UNIVERSITÀ DEGLI STUDI DI MILANO

PhD course in Molecular and Cellular Biology

Department of Biosciences

Exploring the molecular and biophysical mechanisms of proteotoxic immunoglobulin light chains in AL amyloidosis

Scientific tutor: Prof. Stefano Ricagno

Oberti Luca

R11395

ABSTRACT (ITA)

L'Amiloidosi da catene leggere (AL), è la più comune amiloidosi sistemica. Questa patologia porta all'aggregazione delle catene leggere delle immunoglobuline (CL), le quali esplicano proteo-tossicità in organi target. Le CL sono fra di loro caratterizzate da alta variabilità di sequenza, generatasi a livello genomico, in risposta alla ricombinazione VDJ. In questo senso, è possibile affermare che in ogni paziente la malattia venga causata da una determinata CL (Merlini, 2017). Quest'alto livello di complessità rende ad oggi poco chiari i determinanti molecolari che causano l'aggregazione e la proteo tossicità delle CL. Di conseguenza, studi su questi meccanismi richiedono l'analisi di un ampio gruppo di casi per consentire la generalizzazione, con valenza statistica, delle osservazioni. Il nostro approccio considera diverse proprietà biofisiche e strutturali delle CL mediante l'utilizzo di varie tecniche spettroscopiche, proteolisi limitata e diffrazione ai raggi X. Inoltre, in questo lavoro si fa riferimento ad un set di tredici proteine, con sequenze molto diverse tra loro. Fanno parte del gruppo di CL: otto CL amiloidogeniche, le quali provocano gravi sintomi cardiaci in pazienti AL e cinque non amiloidogeniche. In particolare, le CL non tossiche sono isolate e purificate, da pazienti affetti da mieloma multiplo nel quale le CL sono over-espresse ma non mostrano nessun segno d'aggregazione o tossicità.

I dati qui presentati, mostrano che la bassa stabilità di folding e l'alta flessibilità tendono a correlare con la proprietà amiloidogenica delle CL studiate. Le caratteristiche strutturali e di idrofobicità, invece, non sembrano avere un ruolo nel definire la propensione all'aggregazione.

Inoltre, da precedenti studi si sa che la tossicità delle CL *in vivo* è collegata alla presenza di rame. In particolare, la relazione fra rame e CL aumenta la produzione di specie radicali dell'ossigeno e il loro seguente danno mitocondriale (Diomede et al., 2017a). I nostri dati mostrano che le CL amiloidogeniche interagiscono con il Cu^{2+} con una maggiore affinità, rispetto a quelle non amiloidogeniche. In particolare, abbiamo evidenziato che le Istidine possano mediare l'interazione CL – ioni rame. Infatti, il legame fra CL e rame perde di affinità in presenza di Istidine protonate. Inoltre, i nostri dati suggeriscono che le CL acquisiscono flessibilità e perdono stabilità in presenza di rame.

Il nostro approccio ci ha permesso quindi di concludere, che la propensione all'aggregazione delle CL non può essere descritta da un singolo parametro ma dalla concomitante azione di più tratti biofisici. Inoltre, questi dati suggeriscono che la tossicità *in vivo* possa essere dovuta all'alterazione delle proprietà intrinseche delle CL facilitata dalla presenza del rame.

ABSTRACT (ENG)

Herein, immunoglobulin light chains (LCs) native state was studied in the context of the pathology known as light chain amyloidosis (AL). This pathology is characterized by LCs overexpression, which leads to toxicity and aggregation into amyloid fibrils in target organs, with heart being the most affected one. Due to genetic rearrangement and somatic hypermutation, virtually, each AL patient presents a different amyloidogenic LC (Merlini, 2017). Because of such complexity, the fine molecular determinants of LC aggregation propensity and proteotoxicity are, to date, unclear; significantly, their decoding requires investigating large sets of cases. This project is aimed to unravel the molecular determinants linked with LCs toxicity. First we screened several independent biophysical and structural properties of the LCs native state. In particular, we considered hydrophobicity, fold stability, flexibility and 3D structure.

Our experimental approach considered two LCs sets called 'H' and 'M'. The H set is composed of eight LCs from AL patients while the M set by LCs from multiple myeloma (MM) patients. M LCs were chosen as control since they are overexpressed as the toxic H LCs but they do not lead to toxicity or aggregation. To date, the molecular bases leading to LC proteotoxicity remain to be elucidated. Our data show that low fold stability and high protein flexibility correlate with amyloidogenic LCs, while hydrophobicity, structural rearrangements and nature of the LC dimeric association interface (as observed in seven crystal structures here presented) do not appear to play a significant role in protein aggregation.

Additionally, it has been demonstrated that the LCs toxicity *in vivo* is linked to copper (Cu^{2+}) (Diomede et al., 2017a) by increasing the radical oxygen species (ROS) production. We aimed our studied to clarify Cu^{2+} LCs interaction. Moreover, we wanted to assess whether Cu^{2+} is able to alter the biophysical properties of the native state to more aggregation prone states. Our findings reveal that H LCs interacts with Cu^{2+} with a higher affinity than M LCs and that His residues may be involved in Cu^{2+} binding. Indeed the affinity decreases in presence of protonated His residues. Moreover, data suggest that the interaction with Cu^{2+} increases the molecular flexibility and decreases the fold stability.

These observations suggests that protein aggregation cannot be evaluated through one single parameters but by the co-action of several biophysical traits. Moreover, our results suggest that the presence of Cu^{2+} can alter the native LCs properties leading to a higher toxicity *in vivo*.

Aim of the work

This work is aimed to elucidate the molecular determinants of the H LCs aggregation propensity and *in* vivo toxicity. Herein, we considered the LCs biophysics and structure and the effect of Cu^{2+} on H LCs. In order to reach this aim, we analyzed fold stability, hydrophobicity, protein flexibility, molecular interactions and 3D structures. In particular, these properties may be relevant for protein misfolding and subsequently aggregation. However, our approach also account for high sequence variability, making our data generalizable LC traits.

The biophysical and structural characterizations were performed with the final aim to reveal traits specific for H LCs. Then, starting from these data, we elucidate whether and how Cu^{2+} could alter these specific H LCs properties.

1 INTRODUCTION:

1.1 Immunoglobulin and light chains

Immunoglobulins (Ig) are proteins involved in the adaptive immune system which is activated when pathogens succeed in evading the innate immune system (Charles A Janeway et al., 2001). Igs are hetero-tetramers composed by two heavy chains (HC) and two light chains (LC). Specifically, both HC and LC can be functionally separated in two regions: the variable one and the constant one (Figure 1.1). In particular, both LC and HC present a variable region composed by one variable domain (Vl for the LC and Vh for the HC). Instead, the constant region in LC is composed by one constant domain (Cl) and in HC by three constant domains (Ch) (Schroeder and Cavacini, 2010). Remarkably, VL and Vh are characterized by a pronounced sequence variability which enables different Igs to interact with specific antigens (Charles A Janeway et al., 2001).

LCs are produced by clonal plasma cells (Weiss et al., 2016) in the bone marrow (BM). They can be secreted as homodimers or can be assembled with the HC in order to form complete Igs. LCs present a typical oligomerization state according to their isotype. Indeed, these proteins are classifiable in two isotypes called k LCs and λ LCs, which are principally different in the C-terminal portion. Specifically, λ LCs are always secreted as disulfide-linked homodimers (Waldmann et al., 1972). To date, λ LCs are known to be the most represented LCs population in the context of the pathology called light chain amyloidosis (AL). Therefore, if no otherwise stated, herein the term LC will be referred to the λ isotype.

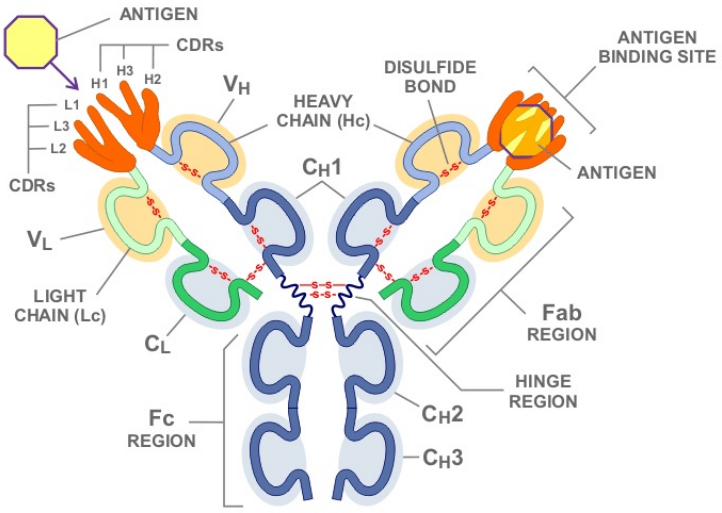


Figure 1.1: Schematic representation of an overall immunoglobulin structure.

1.1.1 Light chain structure

The full length (FL) LC monomer (~ 22KDa), as already reported in the previous paragraph, presents two immunological domains: VL and Cl. Specifically, VL is the N-terminal domain (amino-acids 1-110) followed by CL one (amino-acids 115-215) and they are joined together by a linking region (amino-acids 111-114) (Bourne et al., 2002; Morgan and Kelly, 2016) (Figure 1.2).

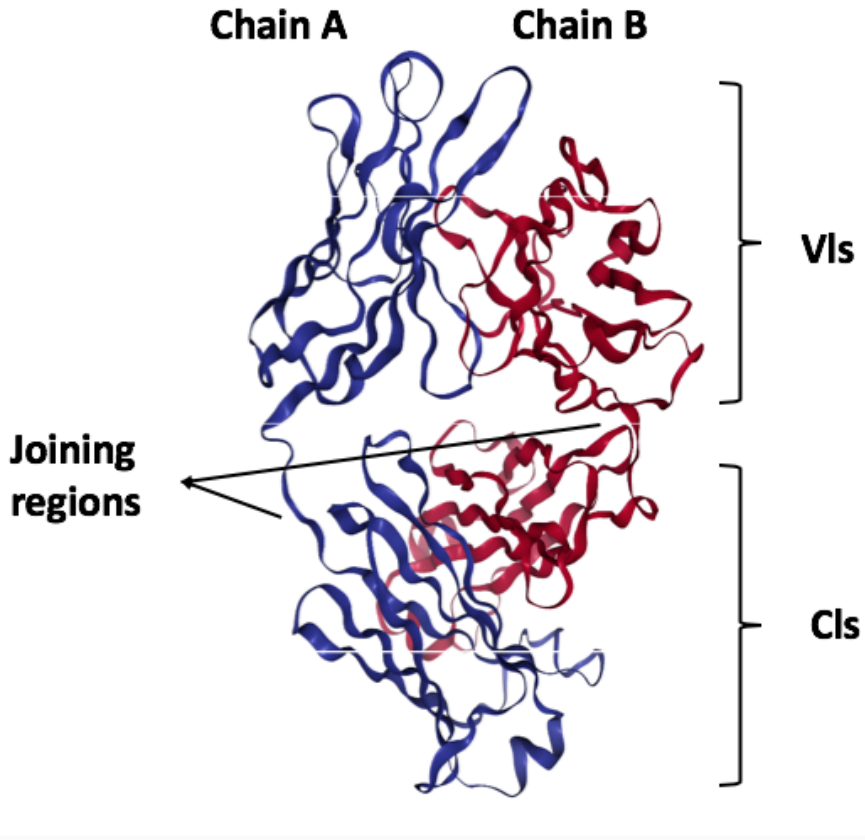


Figure 1.2: LC homodimeric structure. Chain A is represented in blue, Chain B in red. (PDB 1A8J) (Adapted from Edmundson and Manion, 1998).

Both CL and VL share a β-sandwich topology with eight strands (A, B, C, C', D, E, F, and G) arranged in a Greek key motif (Pokkuluri et al., 1999) (Figure 1.3). Each domain is characterized by the presence of an intramolecular disulfide bridge (S-S) which links strands B and F in facing β-sheets stabilizing the domains (Al-Lazikani et al., 1997). In particular, the conservation of the intramolecular S-S bonds in each domain throughout the amino acid sequence divergence, suggests that they play an important structural role (Lapanje and Dorrington, 1973). Moreover, an intermolecular S-S occurs at the C-*terminus* of the Cl. Specifically, this bond link the two Cls in a LC dimer, or the CL

with the Ch in the complete immunoglobulin (Nag et al., 2015). Remarkably, these S-S bonds takes place in stabilizing the overall homodimer quaternary assembly (Leitzgen et al., 1997).

The VL fold consists in nine strands organized in two sheets one composed of four and the other one of five strands. Structurally, the five-stranded sheet is composed by strands C, C',C'', F and G. The other sheet consist in four strands known as A, B, E and D (Figure 1.3).

Vl is directly involved in the interaction with antigens because of the presence of the three complementary determining regions (CDR1, CDR2 and CDR3) (Charles A Janeway et al., 2001). CDRs are known as hypervariable regions because of their sequence diversity, which provides specificity against antigens. Structurally, CDR1 connects strands B-C, CDR2 C'-C'' and CDR3 F-G (Fischmann et al., 1991) (Figure 1.3). The definition and numbering of CDRs is a particular aspect that is well described in literature (*e.g.,* North et al., 2011). In particular, the first attempt on the CDRs characterization was carried out in the mid 1990s by Chothia and coworkers, (Al-Lazikani et al., 1997). Specifically, they used the term "canonical structures" for CDRs, on the base that these regions might only adopt few conformations and structure depending on length and sequence (Chothia et al., 1989). Recently, several studies by North and coworkers, are in contrast with this vision. In 2011 CDRs were classified depending on dihedral angle metric using an affinity-propagation clustering algorithm (North et al., 2011; Adolf-Bryfogle et al., 2015). In particular, the study took in account more than 300 antibody structures previously deposited on the protein data bank (PDB) with the exclusion of those ones with low resolution and high B-factor. To date, this new model is the most used one to identify and numbering CDRs.

The other domain, the Cl, is highly conserved among all the LC in terms of amino acid sequence (KARLSSON and PETERSON, 1972). Its fold consist two sheets as well as the Vl. The first is a three-stranded sheet (strands C, F and G) and it is homologous to the VL five-stranded one but it misses strands C' and C''. The other one is a four-stranded sheet and present the same topology and organization of the VL (Fischmann et al., 1991)(Figure 1.3).

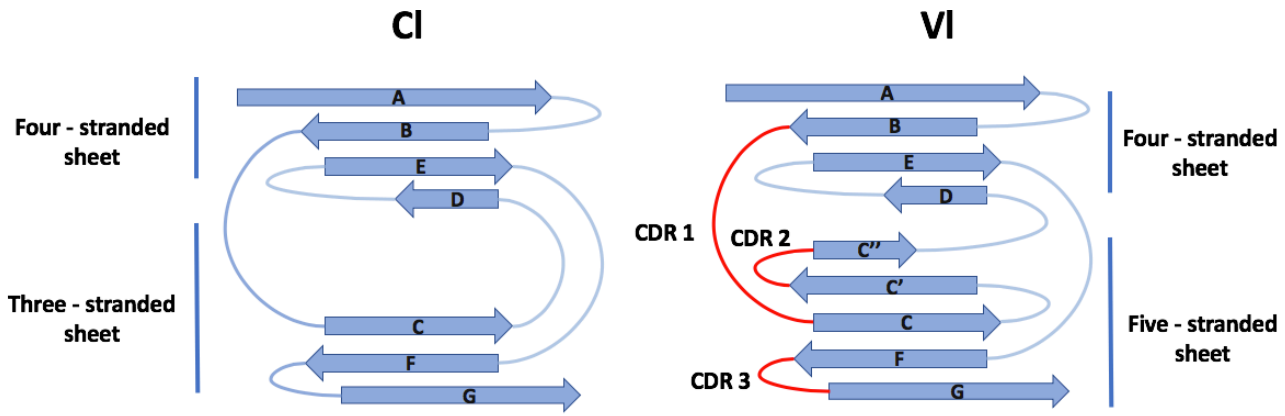


Figure 1.3: Schematic representation of the immunoglobulin light chain VL and CL motifs.

The last part of the LC structure is composed by residues 111-114 which constitute a joining region linking together CL with VL and it is characterized by high flexibility (Bourne et al., 2002).

1.1.2 Sequence variability

The typical LCs high sequence variability arises from genetic and molecular reasons. Lambda locus (IGL) is located on chromosome 22q11.2 and consists of 73-74 IGLV (variable) genes, 7-11 IGLJ (joining) genes and 7-11 IGLC (constant) genes (HUGO Gene Nomenclature Committee, HGNC, Lefranc, 2001) (Figure 1.4). In particular, IGLV genes are divided in three segment: variable (V) and joining (J) (Adolf-Bryfogle et al., 2015). It is worth to know that LCs lack the so called diversity (D) segment. According to this chromosome subdivision each LC can be classified on their V-J (IGLV) and IGLC coding gene, leading to the concept of germline gene. Moreover, the IGLV genes are also sub-grouped in ten families (from IGLV1 to IGLV10) (Jr et al., 1997).

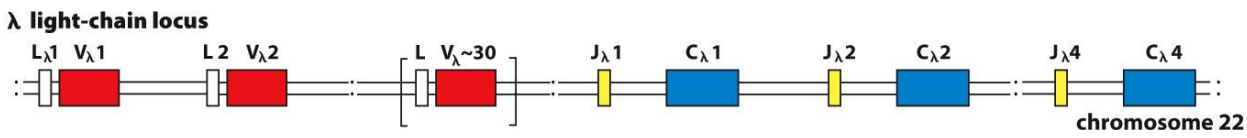


Figure 1.4: Simplified scheme of the LC *locus* on the chromosome 22. (Parham, 2014)

Several studies focused on the definition and nomenclature of the whole repertoire of IGLV genes. In particular, two different works both published in 1995 analyzed the organization of the IGLV leading to two different nomenclature systems (Frippiat et al., 1995; Kawasaki et al., 1995). In 2001, the ImMunoGene Tics database published a study where they proposed the correspondence

between the human IGLV nomenclatures and the one proposed by them (Table 1.1). In particular, they provided the IGLV previous genes according to gene functionality and number of alleles (Lefranc, 2001).

From the molecular point of view, the sequence diversity is the result of the V(D)J recombination, a somatic non-homologous recombination indispensable for B-cells to become mature (González et al., 2007). In particular, this process occurs during the early differentiation of B-cells and generate the primary VL repertoire (González et al., 2007). Specifically, the first recombination event occurs at the variable k locus (IGKV) and if a functional rearrangement is not accomplished, the IGLV undergoes recombination (Gent and Burg, 2007). The process is regulated by the enzymatic machinery V(D)J recombinase (Mansilla-Soto and Cortes, 2003; Schatz and Ji, 2011) and it is controlled by the recombination signal sequences (RSS) flanking the V, D and J gene segments (Ramsden et al., 1994; Schatz and Ji, 2011) (Figure 1.5 A).

V(D)J recombination starts when the V(D)J recombinase binds a RSS creating a single-strand nick between the coding segment (V, D or J) and the first nucleotide of the RSS (Schatz and Swanson, 2011).

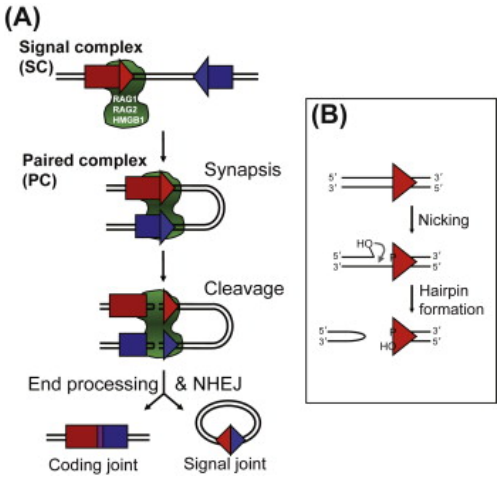


Figure 1.5: (A) Schematic representation of the V-D-J recombination. (B) zoom-in into the hairpin formation (Little et al., 2015).

Table 1.1: Comparison between the human IGLV nomenclatures. (Lefranc, 2001)

IMGT, IGLV gene name [2, 19]	Frippiat et al. [27] Williams, Frippiat et al. [29]	Kawasaki et al. [30]	IMGT, IGLV gene name [2, 19]	Frippiat et al. [27] Williams, Frippiat et al. [29]	Kawasaki et al. [30]
3-1	3r	2-1	1-36	1a	1-11
3-2	3q	2-2P	5-37	5e	4-1
4-3	4c	5-1	(I)-38		1-12P
3-4		2-3P	5-39	5a	
2-5	2a1	1-1P	1-40	1e	1-13
3-6	3a2	2-4P	1-41	1d	1-14P
3-7	3n	2-5P	(VII)-41-1	lambdavgl	
2-8	2c	1-2	(I)-42	V lambda A	1-15P
3-9	3j	2-6	7-43	7a	3-2
3-10	3p	2-7	1-44	1c	1-16
2-11	2e	1-3	5-45	5c	4-2
3-12	3i	2-8	7-46	7b	3-3
3-13	3f	2-9P	1-47	1g	1-17
2-14	2a2	1-4	5-48	5d	4-3
3-15		2-10P	9-49	9a	5-2
3-16	3a	2-11	1-50	1f	1-18
3-17	3g	2-12P	1-51	1b	1-19
2-18	2d	1-5	5-52	5b	4-4
3-19	3l	2-13	(IV)-53		4-5P
(I)-20		1-6P	10-54	10a	1-20
3-21	3h	2-14	11-55		4-6
3-22	3e	2-15	(I)-56		1-21P
(VI)-22-1	lambdavg2		6-57	6a	1-22P
2-23	2b2	1-7	(V)-58		5-3P
3-24	3d	2-16P	(IV)-59		4-7P
3-25	3m	2-17	4-60	4a	5-4
(VI)-25-1	lambdavg3		8-61	8a	3-4
3-26	3b	2-18P	1-62		1-23P
3-27		2-19	(I)-63		1-24P
2-28	2b1	1-8P	(IV)-64		4-8P
3-29	3c	2-20P	(IV)-65		4-9P
3-30	3o	2-21P	(V)-66		5-5P
3-31	3k	2-22P	(IV)-66-1		
3-32	3i1	2-23P	10-67	10b	1-25P
2-33	2f	1-9	(I)-68		1-26P
2-34		1-10P	4-69	4b	5-6
7-35	7c	3-1P	(I)-70		1-27P

This reaction results in the formation at the 3' of a reactive free hydroxyl group which is then positioned, by the V(D)J recombinase activity, to the phosphodiester bond of the other strand forming a hairpin on the coding segment (Curry et al., 2005; Schatz and Swanson, 2011) (Figure 1.5 B). This process is known as recombination center formation (Agrawal and Schatz, 1997).

The coding ends are then ligated in a process which leads to junctional diversity (Lewis, 1994). First, the endonuclease artemis (Helmink and Sleckman, 2012) opens the coding end hairpins in an "off-center" way (Ma et al., 2005) giving rise to a shorter segment which needs to be elongated in order to be further joined. In particular, for this process palindromic (P) nucleotides are used. P nucleotides are complementary to the ones of the other segment and creates diversity among different recombination events (Lu et al., 2007). This crucial step is one of the two steps, which generate the high variability among LCs. Indeed, the addition of random nucleotides are reflected in the formation of virtually infinite amino acid combination after the translation process. In particular, this results in the amino-acid sequence diversity of the further synthetized proteins. Moreover, the diversity is generated by a second step of random amino acids addition. Indeed, the DNA ends elongated with P nucleotides are aligned and successively prolonged using non-templated (N) nucleotides (up to 20) (Gauss and Lieber, 1996). This second recombination event increases the probability to generate differences in LCs sequences. Remarkably, both the random addition of N

and P nucleotides could generate an off-frame DNA, which can lead to stop codons formation. The process ends with the ligation of the two DNA prolonged ends.

1.2 Amyloidosis

Amyloidosis indicates a group of protein-misfolding diseases characterized by the deposition of insoluble structured fibrils know as amyloid fibrils. Among such diseases, systemic amyloidoses are named the pathologies where the protein misfolds and aggregates elsewhere from where it is produced. Systemic amyloidosis can be grouped in several types: primary amyloidosis, usually with no evidence of preceding or coexisting disease, paraproteinemia, or plasma-cell dyscrasia; secondary amyloidosis, with evidence of coexisting previous chronic inflammatory or infectious conditions; hereditary amyloidosis which are associated to certain genotypes.
On the contrary, localized amyloidosis are characterized by the presence of amyloid deposits in the same organ producing the protein involved (Hazenberg, 2013). This kind of amyloidoses is representative of the 11% of the total cases. In Figure 1.6 is reported the distribution of several types of amyloidosis.

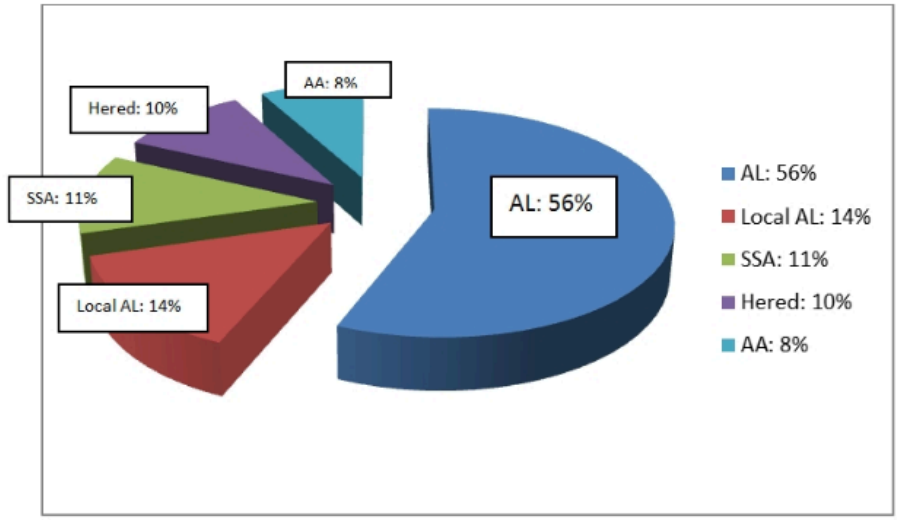


Figure 1.6: Distribution of several amyloidosis types among patients in 2012 (Amyloidosis Patient Information Site).

1.2.1 Amyloid fibrils and amyloidogenesis

Amyloids are insoluble proteinaceous aggregates which acquire a cross ß-sheet enriched folded state (Rambaran and Serpell, 2008). They are found in several pathologies such as Alzheimer disease (AD), diabetes type II and spongiform encephalopathy. Amyloid fibrils structure is studied by several experimental techniques such as X-ray diffraction and electron transmission microscopy (TEM) (Figure 1.7 A). Through these experimental approaches it was possible to asses that amyloid fibrils have a common structure (Blake and Serpell, 1996).

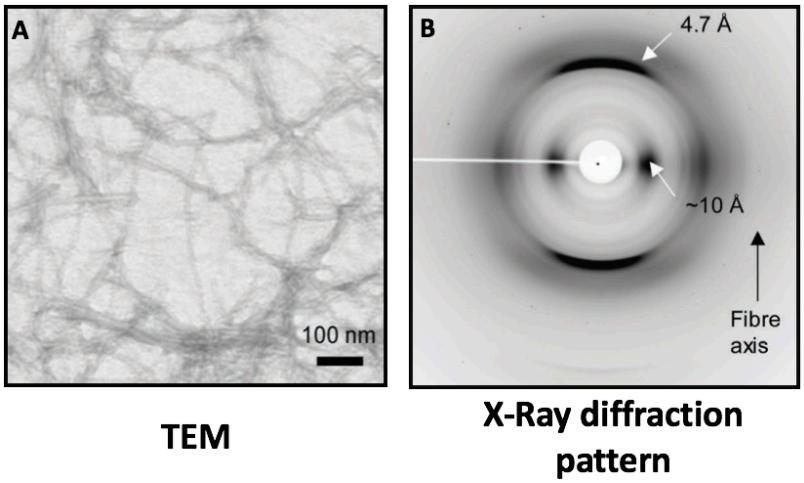


Figure 1.7: A. Amyloid fibrils TEM imaging; **B.** Amyloid fibril X-Ray diffraction pattern where two main reflections (4.7 Å and 10 Å) are highlighted.

Specifically, it consists in one or more filamentous subunits (Figure 1.9 B and C). Each subunit is composed of long polypeptides characterized by a cross-beta sheet quaternary structure, which are perpendicular oriented to the fibril axis (Sipe and Cohen, 2000) (Figure 1.9 C, D and E). In particular, this disposition allows the formation of hydrogen bonds increasing the overall stability of the fibril (Kirschner et al., 1986). The X-ray diffraction pattern present two reflections, the first is sharp and meridional: it is at 4.7 Å resolution (Figure 1.7 B) and it corresponds to the distance between the hydrogen-bonded ß-strands (Figure 1.9 D). The other reflection is strong and more diffuse on the fibril equator (Guijarro et al., 1998) (Figure 1.7 B) and accounts for the distance between facing β-sheets (~10 Å) (Figure 1.9 D) (Rambaran and Serpell, 2008) .

Amyloid fibrils can vary in terms of morphology presenting different width and pitch values. In particular, width values are generally in the order of ~10 nm while the periodic helical twisting (pitch) is highly variable (Figure 1.8)

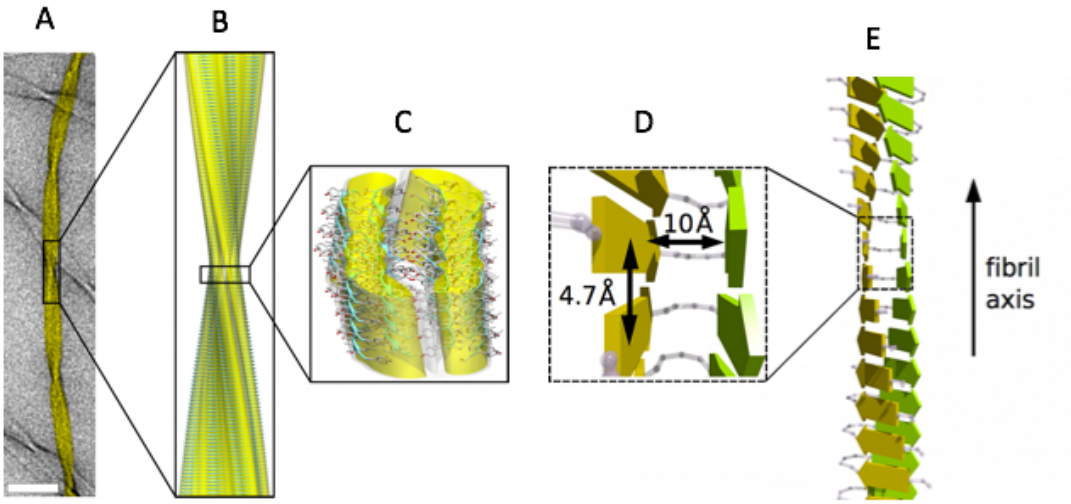


Figure 1.7: (A) Amyloid fibril picture by electron transmission microscopy, **(B)** schematic representation of the fibril, **(C)** Amyloid fibril structural organization, **(D)** Distance between the hydrogen-bonded ß-strands and **(E)** amyloid fibril axis (Fitzpatrick et al., 2013).

