light chain amyloidosis and proximal tubulopathy: Insights into different aggregation behavior-A case report.** EJHaem *3 (4)*, 1377-1380, doi: 10.1002/jha2.555.

Publication 1 is based on the results of chapters 3.2.1 and 3.2.2 of this dissertation. The publication includes 25 of the AL_K patients of this thesis, studied as a preliminary data set. My personal contribution to this work was the compilation of the AL_K clinical parameter and the data generation and analysis of the AL_K sequences. I was also involved in the editing of the manuscript.

Publication 2 contains two of the AL_K sequences already included in **publication 1**. My personal contribution to this work was the compilation of the clinical parameter and the analysis and data generation and analysis of the AL_K sequences. I was also involved in the editing of the manuscript.

Publication 3 contains four of the AL_K sequences already included in **publication 1**. My personal contribution to this work was the compilation of the clinical parameter and the data generation and analysis of the AL_K sequences. I was also involved in the editing of the manuscript.

Publication 4 is based on the results of chapters 3.1.1, 3.1.3, 3.1.4.1 - 3.1.4.5 and 3.1.6 of this dissertation. The publication includes the analysis of 41 kappa AL (8 AL_H, 14 AL_K, 19 AL_D) and 83 MM patients. In addition to the overview of the patient characteristics, the publication focuses on the analysis of the *IGKV/J* gene usage of the AL and MM cohorts and considers the clinical organ presentation of the AL patients as well as the influence of a HC on the most common *IGKV*- subgroups. It also includes a detailed mutation analysis, as well as an analysis of the AA composition and the pI of the *IGKV1/D-33*-subfamily. As well as the analysis of a potential N-glycosylation site of kappa and MM cohort. My personal contribution was the data generation, analysis and interpretation of the underlying AL sequences. The MM sequences were provided by Dr. rer. nat Alexandra Poos and my contribution started with the translation and verification of the sequences. I completely analysed and interpreted them. I performed the statistical

analysis and visualisation of the whole data and wrote the manuscript draft, including the introduction, materials and methods, results and discussion.

Publication 5 describes the work of a member of the research group (FOR2969) in which cryo-EM analysis of a patient derived AL fibril was carried out. My part to this work was the sequence analysis of the underlying LC sequences (together with Natalie Berghaus) and I was involved in editing the manuscript.

Publication 6 contains the sequence of one patient with a concurrent AL and a coexisting monoclonal gammopathy of renal significance, light chain proximal tubulopathy. My personal contribution to this work was the sequencing and mutation analysis of the patient-derived LC. I also wrote the respective part of the manuscript.

9 Appendix

Extracted cDNA sequences with respective translated AA sequences of AL-LCs:

The sequences were ordered according to their *IGK/LV* germline usage.

***IGKV1-5*:**

>FOR1009_cDNA
CAGTCTCCTTCCWCCYTGTCTKCATCTGTGGGAGACAGAGTCATCATCACTTGCCGGGC
CAGTCAGAGTATTAGTAGGTGGTTGGCCTGGTATCAGCAGAAACCAGGAAAAGCCCCTA
AGCTCCTGATCYRTGATGCCTCCAGTTTGGAAAGTGGGGTCCCATCAAGGTTCGGCGGC
AGTGGATCTGGGACAGAATTCACTCTCACCATCAGCAGCCTGCAGCCTGATGATTTTGC
AACTTATTACTGCCAACAGTATGATAGTTATTCTCCAGACACTTTTGGCCAGGGGACCA
AGSTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGAT
GAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAG
AGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGA
GTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTG
AGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCT
GAGCTCGCCCGTCACAAAGAGCTTCA
>FOR1009_AA
QSPSXLSXSVGDRVIITCRASQSISRWLAWYQQKPGKAPKLLIXDASSLESGVPSRFGG
SGSGTEFTLTISSLQPDDFATYYCQQYDSYSPDTFGQGTKXEIKRTVAAPSVFIFPPSD
EQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL
SKADYEKHKVYACEVTHQGLSSPVTKSF
>FOR1014_cDNA
TCTCCTTCCACCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCCGGGCCAG
TCAGAGTATTAATAGCTGGTTGGCCTGGTACCAACAGAAAGCAGGGAAAGCCCCTAAGT
TCCTGATCTTTGATGCCTCCAATTTGGAAGGTGGGGTCCCATCAAGGTTCAGCGGCAGT
GGATCTGGGACAGAGTTCAGTCTCACCATCAGCAACCTGCAGCCTGATGATTTTGCAAC
TTATTACTGCCAACAGTACTATAGTGATTCTCCGACTTTTGGCCAGGGGACCAAGCTGG
AGATCAAGCGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAG
TTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
CAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCA
CAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAA
GCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTC
GCCCGTCACAAAGAGCTTCAACAGG
>FOR1014_AA
SPSTLSASVGDRVTITCRASQSINSWLAWYQQKAGKAPKFLIFDASNLEGGVPSRFSGS
GSGTEFSLTISNLQPDDFATYYCQQYYSDSPTFGQGTKLEIKRTVAAPSVFIFPPSDEQ
LKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK
ADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1027_cDNA
CAGTCTCCTTCCACCCTGTCTGCATCTGTTGGAGACAGAGTCACCGTCACTTGCCGGGC
CAGTCAGAGTATTAGTAGTTGGTTGGCCTGGTATCAGCAGAAACCAGGGAAAGCCCCTA
ACCTCCTGATCTATAAGGCATCTAATTTAGCAAGTGGGGTCCCCTCAAGGTTCAGCGGC
AGTGGATCTGGGACAGAATTCACTCTCACCATCAGCAGCCTGCAGCCTGATGATTTTGC
AACTTATTACTGCCAACAGTATGTGAGTTATCCGCTCACTTTCGGCGGAGGGACCAAGG
TGGAGATCAAGCGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG

CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAGGGGG
>FOR1027_AA
QSPSTLSASVGDRVTVTCRASQSISSWLAWYQQKPGKAPNLLIYKASNLASGVPSRFSG
SGSGTEFTLTISSLQPDDFATYYCQQYVSYPLTFGGGTKVEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKLYACEVTHQGLSSPVTKSFNRG
>FOR1040_cDNA
CCAGTCTCCTTCCACCCTGTCTGCATCTGTAGGAGACAGAGTCACTCTCACTTGCCGGG
CCAGTCAGACCATTAGTAATTGGTTGGCCTGGTATCAGCAGAAACCAGGCAAAGCCCCA
AACCTCCTGATCTATGATGCCTCCAGTTTGGAGAGTGGGGTCCCATCAAGGTTCAGCGG
CAATAGATCTGGGACAGAATTCACTCTCACCATCAGCAGCCTGCAGCCTGATGATTTTG
CAACTTATTACTGCCAACAATATAATAGTTATCCGTACACTTTTGGCCCGGGGACCAAC
CTGGACATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGA
GCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAG
AGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGT
GTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAG
CAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGA
GCTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1040_AA
QSPSTLSASVGDRVTLTCRASQTISNWLAWYQQKPGKAPNLLIYDASSLESGVPSRFSG
NRSGTEFTLTISSLQPDDFATYYCQQYNSYPYTFGPGTNLDIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1041_cDNA
CCAGTCTCCTTCCACCCTGTCTGCGTCTGTAGGAGACAGAGTCACTATGACTTGCCGGG
CCAGTCAAAGTCTTAATAGGTGGTTGGCCTGGTATCAGCAGAAACCAGGGCAAGCCCCT
AATCTCCTGATTAAGAAGGCGTCTAGTTTAGAAAGTGGGGTCCCAGCAAGGTTCAGCGG
CAGTGGATCCGGGACAGAATACACTCTCACCATCAGCAGCCTGCAGCCTGATGACTTTG
CAACTTATTACTGCCAACAGTATGACAGTTACCCGTACACTTTTGGCCAGGGGACCAAG
CTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGA
GCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAG
AGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGT
GTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAG
CAAAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCATCAGGGCCTGA
GCTCGCCCGTCACAAAGAGCTTCAACAGG
>FOR1041_AA
QSPSTLSASVGDRVTMTCRASQSLNRWLAWYQQKPGQAPNLLIKKASSLESGVPARFSG
SGSGTEYTLTISSLQPDDFATYYCQQYDSYPYTFGQGTKLEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKLYACEVTHQGLSSPVTKSFNR

***IGKV1-9*:**
>FOR1017_cDNA
TCCAGATGRACCCAGTCTCCATCCTCCCTGTCTGCTTCTGTAGGAGACAGAGTCACCAT
CTCTTGCCGGGCCAGTCAGGTGATTACCAGTTATTTAGTTTGGTATCAGCAAAAACCAG
GGGAAGCCCCTAAGCTCCTGATCTATGCTGCATCCACTTTGCAAAGTGGGGTCCCATCA
AGGTTCAGCGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTGCAGCC

TGAAGATTTTGCAACTTATTACTGTCAACAGCTTAAAACTTACCCGATCACCTTCGGCC
AAGGGACMCGAGTGGATATTAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCG
CCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTT
CTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACT
CCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACC
CTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCA
TCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCA
>FOR1017_AA
SRXTQSPSSLSASVGDRVTISCRASQVITSYLVWYQQKPGEAPKLLIYAASTLQSGVPS
RFSGSGSGTDFTLTISSLQPEDFATYYCQQLKTYPITFGQGTRVDIKRTVAAPSVFIFP
PSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST
LTLSKADYEKHKVYACEVTHQGLSSPVTKSF

IGKV1-12:
>FOR1029_cDNA
CAGTCTCCATCTTCCGTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGTCGGGC
GAGTCAGGGTGTTAACAGCTGGTTAGCCTGGTATCAGCAGAAACCAGGGAAAGCCCCTA
AACTCCTGATCTATTCTGCATCCAGTTTACAAAGTGGGGTCCCATCAAGGTTCAGCGGC
AATGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTGCAGCCTGACGATTTTGC
AACTTACTATTGTCAACAGGCTTCCAATTTCCCGTACACTTTTGGCCAGGGGACCAAGC
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1029_AA
QSPSSVSASVGDRVTITCRASQGVNSWLAWYQQKPGKAPKLLIYSASSLQSGVPSRFSG
NGSGTDFTLTISSLQPDDFATYYCQQASNFPYTFGQGTKLEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN

***IGKV1/D-13*:**
>FOR1037_cDNA
CCAGTCTCCWTCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCCGGG
CNAGTCAGRRMATTAGYAAYTRTTTRRYCTGGTATCAGCAGAAACCAGGGAAAGCCCCT
AAGCTCCTGATCTATGVTGCATCCASTYTGSAAAGTGGGGTCCCATCAAGGTTCAGCGG
CAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTGCAGCCTGAAGATTTTG
CAACTTATTAYTGTCAACAGWATADYARTTTYCCBYNSACTTTYGGCCRAGGGACCAAG
GTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGA
GCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCASAG
AGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCASGAGAGT
GTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAG
CAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGA
GCTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1037_AA
QSPSSLSASVGDRVTITCRASQXISNXLXWYQQKPGKAPKLLIYXASXLXSGVPSRFSG
SGSGTDFTLTISSLQPEDFATYYCQQXXXFXXTFGXGTKVEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPXEAKVQWKVDNALQSGNSXESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN

***IGKV1-16*:**

>FOR1025_cDNA
CAGTCTCCATCTTCACTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGTCGGGC
GAGTCAAGACATTAGCGCTCATTTAGCCTGGTTGCAGCAGAGACCAGGGAAAGCCCCTA
AGTCCCTGATCTTTGCTGCATCCAGTTCGCATAGTGGGGTCCCATCAAAATTCAGCGGC
AATGGATCTGTGACAGTCTTCACTCTCACCATCAGCAACCTGCAGTCTGAAGATTTTGC
GACTTACTACTGTCAACAGTATAAAGTTTACCCGTACACCTTTGGCCCGGGGACCAAGG
TGGAGGTCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1025_AA
QSPSSLSASVGDRVTITCRASQDISAHLAWLQQRPGKAPKSLIFAASSSHSGVPSKFSG
NGSVTVFTLTISNLQSEDFATYYCQQYKVYPYTFGPGTKVEVKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1031_cDNA
CAGTCTCCAGCCTCACTGTCTGCATCTGTGGGAGACAGAGTCACCATCAGTTGTCGGGC
GAGTCAGGCCATTAGCAATTATTTAGCCTGGTATCAGCAGAAACCAGGGCAAGCCCCTA
AGTCCCTGATCTATGCTGCATCCCGTTTGCAAAGTGGGGTCCCATCAAAGTTCCGCGGC
AGTGGATCAGGGACAGATTTCACTCTCACCATCAGCAGCCTGCAGCCTGAAGATTTTGC
AACTTATTACTGCCAACAGTATAATTATTACCCACTCACTTTCGGCGGAGGGACCAAGG
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1031_AA
QSPASLSASVGDRVTISCRASQAISNYLAWYQQKPGQAPKSLIYAASRLQSGVPSKFRG
SGSGTDFTLTISSLQPEDFATYYCQQYNYYPLTFGGGTKVEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1033_cDNA
CAGTCTCCATCCTCACTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGTCGGGC
GAGTCAGGACATTAGTAATTATTTAGCCTGGTTTCAGCAGAAACCAGGGAAAGCCCCTA
AGTCCCTGATCTATGCTGCATCCAGTTTGCACAGTGGGGTCCCATCAAACTTCAGCGGC
AGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTGCAGCCTGAAGATTTTGC
AACTTATTACTGCCAACAATATAATACTTACCCGTGGACGTTCGGCCAAGGGACCAAGG
TGGACATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1033_AA
QSPSSLSASVGDRVTITCRASQDISNYLAWFQQKPGKAPKSLIYAASSLHSGVPSNFSG
SGSGTDFTLTISSLQPEDFATYYCQQYNTYPWTFGQGTKVDIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS

KADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1035_cDNA
CAGTCTCCATCCTCACTGTCTGCATCTATAGGAGACAGAGTCACCCTTACTTGTCGGGC
GAGTCAGGGCATTAGTAATTCTTTAGCCTGGTTTCAGCAGAAACCAGGGAGAGCCCCTA
AGTCCCTCATCTATGCTGCATCCAGTTTGCACAATGGGGTCCCTTCAAGATTCAGCGGC
AGTGGATCAGGGACAAATTTCTCTCTCACCATCACCACCCTGCAGTCTGAAGATTTTGC
AACTTATTACTGCCAACAGTATAATAGTTACCCTCTCTCTTTCGGCGGAGGGACCAACC
TGGACATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1035_AA
QSPSSLSASIGDRVTLTCRASQGISNSLAWFQQKPGRAPKSLIYAASSLHNGVPSRFSG
SGSGTNFSLTITTLQSEDFATYYCQQYNSYPLSFGGGTNLDIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN

***IGKV1/D-33*:**
>FOR1001_cDNA
TCCAGATGACCCAGTCTCCATCCTCCCTGTCCTCTTCTGTGGGAGACAGAGTCACCATC
ACTTGCCAGGCGAGTCAGGACATTGGCAACTATTTAGTTTGGTATCAGCAGAAACCAGG
GAAAGCCCCTAATCTCCTGATCTACGATGCATCCAATTTGGAAACAGGGGTCTCATCAC
GGTTCAGTGGAAGTGGATCTGCGACACATTTTAGTCTCACCATCAGCAGCCTGCAGACT
GAAGATGTTGCAACATATTACTGTCAACAGTATGAAAGTCTCCCGTACACTTTTGGGCA
GGGGACCAAGGTGGAGCTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCT
TCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAAT
AACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGG
TAACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCA
GCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTC
ACCCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGG
>FOR1001_AA
QMTQSPSSLSSSVGDRVTITCQASQDIGNYLVWYQQKPGKAPNLLIYDASNLETGVSSR
FSGSGSATHFSLTISSLQTEDVATYYCQQYESLPYTFGQGTKVELEIKRTVAAPSVFIF
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSS
TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1006_cDNA
CCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTTGGAGACAGAGTCACCATCA
CTTGCCAGGCGAGTCAGGACATTGGCAACTATTTAAATTGGTATCAGCAGAAACCAGGG
AAAGCCCCTAACCTCCTGATCTACGATGCATCCGATTTGAAAACAGGAGTCCCATCAAG
GTTCAGTGGAAGCGGATCTGGGACAGATTTTACTTTCACCATCAGCAGCCTGCAGCCTG
AAGATATTGCAACATATTATTGTCAACAGTATGATAATTTCTTATTCACTTTCGGCCCT
GGGACCAAAGTGGATATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCC
ATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCT
ATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCC
CAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCT
GACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATC
AGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGG
>FOR1006_AA
QMTQSPSSLSASVGDRVTITCQASQDIGNYLNWYQQKPGKAPNLLIYDASDLKTGVPSR