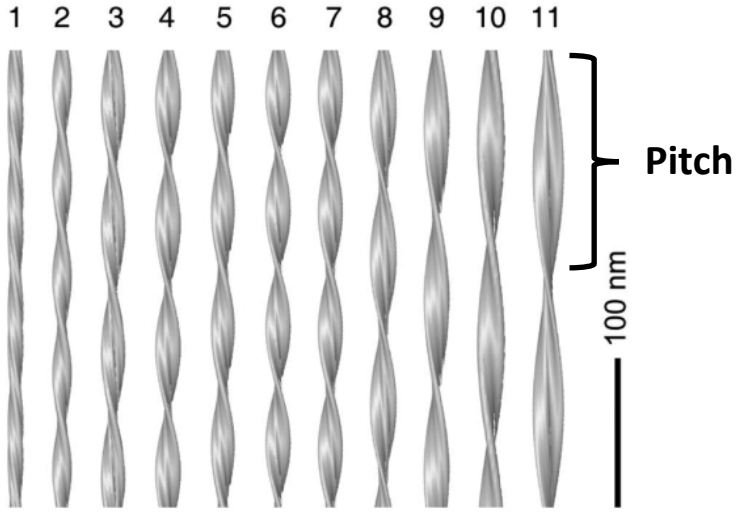


Figure 1.8: Different amyloid fibrils morphologies. From 1 to 11 the pitch decreases (Fändrich et al., 2009).

Amyloid fibrils are the result of the aggregation process known as amyloidogenesis (or fibrillogenesis). This process starts by virtue of protein misfolding. In specific, in the case of

structured proteins, it occurs after an unfolding / partially unfolding event. Misfolding generates protein conformations distinct from the native state, facilitating to protein self-aggregation. Typically, in these aberrant conformations highly "sticky" hydrophobic patches usually buried in the protein core become exposed solvent and mediate monomer-monomer oligomerization (Moreno-Gonzalez and Soto, 2011).

Misfolding can occur because several non-physiologic conditions such as gene mutations, aberrant transcription process, failure of the chaperon machinery, incorrect post-translational modifications and high protein concentration (DeArmond and Prusiner, 1995; Duyckaerts et al., 2008; Matveyenko and Butler, 2006; Soto, 2001). Misfolded proteins form polymers of different size such as oligomers, protofibrils and fibrils (Caughey and Lansbury, 2003). Soluble oligomers are usually arranged in size usually from dimers to 24-mers (Glabe, 2006). To date, it is well accepted the subdivision of soluble oligomers in two principal categories: on-pathway and off-pathway. In particular, on-pathway oligomers are the ones leading to the assembly of mature amyloid fibrils; while off-pathway oligomers remain in equilibrium with the monomeric state of the protein (Lorenzen and Otzen, 2014) (Figure 1.8).

The kinetics of the amyloidogenetic process is usually described as a sigmoidal growth (Ferrone et al., 1985). Basically, it can be divided in three macro-events: lag phase, growth phase and final plateau (Arosio et al., 2015). During the lag phase the microscopic step known as primary nucleation occurs (Figure 1.10 A). This process is aimed to the production of *nuclei* suitable for amyloid growth, starting from soluble monomers. In particular, the *nuclei* are the species characterized by the highest free energy in the fibrillation process (Figure 1.9).

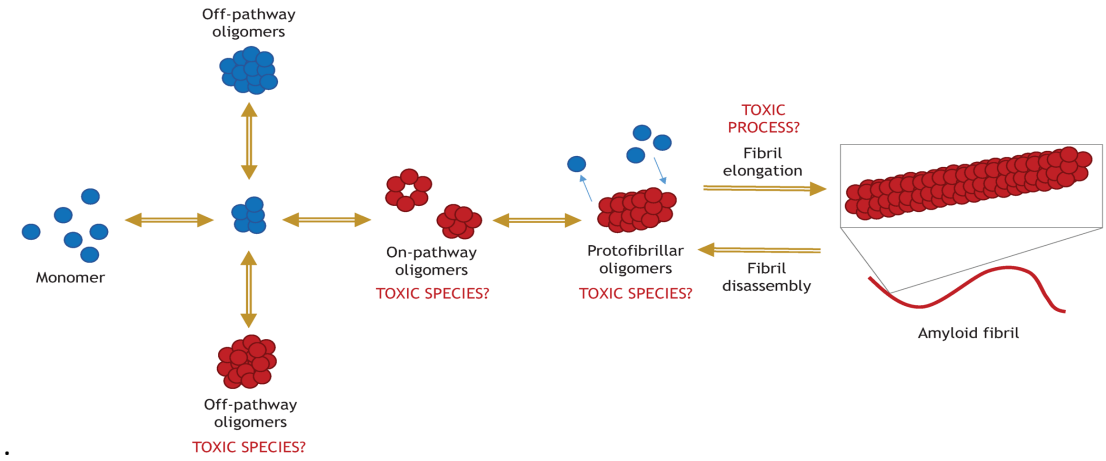


Figure 1.8: Amyloidogenesis. Schematic representation of the amyloid fibril formation with underlined the potential toxic species.

The second step is known as elongation and starts immediately after the primary nucleation, giving rise to the growth phase (Figure 1.10 B). It consists in the addition of monomers to the ends of existing aggregates promoting the growth of oligomers resulting in mature fibrils (Arosio et al., 2015). The elongation rate is usually higher than the primary nucleation. Therefore, the fibrils appear almost immediately after the first nucleation, becoming a catalytic surface for the formation of secondary nuclei (Arosio et al., 2015).

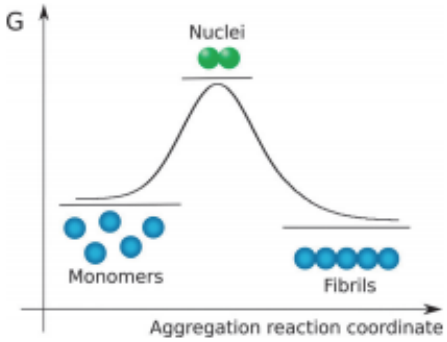


Figure 1.9: Schematic represented the different free energies among the species detectable during the fibrillation process.

Indeed, during the growth phase the secondary nucleation occurs (Castello et al., 2017). This step is the process whereby *nuclei* formation is catalyzed by existing aggregates composed of the same protein (Figure 1.10 C). Therefore secondary nucleation totally depends from the presence of already formed aggregates, opposite to what happens for the primary nucleation (Törnquist et al., 2018). Finally fibrils undergo spontaneous fragmentation, which increase the rate of fibril elongation (Figure 1.10 D).

Over the years, the toxic potential of fibrillation intermediates has been characterized (Figure 1.8). To date, it is accepted that the elements exerting proteo-toxicity are different in different pathologies. For instance, in AD it was demonstrated that the toxicity is related to the amyloid deposits and the circulating soluble oligomer. Instead, in AL it has been observed that the soluble monomeric LCs are toxic *in vivo* (Diomede et al., 2017b) but, to date, there are no evidences of an oligomeric involvement.

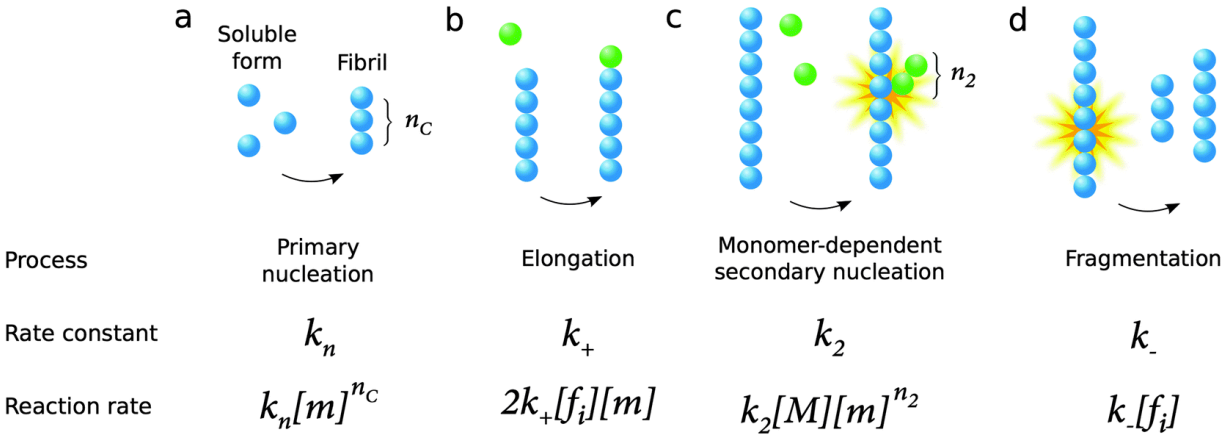


Figure 1.10: (A) primary nucleation from monomers, **(B)** fibril elongation from existing aggregates, **(C)** secondary nucleation from monomers in solution, **(D)** fragmentation (Arosio et al., 2015).

1.2.2 Light chain amyloidosis

As already mentioned above, AL is the most common systemic amyloidosis and it counts for approximately the 70% of all patient suffering from a systemic amyloidosis (Palladini and Merlini, 2016). This pathology is caused by a plasma cell clone that infiltrates into the BM which sets off a devastating multi-organ damage caused by LCs over-production (Merlini and Stone, 2006). Specifically the overexpression is clinically evaluated through the parameter serum free light chain (FLC). Indeed, in pathology the FLC k/ λ ratio is altered (~0.26 mg/L) compared to the normal one (1.85 mg/L). Specifically, the over-produced LCs aggregate and deposit in tissue as amyloid fibers (Merlini, 2017). All organs, with the exception of the central nervous system, can be compromised (Milani et al., 2018) with heart being the most affected one (Milani et al., 2018) (Figure 1.11). It is also worth to know that AL is an aging pathology: indeed, the LCs overexpression leads to the deposit formation over the years (Picken, 2007).

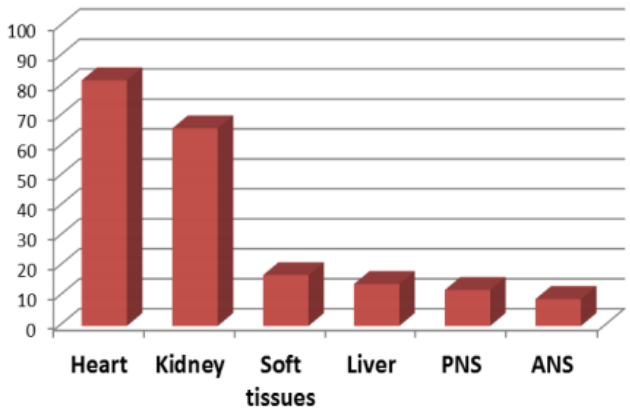
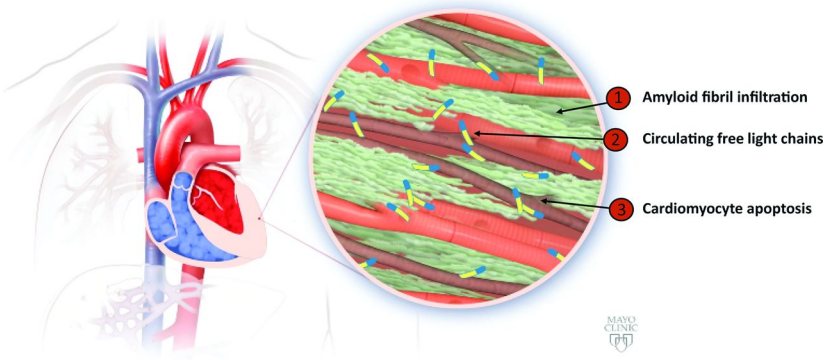


Figure 1.11: Organs involvement in AL. (Merlini, 2017)

The AL etiology is not fully understood but to date several progresses in understanding the biology of plasma cells helped unraveled the molecular pathological mechanisms. In particular, it has been observed that the plasma cell clone is often characterized by several chromosomal abnormalities (Milani et al., 2018). The most representative one is the translocation t(11;14) which is typical of the 50% of patients and due to cyclin D1. Another alteration is the gain of 1q21 which is observed in less than 20% of cases (Bochtler et al., 2008). Interestingly, the phenotypic and copy number alteration of the plasma cell clone in AL is similar to those ones found in MM. However, in AL, the gene expression profile is comparable to the one of normal plasma cells (Paiva et al., 2016). Remarkably, through a genome wide association study (GWAS) it was observed that AL and MM present shared genetic susceptibility. In particular, it was underlined the preferential role of cyclin D1 in driver AL rather than MM. Anyway the most relevant property of such cells is the ability to overproduce LCs, which leads to high soluble protein concentration in the blood flow. In particular, this scenario generates extremely favorable conditions for protein aggregation.

As already explained, all organs can be affected by AL but the tropism remains to be elucidated. In particular, a study published in 2016 reveals the tendency of an organ to be affected, depending on the IGLV clone (Weiss et al., 2016). Other studies (*e.g.,* Abraham et al., 2003; Perfetti et al., 2004) had previously identified IGLV2-14, IGLV6-57, and IGLV3-1 germline genes as the ones contributing in encoding the majority of λ LCs. From these observations, in 2017 It was found that the germline gene IGLV6-57 is commonly found in AL and it is associated with renal involvement, but remains extremely rare in normal B-cells (Kourelis et al., 2017). Moreover, the germline gene IGLV1-44 was found to be linked to predominant cardiac involvement (Perfetti et al., 2012).

Heart failure is typically caused by the alteration of the cell metabolism and oxidative equilibrium and the contraction disfunction. First, the LCs interaction with proteins involved in vitality and metabolism leads to aberrant oxidative balance (Lavatelli et al., 2015) suggesting that LCs may be internalized (Grogan et al., 2017). Moreover, the contraction mechanics are disrupted because of the infiltration of amyloid deposits into the myocardium, which interfere with the cell integrity and cell-cell coupling (Mohty et al., 2013) (Figure 1.12). Specifically, the contraction dysfunction mechanisms are linked to the non-canonical p38α MAPK pathway (Shi et al., 2010) and this molecular signaling promotes the expression of the cardiac stress hormone brain natriuretic peptide (BNP) (Grogan et al., 2017).



Reproduced with permission from the Mayo Clinic

Figure 1.12: Schematic representation of the heart damage due to the presence of amyloid deposits.

BNP secretion is caused by an excessive stretching of the heart walls, which lead to an increase in the ventricular blood volume. It is synthetized as a 134-amino-acids in the pre-pro-hormonal form (preproBNP). This molecule is then converted in the prohormone (proBNP) by a cleavage at its N-terminal signal peptide. Successively, NT-proBNP (the proBNP N-terminal fragment) and in the biological active BNP are produced by a proBNP cleavage between R102 and R103 (Figure 1.13) (Schellenberger et al., 2006). Along this line, BNP and in particular NT-proBNP were extensively characterized (Merlini et al., 2016) and to date, it is routinely used as clinical biomarkers for AL severity of heart involvement. In specific, values of BNP <100 pg/mL and NT-proBNP <300 pg/mL can exclude a cardiac involvement. It is worth to underline that the biological BNP response is also linked to the presence of monomeric LCs, suggesting that they may drive proteo-toxic mechanisms (Merlini et al., 2016) (see results 1.2.2.3).

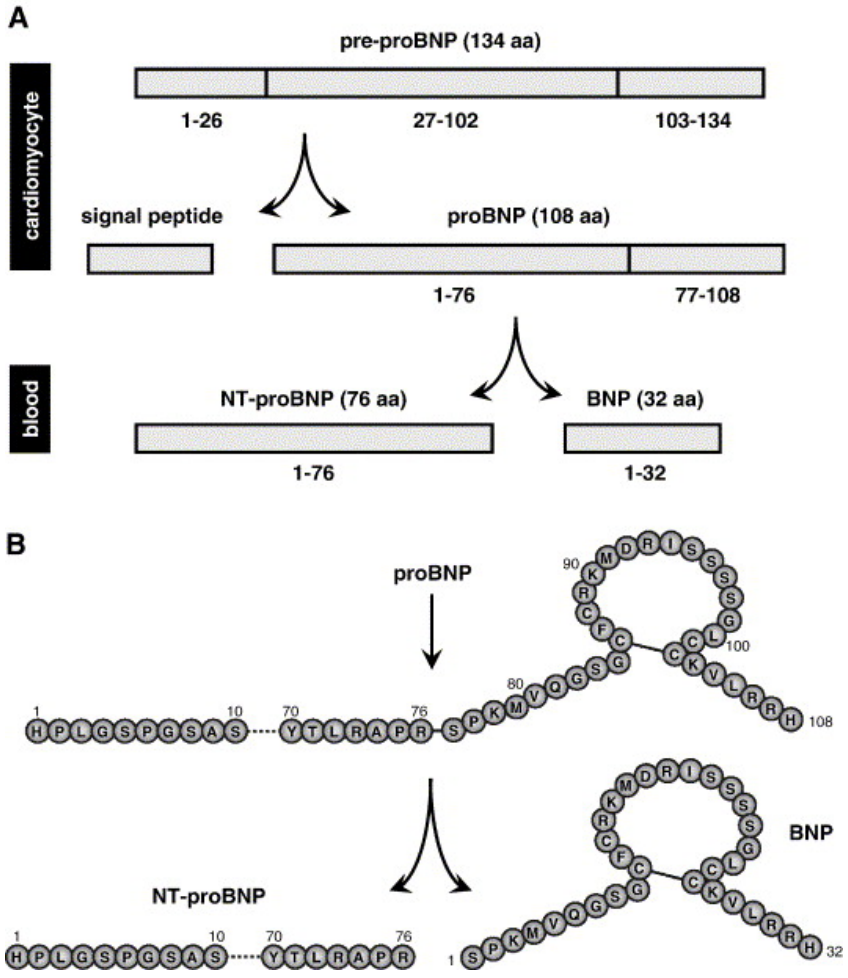


Figure 1.13: (A) BNP biosynthesis.**(B)** detailed on the proBNP cleavage in NT-proBNP and BNP (Jarolim, 2006)

1.2.2.1 Epidemiology, symptoms, diagnosis and therapy

Epidemiology. AL has an incidence of ~10 per million person-years (Falk et al., 1997). Patients are preferential male with a ratio male-female of 3:2 and a mean age at diagnosis of 63 (Baker and Rice, 2012; Quock et al., 2017). In AL the LC λ isotype is the most abundant one found in patients with the k/λ ratio < 0.26 in terms of FLC (Rajkumar, 2016) although in healthy individuals is the reverse (Jones, 1848). However, it must be noted that the pathology could also be caused by the overproduction of k LCs. In those patients, the k/λ ratio is different and higher with respect the physiological one (~ 0.6) (Katzmann et al., 2002).

Symptoms. Symptoms may be very heterogeneous and in the early phases of the pathology can be minor and similar to the ones of other pathologies. By time the disease is becoming progressively more severe on the organ(s) affected. However, there is a common set of symptoms that may have

been present for some time before diagnosis, such as chronic fatigue and weakness (Amyloidosis foundation). The heart involvement is the most life threatening (see introduction 1.2.2).

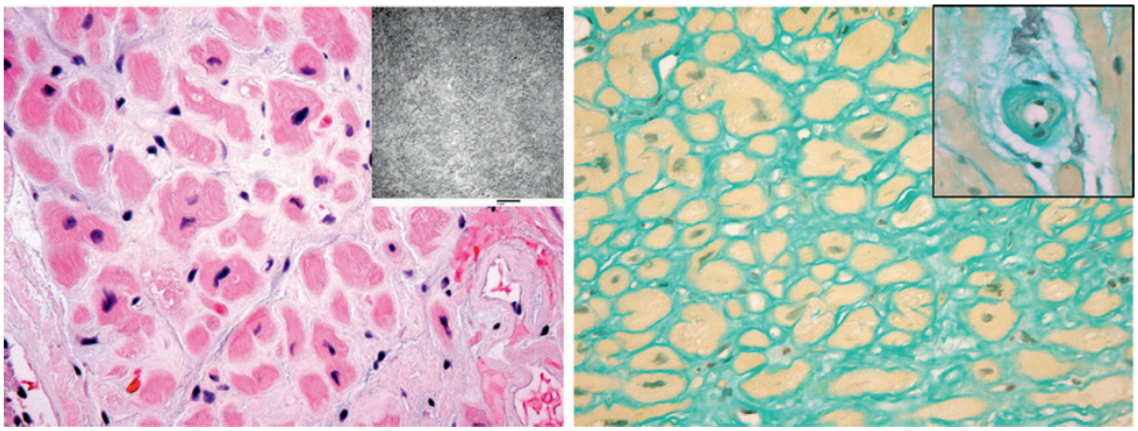


Figure 1.14: Histological sections showing the amyloids infiltrates (Falk et al., 1997)

In particular, deposits lead to the expansion of the extracellular space resulting in the stiffening of cardiac tissue without producing a compensatory dilatation (Figure 1.14). This process results in a restrictive pathophysiology known as "stiff-heart syndrome" (Falk et al., 1997). Specifically, this condition leads the 25-33% of patients to heart failure within six months (Dispenzieri et al., 2004). Another cardiac symptom is the arrhythmia, which is present when amyloids affect the electrical heart system. One very typical condition for AL patients is proteinuria, which is a very useful for the diagnosis. The urine-derived LCs are known as Bence-Jones (BJ) LCs (Jones, 1848).

Diagnosis. The AL diagnosis is usually done in several ways but the main critical step is the clinical suspicion (*e.g.,* altered BNP and pro-BNP values). The 90% of the cases are identified through the combination of serum protein electrophoresis with immunofixation electrophoresis (SPEP/IFE) and a 24-hours urine collection monitored by electrophoresis. In 2005, the addiction of the quantification of the FLC increased efficiency of diagnosis since this clinical parameter is abnormal in 99% patients (Gertz et al., 2005). Although these tests are effective tools for the AL diagnosis, tissue biopsy confirmation is required: in particular, abdominal fat pad aspiration (sensitivity around 85%), rectal biopsy (75-85%), and bone marrow biopsy (around 50%) (van Gameren et al., 2006; Schatz and Swanson, 2011).

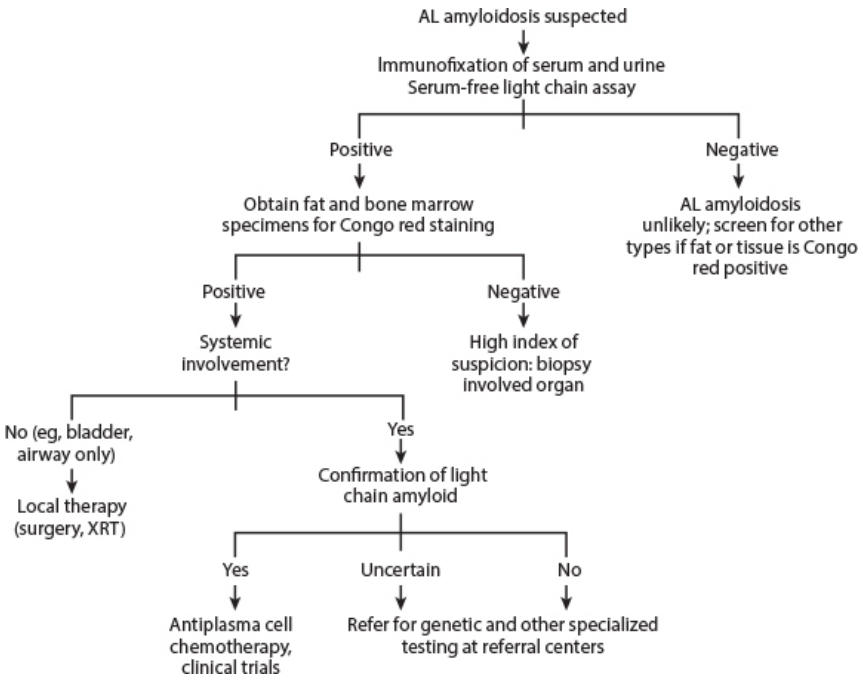


Figure 1.15: AL diagnosis criteria (Syed A. Abutalib and Anita D'Souza).

Therapy. AL therapy is a challenge for physicians since both the bone marrow disorder and the involved organ disfunctions have to be taken in account (Milani et al., 2018). The specific organ treatment should be referred to specialized centers.

The amyloid clone is treated with chemotherapy aimed to its neutralization (Muchtar et al., 2017). The Amyloidosis foundation reports the most used drugs to be Melphalan, Cyclophosphamide and Bendamustine. In patients with a severe cardiac involvement it is not possible the use high chemotherapeutical agents doses since their heart do not withstand the therapy (Merlini et al., 2016). However to date, the ones above are the election therapy for high-risk patients. Low-risk patients, who represent the 15% of the cases, are good candidates for autologous stem cell transplantation (ASCT). This approach is associated with a higher risk of early mortality with respect its use on MM patients. However, the patients selection criteria nowadays are improved leading to a decrease in the risk of mortality over time (D'Souza et al., 2015).

Immunomodulatory drugs (ID) and proteasome inhibitors are the other drug-classes used in AL treatment. However, the choice of a given drug depends on the organs involved. For instance, the proteasome inhibitor carfilzomib leads to a hematological response of 63% but it exerts cardiac toxicity hence it not suitable for patients with cardiac involvement. Among the ID, ixazomib is one of the most promising one. This drug is highly active in pre-treated patients therefore it seems

suitable for drugs combinations. For these reasons ixazomib is now in a randomized phase III trial in relapsed / refractory patients (Milani et al., 2018).

1.2.3 Biochemical and structural characterization of the AL deposition.

The biophysical mechanisms of proteotoxic LCs in AL amyloidosis remains to be fully elucidated. Indeed, to date, the comprehension of the molecular pathophysiology misses some crucial aspects. First, it is not known why only a subset of LCs aggregate. To address this important point, several reports considered the aminoacidic sequence (*e.g.,* Dwulet et al., 1985; Hurle et al., 1994). In specific, in literature is indicated that it may have a role in the aggregation propensity but, to date, a direct correlation between sequence and amyloidogenicity has not yet been assessed. Another important point that needs to be highlighted regards the species that aggregates. Indeed, although several reports identified VL as being the main constituent of fibrils, amyloid can also comprise CL and FL LCs In other words there is no consensus on which species aggregate and which exert the tissue toxicity. In this regard, a recent study by Marin-Argany et al., suggested that LC fragments are toxic, even at low concentrations. They can aggregate and recruit soluble proteins, resulting in elongation of amyloid fibrils and cellular toxicity (Marin-Argany et al., 2016).
All these opened-questions underline the high molecular and biophysical complexity of the LCs aggregation propensity. Herein, it will be discussed the studies conducted on VL and on the FL LCs.

1.2.3.1 VL structural and biophysical studies

As mentioned above, for several reasons VL has been considered for long time as the main component of the amyloid deposits (Buxbaum, 1992), thus, the majority of the structural and biophysical studies focus on VL domains and their amyloid propensity. First, it was established that a low VL thermodynamic stability is associated with an increased propensity to form amyloid fibrils (Baden et al., 2008a, 2008b; Wall et al., 1999). This properties has been correlated to the production of partially unfolded states which may necessary start the aggregation reaction (Wetzel, 1997). Indeed, when they populate amyloidogenic experimental conditions, the fibrillogenesis process is accelerated (*e.g.,* Chiti and Dobson, 2006). This evidence has been further confirmed by other studies. Importantly, it was noticed that unfolded or native LCs conformations do not lead to fibril formation. Indeed, variants with very high or very low thermal stability that is they are in fully folded or totally unfolded states do not produce fibrils (M. Blancas-Mejia et al., 2018). Moreover, Khurana

and coworkers reported that native-like intermediates, generated at relatively low pH values (4.0 - 6.0), resulted to be prone to generate amorphous aggregates. While, relatively unfolded but compact intermediates, observed at pH values below 3.0, are prone to generate amyloid fibrils.
These studies has been also structurally supported. Indeed, Eisenberg and coworkers proposed three models for the VL fibrillogenesis based on the VL-VL interface and the oligomerization state (Figure 1.16). In particular, these models were named: canonical dimer model, non-canonical dimer model and monomer model.