FSGSGSGTDFTFTISSLQPEDIATYYCQQYDNFLFTFGPGTKVDIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1007_cDNA
CCAGATGACCCAGTCTCCATCCTCCCTGTCTGCCTCTGTAGGAGACAGAGTCACCATCA
CTTGCCAGGCGAGTCAAGACATTGGCAACTATTTAAATTGGTATCAGCAAAAATCAGGG
AAAGCCCCTAACCTCCTGATCTACGATGCGTCCAATTTGCAAACAGGGGTCCCATCAAG
GTTCAGTGGAAGTGGATCTGGGACAGATTTTACTTTCACCATCAGCAACCTGCAGCCTG
AAGATACTGCAACATATTACTGTCAACAGTATGATAATCTCCCTAGGTACACTTTTGGC
CAGGGGACCAAGCTGGAGATCAAGCGAACTGTGGCTGCACCATCTGTCTTCATCTTCCC
GCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACT
TCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAAC
TCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCAC
CCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCC
ATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1007_AA
QMTQSPSSLSASVGDRVTITCQASQDIGNYLNWYQQKSGKAPNLLIYDASNLQTGVPSR
FSGSGSGTDFTFTISNLQPEDTATYYCQQYDNLPRYTFGQGTKLEIKRTVAAPSVFIFP
PSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST
LTLSKADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1010_cDNA
CCAGATGACCCAGTCTCCATCCTCCCTGTCGGCATCTGTGGGAGACAGAGTCACCTTCA
CTTGCCAGGCGAGTCAGGCCATTAGAAATTTTGTAAATTGGTATCAGCAGAAGCCAGGT
AAAGCCCCTAAGCTCCTGATCCACGATTCGTCCACTTTGGAAACAGGGGTCTCATCAAG
GTTCAGTGGAAGTGGATCTGGGACCAATTTTTCTTTCACCATCAGCAGCCTGCAGCCTG
AAGATTTTGCAACATATTACTGTCAACAATATGATAATCTTCCTCCGTACACTTTTGGC
CAGGGGACCAAGGTGGAGATGAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCC
GCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACT
TCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAAC
TCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCAC
CCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCC
ATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGG
>FOR1010_AA
QMTQSPSSLSASVGDRVTFTCQASQAIRNFVNWYQQKPGKAPKLLIHDSSTLETGVSSR
FSGSGSGTNFSFTISSLQPEDFATYYCQQYDNLPPYTFGQGTKVEMKRTVAAPSVFIFP
PSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST
LTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1013_cDNA
CCAGATGACCCARTCTCCATCCTTCCTGTCTGCATCTATAGGCGACAGAGTCACCATCT
CCTGCCAGGCGAGTCAGGACATTASTAAYTAYTTAAATTGGTATCAGCACAAACCAGGT
AAAGCCCCTAAACTCCTGATCTACGATGCATCCGATTTGGAAACAGGGGTCCCATCAAG
GTTCAGTGGAAATGGATATGGGACARATTTTACCTTCACCATCAACAGCCTGCAGCCTG
AAGATTTTGCAACATATTACTGTCAACAGTATCATCCTCCCCCGACTTTCGGCGGAGGG
ACCAAGGTGGACATGAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATC
TGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATC
CCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAG
GAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGAC
GCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGG
GCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTG
>FOR1013_AA
QMTQSPSFLSASIGDRVTISCQASQDIXNYLNWYQHKPGKAPKLLIYDASDLETGVPSR

FSGNGYGTXFTFTINSLQPEDFATYYCQQYHPPPTFGGGTKVDMKRTVAAPSVFIFPPS
DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT
LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE
>FOR1018_cDNA
TCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCATTATC
ACTTGCCAGGCGAGTCAGGACATTAGTAATTATTTAAGTTGGTTTCAGCAGAAACCAGG
AAAAGCCCCTAAGCTCCTGATCGACGATGCATCCAATTTGAAAACAGGGGTCCCATCAA
GGTTCAGTGGAGGTGGATTTGGGACAAATTTCACTCTCACCATCAACAACCTGCAGCCT
GAAGATGTTGGAATGTATTACTGCCAACAGTTTTTGAACACTTTTGGCCAGGGGACCAG
GCTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATG
AGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGA
GAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
TGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGA
GCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTG
AGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGA
>FOR1018_AA
QMTQSPSSLSASVGDRVIITCQASQDISNYLSWFQQKPGKAPKLLIDDASNLKTGVPSR
FSGGGFGTNFTLTINNLQPEDVGMYYCQQFLNTFGQGTRLEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFNRGE-
>FOR1020_cDNA
TCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCGTCTGTTGGAGACAGACTCACCATC
ACTTGTCAGGCGAGTCAGGACATTAGCACCTATTTAAATTGGTATCAGCAGAAACCAGG
GAAAGCCCCTAAGATCCTGATCTACCATGCGTCCAGTTTGGAAACAGGGGTCCCATTAA
GGTTCAGTGGAAGTGGATCTGGGACAAATTTTACTTTCACCATCACGAGCCTGCACCTT
GAAGATATTGCAACATATTACTGTCAACAATTTGATAAAGACCCGTACACTTTTGGCCA
GGGGACCAACCTGGAACTCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGC
CATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTC
TATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTC
CCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCC
TGACGCTGAGCAAAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGG
>FOR1020_AA
QMTQSPSSLSASVGDRLTITCQASQDISTYLNWYQQKPGKAPKILIYHASSLETGVPLR
FSGSGSGTNFTFTITSLHLEDIATYYCQQFDKDPYTFGQGTNLELKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKLYACEVTHQGLSSPVTKSFNR
>FOR1021_cDNA
CAGTCTCCTTCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCCAGGC
GAGTCAGGACATTAGCAACTATTTAAATTGGTATCAGCAGAAACCAGGGAAACCCCCTA
AGTTCCTGATCTACGATGCATCTAATTTGGAGACAGGGGTCCCATCAAGGTTCAGTGGA
AATGGATCTGGGACAGATTTTACTTTAACCATCAGCAGCCTGCAACTTGAAGATGTTGC
AACATATTACTGTCAACAGTATGATATTCTCCCCTACACTTTTGGCCAGGGGGCCAAGC
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1021_AA
QSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKPPKFLIYDASNLETGVPSRFSG

NGSGTDFTLTISSLQLEDVATYYCQQYDILPYTFGQGAKLEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1023_cDNA
TCCAGATGACCCAGTCTCCATACTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATC
ACTTGCCAGGCGACTCAGGACATTAGCAACTCTTTAAATTGGTATCAGCAGAAACCAGG
GAAAGCCCCTAAGGTCCTGATCTACGATGTGTCCAATTTGGAAATAGGGGTCCCATCAA
GGTTCAGCGGAGGTGGATCTGGGACACATTTTACTTTCACCATCAGCAGCCTGCAGCCT
GAAGATGTTGCAACATATTACTGTCAACAATATGATAATCTCCCGTGGACGTTCGGCCA
AGGGACCAAGGTGGAAATGAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGC
CATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTC
TATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTC
CCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCC
TGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGG
>FOR1023_AA
QMTQSPYSLSASVGDRVTITCQATQDISNSLNWYQQKPGKAPKVLIYDVSNLEIGVPSR
FSGGGSGTHFTFTISSLQPEDVATYYCQQYDNLPWTFGQGTKVEMKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1024_cDNA
CAGTCTCCATCCTCCCTGTCTGCATCTATAGGAGGCAGAGTCACCATCACTTGCCAGGC
GAGTCAGGACATTAGCGACTTTTTAAATTGGTATCAACAGAGACCGGGGAAAGCCCCTA
AACTCCTGATCTATGATGTTTCCAATCTGGAAACAGGGGTCCCCTCAAGGTTCAGTGGA
AGAGGGTCTGGGACAGACTTTACTTTCACCATCAGCAACCTGCAGTCTGAAGATATTGC
AACATATTACTGTCAACAGTTTAGTGCTCTCCCTCTCACTTTCGGCGGGGGGACCAAGG
TGGAGATCAACCGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1024_AA
QSPSSLSASIGGRVTITCQASQDISDFLNWYQQRPGKAPKLLIYDVSNLETGVPSRFSG
RGSGTDFTFTISNLQSEDIATYYCQQFSALPLTFGGGTKVEINRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1028_cDNA
CAGTCTCCATCCTCCCTGTCTGCATCTGCAGGAGACAGGGTCACCATCACTTGCCAGGC
GAGTCAAGACATTAATTACTATTTAAATTGGTATCAACAGAAGCCAGGGAAAGCCCCTA
AACTCCTGATCTACGATGCATCCAATTTGGAAACAGGGGTCCCATCCAGGTTCAGTGGA
AATGGATCTGGGACAGATTTTACTTTCACCATCAGTAGCCCCCAGTCTGAAGATATTGG
AACATTTTACTGTCAACAATATGATAGTCTCCCGTACACTTTTGGCCTGGGGACCAAGC
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTC
>FOR1028_AA
QSPSSLSASAGDRVTITCQASQDINYYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSG

NGSGTDFTFTISSPQSEDIGTFYCQQYDSLPYTFGLGTKLEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSF
>FOR1036_cDNA
CAGTCTCCATCCTCCCTGTCTGCATCTCTAGGTGACAGAGTCACCATCACTTGCCAGGC
GAGTCAGGACATAAGCAACAATTTAAATTGGTATCAGCAGACATTAGGGAAAGCCCCTA
AATTTCTGATCTACGATGCATCCAATTTGCTAACAGGGGTCCCATCAAGGTTCAGTGGA
AATGGATCTGGGACAGATTTTACTCTCACCATCAGCAACCTGCAGCTTGAAGATATTGC
AACATATTACTGTCAACAGTATCAACAATATGATAATCTTCCTCCGTCAATCACCTTCG
GCCAAGGGACACGACTGGAGATTAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTC
CCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAA
CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTA
ACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGC
ACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCAC
CCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1036_AA
QSPSSLSASLGDRVTITCQASQDISNNLNWYQQTLGKAPKFLIYDASNLLTGVPSRFSG
NGSGTDFTLTISNLQLEDIATYYCQQYQQYDNLPPSITFGQGTRLEIKRTVAAPSVFIF
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSS
TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1039_cDNA
CCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCCAGG
CGAGTCAGGACATTAGCAACTATTTAAGTTGGTATCAGCAGAAACCAGGGAAAGCCCCT
AAGCTCCTGATCTACGATGCATCCAATTTGCAAACAGGGGTCCCATCAAGGTTCAATGG
AAGTGGATCTGGGACAGATTTTACTTTCACCATCAGCAGCCTGCAGCCTGAAGATATTG
CAACATATTACTGTCAACAGTATGATAGTCTCCGTCCGCTCACTTTCGGCGGAGGGACC
AAGGTGGAGAGCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGA
TGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCA
GAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAG
AGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCT
GAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCC
TGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGG
>FOR1039_AA
QSPSSLSASVGDRVTITCQASQDISNYLSWYQQKPGKAPKLLIYDASNLQTGVPSRFNG
SGSGTDFTFTISSLQPEDIATYYCQQYDSLRPLTFGGGTKVESKRTVAAPSVFIFPPSD
EQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL
SKADYEKHKVYACEVTHQGLSSPVTKSFNR

***IGKV1/D-39*:**
>FOR1002_cDNA
CCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCA
CTTGCCGGGCAGGTCACAGCATTAATAGAAATTTAAATTGGTATCAGCAGAAATCAGGG
AAAGCCCCTAACCTCCTGATCTATGCTGCATCCACTTTGCAAAGTGGAGTCTCATCAAG
GTTCAGTGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAAGCTG
AAGATTCTGCAACTTACTACTGTCATCAGACTTACAGTACCCCTCGGACTTTTGGCCAG
GGGACCAAGCTGGAGATCCAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCC
ATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCT
ATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCC
CAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCT
GACGCTGAGCAAAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCATC
AGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAG

>FOR1002_AA
QMTQSPSSLSASVGDRVTITCRAGHSINRNLNWYQQKSGKAPNLLIYAASTLQSGVSSR
FSGSGSGTDFTLTISSLQAEDSATYYCHQTYSTPRTFGQGTKLEIQRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKLYACEVTHQGLSSPVTKSFN
>FOR1004_cDNA
CCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCA
CTTGCCGGGCAAGTCAGAGCATTAGCAAYTATTTAAATTGGTATCAGCAGAAACCAGGG
AAAGCCCCTAARCTCCTGATCTATGMTGCATCCASTTTGCAAAGTGGGGTCCCATCAAG
GTTCAGYGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCARCCTG
AAGATTTTGCAACTTATTACTGTCAACAGWRTWACARTWYCCCKTKCACTTTCGGCCAR
GGGACCAAGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCC
ATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCT
ATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCC
CAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCT
GACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATC
AGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACA
>FOR1004_AA
QMTQSPSSLSASVGDRVTITCRASQSISNYLNWYQQKPGKAPKLLIYXASXLQSGVPSR
FSGSGSGTDFTLTISSLQPEDFATYYCQQXXXXPXTFGQGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1011_cDNA
TCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTTGGAGACAGAGTCACCATC
ACTTGCCGGGCAAGTCAGGACATTAGTATGTATTTAAATTGGTATCAGCAGAAAGCAGG
GAAAGCCCCTAAGTTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGGGGTCCCATCAA
GGTTCAGTGGCAGTGGATCCGTGACAGACTTCACTCTCACCATCAGCAATCTGCAGCCT
GAAGATTTTGCAACTTACTACTGTCAACAGAGTTACAGTGCCCCTCCGACGTTCGGCCA
GGGGACCAAGGTGGAAATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGC
CATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTC
TATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTC
CCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCC
TGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAG
FOR1011_AA
QMTQSPSSLSASVGDRVTITCRASQDISMYLNWYQQKAGKAPKFLIYAASSLQSGVPSR
FSGSGSVTDFTLTISNLQPEDFATYYCQQSYSAPPTFGQGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1012_cDNA
TCCAGATGACCCAGTCTCCATCGTCCCTGTCTGGATCTGTAGGAGACAGCGTCACCATG
ACTTGTCGGGCAAGTCAGAACATTAACAACCATTTGAATTGGTATCAGCAGAAATCAGG
GAAAGCCCCTAACCTCCTGATCTTTGGTGCCTCCAATCTACAAAGTGCGGTCCCATCAA
AGTTCAATGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGTCTGCAACCT
GAAGATTTTGGAACTTACTATTGTCAACAGAGTTACATTACCCCTCTCACTTTCGGCGG
GGGGACCAAGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGC
CATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTC
TATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTC
CCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCC
TGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAG

>FOR1012_AA
QMTQSPSSLSGSVGDSVTMTCRASQNINNHLNWYQQKSGKAPNLLIFGASNLQSAVPSK
FNGSGSGTDFTLTISSLQPEDFGTYYCQQSYITPLTFGGGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTL
TLSKADYEKHKVYACEVTHQGLSSPVTKSFN
>FOR1019_cDNA
AGTCTCCATTGTCCCTGTCTGCATCTGTACGAGACAGAGTCACCATCACTTGCCGGGCA
AGTCAGAACATTAATAACTATTTAAATTGGTATCAGCAGAAACCAGGGGAAGCCCCTAA
AGTCCTGATCTATGCTGCTTCCATTTTGCAAAGTGGGGTCCCATTATGGTTCAGTGGCA
GTAGATCTGGGACAAATTTCACTCTCACCATCAGCAATCTGCAACCTGAAGATTTTGCA
ACTTACTACTGTCAACAGAGTTTCACTACCCCGTGGACGTTCGGCCAAGGGACCAAGGT
GGAACCCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGC
AGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAG
GCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGT
CACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCA
AAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGC
TCGCCCGTCACAAAGAGCTTC
>FOR1019_AA
SPLSLSASVRDRVTITCRASQNINNYLNWYQQKPGEAPKVLIYAASILQSGVPLWFSGS
RSGTNFTLTISNLQPEDFATYYCQQSFTTPWTFGQGTKVEPKRTVAAPSVFIFPPSDEQ
LKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK
ADYEKHKLYACEVTHQGLSSPVTKSF
>FOR1026_cDNA
CAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGMCGGGC
GAGTCAGRRCATTAGCAMYTRTTTRRHYTGKKATCAGCAGAAACCAGGGAAAGCCCCTA
AGCTCCTGATCTATGMTGCRTCYASTTTGCAAAGTGGGGTCCCATCAAGGTTCAGYGGC
AGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGYCTGCARCCTGAAGATTTTGC
AACTTATTACTGTCAACAGWATDAYAGTTTCCCKYDSACTTTCGGCSRAGGGACCAAGG
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAGG
>FOR1026_AA
QSPSSLSASVGDRVTITXRASQXISXXLXXXQQKPGKAPKLLIYXASXLQSGVPSRFSG
SGSGTDFTLTISSLQPEDFATYYCQQXXSFPXTFGXGTKVEIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1032_cDNA
CAGTCTCCATCTTCCCTGTCTGCTTCTGTCGGAGACACAGTCAGCATCACTTGCCGGGC
TAGTCAGGCCATTGGCACTTCTTTACATTGGTATCAGCAGAAACCAGGGAAAGCCCCTA
GCCTCCTAATCTACGCTGCATCCACTTTGCAAAGTGGGGTCCCATCGAGATTCAGTGGC
AGTGGATCTGGGGCAGATTTCACTCTCACCATCACCAGTCTGCAGCCTGAAGATTTTGC
GACTTATCACTGTCAAGAGACCTTCGGTAATCCCTTCACCTTCGGCCAAGGGACACGAC
TGGACATTAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAGG