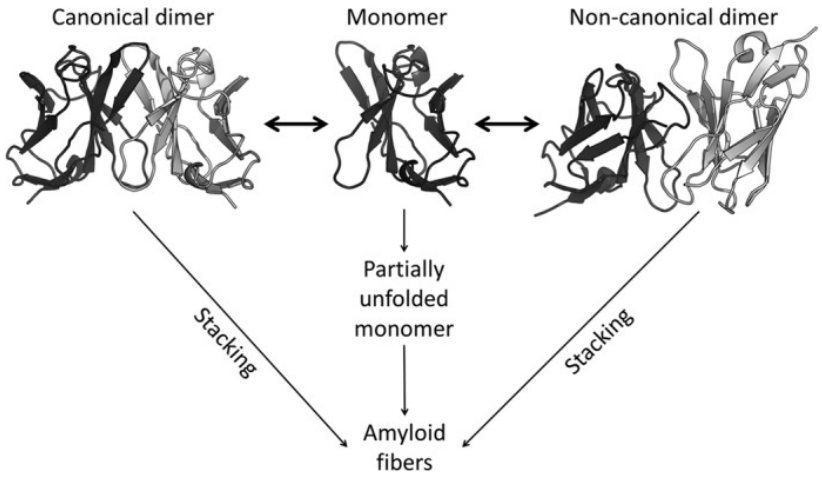


Figure 1.16: Vl-based fibrillogenesis model depending on the VL-VL interface (Eisenberg et al., 2014)

The first one assumes that the amyloid formation is reversible. In the second one the VL-VL dimers quaternary structure is altered but maintains the hydrophobic environment of the cavity in correspondence of the VL-VL interface (Brumshtein et al., 2014). While, the third one considers soluble monomeric Vls. Interestingly, they highlighted that VL-VL dimer, especially the non-canonical one, may be partially unfolded and therefore destabilizing mutations would lead the VL to assume an amyloid-prone tertiary structure while stabilizing mutations would inhibit the fibrillogenesis (Brumshtein et al., 2014; Edmundson et al., 1976). Then, to stress the role of partially unfolded states, it has to be considered that amyloid fibrils *in vitro* can also been produced by monomeric Vl. Indeed, in Brumshtein et al., it is reported that soluble VL-VL dimers dissociate in monomers which partially unfold leading to thermodynamic destabilization which results in amyloid fibrils formation (Qin et al., 2007). Taken together, these data support the hypothesis that the ordered self-assembly of partially folded species are crucial as trigger of amyloidogenesis becoming

the soluble precursors of fibrils (*e.g.,* Chiti and Dobson, 2006; Ramirez-Alvarado et al., 2000 Khurana et al., 2001).

These studies are particularly interesting, and stressed by other reports (*e.g.,* Alvarado, 2018), because the stabilization of these partially folded species may results in the identification of molecules with anti-aggregation properties. Indeed, on this rationale, Eisenberg and colleagues identified ligands that inhibit amyloid formation *in* vitro by stabilizing the VL-VL dimer (Brumshtein et al. 2015).

Another interesting aspect that was considered over the years is the role of sequence variability. Indeed, VL is the LC region that accumulates somatic hypermutations. This fact is stressed by the presence of CDRs, that present the highest sequence variability in order to achieve high-affinity antigen binding.

In 1985 a detailed primary structure analysis on a big number of VL sequences revealed that two residues insertions (positions 68 and 69) may be linked with LC aggregation propensity (Dwulet et al., 1985). Moreover, In 1994, Wetzel and coworkers reported that in amyloidogenic VL in specific positions given amino acids are found at markedly high frequency (Hurle et al., 1994). More recently, it has been reported that the amino-acid sequence has a structural role in the VL-VL interface. Indeed, It was underlined that LCs populate at least 3 different dimer interfaces in a dynamic equilibrium that depends on the specific amino acid substitutions in the dimer interface residues (M. Blancas-Mejia et al., 2018).

Finally, in literature was inspected the role of co-factors in facilitating protein aggregation. For instance, glycosaminoglycans (GAGs), which are negatively charged, unbranched, long hetero-polysaccharides (M. Blancas-Mejia et al., 2018) were found associated to amyloid fibrils (Ramirez-Alvarado, 2012). In particular, their charge is an important factor which contributes to accelerate amyloid formation *in vitro* (Martin and Ramirez-Alvarado, 2011).

1.2.3.2 FL LCs: biochemical studies

As mentioned above, VL plays a crucial role in amyloid formation through its intrinsic instability and its consequent tendency to form partially unfolded states. However, it must be considered that isolated soluble Vls are not detectable *in vivo* (Lavatelli et al., 2011) and the relevance of studying VL domains in the context of the disease is under debate.

Moreover, over the years, it has been demonstrated that not only VL populates amyloid deposits. Indeed, mass spectrometry analysis of biopsy tissues reported the presence of the FL LC in patients'

samples (Lavatelli et al., 2008). Moreover, FL LCs exert cardiotoxicity also without the production of fibrils (*e.g.,* Diomede et al., 2014; Malpasso et al., 2013), especially in the case of AL with cardiac involvement. Remarkably, Merlini and coworkers underlined a solid correlation between the blood concentration of soluble LCs and the severity of heart symptoms (Milani et al., 2018). Indeed, toxic LCs injected *in vivo* in the heart of wt mice lead to end-diastolic pressure (Palladini et al., 2010). Recently Diomede and co-workers established a model to recapitulate LC toxicity *in vivo*: *C. elegans* fed by solution of LCs shows a impairment in the larynx pumping rate in presence of cardiotoxic LCs and not with non-cardiotoxic or M LCs (ref). This well describes what seen in patients. Moreover, they also identified the role of Cu^{2+} in enhancing the FL LCs toxicity by altering the RED-OX balance (Diomede et al., 2017b). In details, it was observed that the LCs and Cu^{2+} co-administration results in a higher toxic effect rather than the LCs alone, opening new routes for the understanding of LCs toxic mechanisms.

Even though the identification of molecular species which plays the toxicity trigger remains a crucial issue to be elucidated, all these observations strongly highlight the importance of characterizing FL LCs being the relevant pathologic molecular species.

Biophysical characterization of FL LCs aggregation propensity. Considering the overall stability of Vl, CL or FL LC, CL is the one characterized to be the highest thermal stable and the VL the less one (Klimtchuk et al., 2010). Consequently, in a FL LC, CL plays the crucial role to stabilize VL (and VL to destabilize CL through intermolecular interactions. In particular, Klimtchuk *et al*., underlines that the VL lower thermodynamic stability correlates with the lower kinetic stability of the FL protein. This implies that, in FL LCs, increasing the population of the unfolded VL region destabilizes the CL region, possibly via the direct interactions between these partially unfolded regions facilitated by their linkage via the joining region. These interdomain interactions result in a FL LC conformational shift toward partially unfolded states that are prone to irreversible aggregation (ref?).

The role of CL domains in stabilizing VL and the whole LC molecules is also stressed by Kelly and coworkers. They observed that amyloidogenic FL LCs unfolds faster than non-amyloidogenic LCs ones. However, they established that this instability is not sufficient by itself to cause amyloid formation indeed they did not observe fibril formation from FL LCs. Thus it has been hypothesized that when covalently bound to VL the CL domain prevents aggregation, through a mechanism that remains to be determined (Morgan and Kelly, 2016).

The lower VL overall stability has also been observed in other studies (*e.g.,* Blancas-Mejía et al., 2017a). In particular, this domain resulted not able to be completely refolded. Indeed, Blancas-Mejía et al. noted that the refolding of FL LCs consists in 2-step transitions while the unfolding present a 3-step unfolding transition. These results suggested that one of the two domains does not completely refold (possibly the Vl) giving rise to a strong kinetic control of the LCs unfolding process. Together, these findings revealed that the FL protein is likely more prone to form amyloid fibrils through partial unfolded states dependently to the kinetic barrier which counterbalance the aggregation propensity. Indeed, in this model the rate of FL LC aggregation is dependent to the height of this kinetic barrier, which is partially regulated by the VL thermodynamic stability. The role of kinetic stability was also considered by Kelly and coworkers: this parameter was shown to influence the FL LCs propensity to form amyloid (Morgan and Kelly, 2016). Their findings underlined that amyloidogenic FL LCs are kinetically unstable in comparison to the non-amyloidogenic ones. In addition, they also underlined endoproteolytic release of fragments from the FL LC as another important aspect in aggregation. Indeed, they observed that amyloidogenic low kinetically stable FL LCs can be readily cleaved into their component domains by proteases, releasing the Vl, which is, in isolation, highly amyloidogenic.

In summary, the above observations stress the relevance of a research project that by biophysical and structural approaches aims to more systematically characterize the specific properties of amyloidogenic and toxic LCs. Given the their relevance *in vivo*, FL proteins should be used rather than isolated VL and in order to take in account the exceptional sequence variability which characterizes this disease several sequence diverse LCs should be taken in account.

By identifying the molecular bases of LC aggregation propensity and toxicity, the long-term aim of such project will be to set the bases for designing strategy to control and ultimately abrogate LC amyloid aggregation.

2 Results

This is a comprehensive study aimed to investigate possible molecular mechanisms driving cardiotoxic LCs to protein aggregation and *in vivo* toxicity. Herein, we considered two interesting points: i) the LCs biophysics and structure and ii) the effect of copper (Cu^{2+}) on cardiotoxic LCs. In order to reach this aim, we focused our studies on several traits by analyzing fold stability, hydrophobicity, protein flexibility, molecular interactions and 3D structures, all properties, which can play a role in protein misfolding.

2.1 Evaluation of cardiotoxic LCs discrepancy to non-cardiotoxic: study set up

The cardiotoxic LCs biophysics and structures have been evaluated in comparison with the non-cardiotoxic LCs. In particular, we analyzed the possible differences, which may account for different aggregation propensity. To this intent, given the high sequence variability, we considered a pool of thirteen proteins divided in two sets: one set is composed of eight LCs isolated from AL patients with severe cardiac involvement and they are referred as 'H' LCs (H3, H6, H7, H9, H10, H15, H16 and H18). The other set ('M' LCs) were non-toxic and isolated from patients suffering of multiple myeloma (MM) (M2, M7, M8, M9 and M10). In specific, as reported in the introduction, H LCs are overexpressed in blood resulting in organ toxicity and amyloid deposition. On the contrary, in MM patients non-toxic LCs are overexpressed but no toxicity and no amyloid formation occur, making M LCs a proper control. Anyway, MM patients producing these proteins were carefully screened and monitored over the years, to rule out the presence of amyloid deposits.

In details, our experimental activities were performed on the native LCs since it is the starting species leading to amyloidogenicity. Moreover, only FL LCs were considered because their concentration correlates with the severity of heart symptoms. It should be noted that a pool composed of a such high number of protein represents an innovation in the characterization of AL proteins. Moreover, all the proteins belong to different plasma-cell families in order to increase the sequence variability (Table 2.1) and to avoid to obtain results specific for given LC families. This pool, large in biochemical terms, allows considering our data representative of general LC traits.

The proteins were purified from two different sources: i) patients' urine (BJ: H9, H10, H15, H16, H18, M2, M7, M8, M9, M10) ii) *Escherichia coli* (recombinant, rec: H3, H6, H7, H9, M7, M8, M9) (Table 2.1).

Table 2.1: LCs ID; Source of purification (BJ or recombinant); Germline gene and family on brackets.

LC ID	REC/BJ	Germline
H3	REC	1c (IGLV1-44)
H6	REC	1b (IGLV1-51)
H7	REC	1b (IGLV1-51)
H9	REC/BJ	2c (IGLV2-8)
H10	BJ	1a (IGLV1-36)
H15	BJ	6a (IGLV6-57)
H16	BJ	2a2 (IGLV2-14)
H18	BJ	3l (IGLV3-19)
M2	BJ	2b2 (IGLV2-23)
M7	BJ	3l (IGLV3-19)
M8	REC/BJ	2b2 (IGLV2-23)
M9	REC/BJ	2b2 (IGLV2-23)
M10	BJ	2a2 (IGLV2-14)

2.1.1 Sequence analysis

The LCs amino acidic sequences significantly vary because of their genetic recombination. Specifically, the sequence variability allows each LC to interact with a specific antigen (Ag). Indeed, the VL sequence identity referred to the LCs pool here presented is 25% (as determined by MUltiple Sequence Comparison by Log- Expectation, MUSCLE) (Edgar, Robert C., 2004). As expected, the highest degree of variation was found in the three CDRs regions where there is no sequence conservation. On the contrary, CL sequence is shown as highly conserved among all the presented LCs (identity 95%, MUSCLE alignment) with the exception of six conservative mutations which do not alter the overall fold (Figure 2.1).

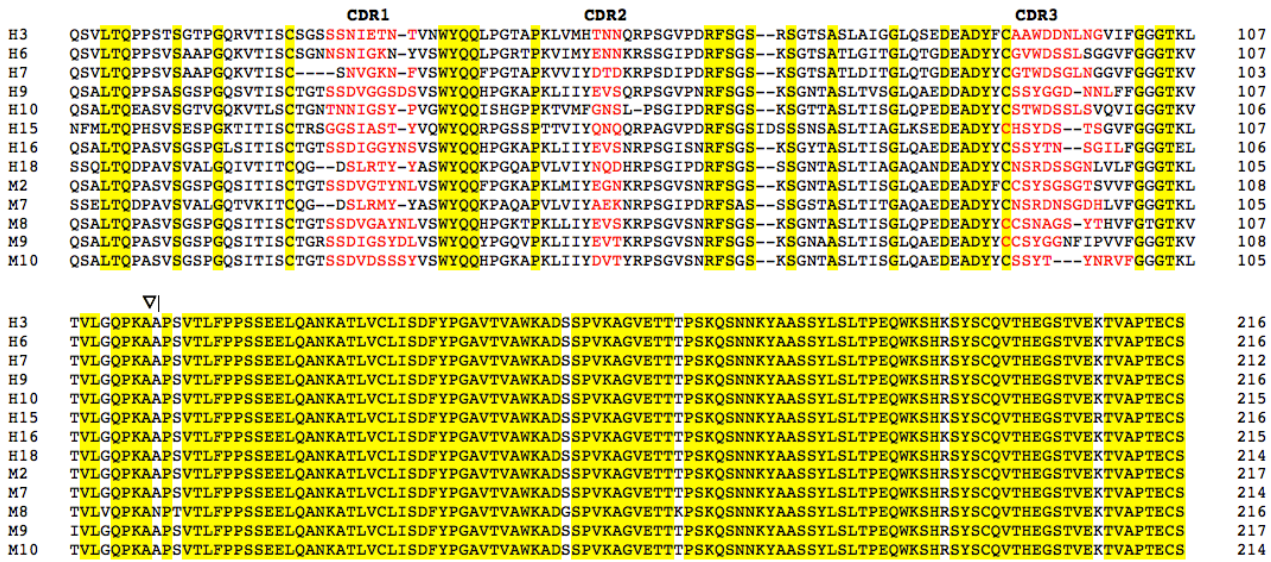


Figure 2.1: Multialinegment of the thirteen LCs: 8 H LCs (H3, H6, H7, H9, H10, H15, H16, H18) AND 5 M LCs (M2, M7, M8, M9, M10). ∇: End of the Vl.

2.1.2 Purification protocol and optimization

With the aim to increase the amount of protein suitable for our studies, we set expression and purification protocols as recombinant proteins for H3, H6, H7, H9, M7, M8 and M9. Indeed, BJ LCs (herein: H9, H10, H15, H16, H18 M2, M7, M8 and M9 purified at the Amyloidosis Research and Treatment Center, San Matteo hospital, Pavia) can be obtained in small amount and for limited amount of time.

The LCs genes were cloned in the PCR-TOPO-2.0 cloning plasmid and successively in the pASK-IBA33 plus (H3, H6, H7 and M9, previously cloned by collaborators in Pavia) or pET21b (H9, M7 and M8) for the *E. coli* expression. LCs cloned in pASK-IBA33 plus were expressed and purified according to previously published protocols (Rognoni et al., 2013). In this case, proteins were produced using the BL21 *E. coli* strain. The overexpression was obtain by the induction with 0.43 µM anhydrotetracycline for four hours at 37 °C.

On the contrary, the expression conditions for the LCs cloned in pET21b were set up from scratch. First, a multi expression test was performed. Different *E. coli* strains (BL21, Rosetta, Rosetta-gami, origami and shuffle) were tested, with the BL21 resulting the most efficient one. We used isopropyl-β-D-1-tiogalattopiranoside (IPTG) as protein overexpression induction system. In particular, the best induction conditions found were 0.5 mM IPTG for four hours at 37°C.

After the expression, LCs were always detected as inclusion bodies in the insoluble *E. coli* fraction. We refolded LCs by dilution in presence of oxidized and reduced glutathione according to their isoelectric point (pI) since the first chromatographic step consisted in an ionic exchange. Indeed, the pI must be considered for the choice between anionic or cationic exchange.

Specifically, H3, H7, H9, M7 (pI ~6.0) and M9 (pI 6.8) were purified by an anionic exchange chromatography in Tris 10mM pH 8.0, while H6 and M8 (pI ~8.0) by a cationic one in 10mM sodium acetate pH 6.0. In each case LCs eluted at a salt concentration in a range between 160 mM and 240 mM sodium chloride (NaCl). The second chromatographic step consisted in a size exclusion chromatography (SEC). This step was carried out in order to: i) eliminate aggregated protein and ii) lower NaCl concentration. LCs eluted in 50 mM sodium phosphate (NaPi) pH 7.4 or 10 mM tris pH 7.4 depending on the requirement of the different techniques. Interestingly, SEC profiles for same LCs were not always comparable both in terms of elution profiles and of yield. First, the differences in the elution profile are probably due to a low refolding efficiency. Indeed, protein refolding is a complicated process difficult to be reproduced. In particular, for instance it is responsible of the right secondary structures acquisition or the right disulfide bridges (S-S) formation. Moreover, the different yields could be accounted for the refolding efficiency too, but they also depend on the different starting material amount used for different purifications.

2.2 Biophysical characterization

The LCs pull was biophysical characterized with the aim to screen different properties of their native state. Herein, we considered fold stability, hydrophobicity and flexibility/dynamics.

2.2.1 Fold stability

With the objective to assess protein stability samples were tested using Circular Dichroism (CD) spectroscopy and fluorescence.

First, LCs were analyzed by CD in Far-UV region (195-280 nm). This technique provides information about the secondary structures content. LCs spectra were collected and compared, showing that LCs fold in solution is comparable (Oberti et al., Figure S1). In particular, the overall shape is conserved and moreover each spectrum present a minimum at 218 nm, which is specific for protein

enriched in β structures. All spectra are highly similar with the exception of H7, which displays a spectrum not comparable with any other proteins. However, from other experiments (fold stability and structure determination), this LC resulted to be in its properly folded native state. Therefore the reasons why it presents a different CD spectrum remains unclear. From these data we assessed LCs fold in solution to be highly similar and no specific features to be correlated with aggregation propensity could be identified.

Successively, the samples melting temperatures (Tm) were assessed by monitoring the CD signal at 202 nm, while the temperature was increased up to 80 °C (Figure 2.2 A). Our data showed H LCs to be less thermal stable than M LCs (Table 2.2). In particular, the averaged H LCs Tm values (53.6 ± 1.5 °C) is lower than the average of the ones measured for M LCs (62.8 ± 7.1 °C) (Figure 2.12). Specifically, H LCs Tm values cover a range between 52.9°C to 55.6°C with H6 and H15 being the least and the most stable, respectively. Instead, M LCs Tms range between 54.7 °C and 72.6 °C with M7 being the least stable and M8 the most one. Among the highly stable M LCs, M7 and M10 represented two exceptions since their Tm values (54.7°C and 58.0 °C respectively) were similar to the ones measured for H LCs. From these experiments we concluded that H LCs typically possess a lower fold stability than M LCs even though the correlation is not 100%.

LCs have also been evaluated by Near-UV CD (250-350 nm). This technique provides information on protein tertiary structure by monitoring at the circular polarized light absorption of the aromatic amino acids. As our findings report, the majority of the LCs tested (H6, H9, H10, H15, H16, M2, H18, M2, M7, M8 and M10) present comparable spectra underlying that the overall tertiary structure is conserved (Oberti et al., Figure S1). However, since this technique considers aromatic amino acids, it provides only information in their local environment. Indeed H3, H7 and M9 could not be compared due to differences in their aromatic amino acids chemical environment.

We also measured the Tms by using Near-UV CD. As reported in Table 2.2 H3, H7, H9, H10, H16, H18, M7, and M9 thermal unfolding monitored by near-UV analyses confirmed the Tm measured by the Far-UV CD. Anyway, the concentration required for this technique (1mg/mL) may lead to protein aggregation (Figure 2.2 B). This prevented the Tm assessment for H6, H15, M2, and M10.

From these analyses we concluded that the loss of the secondary and the tertiary structure are simultaneous processes, indicating that LC dimers unfold through a cooperative process. Moreover, we assessed that although it seems to have a role, the aggregation propensity cannot be interpreted by the fold stability itself.

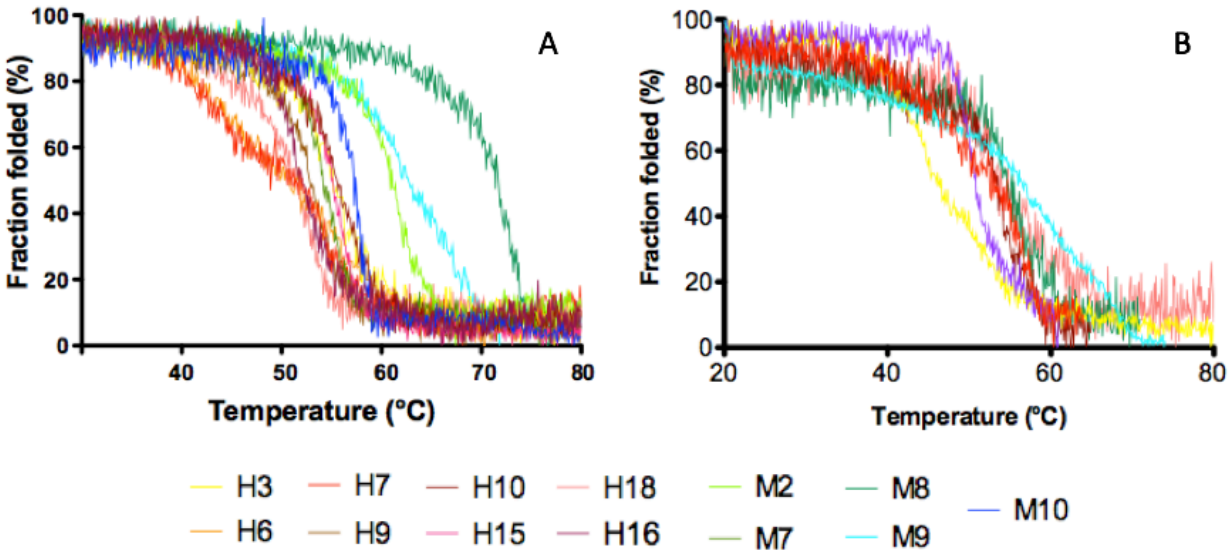


Figure 2.2: CD thermal ramps of H and M LCs: Far-UV **(A)** and Near-UV **(B)**. Labels report the LC ID and the family where they belong.

2.2.2 Surface Hydrophobicity and stability

Aiming to investigate on LCs hydrophobicity we carried out both computational and experimental methods (fluorescence).

First, the hydrophobic amino acids content was computationally measured revealing that it is very comparable between H and M LCs sequences (~23%). Then, surface hydrophobicity was experimentally evaluated monitoring 8-anilinonaphthalene-1-sulfonic acid (ANS) fluorescence. This technique provides information about hydrophobicity since ANS fluorescence increases when the compound interacts with to hydrophobic regions. We titrated LCs using ANS up to 100 µM. Remarkably, we found the titration spectra to be highly comparable among all the LCs (Figure 2.3). In particular, the increment in ANS fluorescence is not significant and therefore these data indicate that none of the LCs display hydrophobic patches on their surfaces.

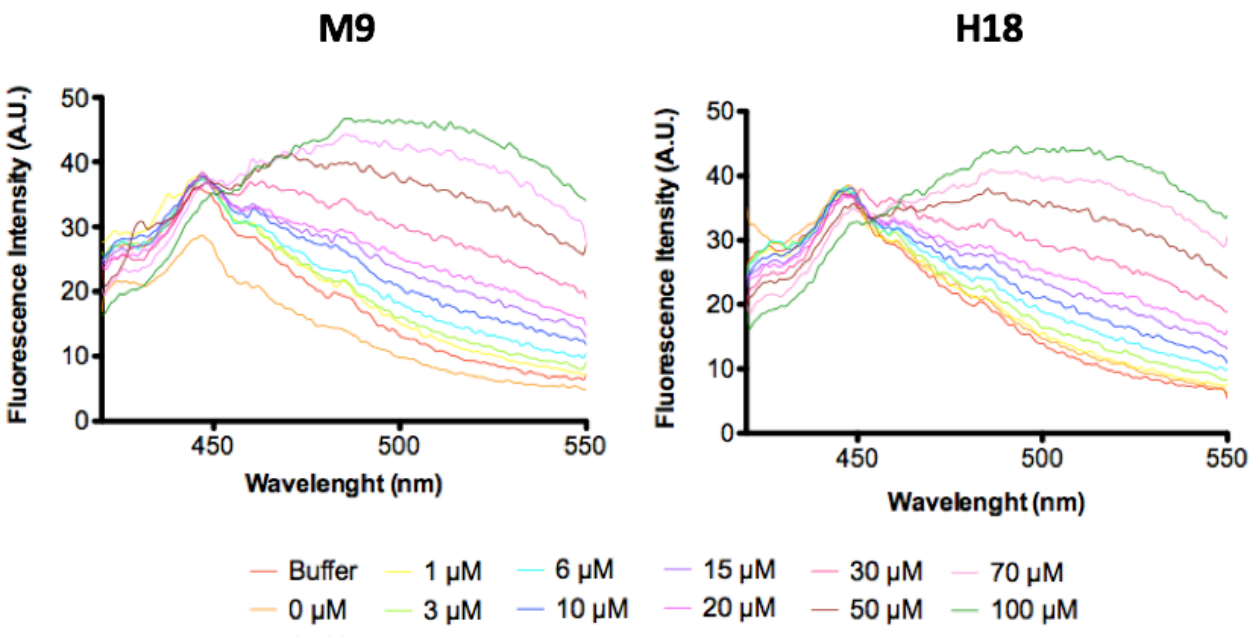


Figure 2.3: Representative ANS titration spectra. All the LCs presented the same spectra' shape and therefore here only H18 and M9 are shown. Labels report ANS concentration.

Then, temperature ramps were recorded in presence of 100 µM ANS for each LC (Figure 2.4). This allows monitoring protein unfolding by the increase in ANS fluorescence upon the exposure of the hydrophobic residues present in the protein core. These findings confirmed our assessments on protein stability. In particular, H LCs tend to be less stable than M LCs and therefore they can quickly expose their hydrophobic core. It must be noted that the unfolding process presented a single denaturation event suggesting that the two domains do not present independent unfolding events. This fact agrees with our results obtained by CD although for some LCs (H3, H6, H7, M9) a two-step unfolding was observed. However, in these cases, we assessed that the exposure of the hydrophobic core corresponds to the second unfolding step recorded by CD. Therefore, further experiments will be performed to clarify the presence of the other step.

Interestingly, the onset of ANS signal correlates with the Tm obtained by CD rather the inflection point of the curve. This fact suggested that the hydrophobic core is exposed only at an advanced state of unfolding leading to an overestimation in ANS thermal unfolding.

Taken together, these data revealed that hydrophobicity is highly similar in all tested LCs. Therefore, it could not be correlated to protein aggregation.

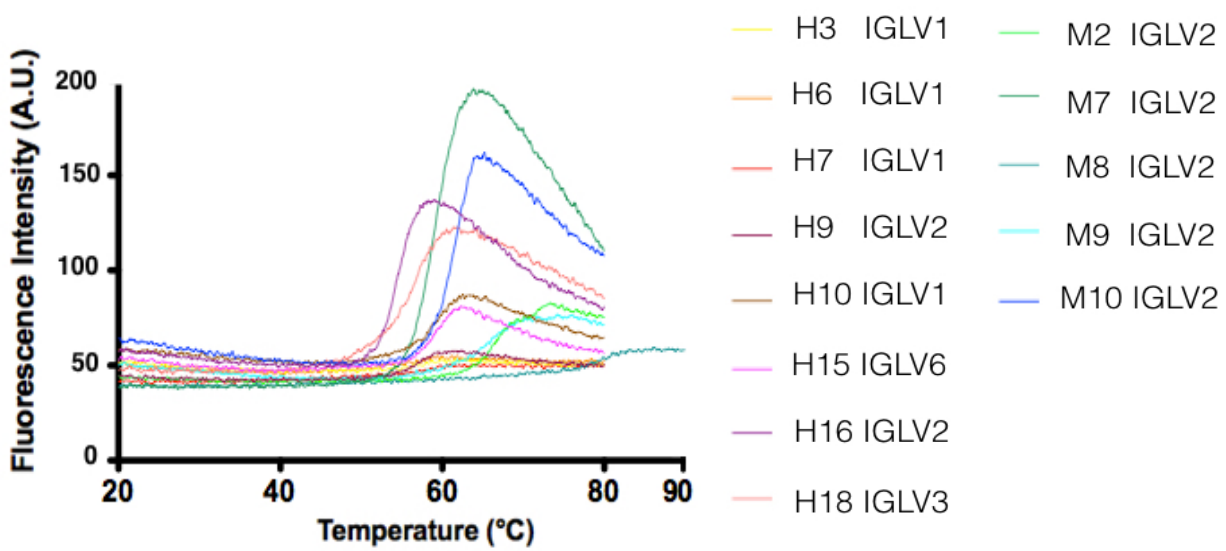


Figure 2.4: Thermal stability assessed monitoring ANS fluorescence of H and M LCs. Labels report the LC ID and the family where they belong.

Table 2.2: LCs Tm recorded in the far-UV and in the near-UV regions for each LC; Temperature values corresponding to the ANS fluorescence onset for each LC.

LC ID	Tm Far-UV (°C)	Tm Near-UV (°C)	Onset ANS fluorescence (°C)
H3	54.4	53.1	54.8
H6	52.9	-	55.3
H7	54.7	56.2	54.8
H9	52.9	54.0	55.8
H10	54.9	54.3	58.3
H15	55.6	-	54.7
H16	51.3	51.9	54.5
H18	52.4	51.1	51.5
M2	61.5	-	63.9
M7	54.7	55.3	58.8
M8	72.6	-	73.5
M9	67.0	67.7	62.8
M10	58.0	-	56.1

2.2.3 Protein flexibility

Then, protein flexibility was considered by limited proteolysis analyses. This molecular property can be studied through this approach since it is correlated to the kinetics of proteolysis. In specific, highly flexible proteins are characterized by fast kinetics.

First, we carried out experiments with trypsin, a protease with cleavage specificity in correspondence of Arg and Lys residues. Our findings showed that the kinetics of proteolysis is faster in the case of H LCs as reported in Figure 2.5 and 2.6. The degree of flexibility was measured, in function of time, as the uncleaved protein fraction evaluated by densitometry after SDS-PAGE analyses. In particular, 50% of the total M LCs amount was still uncleaved after 3 hours of trypsin treatment while H LCs uncleaved fraction was lower than 50%. We divided H LCs in two subgroups: the one composed by H10, H15, and H18 with an uncleaved protein fraction from 20 to 50 % and the other one composed by H3, H6, H9, H7, and H16 with a final amount of uncleaved protein lower than 20%.

These experiments were conducted in presence of a sub-denaturing urea concentration in order to obtain a sufficiently fast kinetics of proteolysis to be studied.

It is worth noting that LCs were all characterized by different proteolytic patterns (Figure 2.5). In no one of the cases was possible to observe the presence over time of protease-resistant fragments. In particular, this feature may be correlated to LCs different amount of trypsin cleavage sites. Indeed, using specific proteases, the presence of recognized sites and their abundance in the protein sequence must be considered. In order to address this issue, computational analyses using Peptide Cutter were performed (ExPASy Peptide-Cutter tool: Gasteiger et al., 2005). Specifically, we observed a comparable number of cleavage sites between H and M LCs with some M LCs presenting more than H LCs. Therefore, we concluded that the faster kinetics of proteolysis is not correlated with the number of trypsin cleavage sites and therefore it may be depended on intrinsic properties of the LCs native state.

Taken together, these data highlighted H LCs to be more flexible compared to the M one. Remarkably, these findings are consistent with what previously reported in literature. For instance, Morgan et al. reported three amyloidogenic LCs presenting a fast kinetics of proteolysis due to kinetic instability of the native state. Indeed, higher flexibility increases the probability to experience an unstable/less stable conformation characterized by high aggregation propensity. It is also remarkable that fold stability and kinetics of proteolysis do not necessarily correlate. For example,

M7 displayed the lowest Tm among the M LCs, but also very slow proteolysis kinetics; while H3 showed relatively high Tm, within the H group, but it is quickly proteolysed by trypsin.

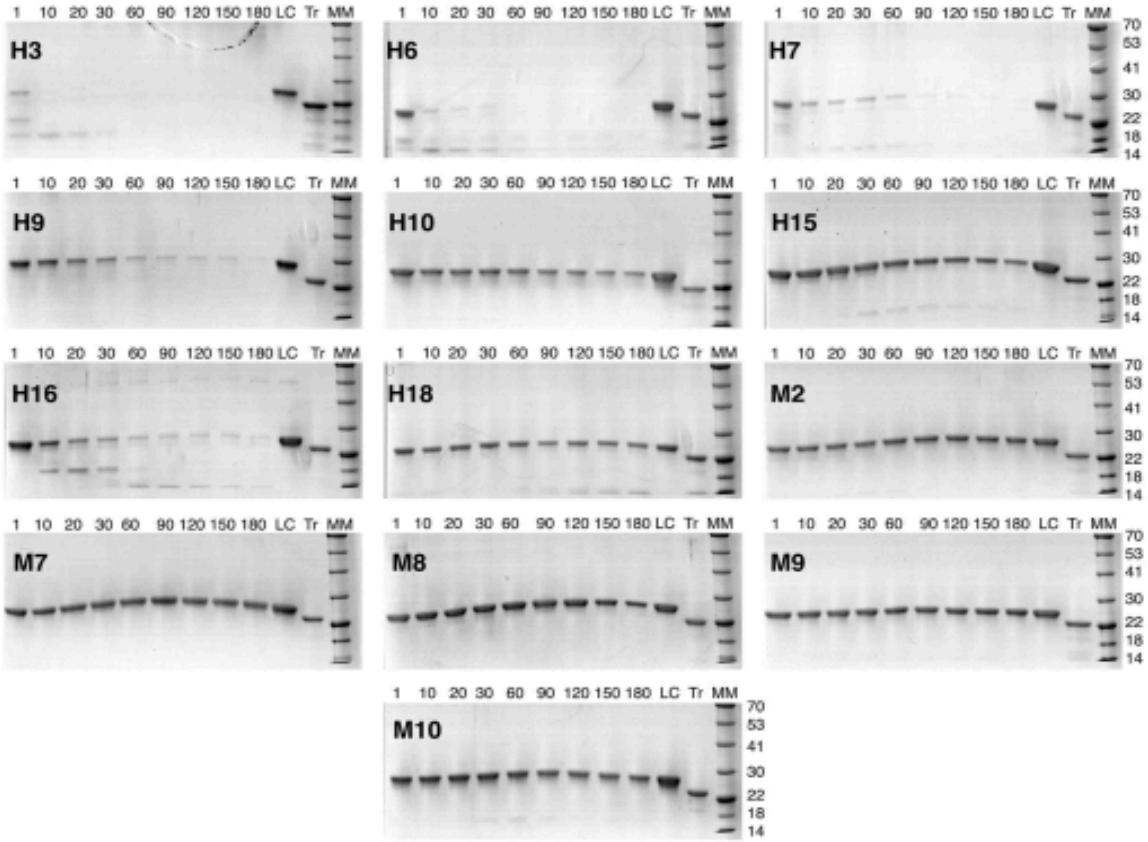


Figure 2.5: SDS-PAGE monitoring the limited proteolysis of H and M LCs by trypsin. The first sample was taken one minute after trypsin addition (1) and then at 10, 20, 30, 60, 90, 120, 150 and 180min of reaction time. LC is a standard amount of pure LC without trypsin (Tr). MM indicates molecular markers their mass is expressed in kDa.

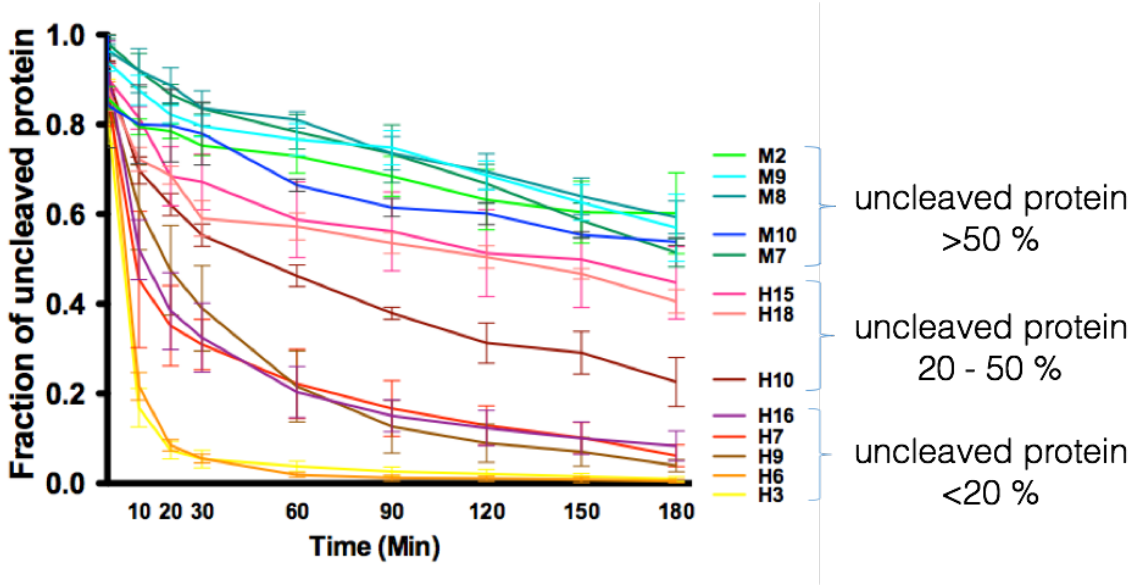


Figure 2.6: Limited proteolysis kinetics of H and M LCs.

To further confirm the results obtained using trypsin, limited proteolysis experiments were performed also using proteinase K, an unspecific protease. Using such protease allows to spot the most flexible regions of a protein regardless the protein sequence thus providing a distinct assessment of protein global/local flexibility from the experiments described above.

Our preliminary findings are consistent with trypsin experiments revealing H LCs to be proteolysed faster than M LCs (Figure 2.7). Usually, H LCs showed a final amount of uncleaved protein lower than 20% while for M7 was more the 40%. However these experiments need to be performed for all the LCs in the pool and repeated in triplicate.

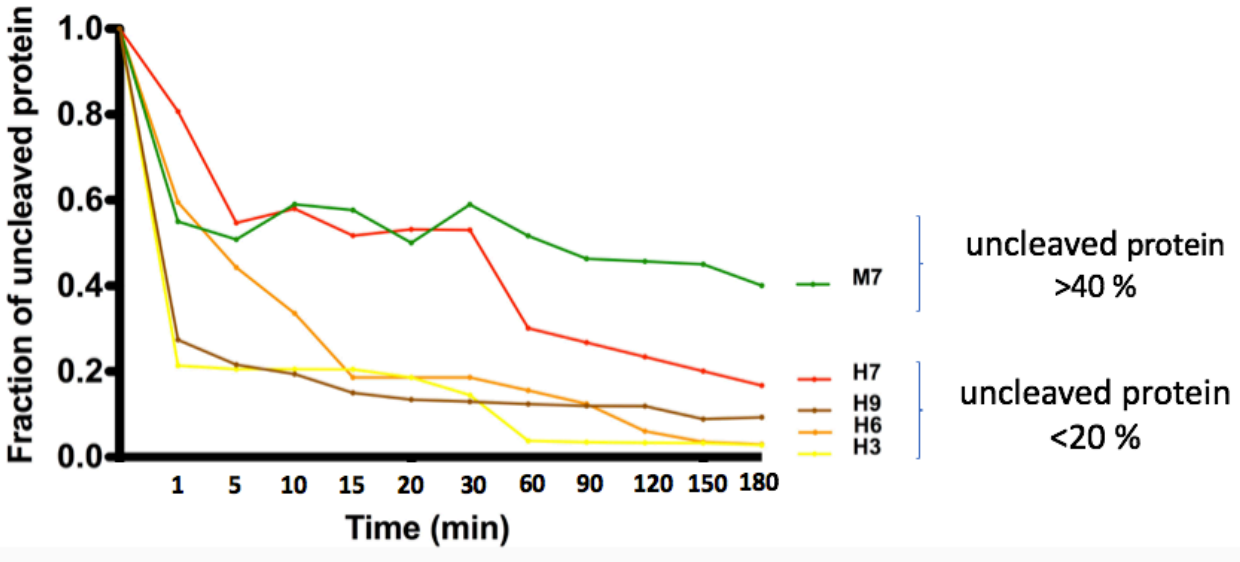


Figure 2.7: LCs Limited proteolysis kinetics using PK.

2.2.4 Structural characterization

2.2.4.1 LCs structure

In order to structurally characterize H and M LCs, we carried out crystallization trials. Specifically, we achieved crystals of five H LCs (H3, H6, H7, H9, H10) and of two M LCs (M7, M8) (Table 2.3) and their 3D-structures were determined.

Table 2.3: LCs crystallization conditions. ***** M7 was crystallized by micro-seeding using M8 crystals as seeds.

LC ID	Buffer pH	Additive	Precipitant	Purchased from
H3	6.5	0.1M Sodium cacodylate	27% w/v PEG 2000 MME	Stura screen (molecular dimensions)
H6	9.0	0.1M Bicine	10% w/v PEG 20K, 2% w/v 1,4-Dioxane	Crystal screen I/II (Hampton)
H7	7.5	0.1M HEPES	10% 2-propanol, 20% w/v PEG 4K	JBS screen (Jena bioscience)
H9	5.5	0.1M Sodium citrate	16%w/v PEG 4000, 10% v/v 2-Propanol	Stura screen (molecular dimensions)
H10		0.05M KBr	30% PEG 2000	JCSG screen (molecular dimensions)
M7 *	4.0	0.1M MMT (Malic acid, MES and Tris-based buffer)	25% w/v 1.5K	PACT screen (molecular dimensions)
M8	4.6	0.2M Sodium acetate, 2.0M NaCl		Crystal screen I/II (Hampton)

X-ray diffraction analyses revealed that LCs crystallized in different space groups with different crystal packing and ranged in resolution between 1.65 Å (H9) and 2.70 Å (H7) (Oberti et al., 2017 Table S2). As expected, in all cases the LCs quaternary arrangement was dimeric (Figure 2.8 A). Indeed, several experimental approaches, such as mass spectrometry, confirmed LCs to be covalent dimers joined by a S-S located at the *C-terminus* (Kaplan et al., 2009), however due to local flexibility, this S-S was visible only in the structure of H9.

Vl and CL showed different structural properties. CL domains are always characterized by a well-defined electron density. On the contrary, electron density was not always entirely detectable for

Vl. In particular, its quality in the CDR loops is markedly dependent on the LC crystal structure and on the crystal packing. In fact, the six CDRs (three for each monomer) were completely traceable in H6 and M8 structures; in the ones of H9 and H10 five out of six CDRs were modeled; three in H7; just two in H3 and M7. Generally CDRs tend to be solvent exposed in all the LC structures. These differences among the CDRs electron density are probably due to the different crystal packing and crystal growth conditions.

In summary the electron density for CL domains is of similarly good quality in all the seven structures and suggest that CL domains are highly rigid. ,Conversely, VL domains present different level of flexibility and CDR conformations. However, the presence of well-defined CDRs, the ability to form crystals or the quality of crystals – in terms of diffraction - do not correlate with the fold stability flexibility in solution or aggregation propensity. Indeed, the low stable H6 presents six out of six traceable CDRs while the more stable H9 and H10 just five. This means that crystallographic studies on CDR do not allow to fully appreciate the high flexibility of these regions.

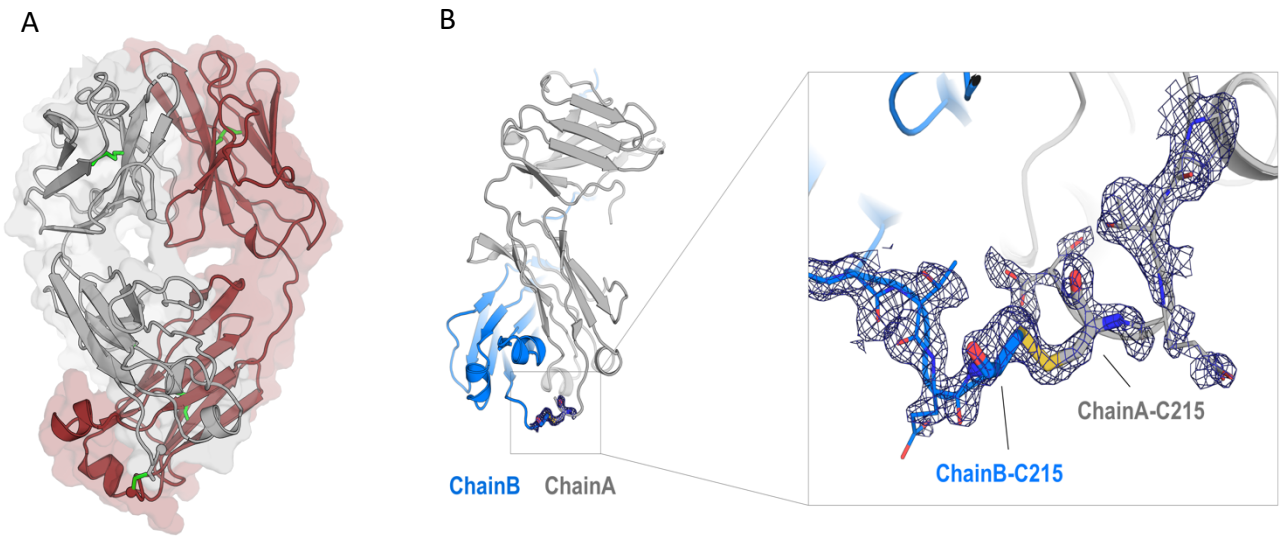


Figure 2.8: **(A)** Cartoon represented the conserved overall quaternary structure of H9 as example of LC structure. **(B)** Zoom into the well-defined electron density of C215-C215 disulphide bond (sticks representation) and the surrounding C-terminal residues (lines representation).

2.2.4.1 Structural comparison studies

With a focus on possible differences between H LCs and M LCs structures, a detailed structural comparison was performed.

First, we compared the LCs FL structures. We observed that CL and VL domains are differently oriented with respect to each other due to the bending of the joining region. In particular, this angle is known as elbow angle and the values spanning from 107.99° (H7) to 157.36° (M8). Because of such a high structural variability the root mean square deviation (r.m.s.d) analyses on the FL LCs are not meaningful. Therefore, the comparison analyses were carried out separately on VL-VL dimers and CL-CL dimers.

The superposition on CL-CL LCs domains suggested that they are conserved not only in the sequence but also from the structural point of view (Figure 2.9 A). Indeed, their comparisons showed r.m.s.d values generally around 1.0 Å over the entire dimer.

The superposition of the VL domains s revealed that the VL monomers match closely, while the CDRs conformations are rarely superposable (Figure 2.9 B). In particular, we observe CDR3 to be the most differently oriented in the space.

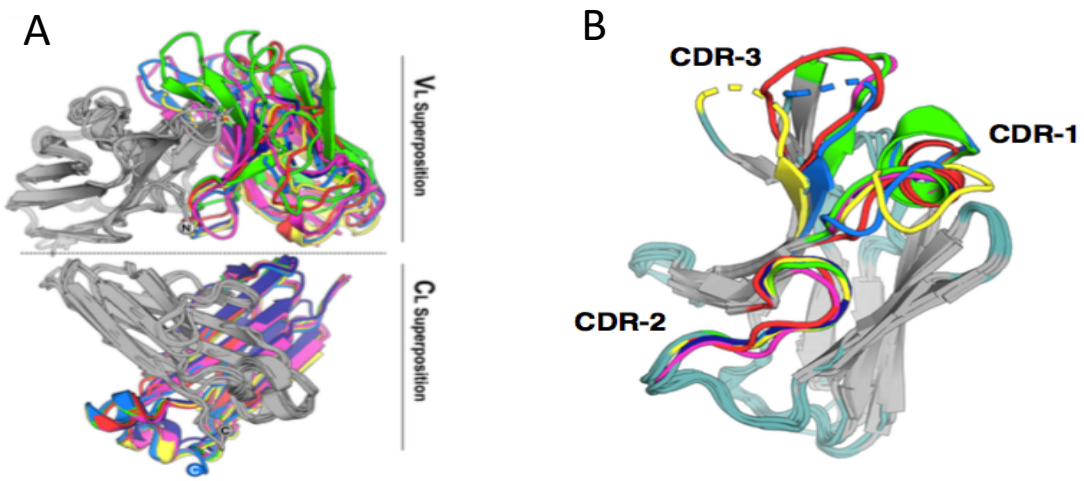


Figure 2.9: (A) Superposition analyses of the LCs structures, VL-VL in the upper part and CL-CL in the bottom. The fixed chain is reported in gray. **(B)** Superposition of VL monomers with the three CDR highlighted.

The VL-VL domains interface was also evaluated by means of structural comparisons and Protein Interfaces, Surfaces and Assemblies (PISA) analyses (Krissinel, 2015). First, comparison analyses revealed a conserved sub-set of amino acids involved in the interface (Figure 2.10 A). Moreover, the same approach showed that in all our H structures and in M7, the interface is symmetric and takes place on the five-stranded β sheet. In particular strands 3, 4, 7, 8 build the intermolecular contacts. Only M8 VL-VL presented a different interface because one of the two VL domains is rotated.

Specifically, this rotation leads the M8 interface to be asymmetric (Figure 2.10 B). Indeed, on the rotated VL domain only strands 1 and 8 provide intermolecular contacts. Interestingly, asymmetric VL-VL interface was already described in literature but only on the crystal structures of isolated VL dimers (Blancas-Mejía et al., 2017b) (Figure 2.11).

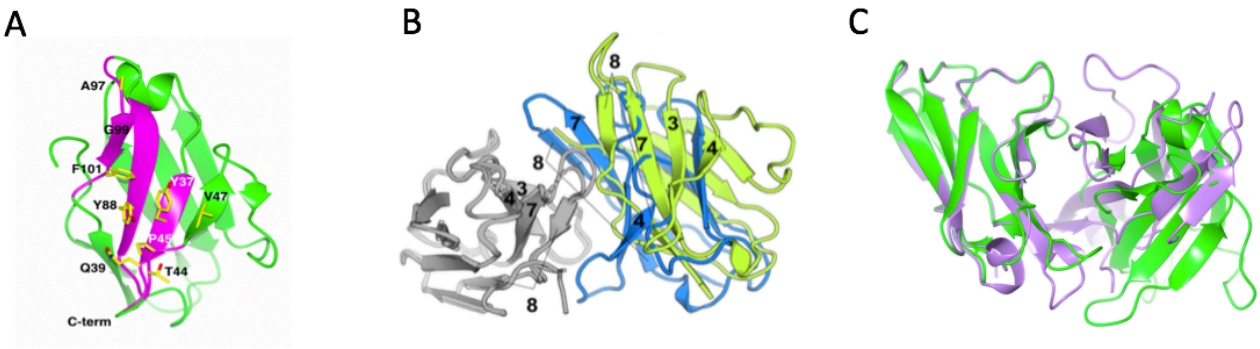


Figure 2.10: **(A)** Monomeric VL showing the strands involved in the interface (purple) with conserve amino-acids subset; **(B)** VL-VL H9 (blue) and M8 (lime) superposition, the interface of M8 is significantly asymmetric with respect to the H9 one; **(C)** VL-VL M8 and VL-VL 5T93 (Blancas-Mejía et al., 2017b) superposition, both the VL-VL dimers present asymmetric interfaces with one M8 VL 180° rotated respect to the 5T93 one.

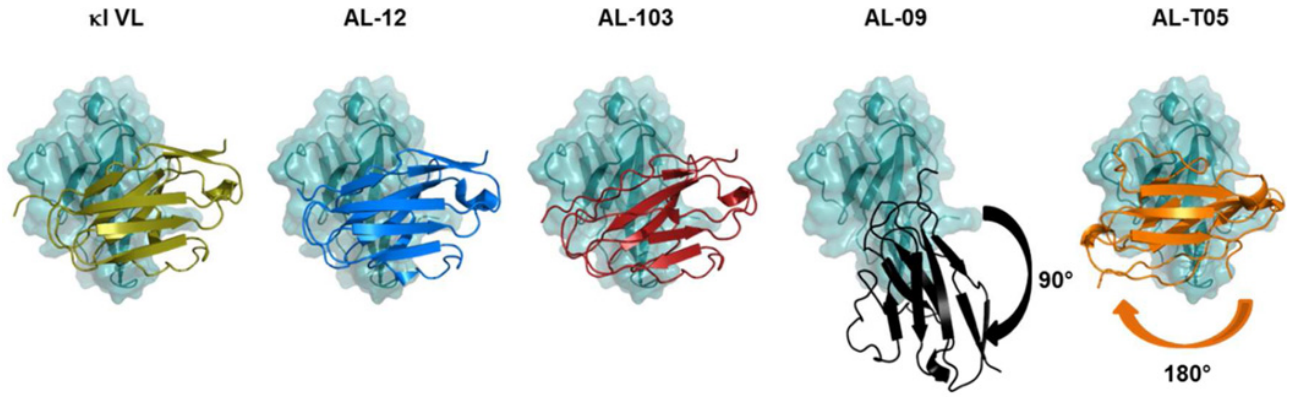


Figure 2.11: Different amyloidogenic LCs dimer interfaces. κI (gold), AL-12 (blue), AL-103 (red), AL-09 (black), and AL-T05 (orange) (PDB codes 2Q20, 3DVF, 3DVI, 2Q1E, and 5T93, respectively). Structures were superimposed using one monomer (surface, cyan) of the homodimer (adapted from Blancas-Mejía et al., 2017b).

We further characterized the interface performing a PISA analysis. This approach allows to compute the VL-VL dimers energy formation and the interface areas. Our findings suggested that there is no

correlation between the free energy gain in VL-VL dimer formation or the interface surface area with protein stability or aggregation propensity. For example, M8, which is the most stable LC against temperature, showed the smallest dimer interface area with the lowest calculated free energy change for quaternary assembly, compared to all other considered LCs.

Together these data revealed the overall quaternary assembly and VL-VL interface to be overall conserved between H and M LCs. Therefore, we concluded that structural traits are not correlated with aggregation propensity.

Table 2.4: Table showing the calculated chain A and chain B i) solvation energy and ii) the dimer interfaces.

LC ID	Solvation energy, chain A (Kcal/mol)	Solvation energy, chain B (Kcal/mol)	Dimer interface, chain A ($Å^2$)	Dimer interface, chain B ($Å^2$)
H3	-9.2	-9.3	1504	1486
H6	-7.3	-7.6	1500	1458
H7	-9.0	-7.6	1316	1352
H9	-7.7	-7.6	1506	1507
H10	-12.1	-11.3	1670	1671
M7	-8.9	-8.9	1300	1307
M8	-6.8	-8.7	1137	1125

2.3 H LCs biophysics and structure: discussion

To date, the molecular properties linking H LCs sequence to aggregation propensity have not been identified. Likely this is because aggregation propensity is complex phenomenon due to several concomitant properties and studies aiming to understand LC aggregation propensity should take in account many parameters. Indeed, our study evaluated several biophysical and structural properties of H LCs. However, the common denominator of this complexity is the high sequence variability among LCs.

From the biophysical point of view, our findings suggested that H LCs typically possess lower fold stability than M LCs (Figure 2.12 A). These data are in agreement with previous findings reported in literature (Costanzo et al., 2014): low fold stability facilitates aggregation because of the low energy required for proteins to unfold compared to proteins characterized by high fold stability. Other works also underlined the same tendency even though the experiments were performed only on isolated VL (*e.g.,* Marin-Argany et al., 2015). However, the correlation we observed between fold stability and aggregation propensity is not 100%, thus this biophysical property may contribute in determining LC aggregation propensity but cannot be considered the only cause of it.

Our findings also indicated that LCs do not present any hydrophobic patches on their surface. Moreover, we demonstrated the hydrophobic core to be exposed in a single-step unfolding process, underling that the two domains do not present independent unfolding pathways. Interestingly, the absence of surface hydrophobicity is not surprisingly. Indeed, in literature several reports indicated that in many cases the aggregation propensity may be due to electrostatic (not hydrophobic) interactions (Déret et al.,1997).

Additionally, we highlighted that H LCs tend to be more flexible than M ones (Figure 2.12 B). This findings are in keeping with other studies where it is well reported that amyloidogenic LCs are less kinetically stable and therefore more susceptible to endoproteolysis rather than non-amyloidogenic ones (Morgan and Kelly, 2016). Interestingly, the low kinetic stability could be due to the non-correct or non-complete LCs folding which leads to protease-sensitive conformations. In particular, this property has been demonstrated to increase the aggregation propensity (Morgan et al., 2017).

From the structural point of view we concluded that there are no major structural determinants, which could anticipate amyloidogenicity. Indeed, we assessed that the tertiary and quaternary assemblies of the seven LCs here presented are highly comparable and conserved. Interestingly, this finding suggested that the biophysical differences observed mainly depend on sequence variability rather than on structural rearrangements.

However, it must be noted that the most thermal unstable and trypsin sensitive LCs of our pool (H3, H6, H7 and H9) have been successfully crystallized. This underlines that despite the low fold stability and high flexibility these proteins presented a properly folded native state.

From these data, we concluded that the *in vivo* aggregation propensity cannot be associated to a specific biophysical parameter. On the contrary, we highlighted that at least two different properties, fold stability and protein flexibility, appear to be related with it. Therefore, our findings suggest that the LCs amyloidogenic behavior may be determined by the interplay of several biophysical properties of native LCs (Figure 2.12 C).

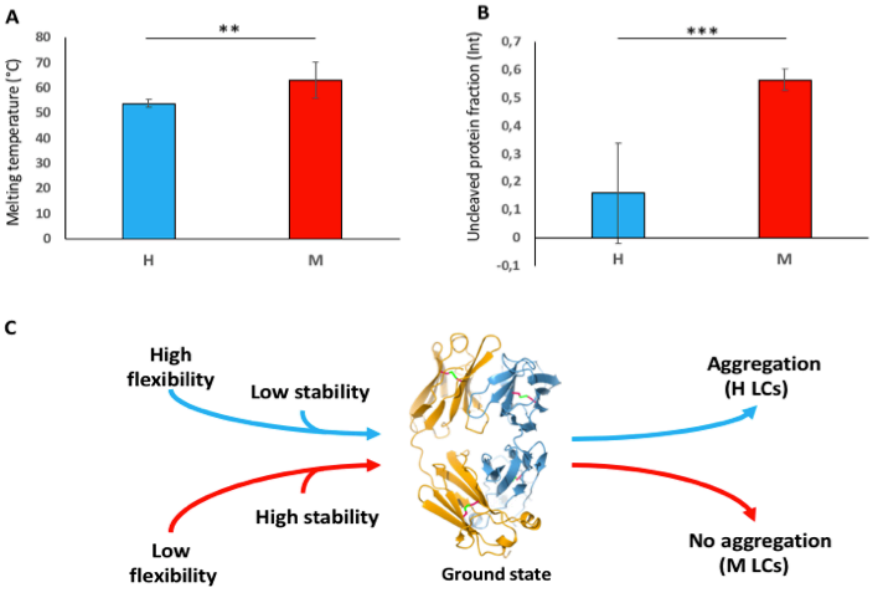


Figure 2.12: (**A)** Average of H LC and M LC melting temperatures. (**B)** Average of the H and M uncleaved protein fraction at 180 min after limited proteolysis; *** *p-value* < 0.0005; ** *p-value* < 0.005. **(C)** Graphical representation of the biophysical properties contributing in the amyloidogenic properties.

2.4 Role of Copper in LCs toxicity and aggregation propensity

2.4.1 Background

2.4.1.1 The role of Cu^{2+} in amyloid-related diseases

To date, several studied aimed to dissect the role of Cu^{2+} in amyloids-related diseases are reported in literature. Among them, it takes particular relevance the neuro-biochemistry research. Indeed, Cu^{2+} activity was extensively studied within Alzheimer disease (AD, amyloid β), Parkinson disease (PD, α-synuclein), Huntington disease (HD, huntingtin), and congophilic amyloid angiopathy (CAA). Indeed, in physiological conditions, Cu^{2+} is present in brain and it is used by a widely number of enzymes as catalyzer. At the contrary, in disease, abnormal levels of Cu^{2+} ions are detected in the brain. Indeed, a high Cu^{2+} concentration is toxic and associated with neurodegeneration (Capanni et al., 2004; Hashimoto et al., 1999; Paik et al., 1999; Uversky et al., 2001; Tiffany-castiglioni et al., 2011).