>FOR1032_AA
QSPSSLSASVGDTVSITCRASQAIGTSLHWYQQKPGKAPSLLIYAASTLQSGVPSRFSG
SGSGADFTLTITSLQPEDFATYHCQETFGNPFTFGQGTRLDIKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1034_cDNA
CAGTCTCCATCCTCCCTGTCTGCATCTATAGGAGACAGAGTCACCATGACTTGCCGGGC
AAGTCAGAGTATTGGCAATTTTTTAAATTGGTATCAGCAGAAACCAGGGAAATCCCCTA
AGCTCCTGATCTATAGTGCATCCAGTTTGCATAGTGGGGTCCCATCAAGGTTCAGTGGC
AGTGCATCTGGGACAGATTTCTCTCTCACCATCAGCGGTCTGCAACCTGAAGATCTTGG
AACTTACTACTGTCAACAGAGTTTCACCCCCCGGACGTTCGGCCAAGGGACCAAGGTGG
AAATCGAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAG
TTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
CAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCA
CAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAA
GCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTC
GCCCGTCACAAAGAGCTTCAACAG
>FOR1034_AA
QSPSSLSASIGDRVTMTCRASQSIGNFLNWYQQKPGKSPKLLIYSASSLHSGVPSRFSG
SASGTDFSLTISGLQPEDLGTYYCQQSFTPRTFGQGTKVEIERTVAAPSVFIFPPSDEQ
LKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK
ADYEKHKVYACEVTHQGLSSPVTKSFN

***IGKV3/D-11*:**
>FOR1038_cDNA
CAGTCAGAGTGTTAGCAGCTACTTAGCCTGGTTCCAACAGAAACCTGGCCAGGCTCCCA
GGCTTCTCATCTTTGATGCATCCAACAGGGCCACTGGCATCCCAGCCAGGTTCAGTGGC
AGTGGGTCTGGGACAGACTTCACTCTCACCATCAGCAGCCTAGAGCCTGAAGATTTCGC
AGTTTATTACTGTCAACAGCGCAGCAACTGGCCTCTCACTTTCGGCGGAGGGACCAAGG
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC
AAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1038_AA
SQSVSSYLAWFQQKPGQAPRLLIFDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFA
VYYCQQRSNWPLTFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPRE
AKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLS
SPVTKSFN

***IGKV3-15*:**
>FOR1005_cDNA
CAGTCAGAGTGTTAGCCGCAACTTAGCCTGGTACCAGCAGAAGCCTGGCCAGGCTCCCA
GGCTCCTCGTCTATGATGCATTCACCAGGGCAACTGGTGTCCCAGCCAGGTTCAATGGC
AGTGGGTCTGGGACAGAGTTCACTCTCACCACCAGCACCCTGCAGTCTGAGGATTTTGC
AGTTTATTACTGTCAGCAGTATAATAACTGGCCGGGCACTTTTGGCCAGGGGACCAAGC
TGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAG
CAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGA
GGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTG
TCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGC

AAAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAG
CTCGCCCGTCACAAAGAGCTTCAACA
>FOR1005_AA
SQSVSRNLAWYQQKPGQAPRLLVYDAFTRATGVPARFNGSGSGTEFTLTTSTLQSEDFA
VYYCQQYNNWPGTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPRE
AKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKLYACEVTHQGLS
SPVTKSFN
>FOR1008_cDNA
GACTCCAGGGGAAAGAGCCACCCTCTCCTGCAGGGCCAGTCAGAATCTTACCAACAACT
TAGCCTGGTATCAGCAGAAACCTGGCCGGGCTCCCAGGCTCCTCATCTATGATGCATCC
ACCAGGGCCACTAATATCCCAGCCAGGTTCAGCGGCAGTGGGTCCGGGACAGAGTTCAC
TCTCACCATCAGCAGCCTGCAGTCTGAAGATTTCGCAGTTTATTTCTGTCAGCAGTATA
TTGCCTGGCCTCCGATCACCTTCGGCCAGGGGACACGACTGGACATTAAACGAACTGTG
GCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGC
CTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGG
TGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAG
GACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACA
CAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCT
TCAACAG
>FOR1008_AA
TPGERATLSCRASQNLTNNLAWYQQKPGRAPRLLIYDASTRATNIPARFSGSGSGTEFT
LTISSLQSEDFAVYFCQQYIAWPPITFGQGTRLDIKRTVAAPSVFIFPPSDEQLKSGTA
SVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
KVYACEVTHQGLSSPVTKSFN
>FOR1022_cDNA
TCTGCAGGGCCAGTCAGAGTTTTACCAACAACTTAGCCTGGTATCAGCAGAAACCTGGC
CAGGCTCCCAGGCTCCTCATCTTTGGTGCATCCACCAGGGCCACTGGTATCCCTGCCAG
GTTCAGTGGCAGTGGGTCTGGGACAGAGTTCACTCTCACCATCAGCAGCCTGCAGTCTG
AAGATTTTGCAGTTTATTACTGTCAGCAGTATAATGCCTGGCCTCTGACGTTCGGCCAA
GGGACCAAGGTGGAAAGCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCC
ATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCT
ATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCC
CAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCT
GACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATC
AGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAG
>FOR1022_AA
CRASQSFTNNLAWYQQKPGQAPRLLIFGASTRATGIPARFSGSGSGTEFTLTISSLQSE
DFAVYYCQQYNAWPLTFGQGTKVESKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY
PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQ
GLSSPVTKSFN

***IGKV4-1*:**
>FOR1015_cDNA
TCTCCAGACTCCCTGGCTGTGTCTCTGGGCGAGAAGGCCACCATCAACTGCAAGTCCAG
CCAGACTCTTTTATACAGTTCCAATGATAAAAACTACTTACTTTGGTATCAGCAGAGAC
CAGGACAGTCTCCTAAGCTACTCATTTCCTGGGCATCTACCCGGGAATCCGGGGTCCCT
GACCGATTCAGTGGCAGCGGATCTGGGACAGATTTCACTCTCACCATCACCAGCCTGCA
GGCTGAAGATGTGGCAGTTTATTACTGTCAGCAATATTTTAGTAGTCCTTTCACTTTCG
GCCCTGGGACCAAAGTGGATATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTC
CCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAA
CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTA

ACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGC
ACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCAC
CCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGG
>FOR1015_AA
SPDSLAVSLGEKATINCKSSQTLLYSSNDKNYLLWYQQRPGQSPKLLISWASTRESGVP
DRFSGSGSGTDFTLTITSLQAEDVAVYYCQQYFSSPFTFGPGTKVDIKRTVAAPSVFIF
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSS
TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR
>FOR1030_cDNA
CAGTCTCCAGACTCCCTGGCTGTGTCTCTGGGCGAGAGGGCCACCATCAACTGCAAGTC
CAGCCAGAGTGTTTTATACAGTTCAAATAAGAATTACTTAGCTTGGTACCAGCAGAAAA
CAGGACAGTCCCCTAGGCTGCTCATTTATTGGGCATCTACCCGGGAATCCGGGGTCCCT
GACCGATTCAGTGGCAGCGGGTCTGGGACAGATTTCACTCTCACCATCAGCAGCCTGCA
GGCTGAAGATGTGGCAGTTTATTACTGTCAGCAATATTATAGTACTCCTCTCACTTTCG
GCGGAGGGACCAAGGTGGAGATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTC
CCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAA
CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTA
ACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGC
ACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCAC
CCATCAGGGCCTGAGCTCGCCCGTCACAAARAGCTTCAACA
>FOR1030_AA
QSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAWYQQKTGQSPRLLIYWASTRESGVP
DRFSGSGSGTDFTLTISSLQAEDVAVYYCQQYYSTPLTFGGGTKVEIKRTVAAPSVFIF
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSS
TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFN

***IGKV5-2*:**
>FOR1003_cDNA
TCATGTCAGCGACTCCAGGAGACACAGTCAGCATCTCCTGCACAGCCAGCCAAGACATT
GATGATGATATGAAYTGGTACCARMRGAAACCRGGAGAASCTGYTHTTYTCWTTATTCA
AGAASMYAMTACTCTCGTTCCTGGAATCCCACCTCGATTCAGTGGCAGCGGSTATGGAA
CAGATTTTACCCTCACAATTAATAAYATAGARKCTGAGGATGCTGCATWTTACTTCTGT
CTRCAACATGATAATTTCCCYYWSACGTTCGGCCARGGGACCMAGGTGGARRTCAAACG
AACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTG
GAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAGTACAG
TGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGAGCAGGA
CAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACG
AGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACA
AAGAGCTTCAACAG
>FOR1003_AA
MSATPGDTVSISCTASQDIDDDMNWYPXKPGEXXXXXIQEXXTLVPGIPPRFSGSGYGT
DFTLTINNIEXEDAAXYFCLQHDNFPXTFGQGTXVEXKRTVAAPSVFIFPPSDEQLKSG
TASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
KHKVYACEVTHQGLSSPVTKSFN

n/a:
>FOR1016
-

***IGLV1-40*:**

>FOR145_cDNA
TGGTCCTGGGCTCAGTCTGTGCTGACGCAGCCGCCCTCAGTGTCTGGGGCCCCAGGGCA
GAGGATCACCATCTCCTGCACTGGGAGCAACTCCAACGTCGGGGCAGGTTATGATGTAC
ACTGGTACCAGCAGATTCCGGGASCAGCCCCCAAACTCSTCATCTATGMTAACAGGAAT
CGGCCCTCAGGGGTCCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCT
GRCCATCACTGGTSTCCAGRCTGAGGATGAGGCTGATTATTCCTGYCAGTCGTTKGACA
GCAGCMKGAATGATTGGGTTTTCGGCGGAGGGACYAAGCTGACCGTCCTASGTCAGCCC
AAGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAA
GGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGA
AGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGC
AACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCA
CAAAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR145_AA
WSWAQSVLTQPPSVSGAPGQRITISCTGSNSNVGAGYDVHWYQQIPGXAPKLXIYXNRN
RPSGVPDRFSGSKSGTSASLXITGXQXEDEADYSCQSXDSSXNDWVFGGGTKLTVLXQP
KAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQS
NNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTV
>FOR156_cDNA
TGGTCCTGGGCTCAGTCTGTGCTGACGCAGCCGCCCTCAGTGTCTGGGGCCCCAGGGCA
GAGGGTCACCATCTCCTGCACTGGGAGCAGCTCCRACATCGGGACAGGTTATGATGTAC
ACTGGTACCAGCACCTTCCAAGAACAGCCCCCAAACTCCTCATCTATGATAACAGGAAT
CGGCCCTCAGGGGTCCCTGACCGCTTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCT
GGACATCACTGGGCTCCAGACTGACGATGAGGGTGATTATTACTGCCAGTCCTATGACG
GCAGCCTGGGTGGTTGGGTGTTCGGCGGTGGGACCAAGGTGACCGTCCkAGGTCAGCCC
AAGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGYCAACAA
GGCCACACTGGTGTGTYTSATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGA
AGGCAGATAGCAGCCCCGTCAAGGCGKGAGTGGAGACCACCACACCCTCCAAACAAAGC
AACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCA
CAAAAGCTACAGC
>FOR156_AA
WSWAQSVLTQPPSVSGAPGQRVTISCTGSSSXIGTGYDVHWYQHLPRTAPKLLIYDNRN
RPSGVPDRFSGSKSGTSASLDITGLQTDDEGDYYCQSYDGSLGGWVFGGGTKVTVXGQP
KAAPSVTLFPPSSEELQXNKATLVCXISDFYPGAVTVAWKADSSPVKAGVETTTPSKQS
NNKYAASSYLSLTPEQWKSHKSYS

***IGLV1-44*:**

>FOR107_cDNA
TGGTCCTGGGCTCAGTCTGTGCTGACTCAGCCACCCTCAGCGTCTGGGACCCCCGGGCA
GAGGGTCACCATCTCTTGTTCTGGAAGCGGGTCCAACATCGGAGGTAATTCTGTAAATT
GGTACCAGCACCTCCCAGGAGCGGCCCCCAAATTTCTCATCTATAGTAATAATCAGCGG
CCCTCAGGGGTCCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGGC
CATCAGTGGGCTCCAGTCTGAGGATGAGGCTGATTACTACTGTGCAACATGGGATGACA
GCCTGAATGGAGTGATATTCGGCGGGGGGACCAAGCTGACCGTCCTGGGTCAGCCCAAG
GCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGC
CACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGG
CAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAAC
AACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAG
AAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGA
>FOR107_AA
WSWAQSVLTQPPSASGTPGQRVTISCSGSGSNIGGNSVNWYQHLPGAAPKFLIYSNNQR

PSGVPDRFSGSKSGTSASLAISGLQSEDEADYYCATWDDSLNGVIFGGGTKLTVLGQPK
AAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSN
NKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVE
>FOR135_cDNA
TCAGTCTGTGCTGACTCAGCCACCCTCAGCGTCTGGGACCCCCGGGCAGAGGGTCACCA
TCTCTTGTTCTGGAAGCAGTTCCAACATCGGTGGTAATACTATAATCTGGTACCAGCAT
CTCCCAGGAACGACCCCCAGACTCCTCATCCAGAGTAGTGATCAGCGGCCCTCAGGGGT
CCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCTGCCTCCCTGGCCATCAGTGGGC
TCCAGTCTGACGATGAGGCTGATTATTACTGTGCTGCCTGGGATGACAGCCTGAATGGT
CCGGTTTTCGGCGGAGGGACCAAGCTGACCGTCCTAAGTCAGCCCAAGGCTGCCCCCTC
GGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAGCAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGC
CCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGC
GGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCT
GCCAGGTCACGCATGAAGGGAGCACCGTG
>FOR135_AA
QSVLTQPPSASGTPGQRVTISCSGSSSNIGGNTIIWYQHLPGTTPRLLIQSSDQRPSGV
PDRFSGSKSGTSASLAISGLQSDDEADYYCAAWDDSLNGPVFGGGTKLTVLSQPKAAPS
VTLFPPSSEELQASKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTV
>FOR137_cDNA
TCAGTCTGAACTGACCCAGCCACCCTCAGCGTCTGGGACCCCCGGGCAGAGGGTCACCA
TCTCTTGTTCTGGAGGCAGCTCCAACATCGGCAGTAATTTTGTGAACTGGTACCAGCAC
CTCCCAGGAACGGCCCCCAAACTCCTCATCTATACTATTAATCAGCGACCCTCAGGGGT
CCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGGCCATCAGTGGGC
TCCGGTCTGAGGATGAGGCTGATTATTACTGTTCAGCATGGGATGACAGCCTGAATGGT
CCGGTGTTCGGCGGAGGGACCAAGCTGACCGTCCTGGGTCAGCCCAAGGCTGCCCCCTC
GGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGC
CCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGC
GGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCT
GCCAGGTCACGCATGAAGGGAGCACCGTG
>FOR137_AA
QSELTQPPSASGTPGQRVTISCSGGSSNIGSNFVNWYQHLPGTAPKLLIYTINQRPSGV
PDRFSGSKSGTSASLAISGLRSEDEADYYCSAWDDSLNGPVFGGGTKLTVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTV
>FOR142_cDNA
TGGTCCTGGGGCTCAGTCTGTGCTGACTCAGCCACCCTCAGCGTCTGGGACCCCCGGGC
AGAGGGTCACCATCTCTTGTTCTGGAAGCAGCTCCAACCTCGGAAGTAATACTGTTAAC
TGGTACCAGCACCTCCCAGGAMCGGCCCCCAAACTCCTCATCTATAGTAATAATCAGCG
GCCCTCAGGGGTCTCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGG
CCATCGGTGGGCTCCAGTCTGACGATGAGTCTGATTATTACTGTGCAGCATGGGATGAC
AGCCTGAATGGTCCGGTATTCGGCGGGGGGACCAAGCTGACCGTCCTAGGTCAGCCCAA
GGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGG
CCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAG
GCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACArAGCAA
CAACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACA
GAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR142_AA
GPGAQSVLTQPPSASGTPGQRVTISCSGSSSNLGSNTVNWYQHLPGXAPKLLIYSNNQR

PSGVSDRFSGSKSGTSASLAIGGLQSDDESDYYCAAWDDSLNGPVFGGGTKLTVLGQPK
AAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSN
NKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTV
>FOR180_cDNA
TGGTCCTGGGCTCAGTCTGTGCTGACTCAGCCACCCTCAGCGTCTGGGACCCCCGGGCA
GAGGGTCACCATCTCTTGTTCTGGAAGCAGCTCCAATATCGGAAGTAATACTGTAAACT
GGTACCAGCACCTCCCAGGAACGGCCCCCAAACTCGTCATCTATAGTAATAATCTGCGG
CCCTCAGGGGTCCCTGACCGAATCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGGC
CATCAGTGGGCTCCAGTCTGAGGATGAGGCTGATTATTACTGTGCAGCATGGGATGACA
GCCTGAATGGTTCGATGTTCGGCGGAGGGACCAGGCTGACCGTCCTAGGTCAGCCCAAG
GCTGCCCCCTCGGTCACTCTGTTCCCACCCTCCTCTGAGGAGCTTCAAGCCAACAAGGC
CACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGG
CAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAAC
AACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAA
AAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGG
>FOR180_AA
WSWAQSVLTQPPSASGTPGQRVTISCSGSSSNIGSNTVNWYQHLPGTAPKLVIYSNNLR
PSGVPDRISGSKSGTSASLAISGLQSEDEADYYCAAWDDSLNGSMFGGGTRLTVLGQPK
AAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSN
NKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTV
>FOR184_cDNA
TCAGTCTGTGCTGACTCAGCCACCCTCAGCGTCTGAGACCCCCGGGCAGAGGGTCACCA
TCTCTTGTTCTGGAAGCAGGTCCAACATCGGAAGTAATACTGTAAACTGGTACCAGCAC
GTCCCAGGAATGGCCCCCAAACTCCTCATCTATAGCAATAATCAGCGACCCTCAGGGGT
CCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGGCCATCAGTGGGC
TCCAGTCTGAGGATGAGGCTGATTATTACTGTGCAGCATGGGATGACAGTCTTAATGGT
CCGAAGTTCGGCGGAGGGACCAAACTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTC
GGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGC
CCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGC
GGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGT
>FOR184_AA
QSVLTQPPSASETPGQRVTISCSGSRSNIGSNTVNWYQHVPGMAPKLLIYSNNQRPSGV
PDRFSGSKSGTSASLAISGLQSEDEADYYCAAWDDSLNGPKFGGGTKLTVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQ
>FOR213_cDNA
TCAGTCTGTGCTGACTCAGTCACCCTCCGCGTCTGGGACCCCCGGGCAGACGGTTACCA
TCTCTTGTTCTGGCASCAGCTCCAACATCGGAGGTAATACTGTAAAYTGGTACCAGCAC
CTCCCAGGAMCGGCCCCCAAACTCCTCATGTTTATTAATAATCAGCGGCCCTCAKGGGT
CCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGGCCATCAGTGGGS
TCCAGTCTGMGGATGAGGCTGATTATTACTGTSCCGCATGGGATGACAGCMTGAATGGT
CCGGTGTTCGGCGGAGGGACCARGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTC
GGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGC
CCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACARAGCAACAACAAGTACGC
GGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCT
GCCAGGTCACGCATGAAGGGAGCACCGTGGA
>FOR213_AA
QSVLTQSPSASGTPGQTVTISCSGTSSNIGGNTVNWYQHLPGXAPKLLMFINNQRPSGV
PDRFSGSKSGTSASLAISGXQSXDEADYYCXAWDDSXNGPVFGGGTXLTVLGQPKAAPS

VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTV
>FOR217_cDNA
TCAGTCTGTGCTGACTCAGCCACCCTCAACGTCTGGGACCCCCGGGCAGAGGGTCACCA
TTTCTTGTTCTGGAAGCAGCTCCAACATCGGAAGTAATACTGTAAACTGGTACCAGCAG
CTCCCAGGAACGGCCCCCAAACTCCTCATCTATAGTAATAATGAGCGGCCCTCAGGGGT
CCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCTCCCTGGCCATCAGTGGGC
TCCAGTCTGACGATGAGGCTGATTATTACTGTGCAGCATGGGATGACAGCCTGAGTGGA
TCTTGGGTGTTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCC
CTCGGTCACTCTGTTCCCACCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGG
TGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGC
AGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTA
CGCGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAAAAGCTACA
GCTGCCAGGTCACGCATGAAGGGAGCACCGTGG
>FOR217_AA
QSVLTQPPSTSGTPGQRVTISCSGSSSNIGSNTVNWYQQLPGTAPKLLIYSNNERPSGV
PDRFSGSKSGTSASLAISGLQSDDEADYYCAAWDDSLSGSWVFGGGTKLTVLGQPKAAP
SVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKY
AASSYLSLTPEQWKSHKSYSCQVTHEGSTV
>FOR240_cDNA
GGTCCTGGGMTCAGTCTGTGCTGACTCAGCCACCCTCAGCGTCTGGGACCCCCGGGCAG
AGGGTCACCATCTCSTGTTCTGGAASCAGCTCCAACATCGGGGDTAATGATGTTAACTG
GTACCAGCAGSTCCCAGGRACGGCCCCCAAACTCCTCATCTATAATAATDATAAGCGGC
CCTCAGGGGTCCCTGACCGATTCTCTGCCTCCAAGTMTGGCACCTCAGCCTCCCTGGCC
ATCAGTGGGCTCCAGTMTGAGGATGAGGMTRATTATTACTGTGCAGCRTGGGATRACAG
CMTGAATGGTTGGGTGTTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCAAGG
CTGCCCCCTCGGTCACTCTGTTCCCRCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCC
ACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGC
AGATRGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACARAGCAACA
ACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCCAGTGGAAGTCCCACAG
AAGC
>FOR240_AA
SWXQSVLTQPPSASGTPGQRVTISCSGXSSNIGXNDVNWYQQXPGTAPKLLIYNNXKRP
SGVPDRFSASKXGTSASLAISGLQXEDEXXYYCAAWDXSXNGWVFGGGTKLTVLGQPKA
APSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADXSPVKAGVETTTPSKQSNN
KYAASSYLSLTPEPVEVPQK

***IGLV1-47*:**

>FOR211_cDNA
ATCAGTCTGTGVTGRATCAGCCACCMTCAKCGTCTGCGACTCCAGGGCAGAKGGTCAYC
ATCTCCTGCTSTRGMAGCAGCWCCAACGTTGGWAGTAATTATSWTTATTGGTACCAGCA
GCAGCCAGGACMGGCCCCYGTRCTSGTCATCTATGATRATAACRASCGGCCCTCAGGGR
TCCCTGACCGATTCTCTKGCTCCAACTCDGGSACCACRGCCWCCNTSACCATCAKYAGG
GTSSAGGCGGAGGAKGAGGCTGATTATTAYTGTCAVKCRTGGGAYAGCAGYAGTAAYCA
TKDGGTRTTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCT
CGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTG
TGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAG
CCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACG
CGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAG
>FOR211_AA
QSVXXQPPSXSATPGQXVXISCXXSSXNVGSNYXYWYQQQPGXAPVLVIYDXNXRPSGX

PDRFSXSNSGTTAXXTIXRVXAEXEADYYCXXWDSSSNHXVFGGGTKLTVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSY

***IGLV1-51*:**

>FOR109_cDNA
CTCAGTCTGTGTTGACGCAGCCGCCCTCAGTGTCTGCGGCCCCAGGACAGACGGTCACC
ATCTCCTGCTCTGGAAGCAGCTCCAACATTGAGAATAATTATGTAACCTGGTACCAGCA
GCTCCCAGGAACAGCCCCCAAACTCCTCATTTATGACAATAATGAGCGACCCTCAACAA
TTCCTGACCGATTCTCTGGCTCCAAGTCTGGCGCGTCAGCCACCCTGGGCATCACCGGA
CTCCAGACTGGGGACGAGGCCGATTATTATTGCGGAACATGGGATAGCAGCCTGAGTGC
TGGCTGGGTGTTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCC
CCTCGGTCACTCTGTTCCCACCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTG
GTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAG
CAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGT
ACGCGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAAAAGCTAC
AGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGA
>FOR109_AA
QSVLTQPPSVSAAPGQTVTISCSGSSSNIENNYVTWYQQLPGTAPKLLIYDNNERPSTI
PDRFSGSKSGASATLGITGLQTGDEADYYCGTWDSSLSAGWVFGGGTKLTVLGQPKAAP
SVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKY
AASSYLSLTPEQWKSHKSYSCQVTHEGSTVE
>FOR119_cDNA
CTCWGTCTGTGTTGACGCAGCCGCCCTCAGTGTCTGCGGCCCCAGGACAGAAGGTCACC
ATCTCCTGCTCTGGAAGCAGCTCCAATATTGGGAGTACTTATGTGTCGTGGTACCGGCA
ACTCCCAGGAACAGCCCCCAGACTCCTCATTTATGACANTTATAAGCGACCCTCAGGGA
TTCCTGACCGATTCTCTGGCTCCAAGTCTGGCACCTCAGCCACCCTGGGCATCACCGGA
CTCCAGACTGGGGACGCGGCCGATTATTACTGCGGAACATGGGATAGCAGCCTGAGTGT
TTATGTCTTCGGAACTGGGACCAACGTCACCGTCCTAGGTCAGCCCAAGGCCAACCCCA
CTGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTCCAAGCCAACAAGGCCACACTAGTG
TGTCTGATCAGTGACTTCTACCCGGGAGCTGTGACAGTGGCCTGGAAGGCAGATGGCAG
CCCCGTCAAGGCGGGAGTGGAGACCACCAAACCCTCCAAACAGAGCAACAACAAGTACG
CGGCCAGCAGCTACCTGAGCCTGACGCCCGAGCAGTGGAAGTCCCACAGAAGCTACAGC
TGCCAGGTCACGCATGAAGGGAGCACCGTGGAGA
>FOR119_AA
XSVLTQPPSVSAAPGQKVTISCSGSSSNIGSTYVSWYRQLPGTAPRLLIYDXYKRPSGI
PDRFSGSKSGTSATLGITGLQTGDAADYYCGTWDSSLSVYVFGTGTNVTVLGQPKANPT
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTVE
>FOR175_cDNA
TCAGTCTGTGTTGACGCAGCCGCCCTCAGTGTCTGCGGCCCCTGGACAGAAGGTCACCA
TCTCCTGCTCYGGAAGYGGCTCCRATATTCAGAGTCATTATGTATCCTGGTRCCARCAA
CTCCCAGGAASCGCCCCCAAACTTCTCATTTATGACAACAATAAGCGACCCTCAGGCAT
TCCTGGCCGATTCTCTGGCTCCAAGTCTGGCACGTCAGCCACCCTGGASATCTCCGGAC
TCCAGACTGGSGACGAGGCCSAYTATTACTGCGGAACATGGGATGACACCCTGAGTACT
GGGGTATTCGGCGGWGGGACCAAGGTGRCCGTCCTAGGTCAGCCCAAGGCTGCCCCCTC
GGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAAYAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGC
CCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGC
GGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACARAAGCTACAGCT
GCCAGGTCACGCATGAAGGGAGCACCGTGGAGA

>FOR175_AA
QSVLTQPPSVSAAPGQKVTISCSGSGSXIQSHYVSWXQQLPGXAPKLLIYDNNKRPSGI
PGRFSGSKSGTSATLXISGLQTGDEAXYYCGTWDDTLSTGVFGGGTKVXVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHXSYSCQVTHEGSTVE

***IGLV2-8*:**
>FOR210_cDNA
CTCAGTCTGCCCTGACTCAGCCTCCCTCCGCGTCCGGGTCTCCTGGACAGTCAGTCACC
ATCTCCTGCACTGGAACCAGCAGTGACGTTGGTGGTTACAACTATGTCTCCTGGTATCA
AGAGCACCCAGGCAAAGCCCCCAAACTCATGATTTATGAGGTCAGTCAGCGGCCCTCAG
GGGTCCCTGATCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCTGACCGTCTCT
GGGCTCCAGACTGAGGATGAGGCTGAGTATTACTGCAGTTCATATGCAGGCAGTAACAA
TTTGATATTTGGCGGAGGGACTAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCT
CGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTG
TGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAG
CCCCGTCAAGKSGGGAGTGGAGMCCACCACACCCTCCAAACAAAGCAACAACAAGTACG
CGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGC
TGCCAGGTCACGCATGAAGGGAGCACCGTGG
>FOR210_AA
QSALTQPPSASGSPGQSVTISCTGTSSDVGGYNYVSWYQEHPGKAPKLMIYEVSQRPSG
VPDRFSGSKSGNTASLTVSGLQTEDEAEYYCSSYAGSNNLIFGGGTKLTVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKXGVEXTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTV
>FOR233_cDNA
CTCAGTCTGCCCTGACTCAGCCTCCCTCCGCGTCCGGGTCTCCTGGACAGTCAGTCACC
ATCTCCTGCACTGGATCCAGCAATGACGTTGGTGGTTATAACTATGTCTCCTGGTACCA
ACAACACCCAGGCRAAGCCCCCAAACTCATGATTTATGAGGTCAACAAGCGGCCCTCAG
GGGTCCCTGATCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCTGACSRTCTCT
GGGCTCCAGGCTAAAGATGAGGCTGATTATTATTGCAGCTCATATGCARGCACCAACAC
CTTGRTATTSRGCGGAGGGACCWACYTGGCCGTCCTAGGTCAGCCCAAGGCTGCCCCCT
CGGTCWCTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAKCCAACAAGGCCWSTYTGGTG
TGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACASTGGCCTGGAAGGCAGATAGCAG
CCCCGTCAAGGCGGGAGTGKRGWCCWCCACACCCTCCAAACAAAGCAACAACAAGTACG
CGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGC
TGCCAGGTCACGCATGAAGGGAGCACCGTGGAGA
>FOR233_AA
QSALTQPPSASGSPGQSVTISCTGSSNDVGGYNYVSWYQQHPGXAPKLMIYEVNKRPSG
VPDRFSGSKSGNTASLTXSGLQAKDEADYYCSSYAXTNTLXXXGGTXLAVLGQPKAAPS
VXLFPPSSEELQXNKAXLVCLISDFYPGAVTXAWKADSSPVKAGVXXXTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTVE

***IGLV2-14*:**
>FOR108_cDNA
TGGTCCTGGGCTCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACA
GTCGATCACCATCTCCTGCACTGGAACCAGCAGTGACGTTGGTGCTTATAACTATGTYT
CCTGGTTCCAAAACCACCCAGGCAGAGCCCCCAAACTCCTGATTTATACTGTTAGTAAG
CGGCCCTCAGGGATTTMTGATYGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCMT
GAGCATCTSTGGGATCCAGGCTGACGACGAGGCTGTTTATTACTGCAGCTCATATACAA
CCACTAGTAATTTGATATTTGGCGGAGGGACCAACCTGACCGTCCTAGGTCAGCCCAAG
GCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGC

CACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGG
CAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAAC
AACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAG
AAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR108_AA
WSWAQSALTQPASVSGSPGQSITISCTGTSSDVGAYNYVSWFQNHPGRAPKLLIYTVSK
RPSGIXDXFSGSKSGNTASLSIXGIQADDEAVYYCSSYTTTSNLIFGGGTNLTVLGQPK
AAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSN
NKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTV
>FOR166_cDNA
TGGTCCTGGGCTCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACA
GTCGATCACTATCTCCTGCACTGGAACCAGCAGTGATGTTGGTACTTATAACTATGTCT
CCTGGTACCAACAACACCCAGGCAAAGCCCCCAAACTCATTCTTTATGATGTCAGTAAG
CGGCCCTCAGGGGTTTCTAATCGCTTCTCTGGCTCCAAGTCTGGCAMCRYGGMCWMSMY
GRYRATCTCTGGGCTCCAGGCTGAGGACGAGGCTCAATATTACTCCTGCTCATCTACAG
GTAGTGGGTCTTGGGTGTTCGGCGGAGGGACCAWGCTGACCGTCCTAGGTCAGCCCAAG
GCTGCCCCCTCGGTCACTCTGTTCCCACCCTCCTCTGAGGAGCTTCAAGCCAACAAGGC
CACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGG
CAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAAC
AACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAA
AAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR166_AA
WSWAQSALTQPASVSGSPGQSITISCTGTSSDVGAYNYVSWFQNHPGRAPKLLIYTVSK
RPSGIXDXFSGSKSGNTASLSIXGIQADDEAVYYCSSYTTTSNLIFGGGTNLTVLGQPK
AAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSN
NKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTV
>FOR167_cDNA
TGGTCCTGGGCTCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACA
GTCGATCACCATCTCCTGCACTGGAACCAGCAGTGACGTGGGTGGTTATAAGTATGTCT
CCTGGTACCAACAACACCCAGGCAAAGCCCCCAAAGTCATGATTTATGAGGTCAGTAAT
CGGCCCTCAGGGGTTTCTAATCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCT
GACCATCTCTGGGCTCCAGGCTGAGGACGAGGCTGATTATTACTGCAGCTCATATACAA
GCAGCAGCACTTATGTCTTCGGAACTGGGACCGAGGTSACCGTCCTAGGTCAGCCCAAG
GCCRMCCCCWCKGTCACTCTGTTCCCGCCCTCCTCTGASGASCTYCAAGYCAACAAGGC
CACACTRGTGTGTCTSATMAGTGACTTCTACCCGGGAGCYGTGACAGTGGCCTGSAASG
CAGATRGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCAMAMCCTCSAAACARAGCAAC
AACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCCGAGCAGTGGAAGTCCCACAG
AAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR167_AA
WSWAQSALTQPASVSGSPGQSITISCTGTSSDVGGYKYVSWYQQHPGKAPKVMIYEVSN
RPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTSSSTYVFGTGTEVTVLGQPK
AXPXVTLFPPSSXXLQXNKATLVCLISDFYPGAVTVAXXADXSPVKAGVETTXXSKQSN
NKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTV
>FOR174_cDNA
CTCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACAGTCGATCACC
ATCTCCTGCACTGGAACCAGCAGTGACGTTGGTGCTTATAACTATGTCTCCTGGTACCA
ACAGAACCCAGGCRCAGCCCCTATACTTATGATTCATGATGTCAGTAATCGGCCCTCAG
GAGTTTCTAATCGCTTCTCTGGCTCCAAGTCTGGCAMCRCGGMCWMCMYGAYSATCACT
GGGCTCCAGGCTGAGGACGAGGCTGATTATTTCTGCTCCTCATATACAAGCAGCAACAC
TCTGGTATTCGGCGGAGGGACCAWGGTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCT
CGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTG

TGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAG
CCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACG
CGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGC
TGCCAGGTCACGCATGAAGGGAGCACCGTG
>FOR174_AA
QSALTQPASVSGSPGQSITISCTGTSSDVGAYNYVSWYQQNPGXAPILMIHDVSNRPSG
VSNRFSGSKSGXXXXXXITGLQAEDEADYFCSSYTSSNTLVFGGGTXVTVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTV
>FOR198_cDNA
TCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACAGTCGATCACCA
TCACCTGCACTGGAACCAGCAGTGACGTTGGTGTTTATAACTATGTCTCYTGGTACCAR
CAGAASCCAGGCAAAGCCCCCAAATTGACGATTTATGAGGTCAGTTATCGGCCCTCAGG
GGTTTCTGATCGCTTTTCTGGCTCCAAGTCTGGCAACACGGCCTCCCTGACCATCTCTG
GGCTCCACACTGAGGACGAGGGTGATTATTACTGCAGCTCATATACACTCAACAGCACT
CTCGTCTTCGGAACTGGGACCACGGTCACCGTCCTACGTCAGCCCAAGGCCAACCCCAC
TGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTCCAAGCCAACAAGGCCACACTAGTGT
GTCTGATCAGTGACTTCTACCCGGGAGCTGTGACAGTGGCCTGGAAGGCAGATGGCAGC
CCCGTCAAGGCGGGAGTGGAGACCACCAAACCCTCCAAACAGAGCAACAACAAGTACGC
GGCCAGCAGCTACCTGAGCCTGACGCCCGAGCAGTGGAAGTCCCACAGAAGCTACAG
>FOR198_AA
QSALTQPASVSGSPGQSITITCTGTSSDVGVYNYVSWYQQXPGKAPKLTIYEVSYRPSG
VSDRFSGSKSGNTASLTISGLHTEDEGDYYCSSYTLNSTLVFGTGTTVTVLRQPKANPT
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYA
ASSYLSLTPEQWKSHRSY
>FOR200_cDNA
GGCTCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACAGTCGATCA
CCATCTCCTGCACTGGAACCAGTAGCGACGTTGGTACTTATAACTATGTCTCYTGGTAT
CAACAGTCCCCCGGCAAAGCCCCCAAACTCCTGATTTATGAGGTCAGTAATCGGCCCTC
AGGAGTTTCTAATCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCTGACCATCT
CTGGCCTCCAGGCTGAGGACGAGACTTTTTATTACTGCAGTTCATATGCAAGTGGCAGT
ACCCCYTATGTCTTCGGAGGWGGGACCACGGTCACCGTCCTAAGTCAGCCCAAGGCCAA
CCCCWCKGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTCCAAGCCAACAAGGCCACAC
TAGTGTGTCTGATCAGTGACTTCTACCCGGGAGCTGTGACAGTGGCCTGGAAGGCAGAT
GGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCAAACCCTCCAAACAGAGCAACAACAA
GTACGCGGCCAGCAGCTACCTGAGCCTGACGCCCGAGCAGTGGAAGTCCCACAGAAGCT
ACAGCTGCCAGGTCACGCATGAAGGGAGCACCRTRGAKAAGACAGTGA
>FOR200_AA
AQSALTQPASVSGSPGQSITISCTGTSSDVGTYNYVSWYQQSPGKAPKLLIYEVSNRPS
GVSNRFSGSKSGNTASLTISGLQAEDETFYYCSSYASGSTPYVFGGGTTVTVLSQPKAN
PXVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNK
YAASSYLSLTPEQWKSHRSYSCQVTHEGSTXXKTV
>FOR219_cDNA
TCAGTCTGCCCTGACTCAGCCCGCCTCCGCGTCTGGGTCTCCGGGACACTCGATCACCA
TCTCCTGCACTGGAACCAGCAGTGACGTTGGTGGTTATAACTATGTCTCCTGGTACCAG
CAACACCCAGGCAAAGCCCCCAAAGTCATCATSTATAATGTCAGTTATCGGCCCTCAGG
GGTCTCTTTTCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCTGACCATCTCTG
GACTCCGGGCTGAGGACGAGGCTGACTACTACTGCAGCTCTTACACAGGCAGCAATTCT
CTCGTTTTCGGCGGAGGGACCACGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTC
GGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGT
GTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGC

CCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGC
GGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCT
GCCAGGTCACGCATGAAGGGAGCACCGTGGAGA
>FOR219_AA
QSALTQPASASGSPGHSITISCTGTSSDVGGYNYVSWYQQHPGKAPKVIXYNVSYRPSG
VSFRFSGSKSGNTASLTISGLRAEDEADYYCSSYTGSNSLVFGGGTTLTVLGQPKAAPS
VTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
ASSYLSLTPEQWKSHRSYSCQVTHEGSTVE
>FOR237_cDNA
TGGTCCTGGGCTCAGTCTGCCCTGACTCAGCCTGCCTCCGTGTCTGGGTCTCCTGGACA
GTCGATCACCATCTCCTGCACTGGAACCACCTTTGACGTTGGTGGTTTTAACTATGTCT
CCTGGTACCAACACCACCCAGGCAAAGCCCCCAAAGTCATGATTTATGAGGTCAATAAT
CGGCCCTCAGGGGTTTCTAATCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCT
GACCATCTCTGGGCTCCAGGCTGAGGACGAAGCCGATTATTACTGCAGCTCGTATACAA
GCAGCAGCACTCGCGTCTTCGGAACTGGGACCAAGGTCACCGTCCTAGGTCAGCCCAAG
GCCAACCCCACTGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTCCAAGCCAACAAGGC
CACACTAGTGTGTCTGATCAGTGACTTCTACCCGGGAGCTGTGACAGTGGCCTGGAAGG
CAGATGGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCAAACCCTCCAAACAGAGCAAC
AACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCCGAGCAGTGGAAGTCCCACAG
AAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR237_AA
WSWAQSALTQPASVSGSPGQSITISCTGTTFDVGGFNYVSWYQHHPGKAPKVMIYEVNN
RPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTSSSTRVFGTGTKVTVLGQPK
ANPTVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSN
NKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTV

***IGLV3-1*:**
>FOR139_cDNA
TCCTATGAGCCTGACACAGCCACCCTCAGTGTCCGTGTCCCCAGGACAGACAGCCAGCA
TCACCTGCTCTGGAGATGGATTGGGGGATAAATTTGCTTGCTGGTATCAGCAGAAGCCA
GGCCAGTCCCCTGTGTTGGTCATCTATCAAGATACTAAGCGGCCCTCAGGGATCCCTGA
GCGATTCTCTGGCTCCAACTCTGGGAACACAGCCACTCTGACCATCAGCGGGACCCAGC
CTATGGATGAGGCTGACTATTACTGTCAGACGTGGGACAGCAACTCTTATGTCTTCGGA
ACTGGGACCAAGGTCACCGTCCTAGGTCAGCCCAAGGCCAACCCCACTGTCACTCTGTT
CCCGCCCTCCTCTGAGGAGCTCCAAGCCAACAAGGCCACACTAGTGTGTCTGATCAGTG
ACTTCTACCCGGGAGCTGTGACAGTGGCCTGGAAGGCAGATGGCAGCCCCGTCAAGGCG
GGAGTGGAGACCACCAAACCCTCCAAACAGAGCAACAACAAGTACGCGGCCAGCAGCTA
CCTGAGCCTGACGCCCGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTCACGC
ATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR139_AA
PMSLTQPPSVSVSPGQTASITCSGDGLGDKFACWYQQKPGQSPVLVIYQDTKRPSGIPE
RFSGSNSGNTATLTISGTQPMDEADYYCQTWDSNSYVFGTGTKVTVLGQPKANPTVTLF
PPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSY
LSLTPEQWKSHRSYSCQVTHEGSTVEKTV
>FOR165_cDNA
TCCGTGTCCCCAGGACAGACAGCCAGCATCACCTGCTCTGGAGATAAATTGGGGGATAA
ATATGTTTGCTGGTATCAGCAGAAGCCAGGCCAGTCCCCTGTGGTGGTCATCTATCAAG
ATACTCAGCGGCCCCCAGGGATCCCTGAGCGATTCTCTGGCTCCAACTCTGCGAACACA
GCCACTCTGACCATCAGCGGGACCCAGGCTATGGATGAGGCTGACTATTACTGTCAGGC
GTGGGACAGCACCGCTGCAATCTTCGGAACTGGGACCAAGGTCACCGTCCTAGGTCAGC
CCAAGGCCAACCCCACTGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTCCAAGCCAAC

AAGGCCACACTAGTGTGTCTGATCAGTGACTTCTACCCGGGAGCTGTGACAGTGGCCTG
GAAGGCAGATGGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCAAACCCTCCAAACAGA
GCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCCGAGCAGTGGAAGTCC
CACAGAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGG
>FOR165_AA
SVSPGQTASITCSGDKLGDKYVCWYQQKPGQSPVVVIYQDTQRPPGIPERFSGSNSANT
ATLTISGTQAMDEADYYCQAWDSTAAIFGTGTKVTVLGQPKANPTVTLFPPSSEELQAN
KATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKS
HRSYSCQVTHEGSTV
>FOR172_cDNA
CTGACACAGCCACCCTCAGTGTCCGTGTCCCCAGGACAGACAGCCAGCATCACCTGCTC
TGGAGATAAATTTGCTATTAAATATGCTTCCTGGTATCAGCAGAAGCCAGGCCAGTCCC
CTGTGCTGGTCATCTATCAAGATGCCAAGCGGCCCTCAGGGATCCCTGAGCGATTCTCT
GGCTCCAACTCTGGGAACACAGCCGCTCTGACCATCAGCGGGACCCAGGCAATGGATGA
GGCTGACTATTACTGTCAGGCGTGGGACAGAAGCGCTGCAGCGGTCTTCGGAACTGGGA
CCAAGGTCACCGTCCTAGGTCAGCCCAAGGCCAACCCCACTGTCACTCTGTTCCCGCCC
TCCTCTGAGGAGCTCCAAGCCAACAAGGCCACACTAGTGTGTCTGATCAGTGACTTCTA
CCCGGGAGCTGTGACAGTGGCCTGGAAGGCAGATGGCAGCCCCGTCAAGGCGGGAGTGG
AGACCACCAAACCCTCCAAACAGAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGC
CTGACGCCCGAGCAG
>FOR172_AA
LTQPPSVSVSPGQTASITCSGDKFAIKYASWYQQKPGQSPVLVIYQDAKRPSGIPERFS
GSNSGNTAALTISGTQAMDEADYYCQAWDRSAAAVFGTGTKVTVLGQPKANPTVTLFPP
SSEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLS
LTPEQ
>FOR224_cDNA
CTGACACAGCCACCCTCAGTGTCCGTGTCCCCAGGACAGACAGCCAGCATCATCTGCTC
TGGAGATAAATTGGGGGATAAATATGCTTCCTGGTAyCAGCAGAAGCCAGGCCAGTCCC
CTGTGGTGGTCATCTATCAAGATAACAAGCGGCCCTCAGGGATCCCTGAGCGATTCTCT
GGCTCCAAcTctGGGAACACAGCCACTCTGACCATCAGCGGGACCCAGGCTATGGATGA
GGCTGACTATTACTGTCAGACATGGGACAGCAGCACTAATGTCTTCGGAACTGGGACCA
AGGTCACCGTCCTAGGTCAGCCCAAGGCCAACCCCACTGTCACTCTGTTCCCGCCCTCC
TCTGAGGAGCTCCAAGCCAACAAGGCCACACTAGTGTGTCTGATCAGTGACTTCTACCC
GGGAGCTGTGACAGTGGCCTGGAAGGCAGATGGCAGCCCCGTCAAGGCGGGAGTGGAGA
CCACCAAACCCTCCAAACAGAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTG
ACGCCCGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTCACGCATGAAGGGAG
CACCGTGGAGAAGACAGTGA
>FOR224_AA
LTQPPSVSVSPGQTASIICSGDKLGDKYASWXQQKPGQSPVVVIYQDNKRPSGIPERFS
GSNSGNTATLTISGTQAMDEADYYCQTWDSSTNVFGTGTKVTVLGQPKANPTVTLFPPS
SEELQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSL
TPEQWKSHRSYSCQVTHEGSTVEKTV

***IGLV3-9*:**
>FOR227_cDNA
TGACACAGCCACTCTCAGTGTCAGTGGCCCTGGGACAGACGGCCAGGATTACCTGTGGG
GGAAATAACATTGGAGATAAAGGTGTACACTGGTACCAGCAGAAGCCAGGCCAGGCCCC
TGTATTGGTCATCTATAGGGATAGAAACCGGCCCACTGGGATCCCTGAGCAATTCTCTG
GCTCCAACTCGGGGAACACGGCCACCCTGACCATCAGCAGAGCCCAAGCCGGGGATGAG
GCTGACTATTACTGTCAGGTGTGGGACAGCACCACTGGCGATGTGGTTTTCGGCGGAGG
GACCAAGCTGAGCGTCCTGGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCGC

CCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTC
TACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGT
GGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTATCTGA
GCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTCACGCATGAA
GGGAGCACCGTGGAGAA
>FOR227_AA
TQPLSVSVALGQTARITCGGNNIGDKGVHWYQQKPGQAPVLVIYRDRNRPTGIPEQFSG
SNSGNTATLTISRAQAGDEADYYCQVWDSTTGDVVFGGGTKLSVLGQPKAAPSVTLFPP
SSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLS
LTPEQWKSHRSYSCQVTHEGSTVE

***IGLV3-19*:**
>FOR234_cDNA
ACCCTGCTGTGTCTGTGGCCTTGGGACAGACAGTCAGGATCACATGCCAAGGAGACAGC
CTCAGAAACTATTATGCAAGTTGGTACCAACAGAAGCCAGGACAGGCCCCTATTTTTGT
CATGTACGGTAATAATGAGCGGCCCGCAGGGGTCCCAGACCGATTCTCTGGCTCCAGAT
CAGGAATCACAGCTTCCTTGACCATCACTGGGGCCCAGGCGGAAGATGAGGCAGACTAT
TACTGTAACTCTCGTGATGGCAGTGATTACCGGTGGGTGTTCGGCGGAGGGACCAAGCT
GACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCACCCTCCTCTG
AGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGA
GCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCAC
CACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGC
CTGAGCAGTGGAAGTCCCACAAAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACC
GTGGAGAA
>FOR234_AA
AVSVALGQTVRITCQGDSLRNYYASWYQQKPGQAPIFVMYGNNERPAGVPDRFSGSRSG
ITASLTITGAQAEDEADYYCNSRDGSDYRWVFGGGTKLTVLGQPKAAPSVTLFPPSSEE
LQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPE
QWKSHKSYSCQVTHEGSTVE

***IGLV3-21*:**
>FOR121_cDNA
TTGGTASCAGCAGMAGCCAGGCCAGVCCCCTGTGSTGGTCGTCTATGATGATARCGACC
GGCCCTCAGGGATCCCTGAGCGATTCTCTGGCTCCAVCTCYGGGAMCACGGCCACCNTS
ACCATCAGCAGGGTCGAAGCCGAGGATGAGGMYGACTATTAYTRTCAGKNGTGGGATAG
TAGTAGTRATCATGTGGTATTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCA
AGGCTGCCCCCTCGGTCACTCTGTTCCCRCCCTCCTCTGAGGAGCTTCAAGCCAACAAG
GCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAA
GGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCA
ACAACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGC
>FOR121_AA
WYQQXPGQXPVXVVYDDXDRPSGIPERFSGSXSGXTATXTISRVEAEDEXDYYXQXWDS
SSXHVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK
ADSSPVKAGVETTTPSKQSNNKYAASSYLSLT
>FOR129_cDNA
TTGGTASCAGCAGMAGCCAGSMCAGGCCCCTGTSSTGGTCRTCTWTGATGATAACRACC
GGCCCTCAGGGATCCCTGASCGATTCTCTKGCTCCAVCTCHGGGAMCACRGCCACCNTG
ACCATCAGCRGGGTCRAGGCSGAGGATGAGGMBGABTATTAYTGTCABTBGTGGGAYAG
CAGTAGTRATCATKDGGTVTTCGGCGGAGGGACCAAGSTGACCGTCCTAGGTCAGCCCA
AGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAG
GCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAA

GGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCA
ACAACAAGTACGCGGCCAGCAGCTAYCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCAC
AGAAGCTACAG
>FOR129_AA
WYQQXPXQAPVXVXXDDNXRPSGIPXRFSXSXSGXTATXTISXVXAEDEXXYYCXXWDS
SSXHXVFGGGTKXTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK
ADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSY
>FOR141_cDNA
GGGTCCTGGGGSTCAGTSTGTGGTGACGCAGCCGCCMTCAGTGTCTGGGGCCCCAGGGC
AGAGGGTCACCATGTCTTGCACTGGGCGCAGCTCCAACATCGGGGCAGGTTATRATGTT
CTTGGGAACAGCAGACGCCAGGACARGCCCCCRTRCTSGTCATTTATGATGAYATCGAT
CGGCCCTCAGGGRTCCCTGASCGATTCTCTGGCTCCACSTCTGGGACCWCRGCCTCCCT
GACCATCAGCRGRGTCSAGGCCGAGGATGAGGCTGATTATTACTGTCAGTCSTGGGACA
GCAGYAGTGATCWTGKGGTGTTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCC
AAGGCTGCCCCCTCGGTCACTCTGTTCCCRCCCTCCTCTGAGGAGCTTCAAGCCAACAA
GGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGA
AGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGC
AACAACAAGTACGCGGCCAGCAGCTAYCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCA
CAGAAGCTACAGCTG
>FOR141_AA
GSWGSXCGDAAAXSVWGPRAEGHHVLHWAQLQHRGRLXCSWEQQTPGXAPXLVIYDDID
RPSGXPXRFSGSTSGTXASLTISXVXAEDEADYYCQSWDSSSDXXVFGGGTKLTVLGQP
KAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQS
NNKYAASSYLSLTPEQWKSHRSYS
>FOR146_cDNA
TTGGTACCAGCAGCAGCCRSGCCCGGCCCCYGTSCTSGTCRTCTWTDATKATRAYGACC
GKCCCTCAGSGWTCCSTGWNCGWTTCTSTKKCTCCWVCTCYGGGWCCRCGGCCACCNTS
ACCATCASCAKSGTBGAGGCCGAGGATGAGGMYGABTATTAYTRTYASTBSTGGGAYAG
YAGTAGTRRTCATGTGGTRTTCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCA
AGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAG
GCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAA
GGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCA
ACAACAAGTACGCGGCCAGCAGCTAYCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCAC
AGAAGCTACAG
>FOR146_AA
WYQQQPXPAPVLVXXXXXDRPSXXXXRFXXSXSGXXATXTIXXVEAEDEXXYYXXXWDS
SSXHVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK
ADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSY
>FOR170_cDNA
TTGGTASCAGCAGCAGCCAGSVCMSGCCCCYVYVSTSGTCATCTATDATRAYAAYRABC
GGCCCTCAGGGRTCCYTGANCGATTCTCYKGCTCCWACTCYGGSACCACVRCCDCCNYS
AYCATCABCRGGGTSSAGGMSGAGGATGAGGMBGAYTATTAYTRTYANTCNTVGGAYRG
CAGYAGTRRYCATKGGGTGTTCGGCGGAGGGACCAAGSTGACCGTCCTAGGTCAGCCCA
AGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAG
GCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAA
GGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCA
ACAACAAGTACGCGGCCAGCAGCTAYCTGAGCCTGACGCC
>FOR170_AA
WXQQQPXXAXXXVIYXXXXRPSGXXXRFXXSXXXTXXXXXIXXXXXEDEXXYXXXXXXX
XSXHXVFGGGTKXTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK
ADSSPVKAGVETTTPSKQSNNKYAASSXLSLT

>FOR178_cDNA
TTGGTACCAGCAGAAGCCAGGCCAGGCCCCTGGACTGGTCGTCTATGATGATGACGACC
GGCCCTCAGGGATCCCTGAGCGATTCTCTGGCTCCAACTCTATGAACACGGssACTCTG
ACCATCAGCAGGGTCGCAGCCGGGGAkGAGGCCGACTATTATTGTCAGGTGTGGGATAG
TACCAGTGATCATCCCGTGGTATTCGGCGGAGGGACCACGCTGACCGTCTTAGGTCAGC
CCAAGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAAC
AAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTG
GAAGGCAGATAGyAGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAA
GsAACAACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCC
CACAGAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCGTGGA
>FOR178_AA
WYQQKPGQAPGLVVYDDDDRPSGIPERFSGSNSMNTXTLTISRVAAGXEADYYCQVWDS
TSDHPVVFGGGTTLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAW
KADSSPVKAGVETTTPSKQXNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTV

***IGLV3-25*:**
>FOR168_cDNA
TCCTATGAGCTGACACAGCCACCCTCGGTGTCAGTGTCCCCAGGACAGACGGCCACGAT
CACCTGCTCTGGAGATGCATTGGCAAAGCAATTTGGTTATTGGTACCAGCAGAAGCCAG
GCCAGGCCCCTGTCTTGATGATATATAAGGACACTGAGAGGCCCTCAGGGGTCCCTGAG
CGATTCTCTGGCTCCATTTCCGGGACAACAGTCACGTTGACCATCACTGGAGTCCAGGC
AGAGGACGAGGCTGACTATTACTGTCAATCACCAGACAGTAGTAGTCTTTTTGTTCTAT
TCGGCGGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACT
CTGTTCCCGCCCTCCTYTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCAT
AAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCA
AGGCGGGAGWGGAGACCMCCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGC
AGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGT
CACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR168_AA
SYELTQPPSVSVSPGQTATITCSGDALAKQFGYWYQQKPGQAPVLMIYKDTERPSGVPE
RFSGSISGTTVTLTITGVQAEDEADYYCQSPDSSSLFVLFGGGTKLTVLGQPKAAPSVT
LFPPSXEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGXETXTPSKQSNNKYAAS
SYLSLTPEQWKSHRSYSCQVTHEGSTVEKTV

***IGLV6-57*:**
>FOR125_cDNA
TCAGCCCCACTCTGTGTCGGAGTCTCCGGGGCAGACGGTAATCATCTCCTGCACCCGCA
GCAGTGGCAGCATTGCCAGCAATTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTGCC
CCCAGTATCGTAATTTATGAGGATAAACAAAGACGCGCTGGGGTCCCTGAACGGTTCTC
TGGCTCCATCGACAGTTCCTCCAACTCTGCCTCCCTCACCATCTCTGGACTGAACACTG
AGGACGAGGCTGACTATTATTGTCAGTCTTATGATAGCAACAATCATTGGGTGTTCGGC
GGAGGGACCAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTT
CCCACCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTG
ACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCG
GGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTA
CCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAAAAGCTACAGCTGCCAGGTCACGC
ATGAAGGGAGCACCGTGGAGAAAGACAGTGA
>FOR125_AA
QPHSVSESPGQTVIISCTRSSGSIASNYVQWYQQRPGSAPSIVIYEDKQRRAGVPERFS
GSIDSSSNSASLTISGLNTEDEADYYCQSYDSNNHWVFGGGTKLTVLGQPKAAPSVTLF
PPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSY

LSLTPEQWKSHKSYSCQVTHEGSTVEKDS
>FOR143_cDNA
TCTCCTGCACCGGCAGCAGTGACACCATTGACAGCAACTATGTGCACTGGTACCAGCAG
CGCCCGGGCAGTGCCCCCACCACTGTGATCTATGAGGATGACAAAAGACCCTCTGGGGT
CCCTGATCGGTTCTCTGGCTCCATCGACGCCTCCTCCAATTCAGCCTCCCTCACCATCT
CTGGACTGAAGACTGAGGACGAGGCTGACTACTTCTGTCAGTCTTTTGATAGTAGCAAT
CGAGTGATATTCGGCGGAGGGACCAAACTGACCGTCCTCGGTCAGCCCAAGGCTGCCCC
CTCGGTCACTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGG
TGTGTCTCATAAGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGC
AGCCCCGTCAAGGCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTA
CGCGGCCAGCAGCTATCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACA
GCTGCCAGGTCACGCATGAAGGGAGCACCGTGGAGAAGACAGTGA
>FOR143_AA
SCTGSSDTIDSNYVHWYQQRPGSAPTTVIYEDDKRPSGVPDRFSGSIDASSNSASLTIS
GLKTEDEADYFCQSFDSSNRVIFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYS
CQVTHEGSTVEKTV
>FOR151_cDNA
GTCGGAGTCTCCGGGGAAGACGGTGACCATCTCCTGCACCGCAGCAGYGGCAGCATYGC
CAGCAACTATGTGCAGTGGTAMCAGCAGCGSCCGGGCAGYGCCCCCACCACTGTGATYT
TTAARGATAAYCAAAGACCCTCTGGGGTYCYTGATCGBTTCTCTKGCTCCATCGACAGC
TCCTCCAACTCTKCCTCCCTCAMCATCTYTGYACTGAAGAMKGAGGACGAGGCTGACTA
CTACTGTCAGTMTTATGATATCMVCAATCATGTGGTHTTCGGCGGAGGGACCAAGMTGA
CCGTCCTASGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCACCCTCCTCTGAG
GAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAGC
CGTGACAGTGGCCTGGAAGGMAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACCA
CACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCCT
GAGCAGTGGAAGTCCCACAAAAGCTACWGCTGCSACGTCMCGCATGAAGGGAGCACCG
>FOR151_AA
VGVSGEDGDHLLHRSSGSIASNYVQWXQQRPGSAPTTVIFKDNQRPSGLPDRFSGSIDS
SSNSXSLXIXXLKXEDEADYYCQXYDIXNHVVFGGGTKXTVLXQPKAAPSVTLFPPSSE
ELQANKATLVCLISDFYPGAVTVAWKXDSSPVKAGVETTTPSKQSNNKYAASSYLSLTP
EQWKSHKSYXCXVXHEGST
>FOR158_cDNA
TCAGCCCCACTCTGTGTCGGAGTCTCCGGGGAAGACGGTAACCATCTCCTGCATCGGCA
CCGGTGGCAGCATTGGCAGCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTGCC
CCCACCACTGTGATCTATGAGAATGAGAAGAGACCCTCTGGGGTCCCTGATCGTTTCTC
TGCCTCCATCGACAGCTCCTCCAACTCTGCCTCCCTCACCATCTCTGGACTGAAGACTG
AGGACGAGGCTGACTACTACTGTCAGTCTTTTGCTGGCAGCACCAATTATAGGGTGTTC
GGCGGAGGGACAAAACTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCT
GTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAA
GTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAG
GCGGGAGTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAG
CTACCTGAGCCTGACGCCTGAGCAGTGGAAGTCCCACAAAAGCTACAGCTGCCAGGTCA
CGCATGAAGGGAGCACCGTGGA
>FOR158_AA
QPHSVSESPGKTVTISCIGTGGSIGSNYVQWYQQRPGSAPTTVIYENEKRPSGVPDRFS
ASIDSSSNSASLTISGLKTEDEADYYCQSFAGSTNYRVFGGGTKLTVLGQPKAAPSVTL
FPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASS
YLSLTPEQWKSHKSYSCQVTHEGSTV

>FOR183_cDNA
TGTGTCGGAGTCTCCGGGGAAGACAGTGACCATCTCCTGCACCCGCAGCAGTGGCAGCA
TAGCCAGCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTGCCCCCACCACTGTG
ATCTTTGACGATAACCAAAGACCCTCTGGGGTCCCTGATCGGTTCTCTGGCTCCATCGA
CAGCTCCTCCAACTCTGCCTCCCTCGCCATCTCTGGACTGAAGACTGAGGACGAGGCTG
ACTACTACTGTCAGTCTTATGATATCAGCAATAAGGTGATATTCGGCGGAGGGACCAAG
CTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTC
TGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGG
GAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACC
ACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGTTATCTGAGCCTGAC
GCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCA
CCGTG
>FOR183_AA
VSESPGKTVTISCTRSSGSIASNYVQWYQQRPGSAPTTVIFDDNQRPSGVPDRFSGSID
SSSNSASLAISGLKTEDEADYYCQSYDISNKVIFGGGTKLTVLGQPKAAPSVTLFPPSS
EELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLT
PEQWKSHRSYSCQVTHEGSTV
>FOR199_cDNA
TGTGTCGGAGTCTCCGGGGAAGACGGTGACCATCTCCTGCACCCGCAGCAGTGGCAGCA
TTGCCAGCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTTCCCCCACCACTGTG
ATCTATGAGGATGACCAAAGACCCTCTGGGGTCCCTGATCGGTTCTCTGGCTCCATCGA
CGGCTCCTCCAACTCTGCCTCCCTCACCATCTCTGGACTGAAGACTGAGGACGAGGCTG
ACTACTATTGTCAGTCTTATGATAGCAGCAATCAGGTGTTCGGCGGAGGGACCAAGCTG
ACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCACCCTCCTCTGA
GGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGGGAG
CCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACCACC
ACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGACGCC
TGAGCAGTGGAAGTCCCACAAAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCACCG
TGGA
>FOR199_AA
VSESPGKTVTISCTRSSGSIASNYVQWYQQRPGSSPTTVIYEDDQRPSGVPDRFSGSID
GSSNSASLTISGLKTEDEADYYCQSYDSSNQVFGGGTKLTVLGQPKAAPSVTLFPPSSE
ELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTP
EQWKSHKSYSCQVTHEGSTV
>FOR207_cDNA
GGCAGCATTGCCAGCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTGCCCCCAC
CACTGTGATCTATGAGGATAACCAAAGACCCTCTGGGGTCCCTGATCGGTTCTCTGGCT
CCATCGACAGCTCCTCCAACTCTGCCTCCCTCACCATCTCTGGACTGAAGACTGAGGAC
GAGGCTGACTACTACTGTCAGTCTTATGATAVYRGMAVYCCTBKSGTGTTCGGCGGAGG
GACCAAGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCGC
CCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTC
TACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGT
GGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGA
GCCTGACGCCTGAGCAGTGGAAGTCCCAC
>FOR207_AA
GSIASNYVQWYQQRPGSAPTTVIYEDNQRPSGVPDRFSGSIDSSSNSASLTISGLKTED
EADYYCQSYDXXXXXVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDF
YPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSH
>FOR209_cDNA
TGTGTCGGAGTCTCCGGGGAAGACGGTGACCATCTCCTGCACCGGCAGCAGTGGCAGCA
TTGCCAGCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTGCCCCCACCACTGTG

ATCTATGAGGATGACCAAAGACCCTCTGGGGTCCCTGATCGGTTCTCTGGCTCCATCGA
CAGTTCCTCCAACTCCGCCACCCTCACCATCTCTGGACTGAAGACTGAGGACGAGGCTG
AGTACTACTGTCAGTCTTTTGATAGTAGTAACTATGTCGTCTTCGGCGGAGGGACCACG
CTGACCGTCCTCGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCCTC
TGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCCGG
GAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGACC
ACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTGAC
GCCTGAGCAGTGGAAgTCCCACAGAAGCTACAGCTGCCAGGTCACGCATGAAGGGAGCA
CCGTGGAGA

>FOR209_AA

VSESPGKTVTISCTGSSGSIASNYVQWYQQRPGSAPTTVIYEDDQRPSGVPDRFSGSID
SSSNSATLTISGLKTEDEAEYYCQSFDSSNYVVFGGGTTLTVLGQPKAAPSVTLFPPSS
EELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLT
PEQWKSHRSYSCQVTHEGSTVE

>FOR212_cDNA

TCAGCCCCACTCTGTGTCGGGGTCTCCGGGGAAGACAGTGACCATCTCCTGCACCCGCA
GCAGTGGCAGCATTGCCACCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTTCC
CCCACCACTGTTATCTATCAGGATTACCAAAGACCCTCTGGGGTCCCTGATCGGTTCTC
TGGCTCCATCGACAGCTCCTCCAACTCTGCCTCCCTCTCCATCTCTGGACTGAAGACTG
AGGACGAGGCTGACTACTACTGTCAGTCTTATGATAACAACAATTGGGTGTTCGGCGGA
GGGACCAAGCTGGCCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCC
GCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACT
TCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGA
GTGGAGACCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCT
GAGCCTGACGCCTGAGCAGTGGAAGTCCCACAGAAGCTACAGCTGCCAGGTCACGCATG
AAGGGAGCACCGTGGAGAAGACAGTGA

>FOR212_AA

QPHSVSGSPGKTVTISCTRSSGSIATNYVQWYQQRPGSSPTTVIYQDYQRPSGVPDRFS
GSIDSSSNSASLSISGLKTEDEADYYCQSYDNNNWVFGGGTKLAVLGQPKAAPSVTLFP
PSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYL
SLTPEQWKSHRSYSCQVTHEGSTVEKTV

>FOR221_cDNA

CAGCATTGCCAGCAACTATGTGCAGTGGTACCAGCAGCGCCCGGGCAGTKCCCCCACCA
CTGTGATCTATGASGATAACSAAAGACCCTCTGGGGTCCCTGATCGGTTCTCTGGCTCC
ATCGACAGSTCCTCCAACTCTGCCTCCCTCACCATCTCTGRACTGAAGACTGAGGACGA
GGCTGACTACTACTGTCAGTCTTATGATAGCARCMVTTGGGTGTTCGGCGGAGGGACCA
AGCTGACCGTCCTAGGTCAGCCCAAGGCTGCCCCCTCGGTCACTCTGTTCCCGCCCTCC
TCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATAAGTGACTTCTACCC
GGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGCGGGAGTGGAGA
CCACCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTACCTGAGCCTG
ACGCCTGAGCAGTGGAAGT

>FOR221_AA

SIASNYVQWYQQRPGSXPTTVIYXDNXRPSGVPDRFSGSIDXSSNSASLTISXLKTEDE
ADYYCQSYDSXXWVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYP
GAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWK

n/a:

FOR181

–

FOR223

–

```
FOR235
-
FOR239
-
```

Translated AA sequences of MM-LCs:

The sequences were ordered according to their *IGKV* germline usage.