Specifically, the role of these ions has been established to influence protein aggregation: it can increase or decrease the aggregation propensity depending on the system where it is involved.

For instance, Cu^{2+} is linked with a reduced amyloid fibril formation in AD and CAA. In AD Cu^{2+} has been reported to prevent the assembly of amyloid- $\beta_{1\text{-}42}$ ($\alpha\beta_{42}$) into amyloid fibrils both at sub-stoichiometric or molar excess (Mold et al., 2013). Moreover, it has also been demonstrated that $\alpha\beta_{42}$ native form is stabilized through Cu^{2+} and Hydrogen peroxide Induced Cross-linking of Unmodified Proteins (CHICUP) (Williams et al., 2016). In details, the crosslinked stabilized dimer shows an attenuated fibril formation and thus should be further explored as a new therapeutic target (Williams et al., 2016). It is also worth to mention that in AD, as well as in CAA, recent studies (*e.g.,* schrag et al., 2011*)* demonstrated that lower Cu^{2+} leads to a more severe amyloid pathology and therefore that Cu^{2+} acts as a kind of protective agent.

However, for other neurodegenerative diseases Cu^{2+} does not show a protective effect. For example, in HD huntingtin aggregation propensity is increased when Cu^{2+} levels are increased in the brain (Xiao et al., 2013). Moreover, in PD, α-Synuclein aggregation rate has been demonstrated to be dependent on Cu^{2+} concentration (*e.g.,* Ranjan et al., 2017). Indeed, Wang and coworkers reported that Cu^{2+} appears to be important for both aggregation and cellular localization of α-Synuclein. Specifically, high Cu^{2+} levels are associated to an increase in aggregates formation (Wang et al., 2010)

Beyond the neuro-biochemistry, Cu^{2+} has been demonstrated to be related with other amyloid-related pathologies. Among these it is particularly interesting the case of β-2-microglobulin (β2m), a protein causing an amyloidosis form called "dialysis-related amyloidosis". β2m amyloid fibril are found in the musculoskeletal system of patients undergo long-term dialysis treatment (Srikanth et al., 2009) Several works have correlated Cu^{2+} to the β2m aggregation propensity. For instance, Srikanth and coworkers reported that Cu^{2+} binding causes the β2m oligomerization into native-like dimers and tetramers. Moreover, it has widely demonstrated that Cu^{2+} organizes β2m oligomers but it released upon amyloid formation (Antwi et al., 2008; Srikanth et al., 2009).

To conclude, Cu^{2+} presents different effect depending on the amyloidogenic system where it is involved. Its role is almost clarified for the disease reported above. Anyhow, future studies will be pivotal to understand its role in other amyloid-related disease and the reason why its effect can be either pro or anti amyloidogenicity.

2.4.1.2 The role of Cu^{2+} in AL amyloidosis

It is well established that metal ions plays a role in protein aggregation diseases (Hashimoto et al., 1999; Paik et al., 1999; Uversky et al., 2001). In particular, in 2017 our collaborators from the Mario Negri Research Institute demonstrated that copper ions (Cu^{2+}) seem to be relevant for LCs toxicity. Specifically, they assessed that Cu^{2+} drives LCs-mediated ROS toxicity.

Caenorhabditis elegans was used as *in vivo* system. Specifically, this system allows assessing LCs cardiotoxicity in the worm pharynx which is evolutionary related to the human heart (Mango, 2007). In particular, the pharynx pumping rate is evaluated *i.e.* the number of pharynx's terminal bulb contractions.

Using this approach, it was shown that H LCs exerts an increased toxic effect on worms when administrated together with Cu^{2+} rather than alone, as indicated by the marked decrease in the pumping rate in presence of H LCs pre-incubated with Cu^{2+}. On the contrary, M LCs in presence of Cu^{2+} did not show altered pumping rate compared to control worms or the ones fed with H LCs alone (Figure 2.10).

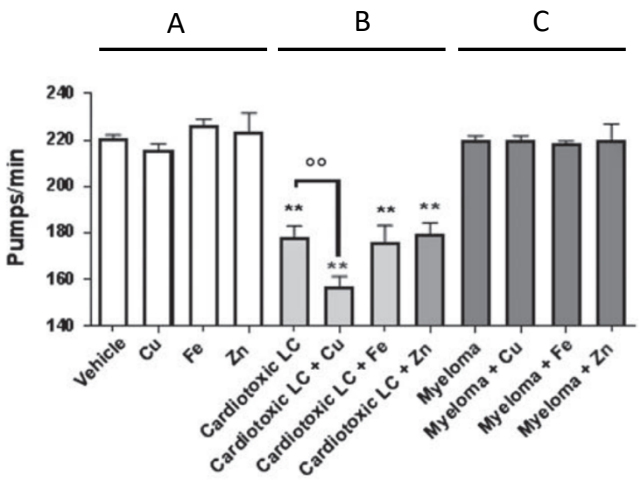


Figure 2.13: Histogram showing the effects on the *C. elegans* pumping rate of **(A)** vehicle and ions administer alone (Cu^{2+}, Fe^{2+} and Zn^{2+}), **(B)** cardiotoxic H LC administered with ions, **(C)** non-cardiotoxic Myeloma M LC administered with ions. Adapted from Diomede et al., 2017.

Moreover, in order to prove a direct role for Cu^{2+} in H LCs toxicity *in vivo*, the proteins were co-administered to worms with specific metal chelator agents. The results highlighted that the toxicity is abolished when Cu^{2+} is not available, as indicated by the increase in pumping rate values.
These observations suggest that H LCs toxic mechanisms are linked with Cu^{2+} and their co-administration results in high LCs toxicity *in vivo.* We aim to understand the biochemical basis of this effect using in our investigation H7 as toxic LC model and M7 as the non-toxic one according to Diomede et al., 2017.

2.4.2 H LCs Cu^{2+} interaction

In order to assess whether H LCs and Cu^{2+} establish a direct and specific complex we tested H7 by means of microscale thermophoresis (MST). This technique is aimed to study biomolecular interactions *e.g.,* protein-protein, protein-ligand and protein-DNA. Specifically, it evaluates by fluorescence the molecules movement along a temperature microgradient, which results in different migration velocities when ligands are bound or non-bound.
Our findings indicated that, among a large set of ions (Cu^{2+}, Mg^{2+}, Fe^{2+}, Ca^{2+}, Zn^{2+}), Cu^{2+} is the only one interacting by H7 showing a *Kd* in the nanomolar range (~700 nM) (Figure 2.14 A).
Moreover, with the aim to assess whether H and M LCs present different affinity for Cu^{2+}, the binding between M7 and Cu^{2+} was also studied by MST. Surprisingly, the *Kd* measured for M7 was a hundred times lower than the one determined for H7, with a value of ~70 µM (Figure 2.14 B).

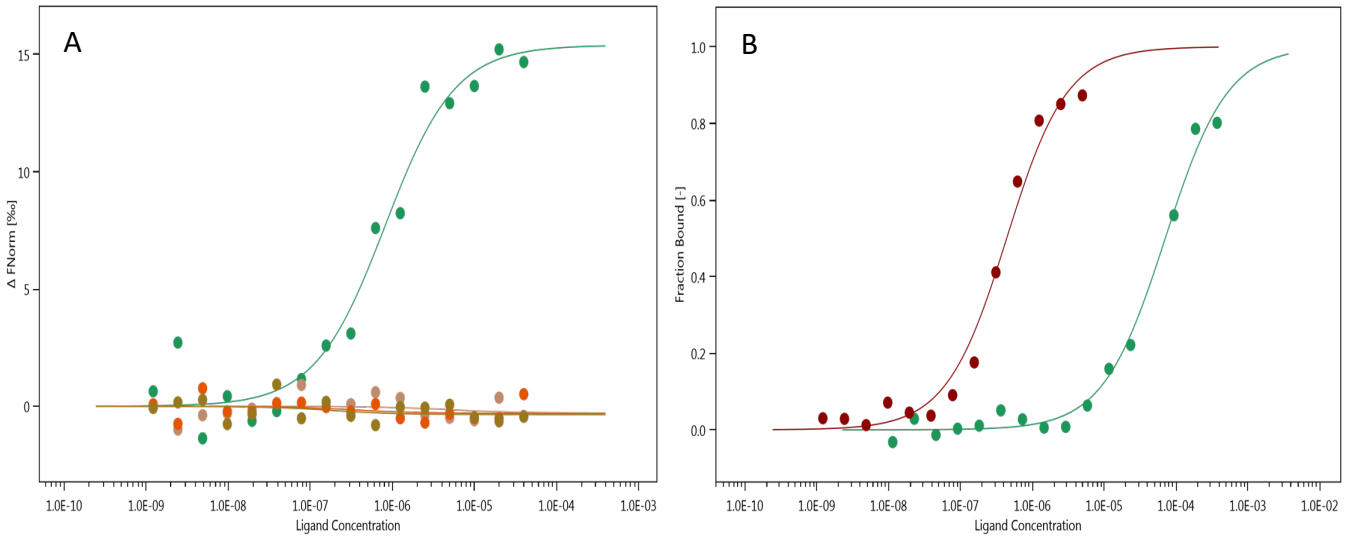


Figure 2.14: Thermophoresis analyses. **(A)** H7-Cu^{2+} binding curve (green) and Mg^{2+}, Fe^{2+}, Ca^{2+}, Zn^{2+}; **(B)** H7-Cu^{2+} binding curve (red) compared with M7-Cu^{2+} (green).

Our data suggest that the increased H LCs toxicity *in vivo* may be the result of a direct and specific interaction between H7 and Cu^{2+}. In particular, we assessed that this interaction occurs with higher affinity in a H LC rather than a M LC**.**

Then, we wanted to shed light on the Cu^{2+} binding site to H LCs. In particular, we considered the role of His residues since in literature their involvement in the metal ion coordination is well established (Lee et al., 2018). Our findings showed that at pH 5.5, when His residues are in their protonated state, the Cu^{2+} binding affinity is markedly decreased. Indeed, the measured *Kd* (~ 70 µM) resulted to be a hundred times lower compared to the one measured at pH 7.5 where His residues are in their non-protonated state (*Kd* ~ 700 nM) (Figure 2.15 A). Moreover, *Kd* value at pH 5.5 is curiously reminiscent of the one measured for M LCs (Figure 2.15 B).

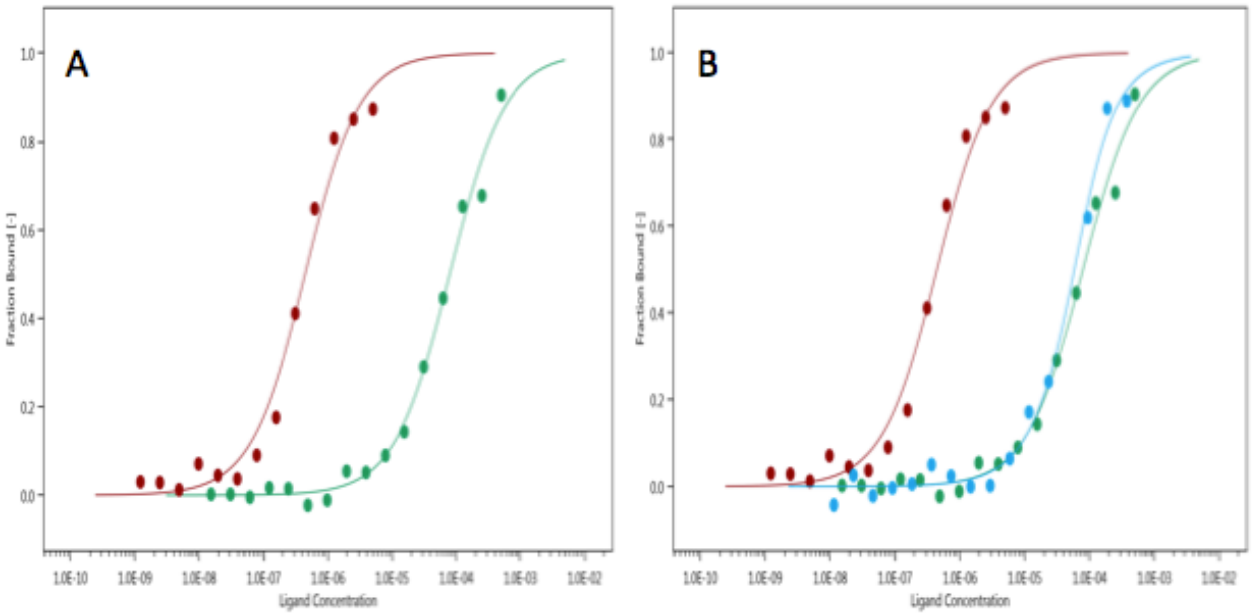


Figure 2.15: Thermophoresis analyses. **(A)** Comparison between H7-Cu^{2+} at pH 7.5 binding curve (red) with H7-Cu^{2+} at pH 5.5. **(B)** Comparison between H7-Cu^{2+} at pH 7.5 (red) with H7-Cu^{2+} at pH 5.5 (green) and M7-Cu^{2+} at pH 7.5 (blue).

These results have been further corroborated by *in vivo* experiments (Mario Negri Research Institute). The pharyngeal pumping rate in *C. elegans* is affected by the co-administration of Cu^{2+} and H7 at pH 7.5 while at pH 5.5 H7 seems to lose the most of its toxicity (Figure 2.16). These data strongly suggest that His residues may be involved in the binding of Cu^{2+} to H7.

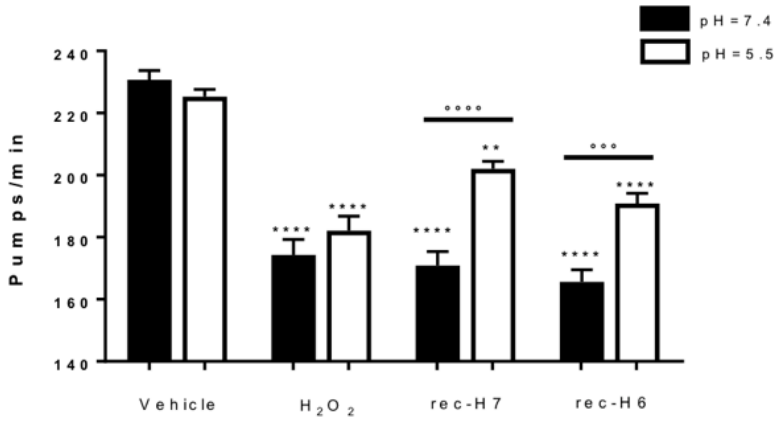


Figure 2.16: Histogram showing the pharyngeal pumping rate in *C. elegans*: toxicity tests performed at pH 5.5 show a markedly decreased protein toxicity compared to the same tests carried out at pH 7.4.

Then, we evaluated the His residues positions in the LCs sequences and their location in the H7 crystal structure. First, all the LCs belonging to our pool present two conserved His residues located in CL domain. H3, H9, H10, H15, H16, H18, M7 and M10 have a further one located in the VL and H15 and M8 present two more His residues in the Vl. Thus, we inspected H7 sequence and analyzed the His residues in H7 crystal structure. Figure 2.17 shows the position of H188 and H197 in the H7 CL domains. In particular, H188 is structurally located in the loop between strands E and F and its side chain in the crystal structure was found to be completely solvent exposed. H197, it is located in the loop between strands F and G and appears to be partially solvent exposed. Thus both H188 and H197 are properly positioned to interact with metal ions; further experiments will be carried out to identify the residues responsible for Cu^{2+} binding.

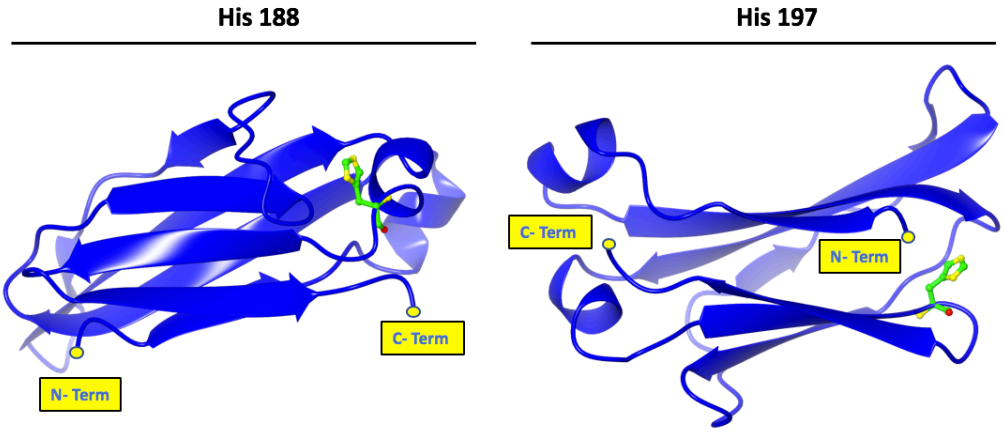


Figure 2.17 H7 CL crystal structure showing the His residues positions. H188 is slightly more solvent exposed than H197.

2.4.3 Fold stability in presence of Cu^{2+}

The effect of Cu^{2+} on H7 and H6 fold stability was evaluated. In particular, this approach allows studying the thermal stability by following Trp fluorescence while temperature increases up to 80°C. Our findings show an overall destabilization of the protein in presence of Cu^{2+}, as indicated by the Tm values (Figure 2.18 and Table 2.3). We performed these experiments at pH 5.5 and at pH 7.5 in presence or absence of Cu^{2+}. These experiments reveal that both H7 and H6 are more stable at pH 5.5 than at pH 7.5 and that Cu^{2+} destabilizes the two proteins under all tested conditions.

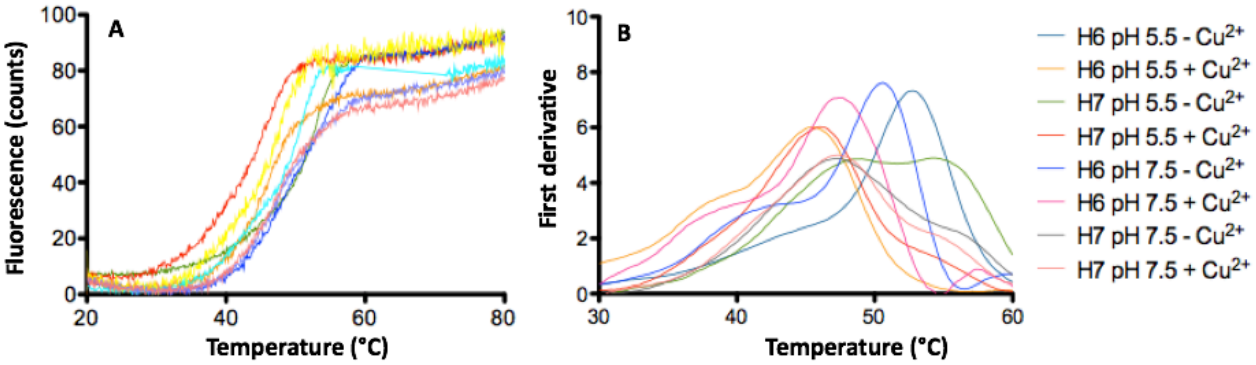


Figure 2.18: (A) Thermal unfolding ramps; **(B)** First derivative of the unfolding curve in A where the maximum indicates the Tm. Cold colors refer to protein in absence of Cu^{2+}, warm color refer to protein in presence of Cu^{2+}. It is worth to note that the warm colored curves are shifted toward lower temperatures underlining the Cu^{2+} destabilization effect.

Table 2.3: Table showing the measured T_m (°C) for H6 and H7 at pH 7.5 and 5.5 in presence or absence of Cu^{2+}.

LC ID	Tm (°C) pH 5.5 -Cu^{2+}	Tm (°C) pH 5.5 +Cu^{2+}	Tm (°C) pH 7.5 -Cu^{2+}	Tm (°C) pH 7.5 +Cu^{2+}
H6	52.6	45.0	50.3	47.0
H7	50.0	46.0	47.0	46.0

2.4.4 Protein flexibility in presence of Cu^{2+}

With the intention of evaluating protein flexibility in presence of Cu^{2+} we analyzed H7 by the spectroscopic technique Red Edge Excitation Shift (REES). REES is an optical phenomenon where low energy excitation of a fluorophore (often Trp for proteins) leads to a red shift in the maximum of the emission intensity (Catici et al., 2016). This phenomenon appears when a fluorophore shows a range of discrete solvation states and therefore it potentially reflects the equilibrium of conformational states (Chattopadhyay and Haldar, 2014). Therefore, a high molecular flexibility degree corresponds to a big red shift and it is described by the ensemble of solvent environments around the Trp probe (Catici et al., 2016).

Our findings suggested that the Cu^{2+} binding to H7 promotes an increase in flexibility. In particular, at pH 7.5 and in presence of Cu^{2+}, H7 acquires a different conformational state as indicated by a shift of ~3 nm in CSM values (Figure 2.19 B compared with Figure 2.19 D), which could be referred

to an increased degree of flexibility. On the contrary at pH 5.5 the spectrum of H7 did not show significant alterations in CSM (shift < 1 nm) indicating that not the presence but the binding of Cu^{2+} affect the protein flexibility (Figure 2.19 A compared with Figure 2.19 C). These observations fit well with the data on pH dependent Cu^{2+}binding and with the pH dependent toxicity of H7.

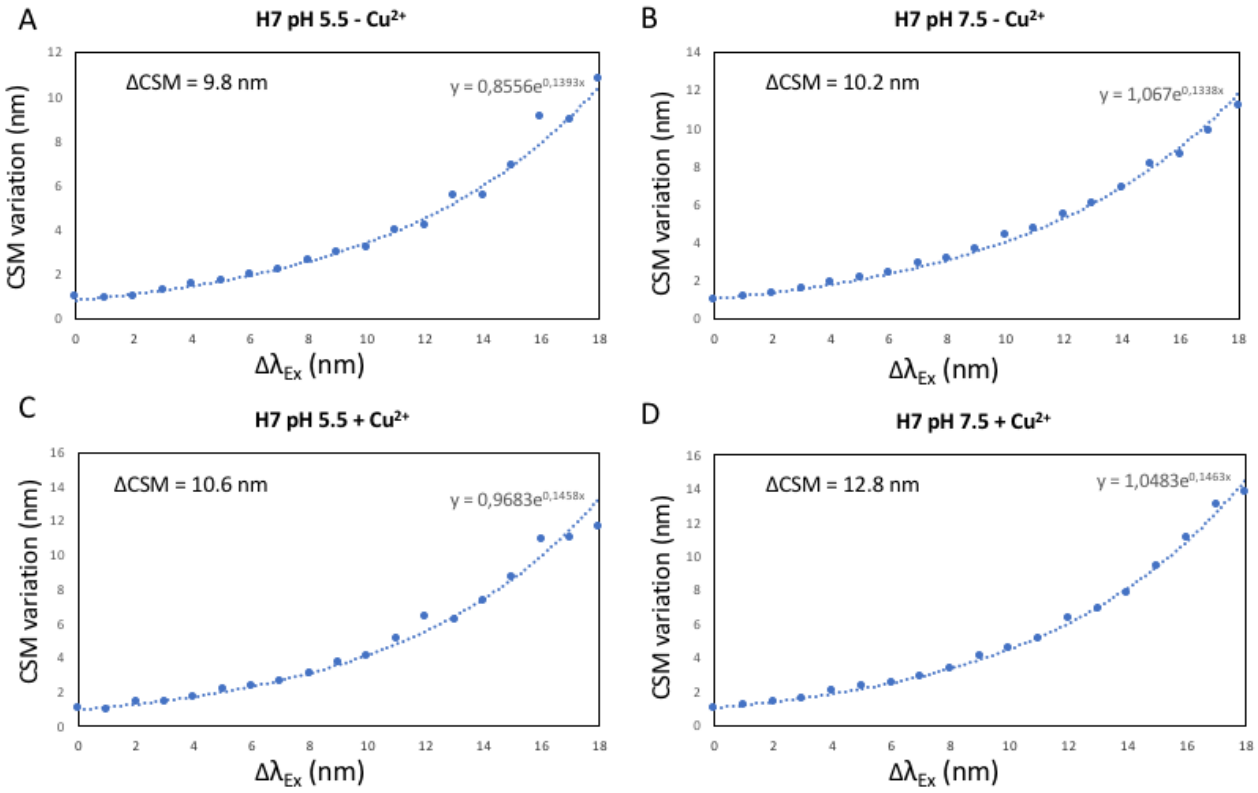


Figure 2.19: REES analyses of H7 in absence of Cu^{2+} at pH 5.5 **(A)** or 7.5 **(B)** and in presence of Cu^{2+} at pH 5.5 **(C)** or 7.5 **(D)**.

2.5 LCs and Cu^{2+}: discussion

As already discussed, Diomede and coworkers highlighted a role for Cu^{2+} in LCs toxicity (Diomede et al., 2017). Our data start to shed light on the molecular bases of such effect. They reveal that Cu^{2+} binds H LCs and we provided a possible molecular mechanism for this interaction. Our findings suggested that this binding may be coordinated by His residues. Indeed, we assessed that in presence of unprotonated His residues the binding occurs and results in a gain in flexibility and toxicity *in vivo*. As proof of this assertion we demonstrated that when His residues are protonated the binding affinity utterly decrease and the toxicity is lowered.

Finally, we assessed that the presence of Cu^{2+} results in a decreased fold stability. Specifically, this effect occurs both when His residues are protonated or not. Likely, this is the result of an unspecific effect probably due to the free Cu^{2+} in solution. Therefore, further experiments will be set up to address this issue.

Interestingly, these results agree with our previous data showing the specific properties of H LCs compared to non-toxic ones. Indeed, we reported that toxic LCs tend to be less stable and more flexible with respect to the control. In this view the *in vivo* toxicity observed in presence of Cu^{2+} may be due to a gain in flexibility and lower fold stability.

Unexpectedly, although M LCs present two His residues in the same location as H LCs, Cu^{2+} binding affinity is a hundred time lower. Likely, this may be referred to different local properties of His residues (*i.e.* flexibility) in H and M LCs. However, further experiments will be aimed to address this issue.

The role of Cu^{2+} in LC toxicity here reported fits well several reports in literature where the role of metal ions is described as connected to the fibrillation process in aggregation diseases. Indeed, in 2004 the human muscle acylphosphatase was used as protein model to study the effect of metal ions on the aggregation propensity. In particular, it was observed that Cu^{2+} produces a 2.5-fold increase in aggregation kinetics (Capanni et al., 2004). Another interesting evidence comes from Lee and coworkers regarding the role of Zn^{2+} to promote the formation of toxic oligomers in Alzheimer Disease. Additionally, other studies found metal ions, including Cu^{2+}, responsible for the *in vitro* aggregation of α-Synuclein, a protein related to Parkinson disease (Hashimoto et al., 1999; Paik et al., 1999; Uversky et al., 2001). Moreover, Cu^{2+} is reported to be involved in the beta-2-microglobulin amyloid fibers formation (Calabrese et al., 2008). Specifically, in this context Cu^{2+} promotes the formation of the aggregation-prone intermediates.

While our data show that the interaction between an H LC and Cu^{2+} leads to I) an increased toxicity mediated by a lower fold stability, II) increased protein flexibility and III) ultimately in a marked protein toxicity; the precise molecular events underlying the increased toxicity are not fully described by our experiments. Two likely hypotheses may be drawn: that Cu^{2+} induces/stabilizes the formation of toxic oligomers or that the destabilization due to Cu^{2+} may facilitates H LCs proteolysis resulting in the formation of toxic peptides. Further experiments will be required to address these crucial issues.

3 FUTURE PERSPECTIVES

Our data suggest that the co-action of several biophysical traits and external ligands play a role in defining the aggregation propensity and protein toxicity.

Starting from data produced during this three-years PhD project, the biophysical properties of H and M LCs will be further investigated. In particular, this will be better assessed by experimental as well as computational techniques.

We plan to point out the pivotal connection between molecular flexibility / thermodynamics and the pathogenic mechanism of AL. Technically, this could be achieved by means of several approaches. First, the production of a rationally designed mutant with lower thermal stability could provide information i) on the role of this biophysical property in toxicity and pathogenesis and ii) on the relation between thermodynamics and protein flexibility.

Then, the study of the first event(s) along the aggregation pathways involved in the pathogenic mechanisms may clarify the role of these two biophysical properties. In particular, it would be interesting to dissect whether the modulation of LCs thermal and kinetic stability might alter the release of toxic and aggregation prone endoproteolytic fragments (Figure 2.20).

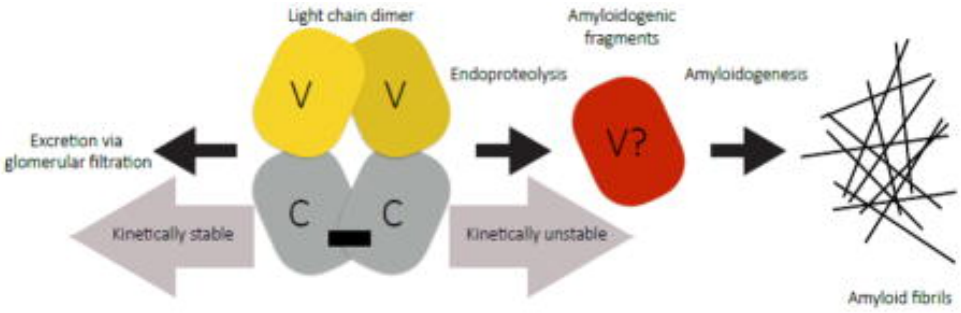


Figure 2.20. Schematic representation of LCs fragments production. (Morgan and Kelly, 2016)

Technically, this could be achieved by the study of isolated and purified soluble LCs fragments. In particular, these fragments will be characterized by means of mass spectroscopy and then will be administered to human fibroblasts or to *C.elegans* to evaluate their toxic potential. Then, the study on their aggregation propensity could provide pivotal information on their role in the first events of AL pathogenesis.

In parallel, the molecular effects of LC – Cu^{2+} interaction will be fully characterized and such understandings will also help in pinpointing out molecular events leading to protein toxicity.

To summarize, our long term goals are to draw the first event(s) along the aggregation pathway ad to understand how non-aggregated LC molecules exert their soluble toxicity: unfolding, aggregation of dimers, release of LC fragments upon proteolysis by extracellular proteases *etc.* may be the player(s) (or some of them) of this complex pathologic molecular misfolding leading to AL amyloidosis.

4 MATERIAL AND METHODS

Patients' samples. Urine and bone marrow plasma cells were isolated from patients at the Amyloid Research and Treatment Center, Foundation IRCCS Policlinico San Matteo (Pavia, Italy). Acquisition, storage and use of biological samples for research purposes were approved by the Institutional Review Board of Fondazione IRCCS Policlinico San Matteo Pavia; all methods were performed in accordance with the relevant guidelines and regulations. Written informed consent was received from participants prior to inclusion in the study. The presence of tissue amyloid deposits and amyloid organ involvement were defined according to the International Consensus Panel Criteria. LC cardiotoxicity was evaluated on the basis of clinical, echocardiography and biochemical parameters (Table S1).

Cloning. All the cloning procedures were performed by Dr. Paola Rognoni (Amyloid Research and Treatment Center, Foundation IRCCS Policlinico San Matteo, Pavia, Italy) and by Dr. Martina Maritan (Department of Bioscience, University of Milan, Milan, Italy).

Bence jones LC purification. LC were purified to homogeneity from 24h urine collection. Urines, immediately combined with 0.1% sodium azide (w/v), were centrifuged at 3000 x g for 30min. Ammonium sulfate was added to the supernatant (65% saturation) and, after overnight incubation, samples were centrifuged at 3000 x g for 30 min. the precipitates were solubilized in 20 mmol/L sodium phosphate, pH 7.0, and dialyzed against the same buffer. All steps were performed at 4 °C. LC were purified by anion exchange chromatography on an AKTÄ Purifier® FPLC system (GE-Healthcare, Piscataway, NJ, USA), using a HiPrep16/10 Q FF column, equilibrated in 20mM sodium phosphate, pH 7.0. Bound proteins were eluted with a 0 up to 1M sodium chloride linear gradient. H6, M2, M9 were purified using a cation exchanger column (HiPrep16/10 SP FF), equilibrated in 20 mM Tris-HCl, pH 8.0, and was eluted with a 0 up to 1M sodium chloride linear gradient. the homogeneity of the isolated species was assessed by 12% SDS-PAGE. The final protein concentration was determined using the Pierce BCA Protein Assay Kit (Termo Scientific, Rockford, IL, USA) and bovine serum albumin as standard.

Recombinant LCs production. Heterologous proteins were produced using BL 21 DE3 *E.coli* strain. 1L culture was inoculated by one single *E.coli* colony and grown until an optical density ~ 0.6. LCs cloned on pASK-IBA33 plus were inducted with 0.43 µM Anhydrotetracycline while those ones cloned on pET21b, with 0.5 mM IPTG. The growth conditions were 4 hours at 37 °C after the induction.

Recombinant LCs purification. LCs were find in the cytoplasm as inclusion bodies. The cells were centrifugated at 6500 rpm for 15 minutes and then resuspended in 70 mL of Buffer A (10 mM TRIS-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM phenylmethylsulfonylfluoride (PMSF)) The cell suspension was sonicated at 200 W for 6 min (Branson Sonifier 250) and then centrifuged at 18000 rpm for 15 minutes at 4°C. The inclusion bodies were resuspended in 15 mL denaturing buffer (Buffer A containing 6 M guanidinium chloride and 5 mM β-mercaptoethanol), and incubated under mild agitation at 4 °C for 4 hours and centrifuged at 30000 g for 15 min. LCs were refolded by a 20-fold dilution in 10 mM Tris-HCl pH 7.4, 1 mM EDTA, 1 mM PMSF with 5 mM reduced L-glutathione and 0.5 mM oxidized L-glutathione. The refolding buffer with the proteins were incubated over-night under agitation at 4°C. Then, the protein solution was centrifuged for 18000 rpm for 15 min at 4°C, and the supernatant was dialyzed overnight against 10 mM TRIS-HCl pH 8.0 (10 mM Sodium acetate pH 6.0 for H6 and M8). The refolding procedure was followed by purification by means of ion exchange and size exclusion chromatography. In specific, H3, H7, H9, M7 (pI ~6.0) and M9 (pI 6.8) were purified by an anionic exchange chromatography by using a Q Sepharose® Fast Flow GE Healthcare 16/10. H6 and M8 (pI ~8.0) were purified by a cationic exchange using a SP Sepharose® Fast Flow GE Healthcare 16/10. The columns were equilibrated with buffer A (10 mM Tris-HCl pH 7.4 (anionic exchange) or 10 mM Sodium acetate pH 6.0 (cationic exchange). LCs were loaded and then eluted by a saline gradient from 0 to 1M NaCl. The proteins always eluted at 160 – 240 mM NaCl. The size exclusion chromatography was performed using a Superdex 75 10/300 GL. The column was equilibrated with 10 mM Tris-HCl pH 8.0 or 50 mM NaPi pH 7.4. LCs eluted at ~ 14 mL. The homogeneity of the protein preparation was assessed by 12% SDS-PAGE, following standard procedures. Recombinant LCs were biochemically characterized by linear MALDI-TOF mass spectrometry and circular dichroism analyses, in order to verify sequence, homogeneity and correct folding. Gel filtration analysis indicates that all LCs used in this work were dimeric in solution.

Circular dichroism spectroscopy. Circular dichroism experiments, in the Far- and Near –UV regions, were carried out on a J-810 spectropolarimeter (JASCO Corp., Tokyo, Japan) equipped with a Peltier system for temperature control. All experiments were carried out in 50mM sodium phosphate pH 7.4. For the experiments in Far-UV region we used 200 µL of 0.2mg/mL LCs in a cuvette with a pathlength of 0.1 cm. Spectra were recorded from 260 to 190nm. Temperature ramps were monitored at 202nm for 1 hour. Temperature increased up to 80 from 20°C with a 60 °C/hour temperature slope. Spectra and temperature ramps were performed in triplicate for each LC except for M8 for which no BJ purified material is available. For the Near-UV region, protein concentration was 1mg/mL in cuvettes with a pathlength 1 cm. Spectra were recorded from 350 to 250nm, whereas temperature ramps were monitored at 288nm from 20 to 80 °C (slope 60 °C/hour). Tm was calculated as the first-derivative minimum of the temperature ramps.

ANS fluorescence. ANS binding experiments were performed in 50 mM sodium phosphate pH 7.4, at 20 °C and at 0.1 mg/mL protein concentration. LCs were titrated with ANS to a final concentration of 100µM. After every addition, ANS fluorescence emission spectra were monitored at 390nm in the 420–550 nm range. Excitation and Emission slits were set at 5nm, with a scanning speed of 50nm/min. When ANS concentration was 100µM, its signal was monitored at 490 nm for 1 hour while temperature increased up to 80 °C (temperature slope 60 °C/hour) in a 1 cm path length cuvette.

Limited proteolysis. LCs at a concentration of 0.8mg/mL, were incubated with a bovine trypsin at 37 °C in 50mM sodium phosphate, 1M urea at pH 7.4, for 3 hours. The molar ratio protease/protein was 1:100. The first sample was collected immediately after trypsin addition and then after 10', 20', 30', 60', 90', 120', 150', 180'. Each sample was diluted in denaturating and reducing sample buffer (NuPAGE, Invitrogen). Then, they were heated for 3 min at 95 °C and analyzed by SDS-PAGE. The uncleaved protein fraction was quantified by densitometry using Chemidoc™ MP System (Bio-Rad). Several LCs presented a very fast kinetic of proteolysis. We used H3, H6, H7, H10 in order to verify whether the presence of 1M Urea was not partially or totally unfolding the proteins. First, the Far-UV spectra are superposable to the ones without urea. Moreover, the temperature unfolding curves monitored by Far-UV indicate that the Urea concentration used is sub-denaturating. Indeed, in presence of Urea the unfolding process start at a temperature beyond 37 °C. These observations

further confirm that the fast kinetics of proteolysis were not artifacts related to excessive urea concentration.
We analyzed H3, H6 and H9, which displayed very fast kinetics of proteolysis, also in the presence of 0.5M and 0M urea, confirming under both conditions a trend for very fast proteolysis.

Crystallization and X-ray structure determination. LCs were crystallized using sitting drops or hanging drops techniques. Proteins at 8.5–10mg/mL were solubilized in 50mM sodium phosphate pH 7.4 at 20°C. Table 2.3 shows the conditions where crystals were obtained. X-ray diffraction data were collected at ESRF (European Synchrotron Radiation Facility of Grenoble–France) at the beam lines: ID29, ID30, ID23-2, BM14. XDS (Kabsch, 2010) and MOSFLM (Leslie et al., 1992)were used to process the diffraction data. The crystal symmetry was evaluated by POINTLESS (Evans, 2006) and the intensities were merged and scaled with SCALA (Project and Number 4, 1994). The crystal structures were determined by molecular replacement using PHASER, BALBES and MOLREP (Long et al., 2008; McCoy et al., 2007; Vagin and Teplyakov, 1997). Only for the molecular replacement of H7 a previously deposited on the PDB LC was used (pdb: 1JVK) as search model. Indeed, for the other molecular replacements LC structures determined in house were used. H9 initial model was generated by ARP-wARP (Perrakis et al., 2001). The initial models were subjected firstly to a rigid-body refinement and then to a restrained and TLS refinement using Phenix Refine, Refmac and Buster (Adams et al., 2010; Murshudov et al., 1997; Smart et al., 2012). Manual model building including water picking and further structure analysis were carried out by means of Coot (Emsley and Cowtan, 2004). Dimer interface analyses was performed using PISA server (Krissinel and Henrick, 2007). Fab elbow angles were calculated with phenix.fab_elbow_angle49. Figures of crystallographic structures were done using PyMOL and CCP4mg (Winn et al., 2011). The atomic coordinates and the structure factors of the seven structures of LCs have been deposited in the Protein Data Bank: 5MTL (H3), 5MUD (H6), 5MUH (H7), 5M6A (H9), 5M76 (H10), 5MVG (M7), 5M6I (M8).

Microscale thermophoresis. LCs lysine residues were labelled, using the covalent dye and protocol provided by the manufacturer (NanoTemper). Two-fold dilution series of Cu^{2+} were prepared, to have a final concentration range from 1.2 nM to 40.0 µM. The final LCs concentration was instead kept constant at 60 nM in all experiments. Assays were performed at 24 °C in 10 mM PBS, pH 8 containing 0.05% Tween20, using standard-treated glass capillaries and the instrument Monolith

NT.115 (NanoTemper) with 20% LED intensity and medium MST power. The Kd of protein-ligand was then calculated with the following equation:

$$F = U + \frac{(B-U)\cdot(\,[ligand] + [protein] + K_d - \sqrt{([ligand] + [protein] + K_d)^2 - 4\cdot[ligand]\cdot[protein]}\,)}{2[protein]}$$

where F indicates the measured fluorescence, while U and B represent the response values of the unbound and bound states, respectively.

Red Edge excitation shift. REES measurement were recorded on a Varian Cary Eclipse fluorimeter connected to a circulating water bath (HAAKE K20) for temperature regulation. These experiments were performed at 25 °C in a 0.5 cm path length cuvette. Protein concentration was adjust for each sample in order to obtain an initial fluorescence of ~ 0.8 mAu and generally it was 1 mg/mL. A 3D emission spectra scan were recorded from 325 to 500nm. The starting excitation wavelength was 292nm and then it was subsequently increased in 1 nm steps for a total of 13 scans. Excitation and emission slits were both set at 5nm and the scan speed at 500 nm/min. The corresponding buffer control was subtracted from the spectra for each experimental condition.
Data were analyzed in function of the CSM. This parameter was calculated using the following equation:

$$CSM = \frac{\Sigma(fi \times \lambda Em)}{\Sigma(fi)}$$

where fi is the measured fluorescence intensity and λEm is the emission wavelength.

5 BIBLIOGRAPHY

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*, 3801–3807.

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Adolf-Bryfogle, J., Xu, Q., North, B., Lehmann, A., and Dunbrack, R.L. (2015). PyIgClassify: a database of antibody CDR structural classifications. Nucleic Acids Res. *43*, D432–D438.

Agrawal, A., and Schatz, D.G. (1997). RAG1 and RAG2 Form a Stable Postcleavage Synaptic Complex with DNA Containing Signal Ends in V(D)J Recombination. Cell *89*, 43–53.

Al-Lazikani, B., Lesk, A.M., and Chothia, C. (1997). Standard conformations for the canonical structures of immunoglobulins. J. Mol. Biol. *273*, 927–948.

Antwi, Kwasi, et al. "Cu (II) organizes β-2-microglobulin oligomers but is released upon amyloid formation." Protein Science 17.4 (2008): 748-759.

Arosio, P., J. Knowles, T.P., and Linse, S. (2015). On the lag phase in amyloid fibril formation. Phys. Chem. Chem. Phys. *17*, 7606–7618.

Baden, E.M., Owen, B.A.L., Peterson, F.C., Volkman, B.F., Ramirez-Alvarado, M., and Thompson, J.R. (2008a). Altered dimer interface decreases stability in an amyloidogenic protein. J. Biol. Chem. *283*, 15853–15860.

Baden, E.M., Randles, E.G., Aboagye, A.K., Thompson, J.R., and Ramirez-Alvarado, M. (2008b). Structural insights into the role of mutations in amyloidogenesis. J. Biol. Chem. *283*, 30950–30956.

Baker, K.R., and Rice, L. (2012). The Amyloidoses: Clinical Features, Diagnosis and Treatment. Methodist DeBakey Cardiovasc. J. *8*, 3–7.

Blake, C., and Serpell, L. (1996). Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous β-sheet helix. Structure *4*, 989–998.

Blancas-Mejía, L.M., Martin, E.B., Williams, A., Wall, J.S., and Ramirez-Alvarado, M. (2017a). Kinetic stability and sequence/structure studies of urine-derived Bence-Jones proteins from multiple myeloma and light chain amyloidosis patients. Biophys. Chem. *230*, 89–98.

Blancas-Mejía, L.M., Misra, P., and Ramirez-Alvarado, M. (2017b). Differences in Protein Concentration Dependence for Nucleation and Elongation in Light Chain Amyloid Formation.

Biochemistry *56*, 757–766.
Bochtler, T., Hegenbart, U., Cremer, F.W., Heiss, C., Benner, A., Hose, D., Moos, M., Bila, J., Bartram, C.R., Ho, A.D., et al. (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*, 4700–4705.
Bourne, P.C., Ramsland, P.A., Shan, L., Fan, Z.C., DeWitt, C.R., Shultz, B.B., Terzyan, S.S., Moomaw, C.R., Slaughter, C.A., Guddat, L.W., et al. (2002). Three-dimensional structure of an immunoglobulin light-chain dimer with amyloidogenic properties. Acta Crystallogr. D Biol. Crystallogr. *58*, 815–823.
Brumshtein, B., Esswein, S.R., Landau, M., Ryan, C.M., Whitelegge, J.P., Phillips, M.L., Cascio, D., Sawaya, M.R., and Eisenberg, D.S. (2014). Formation of amyloid fibers by monomeric light chain variable domains. J. Biol. Chem. *289*, 27513–27525.
Brumshtein, B., Esswein, S.R., Salwinski, L., Phillips, M.L., Ly, A.T., Cascio, D., Sawaya, M.R., and Eisenberg, D.S. Inhibition by small-molecule ligands of formation of amyloid fibrils of an immunoglobulin light chain variable domain. ELife *4*.
van der Burg, M., Tümkaya, T., Boerma, M., de Bruin-Versteeg, S., Langerak, A.W., and van Dongen, J.J. (2001). Ordered recombination of immunoglobulin light chain genes occurs at the IGK locus but seems less strict at the IGL locus. Blood *97*, 1001–1008.
Buxbaum, J. (1992). Mechanisms of disease: monoclonal immunoglobulin deposition. Amyloidosis, light chain deposition disease, and light and heavy chain deposition disease. Hematol. Oncol. Clin. North Am. *6*, 323–346.
Capanni, C., Taddei, N., Gabrielli, S., Messori, L., Orioli, P., Chiti, F., Stefani, M., and Ramponi, G. (2004). Investigation of the effects of copper ions on protein aggregation using a model system. Cell. Mol. Life Sci. CMLS *61*, 982–991.
Castello, F., Paredes, J.M., Ruedas-Rama, M.J., Martin, M., Roldan, M., Casares, S., and Orte, A. (2017). Two-Step Amyloid Aggregation: Sequential Lag Phase Intermediates. Sci. Rep. *7*.
Catici, D.A.M., Amos, H.E., Yang, Y., van den Elsen, J.M.H., and Pudney, C.R. (2016). The red edge excitation shift phenomenon can be used to unmask protein structural ensembles: implications for NEMO-ubiquitin interactions. FEBS J. *283*, 2272–2284.
Caughey, B., and Lansbury, P.T. (2003). Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. Annu. Rev. Neurosci. *26*, 267–298.

Charles A Janeway, J., Travers, P., Walport, M., and Shlomchik, M.J. (2001). The interaction of the antibody molecule with specific antigen. Immunobiol. Immune Syst. Health Dis. 5th Ed.

Chattopadhyay, A., and Haldar, S. (2014). Dynamic insight into protein structure utilizing red edge excitation shift. Acc. Chem. Res. *47*, 12–19.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. *75*, 333–366.

Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., and Tulip, W.R. (1989). Conformations of immunoglobulin hypervariable regions. Nature *342*, 877–883.

Costanzo, J.A., O'Brien, C.J., Tiller, K., Tamargo, E., Robinson, A.S., Roberts, C.J., and Fernandez, E.J. (2014). Conformational stability as a design target to control protein aggregation. Protein Eng. Des. Sel. *27*, 157–167.

DeArmond, S.J., and Prusiner, S.B. (1995). Prion protein transgenes and the neuropathology in prion diseases. Brain Pathol. Zurich Switz. *5*, 77–89.

Diomede, L., Rognoni, P., Lavatelli, F., Romeo, M., di Fonzo, A., Foray, C., Fiordaliso, F., Palladini, G., Valentini, V., Perfetti, V., et al. (2014). Investigating heart-specific toxicity of amyloidogenic immunoglobulin light chains: A lesson from C. elegans. Worm *3*.

Diomede, L., Romeo, M., Rognoni, P., Beeg, M., Foray, C., Ghibaudi, E., Palladini, G., Cherny, R.A., Verga, L., Capello, G.L., et al. (2017a). Cardiac Light Chain Amyloidosis: The Role of Metal Ions in Oxidative Stress and Mitochondrial Damage. Antioxid. Redox Signal. *27*, 567–582.

Diomede, L., Romeo, M., Rognoni, P., Beeg, M., Foray, C., Ghibaudi, E., Palladini, G., Cherny, R.A., Verga, L., Capello, G.L., et al. (2017b). Cardiac Light Chain Amyloidosis: The Role of Metal Ions in Oxidative Stress and Mitochondrial Damage. Antioxid. Redox Signal. *27*, 567–582.

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., et al. (2004). Serum cardiac troponins and N-terminal pro-brain natriuretic peptide: a staging system for primary systemic amyloidosis. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. *22*, 3751–3757.

D'Souza, A., Dispenzieri, A., Wirk, B., Zhang, M.-J., Huang, J., Gertz, M.A., Kyle, R.A., Kumar, S., Comenzo, R.L., Peter Gale, R., et al. (2015). Improved Outcomes After Autologous Hematopoietic Cell Transplantation for Light Chain Amyloidosis: A Center for International Blood and Marrow Transplant Research Study. J. Clin. Oncol. *33*, 3741–3749.

Dubrey, S.W., Cha, K., Anderson, J., Chamarthi, B., Reisinger, J., Skinner, M., and Falk, R.H. (1998).

The clinical features of immunoglobulin light-chain (AL) amyloidosis with heart involvement. QJM Mon. J. Assoc. Physicians *91*, 141–157.
Duyckaerts, C., Potier, M.-C., and Delatour, B. (2008). Alzheimer disease models and human neuropathology: similarities and differences. Acta Neuropathol. (Berl.) *115*, 5–38.
Dwulet, F.E., Strako, K., and Benson, M.D. (1985). Amino Acid Sequence of a λ VI Primary (AL) Amyloid Protein (WLT). Scand. J. Immunol. *22*, 653–660.
Edmundson, A.B., and Manion, C.V. (1998). Treatment of osteoarthritis with aspartame. Clin. Pharmacol. Ther. *63*, 580–593.
Edmundson, A.B., Ely, K.R., Abola, E.E., Schiffer, M., Panagiotopoulos, N., and Deutsch, H.F. (1976). Conformational isomerism, rotational allomerism, and divergent evolution in immunoglobulin light chains. Fed. Proc. *35*, 2119–2123.
Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.
Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. *62*, 72–82.
Falk, R.H., Comenzo, R.L., and Skinner, M. (1997). The systemic amyloidoses. N. Engl. J. Med. *337*, 898–909.
Fändrich, M., Meinhardt, J., and Grigorieff, N. (2009). Structural polymorphism of Alzheimer Aβ and other amyloid fibrils. Prion *3*, 89–93.
Ferrone, F.A., Hofrichter, J., and Eaton, W.A. (1985). Kinetics of sickle hemoglobin polymerization. I. Studies using temperature-jump and laser photolysis techniques. J. Mol. Biol. *183*, 591–610.
Fischmann, T.O., Bentley, G.A., Bhat, T.N., Boulot, G., Mariuzza, R.A., Phillips, S.E., Tello, D., and Poljak, R.J. (1991). Crystallographic refinement of the three-dimensional structure of the FabD1.3-lysozyme complex at 2.5-A resolution. J. Biol. Chem. *266*, 12915–12920.
Fitzpatrick, A.W.P., Debelouchina, G.T., Bayro, M.J., Clare, D.K., Caporini, M.A., Bajaj, V.S., Jaroniec, C.P., Wang, L., Ladizhansky, V., Müller, S.A., et al. (2013). Atomic structure and hierarchical assembly of a cross-β amyloid fibril. Proc. Natl. Acad. Sci. *110*, 5468–5473.
Frippiat, J.P., Williams, S.C., Tomlinson, I.M., Cook, G.P., Cherif, D., Le Paslier, D., Collins, J.E., Dunham, I., Winter, G., and Lefranc, M.P. (1995). Organization of the human immunoglobulin lambda light-chain locus on chromosome 22q11.2. Hum. Mol. Genet. *4*, 983–991.
van Gameren, I.I., Hazenberg, B.P.C., Bijzet, J., and van Rijswijk, M.H. (2006). Diagnostic accuracy of subcutaneous abdominal fat tissue aspiration for detecting systemic amyloidosis and its utility

in clinical practice. Arthritis Rheum. *54*, 2015–2021.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook, J.M. Walker, ed. (Totowa, NJ: Humana Press), pp. 571–607.
Gauss, G.H., and Lieber, M.R. (1996). Mechanistic constraints on diversity in human V(D)J recombination. Mol. Cell. Biol. *16*, 258–269.
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., et al. (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*, 319–328.
Glabe, C.G. (2006). Common mechanisms of amyloid oligomer pathogenesis in degenerative disease. Neurobiol. Aging *27*, 570–575.
González, D., van der Burg, M., García-Sanz, R., Fenton, J.A., Langerak, A.W., González, M., van Dongen, J.J.M., San Miguel, J.F., and Morgan, G.J. (2007). Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. Blood *110*, 3112–3121.
Grogan, M., Dispenzieri, A., and Gertz, M.A. (2017). Light-chain cardiac amyloidosis: strategies to promote early diagnosis and cardiac response. Heart heartjnl-2016-310704.
Guijarro, J.I., Sunde, M., Jones, J.A., Campbell, I.D., and Dobson, C.M. (1998). Amyloid fibril formation by an SH3 domain. Proc. Natl. Acad. Sci. U. S. A. *95*, 4224–4228.
Hashimoto, M., Hsu, L.J., Xia, Y., Takeda, A., Sisk, A., Sundsmo, M., and Masliah, E. (1999). Oxidative stress induces amyloid-like aggregate formation of NACP/alpha-synuclein in vitro. Neuroreport *10*, 717–721.
Hazenberg, B.P.C. (2013). Amyloidosis: a clinical overview. Rheum. Dis. Clin. North Am. *39*, 323–345.
Helmink, B.A., and Sleckman, B.P. (2012). The Response to and Repair of RAG-Mediated DNA Double-Strand Breaks. Annu. Rev. Immunol. *30*, 175–202.
Hurle, M.R., Helms, L.R., Li, L., Chan, W., and Wetzel, R. (1994). A role for destabilizing amino acid replacements in light-chain amyloidosis. Proc. Natl. Acad. Sci. *91*, 5446–5450.
Jarolim, P. (2006). Serum biomarkers for heart failure. Cardiovasc. Pathol. *15*, 144–149.
Jones, H.B. (1848). III. On a new substance occurring in the urine of a patient with mollities ossium. Philos. Trans. R. Soc. Lond. *138*, 55–62.

Jr, P., A.s, G., and Lefranc, M.-P. (1997). A 37-kb restriction map of the human immunoglobulin lambda variable locus, VB cluster, harboring four functional genes and two non-coding V<FONT FACE=Symbol>l sequences. Braz. J. Genet. *20*.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kaplan, B., Ramirez-Alvarado, M., Sikkink, L., Golderman, S., Dispenzieri, A., Livneh, A., and Gallo, G. (2009). Free light chains in plasma of patients with light chain amyloidosis and non-amyloid light chain deposition disease. High proportion and heterogeneity of disulfide-linked monoclonal free light chains as pathogenic features of amyloid disease. Br. J. Haematol. *144*, 705–715.

KARLSSON, A., and PETERSON, P.A. (1972). TWO COMPACT DOMAIKS CONNECTED BY A SMALL SWITCH REGIOK. 9.

Katzmann, J.A., Clark, R.J., Abraham, R.S., Bryant, S., Lymp, J.F., Bradwell, A.R., and Kyle, R.A. (2002). Serum Reference Intervals and Diagnostic Ranges for Free κ and Free λ Immunoglobulin Light Chains: Relative Sensitivity for Detection of Monoclonal Light Chains. Clin. Chem. *48*, 1437–1444.

Kawasaki, K., Minoshima, S., Schooler, K., Kudoh, J., Asakawa, S., de Jong, P.J., and Shimizu, N. (1995). The organization of the human immunoglobulin lambda gene locus. Genome Res. *5*, 125–135.

Khurana, R., Gillespie, J.R., Talapatra, A., Minert, L.J., Ionescu-Zanetti, C., Millett, I., and Fink, A.L. (2001). Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. Biochemistry *40*, 3525–3535.

Kirschner, D.A., Abraham, C., and Selkoe, D.J. (1986). X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-beta conformation. Proc. Natl. Acad. Sci. U. S. A. *83*, 503–507.

Klimtchuk, E.S., Gursky, O., Patel, R.S., Laporte, K.L., Connors, L.H., Skinner, M., and Seldin, D.C. (2010). The critical role of the constant region in thermal stability and aggregation of amyloidogenic immunoglobulin light chain. Biochemistry *49*, 9848–9857.

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*, 299–306.

Krissinel, E. (2015). Stock-based detection of protein oligomeric states in jsPISA. Nucleic Acids Res. *43*, W314–W319.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797.

Lapanje, S., and Dorrington, K.J. (1973). Influence of subunit interaction and intersubunit disulphide bonds on the unfolding of immunoglobulin G by guanidine hydrochloride. Biochim. Biophys. Acta *322*, 45–52.

Lavatelli, F., Perlman, D.H., Spencer, B., Prokaeva, T., McComb, M.E., Théberge, R., Connors, L.H., Bellotti, V., Seldin, D.C., Merlini, G., et al. (2008). Amyloidogenic and Associated Proteins in Systemic Amyloidosis Proteome of Adipose Tissue. Mol. Cell. Proteomics *7*, 1570–1583.

Lavatelli, F., Brambilla, F., Valentini, V., Rognoni, P., Casarini, S., Di Silvestre, D., Perfetti, V., Palladini, G., Sarais, G., Mauri, P., et al. (2011). A novel approach for the purification and proteomic analysis of pathogenic immunoglobulin free light chains from serum. Biochim. Biophys. Acta *1814*, 409–419.

Lavatelli, F., Imperlini, E., Orrù, S., Rognoni, P., Sarnataro, D., Palladini, G., Malpasso, G., Soriano, M.E., Di Fonzo, A., Valentini, V., et al. (2015). Novel mitochondrial protein interactors of immunoglobulin light chains causing heart amyloidosis. FASEB J. *29*, 4614–4628.

Lee, M.-C., Yu, W.-C., Shih, Y.-H., Chen, C.-Y., Guo, Z.-H., Huang, S.-J., Chan, J.C.C., and Chen, Y.-R. (2018). Zinc ion rapidly induces toxic, off-pathway amyloid-β oligomers distinct from amyloid-β derived diffusible ligands in Alzheimer's disease. Sci. Rep. *8*, 4772.

Lefranc, M.-P. (2001). Nomenclature of the Human Immunoglobulin Lambda (IGL) Genes. Exp. Clin. Immunogenet. *18*, 242–254.

Leitzgen, K., Knittler, M.R., and Haas, I.G. (1997). Assembly of Immunoglobulin Light Chains as a Prerequisite for Secretion A MODEL FOR OLIGOMERIZATION-DEPENDENT SUBUNIT FOLDING. J. Biol. Chem. *272*, 3117–3123.

Leslie, A.G., Leslie, A.G.W., Leslie, A.G.W., Leslie, A., Leslie, A.G.W., Leslie Banks, A., and Leslie, A.W.G. (1992). Recent changes to the MOSFLM package for processing film and image plate data.

Lewis, S.M. (1994). The Mechanism of V(D)J Joining: Lessons from Molecular, Immunological, and Comparative Analyses. In Advances in Immunology, F.J. Dixon, ed. (Academic Press), pp. 27–150.

Little, A.J., Matthews, A., Oettinger, M., Roth, D.B., and Schatz, D.G. (2015). Chapter 2 - The Mechanism of V(D)J Recombination. In Molecular Biology of B Cells (Second Edition), F.W. Alt, T. Honjo, A. Radbruch, and M. Reth, eds. (London: Academic Press), pp. 13–34.

Long, F., Vagin, A.A., Young, P., and Murshudov, G.N. (2008). BALBES: a molecular-replacement pipeline. Acta Crystallogr. D Biol. Crystallogr. *64*, 125–132.

Lorenzen, N., and Otzen, D.E. (2014). Oligomers of α-synuclein: picking the culprit in the line-up. Essays Biochem. *56*, 137–148.