***IGKV1-5*:**

>MM1012
MDMRVPAQLLGLLLLWLPGAKCDIHMTQSPSTLSAYVGDRVTITCRASQNINSWLAWYQ
QKPGKAPKLLIYKASSLESGVPSRFSGSGSGTQFTLTISSLQPDDFATYYCQQYMTFWT
FGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1016
MDMRVPAQLLGLLLLWLPGAKCDIRMTQSPSTLSASVGDRVTITCRASQSINSWLAWYQ
QKPGKAPKLLIYKASNLNNGVPSRFSGSGSGTEFTLTISGLQPDDFASYYCQHYNGYSP
SGTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNA
LQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1017
MDMRVPAQLLGLLLLWLPGGKCDIQMTQSPSTLSASVGDRVTITCRASQSINTWLAWFQ
QKPGKAPKLLIYKASSLESGVPSSFSGSGSGTEFTLTISSLQADDFATYYCQQYDDYPW
TFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADY
>MM1032
MDMRVPAQLLGLLLLWLPGAKCDIQMTQTPSTLSASVGDRVTITCRASQGINSWLTWSQ
HKPGKAPKLLIYQASTVDNGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYDSYPW
TFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1035
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTVTCRASHYISSWLAWYQ
QRPGKAPNLLIYKASTLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYNSYPY
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADY
>MM1040
MDMRVPAQLLGLLLLWLPGANGDIQLTQSPSTLSASVGDRVTIACRASQTINSRVAWYQ
QKPGKAPNLLIYRASTLESGVPSRFSGSGSGTEFTLSISGLQPDDFATYYCQQYNTSWT
FGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1041
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTITCRASQNIGSWLAWYQ
QKPGKAPKLLIYKASSLETGVPSRFSGSGSGTEFTLAIDSLQPDDFATYYCQQYNGYSR
TFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1042
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTITCRASQSISTWLAWYQ
QKPGKAPRLLIYKASSLESGVPARFRGSGSGTEFTLTISSLQPDDFATYYCQQYNSLLY
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1045
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTITCRASQTISNLLAWYQ
QKPGKAPKFLIYEASNLEGGVPSRFSGSGSGTEFTLTISNLQPDDFATYYCQQYSSDPY
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ

SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1054
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTITCRASQSISRWVAWYQ
QKPGKAPNLLIYKASSLQSGVPSRFSGSGSGTEFTLNISSLQPDDFATYYCQQYNSYSP
YTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEV
>MM1058
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTITCRASQSIGTWMAWYQ
QRPGKAPNLLIYAASTLESGVPSRFSGSGSGTEFTLTIISLQPDDFATYYCQQYNSSSI
TFGQGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1060
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVTITCRASQSISIWLAWYQ
QKPGKAPKLLISKASSLESGVPSRFSGSGSATEFTLTISSLQPDDSATYFCQQYSHYEV
TFGGGSKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1065
MDMRVPAQLLGLLLLWLPGAKCDIQMTQSPSTLSASVGDRVIITCRASQSINRWLAWYQ
QKPGKAPQFLIEEASILNSGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYDSYPY
SFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1078
MDMRVPAQLLGLLLLWLPGANCDIQMTQSPSTLSASVGDRVTFTCRASQSLGTWLAWYQ
QKPGKAPNLLIYRASTLQSGVPSRFSGSGSGTEFTLTISSVQPDDFATYYCLQYNSYSN
TFGPGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH

***IGKV1-6*:**
>MM1070
MDMRVPAQLLGLLLLWLPGARCAIQMTQSPSSLSASVGDRVTITCRASQDITNDLGWYQ
QKPGKAPKLLIYAASSLQSGVPSRFSGSRSGTDFVLTISSLQPEDFATYFCLQDHTYPY
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE

***IGKV1-9*:**
>MM1019
MDMRVPAQLLGLLLLWLPGARCDIQLTQSPSFLSASVGDRVTITCRASQGINRYLAWYQ
QKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTISSLQPEDFATYYCQQLNSYPF
LTFGPGTKVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1030
MDMRVPAQLLGLLLLWLPGARCDIQLTQSPSSLSASVGDSVTITCRASEAISNYLAWYQ
QKPGKAPNLLIFNAVSLQSGVATRFSGSGSGTEFTLTISSLRPEDFATYYCQQLYDYPR
TFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK

***IGKV1-12*:**
>MM1033
MDMRVPAQLLGLLLLWFPGSRCDSQLTQSPSVVSASVGDRVTISCRASQGISRYLAWYQ
QKPGKAPELLISGASTLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQAYSFLS
FGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS

GNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKV
>MM1037
MDVRVPAQLLGLLLLWFPGSRCDIQMTQSPSFVSASVGDRVTITCRASLGISKWLAWYQ
QTPGKAPKLLIYAASSLQSGVPSRFSGRGSGTDFILTISSLQPEDSATYYCQQANSFPL
TFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH

***IGKV1/D-33*:**
>MM1009
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPTSLSASVGDSVTFSCQASQDISNYLNWYQ
QKPGRAPKLLIHDASNLETGVPSRFSGSGSGTGFTFTISSLQPEDIATYFCLQYDTLPY
TFGQGTKLDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1021
MDMRVPAQLLGLLLLWLSGARCGIRMTQSPSSLSASVGDRVTITCQASQDISNFVNWYQ
QKPGKAPRLLIYDAYNLEEGVPTRFSGTGSGTDFSFTISSLQPEDFATYYCQQYDNPWT
FGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1024
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQ
QKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTISSLQSEDIATYYCQQYKNLPP
MYTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNA
LQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1043
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSAIVGDRVTITCQASQDISSYLNWYQ
QKPGKAPKLLIFETSNLETGVPSRFSGSGSGTDFSLTVSSLQPEDVGTYYCQQYEDLPL
TFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1048
MDMRVPAQLLGLLLLWLSGVRCDIQMTQSPSSLSASVGDRVTITCQASRGIGNYLNWYQ
QKPGKAPKLLIYDASNLETGVSSRFSGSGSGTHFTFAINSLQPEDIATYYCQQYDSLPR
TFGQGTKLDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1050
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSASVGDTVTITCQASQDISSYLNWYQ
HKPGKAPKFLIYDTSNLDSGVPSRFSGRRSGTNFTFTISGLQPEDFATYYCQQYDHLPP
KFGGGTTVEMKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1055
MDMRVPAQLLGLLLLWLSGARCDVQMTQSPSSLSAFVGDRVTITCHASQHITNYLNWYQ
HKAGKAPKLLISDASNLESGVPTRFSGSGFGTDFTFSINGLQPEDVATYYCQQFDSLPP
AFGPGTKVDVKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLT
>MM1056
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSASVGDRVTITCQASQDIAKYLNWYQ
QKPGKPPKLLIYDTSNLETGVPSRFSGGGGGTDFTFTINSLQPEDIATYYCQQYDDFPL
TFGPGTKVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1062
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQ
QKSGKAPKLLIYDASSLQTGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQYDNLPP

TFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1068
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSASVGDRVTITCQASQDIDNFLSWYQ
QKPGKAPKVLIYDASNLETGVPSRFSGSGSGTDFAFTISSLQPEDFATYYCQQYDSLPL
TFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1074
MDMRVPAQLLGLLLLWLSGARCDIQMTQSPSSLSASLGDRVTITCQASQDISSYLNWYH
QKPGKAPKLLIFEASLLEPGVPSRFSGSGSGTDFTFTISTLQPEDIGTYFCQHYDNLPP
SFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE

***IGKV1/D-39*:**
>MM1003
LDMRVPAQLLGLLLLWLRGARCDIRMTQSPSSLSASVGDRVTITCRASQSISSYLSWFQ
QKPGKAPKLLISAASSLQGGVPSRFSGSGSGTDFTLTISNLQPEDFATYFCQQSYSTPH
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADY
>MM1004
MDMRVPAQLLGLLLLWLRGARCDIKMTQSPSSLSASVGDRVTITCRASRNITNYLNWYQ
LKPGKAPKLLIFAASSSHSGVPSRFSGSGSGTDFSLSVSSLQPEDFVTYYCQQSYSTPY
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1005
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASVGDRVTITCRASQSISTYLNWYQ
QRPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQESFSTLM
FTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1006
MDMRVPAQLLGLLLLWLRGARCDIEMTQSPASLSASVGDRVTISCRAHLDVGRYLNWYQ
LKPGKAPRLLIYAASTLQTGVPSRFSGGGRGTDFTLTIDSLQVEDFATYCCQQTFSAPP
WTFGRGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1020
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASVGDRVTLTCRAGQSISTYLNWYQ
AKPGKAPELLIYSASSLHNGVPSRFSGSGSGTDFILTINSLQPEDFATYYCQQTYGPPH
TFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVY
>MM1031
MDMRVPAQLLGLLLLWLRAARCDIQMTQSPSSLSASVGDRVNITCRASQSISTFLNWYQ
QKPGKAPRLLIYIASNLQSGVPSRFSGSGSGTDFTLTINSLQPEDSATYYCQQCYSTSV
FTFGPGTKVDLKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNKFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1034
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASIGDRVTITCRASQSISFNLIWYQ
QKPGKAPKLLIYATSSLQSGVPSRFSGSGSGTDFTLTISSLQPEDYATYYCQQSYTTPL
TFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1036
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASVGDRITITCRTSQTISTYLNWYQ

QKPGKAPKLLIFAASSLENGVPSRFSGSGSGTEFTLTISSLQPEDFATYYCHQTYLIPL
FGGGTRVEVKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1052
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASVGDRVTITCRASQNISRYLNWYQ
HKPGKAPKLLIYVASSLHAGVPSRFSGGGSGTDFTLTISSLQPEDFATYFCQQSYSTRM
YTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1057
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASVGDRVSITCRASESISSYVNWYQ
QKPGKAPKLLIYTASSLQSGVPPRFSGSASGTDFTLTISSLQPEDFATYYCQQSYSTPI
TFGQGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1061
MDMRVPAQLLGLLLLWLRGARCDIQLTQSPSSLSASVGDRVTVTCRTSQTINSHLNWYQ
QRPGKAPKLLIYTASSLQSAVPSRFSGSGSGTDFTLTISSLQPGDFATYYCQQTYSTPV
TFGQGTRVEIRRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1069
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASVGDRVTITCRASQNCGTYLNWYQ
QKPGKAPKVLIYGASSLQSGVPSRFSGNGSGTDFTLTISSLQPEDFATYYCQQTYSLTI
AFGPGTKVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKAD
>MM1079
MDMRVPAQLLGLLLLWLRGARCDIQMTQSPSSLSASIGDRVTITCRASQSISQFLNWYQ
QKPGKAPRPLIYGASNLQSGVPSRFSGSGSGTDFTLTISGLQPEDFATYSCQQSYSRPR
TFGQGTRVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1084
MDMRVPAQLLGLLLLWLRGARCVIQMTQSPSSLSASVGDRISITCRASENIKTYLNWYQ
QTPGKAPKVLIFGASSLHSGVPSRFSGSGSGTDFTLTINSLQPEDFATYYCHQSYILPW
TFGQGSKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE

***IGKV2/D-28*:**
>MM1026
MRLPAQLLGLLMLWVSGSSGVILMTQSPLSLPVTPGEPASISCRSSQSLLHSNGYNFLD
WYLQKPGQSPQLLIFLGSNRASGVPDRFSASASGTDFTLNISRVEADDVGVYYCMQARQ
TPPTFGQGTKVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD
>MM1049
MRLPAQLLGLLMLWVSGSSGDFVLTQSPLSLPVTPGEPASISCRSTQSLLDNNGHNRLD
WYLQKPGQSPQLLIYMGSNRASGVPDRFSGSESDTDFTLRISRVEAEDVGVYYCMQSLQ
TPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1063
MRLPAQLLGLLMLWVSGSSGDIVMTQSPLSLTVTPGEPASISCRSSQTLLHSNGYNYLD
WYVQARAVSTAPDLFGF-
SGLLPDRFSGSGSGTDFTLNISRVEAGDVGVYYCMQALRTLTFGGGTKVEIRRTVAAPS
VFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTY
SLSSTLTLSKADYE

>MM1073
MRLPAQLLGLLMLWVSGSSGDIVMTQSPLSLPVTPGEPASISCRSSESLLHNSGYNYLD
WYLQKPGRSPQLLIYLGSKRASGVPDRFSASGSGTDFTLKISRVEAEDVGVYYCMQALQ
APTFGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNA
LQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD

***IGKV2/D-30*:**
>MM1044
MRLPAQLLGLLMLWVPASTGDVVMTQSPLSLPVTLGQPASISCRSSQSLVDSDGDTYLN
WYHQRPGQSPRRLIYEVSKRDSGVPDRFSGSGSGSDFTLKISRVEAEDVGIYYCMQGTH
WPRGTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVD
NALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADY
>MM1051
MRLPAQLLGLLLLWVPGFSGDVVLIQSPVSLPVTLGQAASISCRSSQSLLYRDGTTYLN
WFHQRPGQSPRRLISKVSNRDSGVPARFSGSGSGTDFTLRISRVEAEDVGLYYCMQGTH
WPQTFGQGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1075
MRLPAQLLGLLMLWVPVSSGDVVMTQSPLSLPVTLGQSASISCRSSQSLVFSDGNTFLH
WFQQRPGQSPRRLVHKVSNRDSGVPDRFSGSGSGTDFTLRITRVEAEDVGIYYCMQGTH
WPFTFGPGTRVDINRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN
ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKA

***IGKV2/D-40*:**
>MM1082
GSSGDIVLTQTPLSLAVTPGEPASISCRSSQSLFRSDDGNTFLDWYLQKPGQSPQLLIY
ALSHRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMHRIEFPWTFGQGTKVEIKR
TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD
SKDSTYSLSSTLTLSKADYE

***IGKV3/D-11*:**
>MM1007
MEAPAQLLFLLLLWLPDITGEIVLTQSPATLSLSPGERATLSCRASQNVGSYVAWYQQK
PGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFVVYYCQQRSNWPGLT
FGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1008
MEAPAQLLFLLLLWLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSNYLAWYQHK
PGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQQRSNWPSFG
PGTKVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGN
SQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1011
MEAPAQLLLLLLLWLPDITGEIVVTQSPATLSLSPGERATLSCRASQSVSSNLAWYQQK
PGQAPRLLIYEASKRATGIPARFSGSGSGTDFTLTISGLEPEDFAFYYCQQSRTWPETF
GQGTKVEVKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1014
MEAPAQLLFLLLLWLPDSTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQK
PGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFAIYYCQLRSNWPPGT
FGPGTRVDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE

>MM1015
MEAPAQLLFLLLLWLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQK
PGQAPRLLIYDASTRATGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQQRSNGPPLT
FGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1018
MEAPAQLLFLLLLWLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVTNYLDWYQQK
PGQAPRLLIFDSSNRATGIPARFSGSGSGTDFTLTVSSLEPEDFAVYYCHQRSNWPWTF
GQGTNVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADY
>MM1022
MEAPAQLLFLLLLWLPVTTGETVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQK
PGQAPRLLIYDTSNRATDIPARFSGSGSGTGFTLNISTLEPEDSAVYYCQQRFNWPWTF
GQGTRVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1023
MEAPAQLLFLLLLWLPDSTGEIVLTQSPATLSLSPGERATLSCRASQSVTTYLAWYQQK
PGQPPRLLIYDTSNRATGVPARFSGSGSGTDFTLTISSLEPEDFAVYYCQQNTNWPRTF
GQGTRVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1046
MEAPAQLLFLLLLWLPDTTGEIVLTQSPATLSLSPGERATLSCRASQSVRTYLAWYQQK
LGQAPRLLIYDASVRATGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQQRSNWLTFG
GGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGN
SQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1071
MEAPAQLLFLLLLWLPAAAERIVLTQSPATLSLSPGERATLSCRASQSVRSYLAWYQHK
PGQAPRLLIYDAGNRATGTPARFSGSGSGTDFILTISSLEPEDSAVYYCQHRSNWFTFG
QGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGN
SQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1077
MEAPAQLLFLLLLWLPDSTGEIVLTQSPATLSLSPGERATLSCRASQSVRTYLAWYQQK
RGQAPRLLIYDASSRATGIPARFSGSGSGTDFTLTISGLEPEDFAVYYCQQRSNWPITF
GQGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1080
MEAPAQLLFLLLLWLPDATGEVVLTQSPATLSLSPGERATLSCRASQSVDTYIAWFQQK
PGQAPKVLIYDASTRATGIPGRFSGSGSGTDFTLTISSLEPEDFAVYYCQQHLNGLFGF
GPGTKLDIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEK

***IGKV3-15*:**
>MM1025
MEAPAQLLFLLLLWLPDTTGEIVMTQSPATLSVSPGEGATLSCRASQNVNSNLVWYQQK
PGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTISSLQSEDFAVYYCQHYNNWPHTF
GRGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1029
MEAPAQLLFLLLLWLPENTGDIVMTQSPATLSVSPGARATLSCRASQSVASSLAWYQHK
PGQAPRLLIYAASTRATGIPARFSGSGSGTDFTLTITSLQSEDFVVYYCQQYDDWPRTF
GQGTKVEMKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG

NSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1067
MEAPAQLLFLLLLWLPDTTGEIVMTQSPATLSVSPGERATLSCRASQTVYSNLVWFQQK
PGQAPSLLIYGATTRATGIPARFSGSGSGTEFTLTIISLQSEDSAVYYCQQYNTRPYTF
GPGTKVNIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1076
MEAPAQLLFLLLLWLPETTGEIVMTQSPATLSVSPGDSAALSCRASQSVHSNLAWYQQK
PGQAPRLLIYGASTRATGIPGRFSGSGSGTEFTLTISSLQSEDFAVYYCQQYDNWPPLT
FGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE

***IGKV3/D-20*:**
>MM1010
METPAQLLFLLLLWLPDTTGEIVLTQSPGSLSLSPGERATLSCRASQTVSSNYVAWYHQ
APGQAPRLLIYAASRRATGIPDRFSGSGSGTDFTLTISRLEPEDFGVYHCQQYGRSITF
GGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKA
>MM1027
METPAQLLFLLLLWLPDTTGEIVLTQSPGTLSLSPGERATLSCRASQSLSSMYLAWYQH
KPGQAPRLLIYAASNRATGVPGRFSGSVSGTDFTLTINRLEPEDFAVYYCQQFGSSTAT
FGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1028
METPAQLLFLLLLWLPDTTGEIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQ
KPGQAPRLLIYATSNRAAGIPDRFSGSGSGTDFTLTISRLGPEDFAVYFCQQYGSSPYT
FGGGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1039
METPAQLLFLLLLWLPDTTGEIVLTQSPGTLSLSPGERATLSCRAGHSVSSSYLAWYQQ
KPGQAPRLLIYGASTRAAGVPDRFSGSASGTDFTLTIGRLEPEDFAMYYCQYYGGSGRW
TFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKA
>MM1047
METPAQLLFLLLLWLPDTTGEIVLTQSPGTLSLSPGERATLSCRASQSVTTTYLHWYQQ
KPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCHQYGYSPQT
FGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1059
METPAQLLFLLLLWLPDTTAEIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQ
KPGQAPRLLIYDASTRATGIPDRFSGSGSGADFILTISSLEPEDFAMYYCQQYGRSPYT
FGPGTKVDLKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1066
METPAQLLFLLLLWLPDTSGEIVLTQSPGTLSLSPGERAILSCRASQSISSIYVAWYQQ
KPGQAPSLLMYAASSRATGIPDRFTGGGSGTDFTLTISRLEPEDFAVYYCQHYDTSPMY
TFGQGTRLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1072
EPWKPQRSFSSSCYSGSQTGEIVLTQSPGTLSLSPGERATLSCRASQSVSGGHLVWYQK
KPAQAPRLVMYEVSIRATGIPDRFSGSGSGTDFTLTITRLESEDFAVYYCQQYDIAPFT

FGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS
GNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1081
METPAQLLFLLLLWLPDATGEIVLRQSPGTLSLSPGERATLSCRASQSVSSAYLAWYQQ
KPGQTPRLLIYAASRRATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGGSPGL
TFGGGTKVETKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQ
SGNSQESVTEQDSKDSTYSLSSTLTLSKADYE

***IGKV4-1*:**
>MM1001
MVLQTQVFISLLLWISGASGDIVMTQSPDSLAVSLGERATINCKSSQSVLYSPNNKNYL
AWYQQKPGQPPKLLIYWAYTRESGVPDRFSGSGSGTDFTLTISSLQAEDVAVYYCHQYY
STTWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVD
NALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1013
MVLQTQVFISLLLWISGANGDIVMTQSPDSLAVSLGERATINCKSSQTVLYDSKNKHYL
AWFQQKPGQPPKLLIYWASTRQSGVPDRFSGSGSGTDFTLTISSLQAEDVAVYYCQQYY
IIPYTFGQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVD
NALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
>MM1038
MVLQTQVFISLLLWISGAYGDIVLTQSPDSLAVSLGERATINCESSQSLLYRSNNRKYL
TWFQQKPGQPPKLLFYWASSRESGVPDRFSASGSGTFFTLTINSLQAEDVAVYYCQQYL
TIPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVD
NALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
>MM1053
MVLQTQVFISLLLWISGVYGDIVMTQSPDSLAVSLGERATINCTSSQSIFYISKNRNYL
AWYQQKPRQPPSLLIYWASTRESGVPDRFSGSGSGTDFTLTISSLQPEDVAVYYCQQYY
RTPYTFGQGTKLEIKGTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVD
NALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
>MM1083
MVLQTQVFISLLLWISGAYGDIVMTQSPDSLAVSLGERATINCKSSQSILYSSNNENYL
AWYQQKPGQPPKLLIYWASTRESGVPDRFSGSGSGTDFTLTISSLQAEDVAVYYCQQYS
TPPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVD
NALQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD

***IGKV5-2*:**
>MM1064
MGSQVHLLSFLLLWISDTRAETTLTQSPAFVSATPGDKVNISCKASQDIDDDVSWYQRK
PGEAAIFIIQEATLLVPGIPPRFSGSGYGTEFTLTINNIESEDAAYYFCLQHDDFPYTF
GQGTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEKH


MM1024_IGKV1_NGS
MM1024_IGKV1_Sanger
MM1026_IGKV2_NGS
MM1026_IGKV2_Sanger
MM1026_IGKV3_NGS
MM1007_IGKV3_Sanger
MM1007_IGKV3_NGS
MM1007_IGKV3_Sanger
MM1001_IGKV4_NGS
MM1001_IGKV4_Sanger

SI Figure 1. Verification of multiple myeloma bulk RNA-sequencing data.
The multiple myeloma RNA-sequencing data were verified using the same Sanger sequencing protocol as for the AL amyloidosis patients. For verification sequences from multiple myeloma patients of all five major *IGKV* families were randomly selected. The *IGKV5* sequence of the Sanger sequencing displayed heterozygous signals, which were defined according to the IUPAC-IUB code (NC-IUB 1985) and the sequence from bulk RNA was detectable in the subsequences. (Schreiner et al. 2024) This figure was adapted and published under the terms of the Creative Commons Attribution (https://creativecommons.org/licenses/by/4.0/#ref-appropriate-credit).

SI Table 1. Calculation of the relative risk.

	Relative risk	p-value
IGKV1-5	0.741	0.536
IGKV1-6	0.683	0.814
IGKV1-9	1.038	0.979
IGKV1-12	1.038	0.976
IGKV1/D-13	6.146	0.263
IGKV1-16	18.44	0.049
IGKV1/D-33	2.452	0.013
IGKV1/D-39	1.186	0.670
IGKV2/D-28	0.228	0.317
IGKV2/D-30	0.293	0.413
IGKV2/D-40	0.683	0.814
IGKV3/D-11	0.173	0.086
IGKV3-15	1.556	0.550
IGKV3/D-20	0.111	0.126
IGKV4-1	0.830	0.819
IGKV5-2	2.075	0.602

MedCalc Software Ltd. Relative risk calculator. https://www.medcalc.org/calc/relative_risk.php (Version 20.218) [05.03.2023].

SI Table 2. Analysis of potential N-glycosylation site.

	potential N-glycosylation site [n]	no potential N-glycosylation site [n]
IGKV1-5		
AL [n= 5]	1	4
MM [n= 14]	3	11
IGKV1-6		
MM [n= 1]	1	-
IGKV1-9		
AL [n= 1]	-	1
MM [n= 2]	-	2
IGKV1-12		
AL [n= 1]	1	-
MM [n= 2]	-	2
IGKV1/D-13		
AL [n= 1]	-	1
IGKV1-16		
AL [n= 4]	3	1
IGKV1/D-33		
AL [n= 13]	8	5
MM [n= 11]	1	10
IGKV1/D-39		
AL [n= 8]	2	6
MM [n= 14]	4	10
IGKV2/D-28		
MM [n= 4]	2	2
IGKV2/D-30		
MM [n= 3]	1	2
IGKV2/D-40		
MM [n= 1]	-	1
IGKV3/D-11		
AL [n= 1]	-	1
MM [n= 12]	1	11
IGKV3-15		
AL [n=3]	2	1
MM [n= 4]	-	4
IGKV3/D-20		
MM [n= 9]	-	9
IGKV4-1		
AL [n= 2]	-	2
MM [n= 5]	1	4
IGKV5-2		
AL [n= 1]	-	1
MM [n= 1]	1	-
IGLV6-57		
AL [n=10]	0	10
IGLV2-23		
MM [n= 8]	0	8

Analysis according to (Julenius 2007). AL= AL amyloidosis, MM= multiple myeloma

SI Table 3. Calculation of prediction probability with VLAmY-Pred- tool.

	Organ involvement	Prediction probability	Classification
IGKV1-5			
FOR1009	Diverse	0.966	No
FOR1014	Diverse	0.629	Yes
FOR1027	Diverse	0.569	No
FOR1040	Heart	1.000	Yes
FOR1041	Diverse	0.569	No
MM1012	-	0.966	No
MM1016	-	0.966	No
MM1017	-	0.629	Yes
MM1032	-	1.000	Yes
MM1035	-	0.569	No
MM1040	-	0.569	No
MM1041	-	0.966	No
MM1042	-	0.966	No
MM1045	-	0.629	Yes
MM1054	-	1.000	Yes
MM1058	-	0.569	No
MM1060	-	0.569	No
MM1065	-	0.629	Yes
MM1078	-	0.569	No
IGKV1/D-33			
FOR1001	Diverse	0.820	Yes
FOR1006	Heart	0.629	Yes
FOR1007	Heart	0.820	Yes
FOR1010	Heart	0.820	Yes
FOR1013	Kidney	0.629	Yes
FOR1018	Heart	0.629	Yes
FOR1020	Kidney	0.820	Yes
FOR1021	Heart	0.629	Yes
FOR1023	Heart	0.820	Yes
FOR1024	Diverse	0.820	Yes
FOR1028	Diverse	0.629	Yes
FOR1036	Diverse	0.820	Yes
FOR1039	Diverse	0.82	Yes
MM1009	-	0.820	Yes
MM1021	-	0.629	Yes
MM1024	-	0.820	Yes
MM1043	-	0.629	Yes

	Organ involvement	Prediction probability	Classification
MM1048	-	0.820	Yes
MM1050	-	0.820	Yes
MM1055	-	0.629	Yes
MM1056	-	0.629	Yes
MM1062	-	0.820	Yes
MM1068	-	0.629	Yes
MM1074	-	0.629	Yes
IGKV1/D-39			
FOR1002	Kidney	0.574	No
FOR1004	Kidney	1.000	Yes
FOR1011	Kidney	0.966	No
FOR1012	Kidney	0.820	Yes
FOR1019	Kidney	1.000	Yes
FOR1026	Diverse	0.569	No
FOR1032	Diverse	0.820	Yes
FOR1034	Diverse	0.966	Yes
MM1003	-	0.569	No
MM1004	-	0.901	Yes
MM1005	-	0.966	No
MM1006	-	0.966	No
MM1020	-	0.966	No
MM1031	-	0.574	No
MM1034	-	1.000	Yes
MM1036	-	0.966	No
MM1052	-	0.966	No
MM1057	-	0.574	No
MM1061	-	0.901	Yes
MM1069	-	0.569	No
MM1079	-	0.569	No
MM1084	-	0.574	No
IGKV1-16			
FOR1025	Diverse	0.569	No
FOR1031	Diverse	0.569	No
FOR1033	Diverse	0.820	Yes
FOR1035	Kidney	0.820	Yes
IGKV3/D-11			
MM1007	-	0.966	No
MM1008	-	0.963	No
MM1011	-	0.963	No
MM1014	-	0.966	No

	Organ involvement	Prediction probability	Classification
MM1015	-	0.966	No
MM1018	-	0.569	No
MM1022	-	0.569	No
MM1023	-	0.963	No
MM1046	-	0.966	No
MM1071	-	0.569	No
MM1077	-	0.569	No
MM1080	-	0.963	No
IGLV6-57			
FOR125	Kidney	0.629	Yes
FOR143	Kidney	0.963	No
FOR158	Kidney	0.963	No
FOR199	Kidney	0.629	Yes
FOR207	Kidney	1.000	Yes
FOR209	Kidney	0.901	Yes
FOR212	Kidney	1.000	Yes
FOR221	Kidney	0.901	Yes
FOR151	Diverse	0.629	Yes
FOR183	Diverse	0.629	Yes
IGLV2-23			
MM107	-	0.966	No
MM114	-	0.569	No
MM117	-	0.966	No
MM125	-	0.966	No
MM132	-	0.569	No
MM133	-	0.569	No
MM135	-	0.569	No
MM149	-	0.569	No

Analysis according to (Rawat et al. 2021) FORx= sequences from AL amyloidosis patients, MMx= sequences from multiple myeloma patients, yes= predicted as amyloidogenic variable light chain sequences, no= predicted as non-amyloidogenic variable light chain sequences.

SI Table 4. NGS analysis of *IGLV3-21*.

Sub-sequences	Organ involvement	Sequence fraction [%]	*IGLV*
FOR129-1	Kidney	2.837	*IGLV3-19*
FOR129-2		2.436	*IGLV3-16/IGLV3-25*
FOR129-3		1.528	*IGLV3-21*
FOR129-4		1.085	*IGLV3-21*

Sub-sequences	Organ involvement	Sequence fraction [%]	*IGLV*
FOR146-1	Diverse	1.441	*IGLV3-19*
FOR146-2		1.324	*IGLV3-25*
FOR146-3		1.195	*IGLV3-19*
FOR146-4		1.148	*IGLV3-16/IGLV3-25*
FOR146-5		1.061	*IGLV3-21*
FOR146-6		1.056	*IGLV3-25/IGLV3-16*
FOR146-7		1.053	*IGLV3-25*
FOR170-1	Diverse	1.980	*IGLV7-43*
FOR170-2		1.258	*IGLV3-21*
FOR170-3		1.144	*IGLV3-21*
FOR170-4		1.085	*IGLV3-21*
FOR178-1	Kidney	12.039	*IGLV3-25*
FOR178-2		5.902	*IGLV3-21*
FOR178-3		2.234	*IGLV2-8/IGLV2-11*
FOR178-4		1.988	*IGLV3-27*
FOR178-5		1.609	*IGLV3-1*
FOR178-6		1.417	*IGLV1-36/IGLV1-44/IGLV1-47*
FOR178-7		1.352	*IGLV8-61*
FOR178-8		1.267	*IGLV3-16/IGLV3-25*
FOR178-9		1.263	*IGLV3-19*
FOR178-10		1.116	*IGLV3-21*
FOR178-11		1.020	*IGLV2-11*

Analysis of sub-sequences of two patients with dominant kidney involvement and two patients classified as diverse. Sequencing was performed using a multiplex primer and sub-sequences were extracted using MiXCR. Sub-sequences with a fraction size >1 % were analysed.

Curriculum Vitae

PERSONAL INFORMATION

Name:	Sarah Ursula Schreiner
Date of Birth:	08.09.1989 in Speyer
Nationality:	German

SCHOOL HISTORY

1996 - 2009	Edith-Stein Gymnasium Speyer
16.03.2009	Abitur

PROFESSIONAL TRAINING

2009-2011	*Training as Biological-Technical Assistant*, Naturwissenschaftliches Technikum, Dr. Künkele, Landau

ACADEMIC HISTORY

WS 11/12 - SS 15	*Bachelor Biotechnology*, University of Applied Science Darmstadt
13.05.2015	Bachelor of Science
WS 15/16 - WS 17/18	*Master Applied Life Science* „Angewandte Bio-, Pharma- und Medizinwissenschaften, University of Applied Science Kaiserslautern
22.12.2017	Master of Science
10/2020	*accepted PhD Student (Dr. sc. hum.)* University Heidelberg

PROFESSIONAL EXPERIENCE

05/2015 - 08/2015	*Research associate*, University Hospital Frankfurt am Main
01/2018 - 11/2019	*Development engineer*, R-Biopharm AG, Darmstadt
12/2019-05/2023	*Research associate* University Hospital Heidelberg
since 06/2023	*Head of laboratory*, SRH Kurpfalzkrankenhaus, Heidelberg

AWARDS AND PRESENTATIONS

09/2022	*Oral presentation* “Sequence diversity of the kappa light chains from patients with AL amyloidosis and multiple myeloma” at the XVIII. International Symposium on Amyloidosis 2022 *DGAK Young Investigator Award*

Danksagung

An erster Stelle möchte ich mich bei meiner Doktormutter Prof. Dr. med. Ute Hegenbart für die Überlassung des Themas dieser Arbeit und ihre Unterstützung bedanken. Danke für die Beantwortung all meiner klinischen Fragen und für dein stets offenes Ohr. Besonders möchte ich mich auch bei Prof. Dr. med. Stefan Schönland bedanken. Danke für dein andauerndes Interesse an meiner Forschung, den wissenschaftlichen Austausch und die Unterstützung während dieses Projektes. Eure Tür stand immer offen für mich trotz hoher Belastung in der Klinik.

Darüber hinaus möchte ich mich bei Dr. sc. hum. Stefanie Huhn für die Möglichkeit diese Arbeit in ihrem Labor durchzuführen bedanken. Danke für deine Unterstützung sowohl bei praktischen Laborthemen als auch in wissenschaftlicher Hinsicht. Danke Dr. rer. nat. Philipp Reichert für deine Hilfsbereitschaft, vor allem bei allen Themen rund um NGS und IT. Dem gesamten Laborteam möchte ich für die schöne Arbeitsatmosphäre danken.

Ich danke der „Deutschen Forschungsgemeinschaft“ für die Finanzierung meines Promotionsprojektes, sowie allen weiteren Mitgliedern des FOR2969. Es war eine schöne Erfahrung mit Forscherinnen und Forschern unterschiedlicher Disziplinen zusammenzuarbeiten.

Ganz besonders möchte ich meiner Kollegin und Freundin Dr. sc. hum. Natalie Berghaus danken. Danke für die schöne gemeinsame Zeit im Labor und darüber hinaus. Wir haben uns immer unterstützt und du hast mir oft einen anderen, wertvollen Blickwinkel, nicht nur in Bezug auf die Forschung, gezeigt.

Ich danke meinen Eltern für ihre bedingungslose Unterstützung auf meinem langen Weg der Ausbildung und des Studiums. Danke für eure aufbauenden Worte und eure Motivation als das Ziel noch so weit weg schien. Ohne euch hätte ich es nicht geschafft!

Christopher, danke für einfach alles!

Eidesstattliche Versicherung

1. Bei der eingereichten Dissertation zu dem Thema

Biological and clinical impact of the immunoglobulin light chains sequence diversity in patients with dominant kidney AL amyloidosis

handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Heidelberg, 17.09.2024 — S. Schreiner

Ort und Datum — Unterschrift