Lu, H., Schwarz, K., and Lieber, M.R. (2007). Extent to which hairpin opening by the Artemis:DNA-PKcs complex can contribute to junctional diversity in V(D)J recombination. Nucleic Acids Res. *35*, 6917–6923.

Ma, Y., Schwarz, K., and Lieber, M.R. (2005). The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. DNA Repair *4*, 845–851.

Malpasso, G., Rognoni, P., Lavatelli, F., Cervio, E., Pellegrini, C., Viarengo, G.L., Merlini, G., and Gnecchi, M. (2013). Amyloidogenic light chains induce human cardiac fibroblast toxicity through alteration of mitochondrial functionality. Eur. Heart J. *34*.

Mango, S. (2007). The C. elegans pharynx: a model for organogenesis. WormBook.

Mansilla-Soto, J., and Cortes, P. (2003). VDJ Recombination. J. Exp. Med. *197*, 543–547.

Marin-Argany, M., Güell-Bosch, J., Blancas-Mejía, L.M., Villegas, S., and Ramirez-Alvarado, M. (2015). Mutations can cause light chains to be too stable or too unstable to form amyloid fibrils. Protein Sci. Publ. Protein Soc. *24*, 1829–1840.

Marin-Argany, M., Lin, Y., Misra, P., Williams, A., Wall, J.S., Howell, K.G., Elsbernd, L.R., McClure, M., and Ramirez-Alvarado, M. (2016). Cell Damage in Light Chain Amyloidosis: FIBRIL INTERNALIZATION, TOXICITY AND CELL-MEDIATED SEEDING. J. Biol. Chem. *291*, 19813–19825.

Martin, D.J., and Ramirez-Alvarado, M. (2011). Glycosaminoglycans promote fibril formation by amyloidogenic immunoglobulin light chains through a transient interaction. Biophys. Chem. *158*, 81–89.

Matveyenko, A.V., and Butler, P.C. (2006). Islet amyloid polypeptide (IAPP) transgenic rodents as models for type 2 diabetes. ILAR J. *47*, 225–233.

M. Blancas-Mejia, L., Misra, P., J. Dick, C., A. Cooper, S., R. Redhage, K., R. Bergman, M., L. Jordan, T., Maar, K., and Ramirez-Alvarado, M. (2018). Immunoglobulin light chain amyloid aggregation. Chem. Commun. *54*, 10664–10674.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Merlini, G. (2017). AL amyloidosis: from molecular mechanisms to targeted therapies. Hematol. Am. Soc. Hematol. Educ. Program *2017*, 1–12.

Merlini, G., and Stone, M.J. (2006). Dangerous small B-cell clones. Blood *108*, 2520–2530.

Merlini, G., Lousada, I., Ando, Y., Dispenzieri, A., Gertz, M.A., Grogan, M., Maurer, M.S.,

Sanchorawala, V., Wechalekar, A., Palladini, G., et al. (2016). Rationale, application and clinical qualification for NT-proBNP as a surrogate end point in pivotal clinical trials in patients with AL amyloidosis. Leukemia *30*, 1979–1986.

Milani, P., Merlini, G., and Palladini, G. (2018). Light Chain Amyloidosis. Mediterr. J. Hematol. Infect. Dis. *10*, e2018022.

Mohty, D., Damy, T., Cosnay, P., Echahidi, N., Casset-Senon, D., Virot, P., and Jaccard, A. (2013). Cardiac amyloidosis: Updates in diagnosis and management. Arch. Cardiovasc. Dis. *106*, 528–540.

Mold, Matthew, et al. "Copper prevents amyloid-β 1–42 from forming amyloid fibrils under near-physiological conditions in vitro." Scientific reports 3 (2013): 1256.

Moreno-Gonzalez, I., and Soto, C. (2011). Misfolded Protein Aggregates: Mechanisms, Structures and Potential for Disease Transmission. Semin. Cell Dev. Biol. *22*, 482–487.

Morgan, G.J., and Kelly, J.W. (2016). The Kinetic Stability of a Full-Length Antibody Light Chain Dimer Determines whether Endoproteolysis Can Release Amyloidogenic Variable Domains. J. Mol. Biol. *428*, 4280–4297.

Morgan, G.J., Usher, G.A., and Kelly, J.W. (2017). Incomplete Refolding of Antibody Light Chains to Non-Native, Protease-Sensitive Conformations Leads to Aggregation: A Mechanism of Amyloidogenesis in Patients? Biochemistry *56*, 6597–6614.

Muchtar, E., Gertz, M.A., Kumar, S.K., Lacy, M.Q., Dingli, D., Buadi, F.K., Grogan, M., Hayman, S.R., Kapoor, P., Leung, N., et al. (2017). Improved outcomes for newly diagnosed AL amyloidosis between 2000 and 2014: cracking the glass ceiling of early death. Blood *129*, 2111–2119.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Nag, M., Bera, K., and Basak, S. (2015). Intermolecular disulfide bond formation promotes immunoglobulin aggregation: Investigation by fluorescence correlation spectroscopy. Proteins Struct. Funct. Bioinforma. *83*, 169–177.

North, B., Lehmann, A., and Dunbrack, R.L. (2011). A new clustering of antibody CDR loop conformations. J. Mol. Biol. *406*, 228–256.

Olsen, K.E., Sletten, K., and Westermark, P. (1998). Fragments of the constant region of immunoglobulin light chains are constituents of AL-amyloid proteins. Biochem. Biophys. Res. Commun. *251*, 642–647.

Padlan, E.A. (1994). Anatomy of the antibody molecule. Mol. Immunol. *31*, 169–217.

Paik, S.R., Shin, H.J., Lee, J.H., Chang, C.S., and Kim, J. (1999). Copper(II)-induced self-

oligomerization of alpha-synuclein. Biochem. J. *340*, 821–828.

Paiva, B., Martinez-Lopez, J., Corchete, L.A., Sanchez-Vega, B., Rapado, I., Puig, N., Barrio, S., Sanchez, M.-L., Alignani, D., Lasa, M., et al. (2016). Phenotypic, transcriptomic, and genomic features of clonal plasma cells in light-chain amyloidosis. Blood *127*, 3035–3039.

Palladini, G., and Merlini, G. (2016). What is new in diagnosis and management of light chain amyloidosis? Blood *128*, 159–168.

Palladini, G., Barassi, A., Klersy, C., Pacciolla, R., Milani, P., Sarais, G., Perlini, S., Albertini, R., Russo, P., Foli, A., et al. (2010). The combination of high-sensitivity cardiac troponin T (hs-cTnT) at presentation and changes in N-terminal natriuretic peptide type B (NT-proBNP) after chemotherapy best predicts survival in AL amyloidosis. Blood *116*, 3426–3430.

Parham, P. (2014). The Immune System, Fourth Edition (Garland Science).

Perfetti, V., Vignarelli, M.C., Palladini, G., Navazza, V., Giachino, C., and Merlini, G. (2004). Insights into the regulation of immunoglobulin light chain gene rearrangements via analysis of the κ light chain locus in λ myeloma. Immunology *112*, 420–427.

Perfetti, V., Palladini, G., Casarini, S., Navazza, V., Rognoni, P., Obici, L., Invernizzi, R., Perlini, S., Klersy, C., and Merlini, G. (2012). The repertoire of λ light chains causing predominant amyloid heart involvement and identification of a preferentially involved germline gene, IGLV1-44. Blood *119*, 144–150.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. Acta Crystallogr. D Biol. Crystallogr. *57*, 1445–1450.

Picken, M.M. (2007). Immunoglobulin Light and Heavy Chain Amyloidosis AL/AH: Renal Pathology and Differential Diagnosis. Kidney Plasma Cell Dyscrasias *153*, 135–155.

Pokkuluri, P.R., Solomon, A., Weiss, D.T., Stevens, F.J., and Schiffer, M. (1999). Tertiary structure of human lambda 6 light chains. Amyloid Int. J. Exp. Clin. Investig. Off. J. Int. Soc. Amyloidosis *6*, 165–171.

Project, C.C., and Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Qin, Z., Hu, D., Zhu, M., and Fink, A.L. (2007). Structural characterization of the partially folded intermediates of an immunoglobulin light chain leading to amyloid fibrillation and amorphous aggregation. Biochemistry *46*, 3521–3531.

Quock, T.P., Yan, J.T., Chang, E., Guthrie, S.D., and Broder, M.S. (2017). Epidemiology of AL Amyloidosis in a US Commercially Insured Population. Blood *130*, 5335–5335.

Rambaran, R.N., and Serpell, L.C. (2008). Amyloid fibrils. Prion *2*, 112–117.

Ramirez-Alvarado, M. (2012). Amyloid formation in light chain amyloidosis. Curr. Top. Med. Chem. *12*, 2523–2533.

Ramirez-Alvarado, M., Merkel, J.S., and Regan, L. (2000). A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. Proc. Natl. Acad. Sci. U. S. A. *97*, 8979–8984.

Ramsden, D.A., Baetz, K., and Wu, G.E. (1994). Conservation of sequence in recombination signal sequence spacers. Nucleic Acids Res. *22*, 1785–1796.

Ranjan, Priyatosh, et al. "Differential copper binding to alpha-synuclein and its disease-associated mutants affect the aggregation and amyloid formation." Biochimica et Biophysica Acta (BBA)-General Subjects 1861.2 (2017): 365-374.

Rognoni, P., Lavatelli, F., Casarini, S., Palladini, G., Verga, L., Pedrazzoli, P., Valentini, G., Merlini, G., and Perfetti, V. (2013). A Strategy for Synthesis of Pathogenic Human Immunoglobulin Free Light Chains in E. coli. PLoS ONE *8*.

Schatz, D.G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. Nat. Rev. Immunol. *11*, 251–263.

Schatz, D.G., and Swanson, P.C. (2011). V(D)J Recombination: Mechanisms of Initiation. Annu. Rev. Genet. *45*, 167–202.

Schellenberger, U., O'Rear, J., Guzzetta, A., Jue, R.A., Protter, A.A., and Stephen Pollitt, N. (2006). The precursor to B-type natriuretic peptide is an O-linked glycoprotein. Arch. Biochem. Biophys. *451*, 160–166.

Schroeder, H.W., and Cavacini, L. (2010). Structure and Function of Immunoglobulins. J. Allergy Clin. Immunol. *125*, S41–S52.

Shi, J., Guan, J., Jiang, B., Brenner, D.A., Monte, F. del, Ward, J.E., Connors, L.H., Sawyer, D.B., Semigran, M.J., Macgillivray, T.E., et al. (2010). Amyloidogenic light chains induce cardiomyocyte contractile dysfunction and apoptosis via a non-canonical p38α MAPK pathway. Proc. Natl. Acad. Sci. *107*, 4188–4193.

Schrag, Matthew, et al. "Iron, zinc and copper in the Alzheimer's disease brain: a quantitative meta-analysis. Some insight on the influence of citation bias on scientific opinion." Progress in neurobiology 94.3 (2011): 296-306.

Sipe, J.D., and Cohen, A.S. (2000). Review: history of the amyloid fibril. J. Struct. Biol. *130*, 88–98.

Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and

Bricogne, G. (2012). Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D Biol. Crystallogr. *68*, 368–380.

Soto, C. (2001). Protein misfolding and disease; protein refolding and therapy. FEBS Lett. *498*, 204–207.

Srikanth, Rapole, et al. "Copper binding to β-2-microglobulin and its pre-amyloid oligomers." Biochemistry 48.41 (2009): 9871-9881.

Tiffany-Castiglioni, Evelyn, Sandra Hong, and Yongchang Qian. "Copper handling by astrocytes: insights into neurodegenerative diseases." International Journal of Developmental Neuroscience 29.8 (2011): 811-818.

Törnquist, M., Michaels, T.C.T., Sanagavarapu, K., Yang, X., Meisl, G., Cohen, S.I.A., Knowles, T.P.J., and Linse, S. (2018). Secondary nucleation in amyloid formation. Chem. Commun. Camb. Engl. *54*, 8667–8684.

Uversky, V.N., Li, J., and Fink, A.L. (2001). Metal-triggered structural transformations, aggregation, and fibrillation of human alpha-synuclein. A possible molecular NK between Parkinson's disease and heavy metal exposure. J. Biol. Chem. *276*, 44284–44296.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an Automated Program for Molecular Replacement. J. Appl. Crystallogr. *30*, 1022–1025.

Vrana, J.A., Gamez, J.D., Madden, B.J., Theis, J.D., Bergen, H.R., and Dogan, A. (2009). Classification of amyloidosis by laser microdissection and mass spectrometry–based proteomic analysis in clinical biopsy specimens. Blood *114*, 4957–4959.

Waldmann, T.A., Strober, W., and Mogielnicki, R.P. (1972). The renal handling of low molecular weight proteins: *II. Disorders of serum protein catabolism in patients with tubular proteinuria, the nephrotic syndrome, or uremia*. J. Clin. Invest. *51*, 2162–2174.

Wall, J., Schell, M., Murphy, C., Hrncic, R., Stevens, F.J., and Solomon, A. (1999). Thermodynamic instability of human lambda 6 light chains: correlation with fibrillogenicity. Biochemistry *38*, 14101–14108.

Wang, Xiaoyan, et al. "Copper binding regulates intracellular alpha-synuclein localisation, aggregation and toxicity." Journal of neurochemistry 113.3 (2010): 704-714.

Weiss, B.M., Wong, S.W., and Comenzo, R.L. (2016). Beyond the plasma cell: emerging therapies for immunoglobulin light chain amyloidosis. Blood *127*, 2275–2280.

Wetzel, R. (1997). Domain stability in immunoglobulin light chain deposition disorders. Adv. Protein Chem. *50*, 183–242.

Williams, Thomas L., Louise C. Serpell, and Brigita Urbanc. "Stabilization of native amyloid β-protein oligomers by Copper and Hydrogen peroxide Induced Cross-linking of Unmodified Proteins (CHICUP)." Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1864.3 (2016): 249-259.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.
Xiao, Guiran, et al. "Huntington disease arises from a combinatory toxicity of polyglutamine and copper binding." Proceedings of the National Academy of Sciences 110.37 (2013): 14995-15000.

Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity

I performed the following experiments:

- **Setting of the expression conditions and purification protocol for the recombinant LCs (H3, H6, H7, H9, M7, M8 and M9);**
- **Large scale production of H3, H6, H7, H9, M7, M8 and M9;**
- **Far-UV CD, Near-UV CD, ANS fluorescence (fold stability and surface hydrophobicity);**
- **Limited proteolysis and Red edge excitation shift (protein flexibility);**
- **Crystallization of H6, H9, H10, M7;**
- **Structure determination of H6, H9, H10, M7;**
- **Microscale thermophoresis: confirmation experiments on M7;**
- **Fluorescence assays (Fold stability in presence of copper).**

SCIENTIFIC REPORTS

OPEN

Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity

Received: 10 May 2017
Accepted: 10 November 2017
Published online: 01 December 2017

Luca Oberti[1], Paola Rognoni[2], Alberto Barbiroli[3], Francesca Lavatelli[2], Rosaria Russo[4], Martina Maritan[1], Giovanni Palladini[2], Martino Bolognesi[1,5], Giampaolo Merlini[2] & Stefano Ricagno[1]

Light chain amyloidosis (AL), the most common systemic amyloidosis, is caused by the overproduction and the aggregation of monoclonal immunoglobulin light chains (LC) in target organs. Due to genetic rearrangement and somatic hypermutation, virtually, each AL patient presents a different amyloidogenic LC. Because of such complexity, the fine molecular determinants of LC aggregation propensity and proteotoxicity are, to date, unclear; significantly, their decoding requires investigating large sets of cases. Aiming to achieve generalizable observations, we systematically characterised a pool of thirteen sequence-diverse full length LCs. Eight amyloidogenic LCs were selected as responsible for severe cardiac symptoms in patients; five non-amyloidogenic LCs were isolated from patients affected by multiple myeloma. Our comprehensive approach (consisting of spectroscopic techniques, limited proteolysis, and X-ray crystallography) shows that low fold stability and high protein dynamics correlate with amyloidogenic LCs, while hydrophobicity, structural rearrangements and nature of the LC dimeric association interface (as observed in seven crystal structures here presented) do not appear to play a significant role in defining amyloid propensity. Based on the structural and biophysical data, our results highlight shared properties driving LC amyloid propensity, and these data will be instrumental for the design of synthetic inhibitors of LC aggregation.

Systemic amyloidoses are protein misfolding diseases caused by deposition of proteins as fibrillar aggregates in target organs[1]. In these disorders, the amyloidogenic protein precursor is produced at sites distant from those of deposition, being then transported to the tissues through blood[2]. Light chain amyloidosis (AL) is the most frequent systemic form, with an incidence of approximately 10 cases per million-persons/year; it is caused by deposition of excess monoclonal immunoglobulin light chains (LCs) produced by a bone marrow plasma cell clone[3]. Since the incidence of AL increases with age, the socio-economic impact of this devastating disease is expected to grow within the ageing population of industrialized countries.

AL is a heterogeneous disease, both in terms of causative proteins and of the pattern of organ involvement. The extreme variability among LCs, caused by genetic rearrangement and somatic hypermutation[4], translates into the fact that virtually every monoclonal LC is unique in its amino acid sequence. The clinical phenotype of AL is also polymorphic, most patients showing multi-organ involvement at presentation[2]: involvement of the heart, in particular, is frequent (~75% of cases) and dictates the prognosis[2,3,5]. A growing body of experimental and clinical evidences from patients with cardiac involvement indicate that damage is not only caused by fibril deposits, but also by pre-fibrillar amyloidogenic LCs, which are themselves directly toxic for target cells[6–9].

Understanding the specific properties of different LCs in their soluble native state, therefore, appears as a rational approach to explore the determinants of amyloid formation, organ tropism and dysfunction *in vivo*.

[1]Dipartimento di Bioscienze, Università degli Studi di Milano, 20133, Milano, Italy. [2]Amyloidosis Research and Treatment Center, Fondazione IRCCS Policlinico San Matteo, and Department of Molecular Medicine, University of Pavia, 27100, Pavia, Italy. [3]Dipartimento di Scienze per gli Alimenti, la Nutrizione e l'Ambiente, Università degli Studi di Milano, 20133, Milano, Italy. [4]Dipartimento di Fisiopatologia Medico-Chirurgica e dei Trapianti, Università degli Studi di Milano, Milano, Italy. [5]CNR Istituto di Biofisica, c/o Università degli Studi di Milano, 20133, Milano, Italy. Correspondence and requests for materials should be addressed to S.R. (email: Stefano.ricagno@unimi.it)

LC code	Germline	Gender, age	Diagnosis	Organs involved	Serum λ FLC (mg/l)
H3	1c (IGLV1-44)	M, 65	AL	H	252
H6	1b (IGLV1-51)	F, 72	AL	H, K, PNS	248
H7	1b (IGLV1-51)	M, 45	AL	H	477
H9	2c GLV2-8)	M, 59	AL	H, ST	699
H10	1a (IGLV1-36)	M, 73	AL	H, L	475
H15	6a (IGLV6-57)	M, 53	AL	H, PNS, ST	839
H16	2a2 (IGLV2-14)	M, 72	AL	H, K	383
H18	31 (IGLV3-19)	M, 69	AL	H, ST, PNS	509
M2	2b2 (IGLV2-23)	M, 65	MM	—	1140
M7	31 (IGLV3-19)	F, 71	MM	—	6130
M8	2b2 (IGLV2-23)	M, 48	MM	—	573
M9	2b2 (IGLV2-23)	M, 61	MM	—	8510
M10	2a2 (IGLV2-14)	M, 55	MM	—	12200

Table 1. Biochemical and clinical features of H and M LCs (for more extended information, with a complete set of clinical data, see Table S1). Abbreviations: M, male; F, female; MM, multiple myeloma; H, Heart; K, Kidney; ST, Soft Tissues; PNS, Peripheral Nervous System; BJ, Bence Jones protein (monoclonal urinary free light chains); FLC, Free Light Chains.

Because of insufficient insight, no treatments exist yet to block fibril formation and prevent tissues damage. In fact, the current therapeutic strategies are based essentially on halting the production of amyloidogenic LCs from the plasma cell clone by means of chemotherapy[5].

Structural characterization has shown that LCs assemble into homodimers, each monomer consisting of two immunoglobulin domains[10], where the N-terminal variable domain (V_L) displays high sequence variability. In particular, the three hypervariable complementarity determining regions (CDR) that target specific antigens are located in the V_L domain. The C-terminal constant domain (C_L), on the contrary, displays highly conserved sequence within the λ and the κ isotypes. To date, most studies dealing with the biophysical and structural properties of amyloidogenic LCs have focused specifically on V_L domains, as they are abundant in fibrils: typically, V_L domains belonging to amyloidogenic LCs are thermodynamically and kinetically unstable, while the three-dimensional structures of amyloidogenic and non-amyloidogenic V_L studied match closely[11–15]. However, many observations stress the relevance of studying full length LCs: (i) although V_L are abundant in fibrils, the full length LCs and C_L domains have also been found in deposits[12,16–18]; (ii) so far, FL LCs and not truncated forms, have been found soluble in blood and serum[19]; iii) recent reports underline the relevance of the C_L domain in determining both the biophysical properties and the aggregation propensity of LCs[20–23].

The biophysical characterizations hitherto reported have mostly focused on the pairwise comparisons of one amyloidogenic with one non-amyloidogenic LC[11,16,20,21,24]. At the light of the mentioned sequence variability[4], the results of such studies may prove difficult to generalize. In an effort to extend our views, here we present a comprehensive biophysical and structural characterization of a pool of eight amyloidogenic LCs originating from different patients (thus endowed with different amino acid sequences). Among all patients who overexpress one LC variant, only a subset develops AL (specific germline genes typically of the λ isotype are overrepresented among amyloidogenic LCs[25–27]), therefore five LCs not displaying amyloid propensity in patients were included as controls in our test pool. We explored the correlation between the aggregation propensity observed in patients and the LC molecular properties that may elicit misfolding, and are held to be associated with proteotoxicity. Our results provide insight on the fundamental molecular properties of pathogenic LCs, and in parallel suggest concepts for the design of therapeutic approaches to AL directly targeting circulating LC molecules.

Results

Set-up of the LC test pool. In order to explore a most general context, while coping with the issue of high LC sequence variability, we devised an experimental approach based on three criteria. (i) Our biophysical and structural characterization covered thirteen LCs from distinct patients. LCs were distributed in two groups: a set of eight amyloidogenic LCs responsible for severe cardiac involvement and symptoms in patients (hereafter, H LC), and a set of five non amyloidogenic LCs isolated from patients affected by multiple myeloma for which no amyloid aggregation or proteotoxicity was observed in patients (hereafter, M LC) (Table 1 and Table S1). (ii) All the selected LCs were of the λ isotype, which is overrepresented among amyloidogenic LCs (approximately 75%)[2,28], all the H LCs (except H10 from germline 1–36) belong to germlines commonly found among AL patients[29]. However, the LCs were distributed across different families and germlines to avoid focusing on family/germline-specific properties (Table 1 and Fig. 1). (iii) All proteins used in the experiments were full length LCs, since the blood concentration of these species in each patient directly correlates with the severity of organ dysfunction, particularly regarding the heart (Table S1)[5,8]. Our experiments focused on the properties of native LCs, since ultimately these represent the circulating reservoir of aggregation prone material; LCs were either expressed recombinantly[30] or purified from patients' urines (Table 2).

LC fold stability. Protein stability within the two LC groups was assessed using three different and complementary approaches. Unfolding temperatures for each LC were monitored by Far-UV circular dichroism (CD),

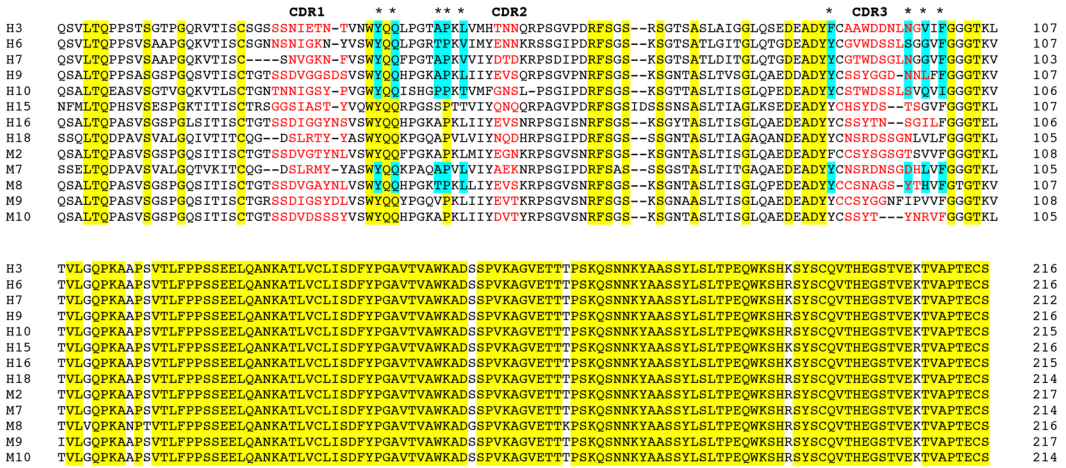


Figure 1. Multialignment of all thirteen LCs used in this work: the residues conserved in all sequences are highlighted in yellow, residues involved in the $V_L - V_L$ interface in all the crystal structures are highlighted in cyan, while the residues belonging to the three CDRs are shown in red.

LC	Source	Tm_{app} Far UV (°C)	Tm_{app} Near UV (°C)[1]	ONSET ANS SIGNAL (°C)	Hydrophobicity (%)	Proteolysis[2]	Trypsin sites
H3	Rec	(40.2) – 54.4 ± 0.8	(44.5)—53.1	54.8	23.1	+++	14
H6	Rec	(43.9) – 54.5 ± 0.8	(45.3)—	55.3	23.1	+++	18
H7	Rec	(42.8) – 54.7 ± 0.5	(48.7) – 56.2	54.8	23.6	++	16
H9	BJP	52.0 ± 0.8	54.0	55.8	21.3	+++	14
H10	BJP	54.9 ± 0.5	54.3	58.3	23.7	++	14
H15	BJP	55.6 ± 0.6	—	54.7	21.8	++	14
H16	BJP	51.3 ± 0.7	51.9	54.5	22.3	+++	13
H18	BJP	52.4 ± 0.5	51.1	51.5	22.9	++	13
M2	BJP	61.5 ± 0.8	—	63.9	23.0	+	15
M7	BJP	54.7 ± 0.5	55.3	58.8	22.4	+	15
M8	BJP	72.6	—	73.5	22.7	+	15
M9	Rec/BJP	(63.2) – 68.2 ± 1.4	(58.8) – 67.7	62.8	24.2	+	15
M10	BJP	58.0 ± 0.6	—	56.1	21.5	+	15

Table 2. Biophysical properties assessed for H and M LCs. Rec, recombinantly expressed LC; BJP, Bence Jones proteins purified from urines. ([1]): some Tm_{app} values could not be measured due to protein aggregation during the temperature ramp. ([2]): +++: < 20%, ++: 20–60%, +: > 60% fraction of uncleaved protein after 180 min.

by Near-UV CD, and by 8-anilino-1-naphthalenesulfonic acid (ANS) binding and fluorescence signal (Fig. 2). LC unfolding monitored through these spectroscopic methods provides information on the loss of secondary structure, of tertiary structure, and on the exposure of buried hydrophobic residues to the solvent, respectively. As previously reported the unfolding of full length LCs is not reversible (data not shown)[22], thus all temperatures corresponding to the unfolding inflection points should be considered apparent melting temperatures (Tm_{app}). The Far and Near-UV spectra of all LCs under native conditions are shown in Fig. S1.

In general, H LCs (warm colours in Fig. 2) displayed lower Tm_{app} compared to M LC (cool colours in Fig. 2), as determined from the Far-UV and Near-UV CD temperature ramps shown in Fig. 2A and B, respectively. However, low Tm_{app} values do not correlate perfectly with LCs amyloidogenicity, since M7 and M10 LCs unfold at temperatures close to those of the most stable H LCs (Table 2). Interestingly, most LCs show a one-step unfolding process, but three H LCs (H3, H6, H7) and one M LC (M9) display a two-step process, observed both in Far- and Near-UV temperature ramps (Fig. 2A,B). At the protein concentration required for Near-UV experiments four LCs (H10, M2, M8, M10) precipitated before the unfolding process was complete, thus Tm_{app} could not be determined. Nevertheless, the Tm_{app} values obtained by monitoring the unfolding both by Near-UV and Far-UV CD, show a good agreement (Table 2), indicating that for these LCs the loss of secondary and tertiary structures are simultaneous processes. Student's test calculated using Tm_{app} values determined by Far-UV indicate that the differences between H and M-LCs are significantly different ($P < 0.005$).

While the Near-UV CD signal is strongly dependent on Trp residues, and may therefore provide rather local information about the unfolding process, an increase in ANS fluorescence indicates the exposure of hydrophobic residues from the protein core. Inspection of Fig. 2C shows that unfolding monitored by ANS fluorescence is compatible with exposure of the hydrophobic core residues in a one-step process in all the analysed LCs. This observation rules out that stepwise unfolding of the two distinct domains occur in the LCs (H3, H6, H7, M9), which display

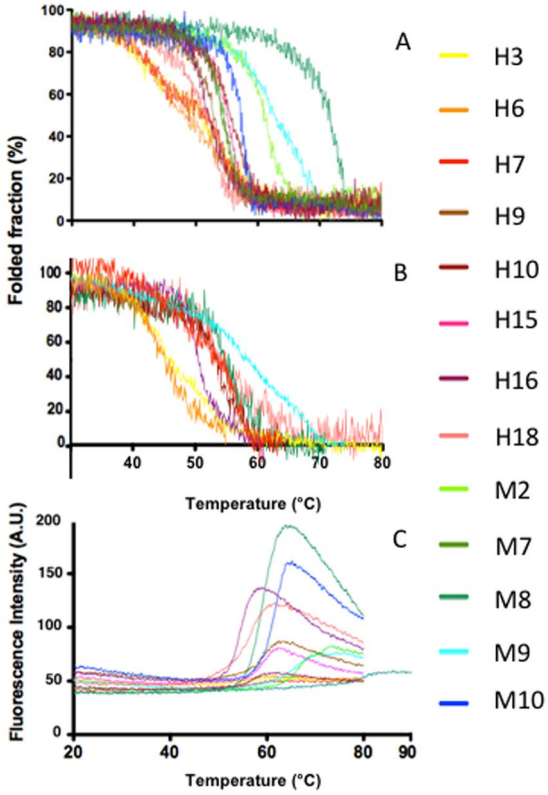


Figure 2. Thermal stability of H and M LCs. Far-UV (**A**) and Near-UV (**B**) CD temperature ramps of the two sets of LCs used in this study. (**C**) Temperature ramps followed using ANS fluorescence. The H and M LCs are shown with warm and cool colours, respectively.

two-step unfolding when monitored by CD signal; thus the temperatures corresponding to the first inflection points are written in parenthesis in Table 2. The onset of the ANS signal, but not the inflection point of the curves, correlates well with the Tm_{app} values calculated from the CD signals; such apparent discrepancy suggests that the hydrophobic core is exposed only at an advanced state of unfolding, which leads to an (apparent) systematic Tm_{app} overestimation in ANS temperature unfolding. Taken together these data suggest that the loss of tertiary and of secondary structures are simultaneous processes, indicating that LC dimers unfold through a cooperative process. Importantly, the ANS spectra recorded under native conditions yield undistinguishable low signals (data not shown), suggesting that under such conditions none of the LCs studied in this work expose relevant hydrophobic patches to the surface.

LC dynamics. In order to obtain an indirect assessment of protein dynamics, limited proteolysis of the different LCs using trypsin was performed. Typically, fast proteolysis kinetics correlate with marked protein dynamics; conversely, protein rigidity correlates with slow proteolysis kinetics. In our hands, many but not all of the LCs belonging to the explored set displayed an unexpected resistance to proteolysis by trypsin and chymotrypsin (data not shown); controlled shaking (*i.e.* simulating shear forces) was tested to increase the kinetics of proteolysis without any effect (data not shown). Thus, in order to perform proteolysis in an adequate time frame, a sub-denaturant concentration of urea (1 M) was added to the protein solutions: under such conditions all LCs displayed CD spectra indistinguishable from those recorded in the absence of 1 M urea (data not shown).

Figure 3 shows the SDS-PAGE monitoring the controlled proteolysis experiments and the plotted fraction of uncleaved protein at different time points. Analysis of the data shows that the kinetics of proteolysis correlate well with amyloid propensity (Fig. 3B): H3, H6, H7, H9, and H16 are almost or totally proteolysed after 60 minutes, and more than 50% of H10, H15 and H18 LCs is cleaved after three hours. This is in stark contrast with the behaviour observed for the five M LCs, which are consistently more resistant to trypsin, in all cases more than 60% of the M LCs remaining uncleaved at the end of the experiment (total of 3 hours). The number of potential trypsin cleavage sites for each LC is reported in Table 2; such numbers vary and do not correlate with the kinetics of proteolysis, suggesting that the observed kinetics underlie genuine differences in protein dynamics or conformational flexibility. Student's test calculated using the percentage of uncleaved LC at the end of the experiment indicates that the kinetics of proteolysis between H and M-LCs are significantly different ($P < 0.0005$).

The pattern of proteolysis varied among different LCs. SDS-PAGE gel analysis showed that for most LCs no bands corresponding to specific LC fragments were present, indicating that the LCs had been directly proteolysed to small peptides; although in a few cases discrete bands were detected, they proved unstable over the entire proteolysis experiment (Fig. 3A). Overall, these observations suggest that the proteolytic path is distinct in different LCs, and that no stable LC domains (*e.g.* V_L, C_L domains or other fragments) can be isolated during proteolysis under the conditions tested.

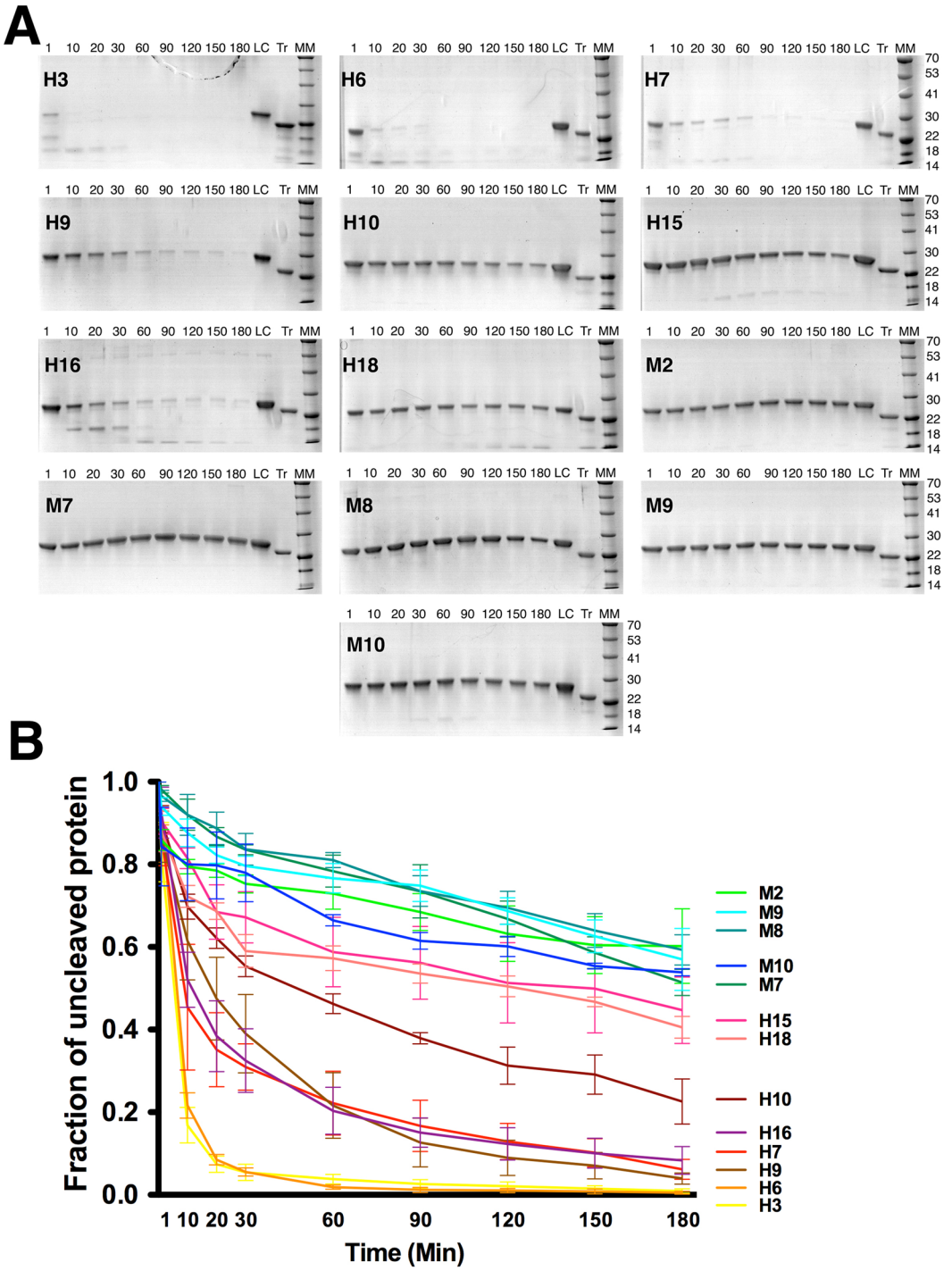


Figure 3. (**A**) SDS-PAGE monitoring the limited proteolysis of H and M LCs by trypsin. The first sample was taken one minute after trypsin addition (1) and then at 10, 20, 30, 60, 90, 120, 150 and 180 min of reaction time. In LC a standard amount of the corresponding LC loaded onto the gel without adding trypsin (Tr). MM indicates molecular markers their mass is expressed in kDa. All SDS-PAGE were run under reducing conditions. Raw images of the all SDS-PAGE are shown in Figure S4 (**B**) Kinetics of LC proteolysis. The intensity of the band corresponding to the uncleaved LC has been quantified at different time points and plotted. As starting point the amount of protein present in the LC sample was chosen. Each curve results from three independent proteolysis experiments. The curves are colour coded as in Figs 2 and 3.

Structural analyses of H and M LCs. In order to gain high-resolution structural insight, all the thirteen LCs were screened for crystal growth. Five H LCs (H3, H6, H7, H9 and H10) and two M LCs (M7 and M8) were successfully crystallized and their 3D-structures solved. Each of the seven LCs crystallized in a different space group, with distinct crystal packing; the crystallographic resolutions achieved ranged from 2.70 Å (H7) to 1.65 Å

(H9) (Table S2). The crystal structures were all refined to satisfactory refinement parameters (Table S2); in all cases the LCs were dimeric (Fig. 4A), with an overall quaternary arrangement closely matching that of the previously reported LC dimers[10]. Mass spectrometry and non-reducing SDS-PAGE indicated that 100% of the LC dimers were covalently linked by a disulphide bond located at the chain C-*terminus* (data not shown). However, due to underlying local flexibility, only for H9 a well-defined electron density for such C-terminal intermolecular disulphide bond is visible (Fig. S2).

V_L and C_L domains display very different structural properties in the seven crystal structures. C_L domains (res 115–214) are typically characterized by high quality electron density, are all very well superposable, and the dimeric C_L–C_L arrangement is clearly conserved in all the analysed structures (root mean square deviations -r.m.s.d.- generally fall below 1.0 Å over the entire C_L–C_L dimer - Table S3A and Fig. 4B-bottom). The linker regions connecting V_L and C_L domains (residues 110–115) allow evident variability in the mutual orientation of V_L and C_L domains, which is described by the elbow angles (Table S2) spanning from 107.99° (H7) to 157.36° (M8). (Interestingly, the greatest elbow angles were found in the M LCs structures, as for previously reported works[10]. For this reason, the r.m.s.d. values calculated for the full length LCs are not meaningful and are not reported.

Contrary to what has been observed for the C_L domains, the electron density quality for the V_L domains (residues 1–109) greatly differs among the seven crystal structures. With the only exception of one of the two V_L domains of the M7 LC dimer, the two β-sheets building the V_L domains are in general easily traceable; however the quality of the electron density for the CDR loops is markedly dependent on the LC crystal structure considered. In H6 and M8 dimers all six CDRs (three for each LC monomer) are traceable in the electron density; in H9 and H10 five out of six are traceable, in H7 three, and in H3 and M7 only two CRDs are traceable. Such different behaviours, which are related to the intrinsic CDR conformational adaptability, do not correlate with LC aggregation propensity or with the flexibility assessed by limited proteolysis. Thus, the observed variety of CDR conformations are likely due to different CDR amino acid sequences, but may also reflect different chemical conditions for crystal growth and different crystal packing.

The above concept is further stressed by the fact the V_L domain tertiary structures match closely (Table S3B and Fig. 4C), while conformations of their CDRs are rarely superposable. The inspection of the V_L dimer provides a relevant example: the fine details of the V_L–V_L association interface are LC-specific, depending on: (i) sequence variability, which likely causes the mutual reorientation of V_L domains by some degrees, as observed in the different structures (Fig. 4B-top); and, (ii) CDR conformational variability (and traceability in the electron density). In some cases (H6, H9, H10) CDR3 (residues 89–97) participates in the interface; both CDR3s in the H7 dimer lack electron density, while in M7 and M8 CDR3 residues are located far from the V_L–V_L interface.

Despite these fine differences, all the structures here presented, except for M8, display an overall conserved V_L–V_L quaternary assembly and association interface (Table S3C and Fig. 4B and E). In the structures of all H LCs and in M7, the association interface is roughly symmetric, and is located on the five-stranded beta sheet (strands 1, 3, 4, 7, 8), where the 3, 4, 7, 8 strands establish the intermolecular contacts. Figure 4D shows the region involved in the V_L–V_L interface together with the most relevant and conserved residues. Only the V_L–V_L interface in M8 is markedly different, one V_L domain being rotated and resulting in an asymmetric V_L–V_L interface (Fig. 4E). As for the other structures, in one V_L domain of M8 strands 3, 4, 7 and 8 participate to the dimeric interface, while only strands 1 and 8 of the facing/rotated V_L domain provide association contacts.

The calculated free energy gain associated with dimer formation, and the resulting dimer interface areas (Table S4), do not correlate with amyloidogenicity. In particular, M8 LC, which is the most stable LC against temperature (Table 2), displays the smallest dimer interface area with the lowest calculated free energy change for quaternary assembly, compared to all other considered LCs (Table S4).

Discussion

The five M LCs presented in this study are found in patients at extremely high concentrations, much higher than the average concentration of the H LCs (Table 1 and S1), and yet, through the years, the M proteins remain soluble and do not aggregate. Our study aimed to explore the biophysical and structural properties that correlate with LC amyloid aggregation; to reach most general conclusions, we analysed a set of thirteen patient-derived LCs, focusing on some of the commonly proposed structural and biophysical determinants of protein aggregation and toxicity[31], such as fold stability, protein overall hydrophobicity, protein dynamics and flexibility, and loss or heterogeneity of 3D-structures. The first general conclusion we can draw from the data here presented is that none of such biophysical properties, taken alone, allows clustering of the H *versus* the M LCs; on the contrary, amyloidogenicity seems to stem from the co-existence of more than one of the biophysical factors explored.

According to our assessments of fold stability through CD and ANS fluorescence, the examined H LCs display Tm_{app} values lower than the M LCs; on average the Tm_{app} for H LCs is 53.6 °C *vs*. an average Tm_{app} value of 62.8 °C for the M LCs. Such a trend is mirrored by the average onset temperature for the ANS signal, which is 55.0 °C for H LCs vs. 63.0 °C for the M LCs. Such an overall observation would be in keeping with the idea that amyloidogenicity correlates with lower fold stability, as previously suggested[32]; however, M7 and M10 are notable exceptions showing Tm_{app} values comparable with some of the H LCs. However, M7 was found present at high concentration *in vivo*, and yet no aggregation or toxicity was observed at presentation and during the later follow-up times (Table 1 and Table S1). This marked variability stresses the importance of using large set of proteins and it likely explains the contained differences in Tm_{app} recently observed by others between amyloidogenic and M LCs[23]: Andrich *et al.* reports an interesting biophysical characterisation of nine full length LCs but only four are of the λ isotypes and thus could be compared[23].

Intriguingly, the low ANS fluorescence observed under native conditions suggests that bulk surface hydrophobicity does not play a main role in determining the H LC toxicity in their native state. Indeed, LC overall hydrophobicity (computed from the amino acid sequences - Table 2) indicates a comparable amount of hydrophobic

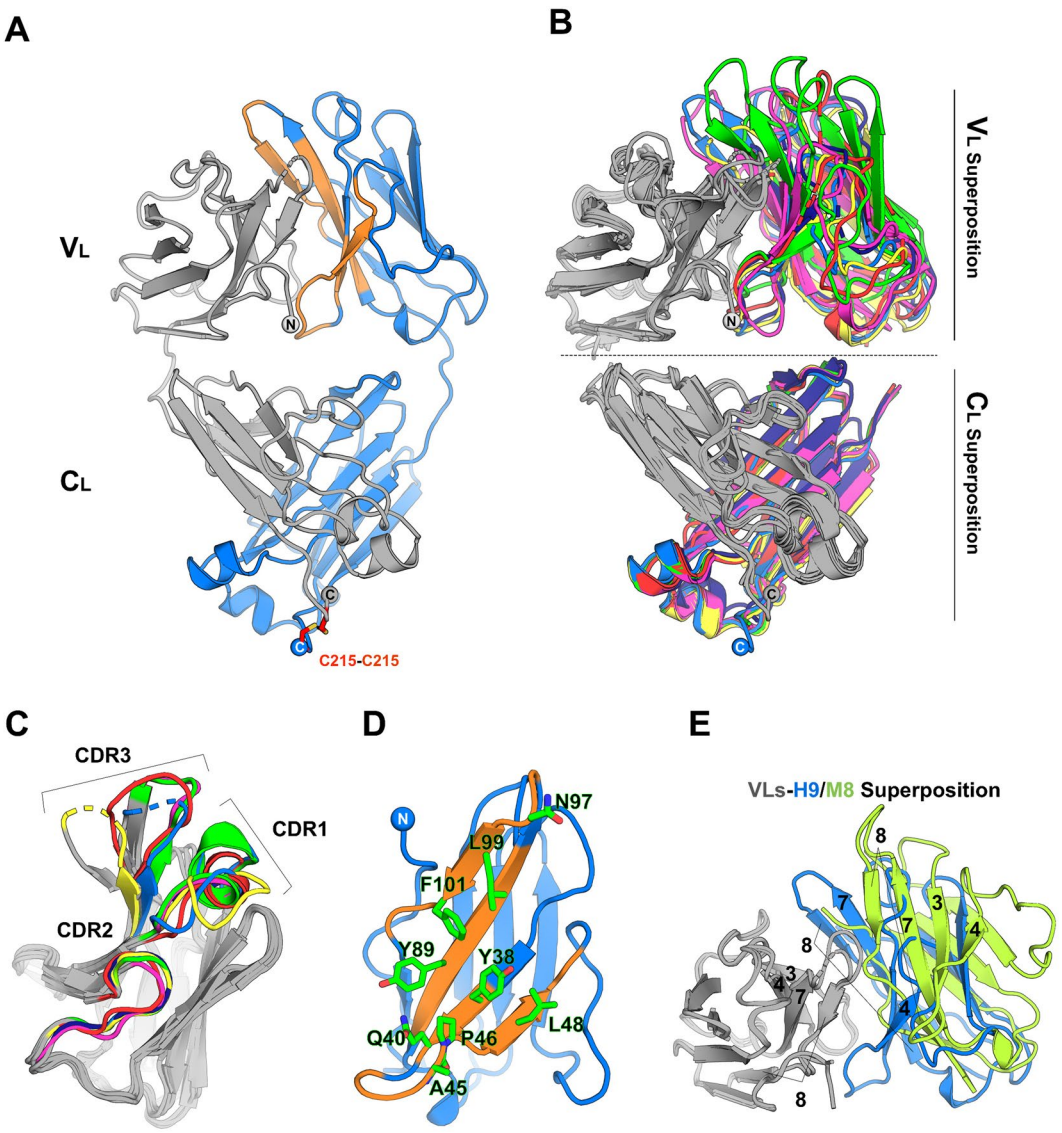


Figure 4. (**A**) Cartoon model of the crystal structure of the H9 homodimer, as representative of the tertiary and quaternary organization of all the LC structures determined in this work. The two LC monomers are coloured in grey and blue. The V_L interface region on the blue monomer is coloured in orange. The spheres indicate the position of N/C *termini*. (**B**) Superposition of the dimeric V_L (top panel) and C_L (bottom panel) domains. One V_L/C_L domain (coloured in grey) was fixed for all the seven structures and the second is coloured according to the different LCs (H3 yellow; H6 green; H7 dark blue; H9 blue; H10 magenta; M7 red; M8 lime green). V_L domain from M8 is shown only in panel E. (**C**) Superposition of a single V_L domain (grey) from each of the seven LC structures. The complementarity-determining regions (CDRs) belonging to different LCs are coloured following the B panel colour code. (**D**) Cartoon model of a single V_L domain from the structure of H9 showing in orange the regions involved in the V_L – V_L interaction. The residues represented as sticks indicate the positions involved in the V_L – V_L interface that are conserved among of all LCs structures. (**E**) Cartoon representation of H9 (grey-blue) and of M8 (grey-lime green) where the grey V_L are superposed and oriented as in B (top panel). The different orientation of the second V_L domain is apparent in the H9 dimer, chosen as an example, compared to the M8 dimer. Labels indicate the β-strand identification number.

residues in all thirteen LCs. The Far-UV CD spectra that on average display a slightly more intense signal for M LCs (Figure S1) suggest that typically the latter contain very regular β structure, while H LCs spectra vary more markedly; however, the inference value of these observations is essentially of qualitative nature.

Contrary to the limited information offered by the spectroscopic analyses, controlled proteolysis provided a robust indication that M LCs are more rigid, or less dynamic, than H LCs. These are however divided into two subsets: H3, H6 H7, H9 and H16 are very efficiently proteolysed, while H10, H15 and H18 show slower cleavage kinetics. Inspection of the SDS-PAGE gels shows that the proteolysis pattern is different in different H LCs, suggesting that the sequence of proteolytic events, and not only their kinetics, may be characteristic of each H LC. It is also remarkable that fold stability and kinetics of proteolysis do not necessarily correlate: M7 displays the lowest

Tm_{app} among the M LCs, but also very slow proteolysis kinetics; H3, H6 and H9 show relatively high Tm_{app}, within the H group, but are almost instantly proteolysed by trypsin. Recently, Morgan *et al*. reported faster kinetics of proteolysis for three amyloidogenic LCs due to kinetic instability of the native state[22]. In the present study we cannot distinguish whether the different kinetics of proteolysis are due to thermodynamic or kinetic instability, which are anyway linked to increased protein dynamics.

Among the experiments here reported, limited proteolysis provided the best correlation between LC behaviour and amyloidogenicity, showing a pronounced proteolytic trend in the H LC group; such consideration may be relevant under different aspects. Increased protein dynamics is considered a risk factor for aggregation propensity[33]. Furthermore, previous studies suggest that the LC fragments released by proteolytic cleavage of full length LCs are more amyloidogenic than the LCs themselves, and may be important species favouring the process of LC aggregation[21,22,34]. Indeed, the existence of LC fragments in amyloid deposits *in vivo* has been uncontrovertibly demonstrated[4,34]. Although the site (*i.e*. whether it occurs in the circulation, in the extracellular space or inside cells) and timing (before or after amyloid formation) of LC proteolysis is still a matter of debate, the strong correlation between amyloidogenicity and the ability to be easily proteolysed is stimulating, and suggests that H LCs may release pathogenic peptides much more efficiently than M LCs.

Although crystals could not be grown for the whole set of LCs, the relevant number of crystal structures solved, allows us to extend the discussion to the relationships between the LC biophysical and structural properties. Firstly, and somewhat unexpectedly, the most temperature- and trypsin-sensitive LCs (H3, H6, H7, H9) could be successfully crystallized. Such results indicate that, although characterized by low stability and high dynamics within the group tested, these LCs display a properly folded native structure devoid of large disordered regions that would hamper crystal growth. Secondly, the high level of conservation of the tertiary and quaternary assemblies in the seven LC structures here reported strongly suggests that the differences observed in the biophysical traits mainly depend on sequence variability, and not on major structural rearrangements. Such a consideration would imply that the intrinsic sequence variability, mainly located in the V_L domains, does not translate into extended conformational changes but is responsible for the LC behaviour at the molecular level (reflected by the biophysical traits and amyloidogenicity).

Intriguingly, plotting percentage of uncleaved LC at the end of the proteolysis experiments against Tm_{app} (determined by Far-UV) shows that H and M LCs could be clearly clustered (Figure S3). These observations suggest basic criteria for the design of ligands that may decrease LC amyloidogenicity. Molecules acting on the LC dimers, stabilizing the quaternary structure upon binding, as is the case of Tafamidis for the transthyretin tetramer (*i.e*. stabilization of the quaternary structure upon binding)[35], would result in an increased LC fold stability, and in a reduction of overall protein flexibility, the two biophysical traits that were shown here to correlate more strongly with LC amyloid propensity. Specifically, within LC dimers, targeting the dimer region encompassing the V_L domains appears as the proper strategy to combat LC amyloidogenicity at its biophysical roots: indeed, small-molecule ligands stabilizing V_L dimers have recently been shown to inhibit amyloid formation[36].

In conclusion, analysis and comparison of the different biophysical properties of H *vs*. M LCs suggests that no single molecular determinant by itself can account for the observed toxicity and aggregation trends, thus stressing the value of biochemical and biophysical studies based on a large pool of proteins rather than on pairwise comparisons. However, fold stability and protein dynamics (as assessed by proteolysis), but not surface hydrophobicity in the native state, or overall 3D-structure rearrangements of the native state, appear to play main roles in determining LC amyloidogenic behaviour. Our study therefore suggests that *in vivo* amyloidogenicity would be the result of concurrent biophysical traits that, as in our pool of proteins, may not necessarily all be present in each toxic LC at the same level.

Materials and Methods

Patients' samples. Urine and bone marrow plasma cells were obtained from patients during routine diagnostic procedures at the Amyloid Research and Treatment Center, Foundation IRCCS Policlinico San Matteo (Pavia, Italy). Acquisition, storage and use of biological samples for research purposes were approved by the Institutional Review Board of Fondazione IRCCS Policlinico San Matteo Pavia; all methods were performed in accordance with the relevant guidelines and regulations. Written informed consent was received from participants prior to inclusion in the study. The presence of tissue amyloid deposits and amyloid organ involvement were defined according to the International Consensus Panel Criteria[37,38]. LC cardiotoxicity was evaluated on the basis of clinical, echocardiography and biochemical parameters[39] (Table S1).

In parallel to amyloidogenic cardiotoxic LC (H), non-amyloidogenic LC from multiple myeloma patients (M) were used. All the monoclonal LCs included in the study, belong to the λ isotype.

Cloning of complete monoclonal free LC nucleotide sequences. Total RNA was extracted from 10^7 bone marrow mononuclear cells using TRIzol reagent (Life Technologies, Paisley, United Kingdom). Monoclonal variable (V_L) region nucleotide sequences were cloned by an inverse-PCR strategy that preserves the original sequence at 5′ and 3′ ends[40]. The PCR fragment was ligated into the pCR®2.1Vector (TA Cloning Kit; Life Technologies) and cloned into the TOP10 *E. coli* cells. After recombinant plasmid purification, insert was sequenced. In order to obtain the original full-length monoclonal LC (variable and constant regions, approximately 650 bp, from codons +1 to +215), standard RT-PCR was employed using the same marrow RNA, a forward patient-specific primer (dictated by codons +1 to +7 of the monoclonal V sequence) and a universal reverse Cλ carboxyterminal cloning primer, corresponding to the last amino acids of the constant region (codons +208 to +215, 5′-TGAACATTCTGTAGGGGCCACTGT-3′). To determine the presumed germline genes of V_L regions, sequence alignment was made with the current releases of EMBL-GenBank, V-BASE (V BASE Sequence Directory, MRC Centre for Protein Engineering, Cambridge, UK) and IMGT sequence directories. The gene sequences of the LCs here discussed have been deposited in the GenBank database: KC433670 (H3), KY471433

(H6), KC433671 (H7), KY471435 (H9), KY471432 (H10), KY471436 (H15), KY471437 (H16), KY471434 (H18), KY471441 (M2), KY471438 (M7), KY471439 (M9), KY471440 (M10).

LC purification from urines. LC were purified to homogeneity from 24 h urine collection. Urines, immediately combined with 0.1% sodium azide (w/v), were centrifuged at 3000 x g for 30 min. Ammonium sulfate was added to the supernatant (65% saturation) and, after overnight incubation, samples were centrifuged at 3000 x g for 30 min. The precipitates were solubilized in 20 mmol/L sodium phosphate, pH 7.0, and dialyzed against the same buffer. All steps were performed at 4 °C. LC were purified by anion exchange chromatography on an AKTÄ Purifier® FPLC system (GE-Healthcare, Piscataway, NJ, USA), using a HiPrep16/10 Q FF column, equilibrated in 20 mM sodium phosphate, pH 7.0. Bound proteins were eluted with a 0 up to 1 M sodium chloride linear gradient. H6, M2, M9 were purified using a cation exchanger column (HiPrep16/10 SP FF), equilibrated in 20 mM Tris-HCl, pH 8.0, and was eluted with a 0 up to 1 M sodium chloride linear gradient. The homogeneity of the isolated species was assessed by 12% SDS-PAGE. The final protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, IL, USA) and bovine serum albumin as standard.

Production of recombinant patient-derived LC. Recombinant LCs of selected patients were produced according to[30]. Briefly, heterologous proteins, produced in the cytoplasm as inclusion bodies, were retrieved and subjected to a renaturation procedure, followed by purification by means of ion exchange and size exclusion chromatography. Recombinant LCs were biochemically characterized by linear MALDI-TOF mass spectrometry and circular dichroism analyses, in order to verify sequence, homogeneity and correct folding. Gel filtration analysis indicates that all LCs used in this work were dimeric in solution (data not shown).

Circular dichroism spectroscopy. Circular dichroism experiments, in the Far- and Near –UV regions, were carried out on a J-810 spectropolarimeter (JASCO Corp., Tokyo, Japan) equipped with a Peltier system for temperature control. All experiments were carried out in 50 mM sodium phosphate pH 7.4. For the Far-UV region, protein concentration was 0.2 mg/mL in a cuvette with a pathlength of 0.1 cm. Spectra were recorded from 260 to 190 nm, whereas temperature ramps from 20 to 80 °C (monitored wavelength 202 nm, temperature slope 60 °C/hour). Spectra and temperature ramps were performed in triplicate for each LC except for M8 for which no BJ purified material is available. For the Near-UV region, protein concentration was 1 mg/mL in cuvettes with a pathlength 1 cm. Spectra were recorded from 350 to 250 nm, whereas temperature ramps from 20 to 80 °C (monitored wavelength 288 nm, temperature slope 60 °C/hour). Tm_{app} was calculated as the first-derivative minimum of the temperature ramps. Spectra recorded on cooled samples after temperature ramps confirmed that LC dimers unfold irreversibly as previously reported[22].

ANS fluorescence. 8-Anilino-1-naphthalenesulfonic acid (ANS) binding experiments were carried out at 20 °C and at 0.1 mg/mL protein concentration in 50 mM sodium phosphate pH 7.4. Each experiment was performed adding ANS to a final concentration of 100 µM. After every addition, ANS fluorescence emission spectra were recorded in the 420–550 nm range with excitation at 390 nm, excitation and emission slits were set at 5 nm, with a scanning speed of 50 nm/min. When ANS concentration was 100 µM, its signal was monitored at 490 nm along a temperature ramp starting from 20 to 80 °C (temperature slope 60 °C/hour) in a 1 cm path length cuvette.

Limited proteolysis. LCs at a concentration of 0.8 mg/mL, were incubated at 37 °C in 50 mM sodium phosphate, 1 M urea at pH 7.4, using a bovine trypsin/LC molar ratio of 1:100. The first sample was collected immediately after trypsin addition and then after 10′, 20′, 30′, 60′, 90′, 120′, 150′, 180′. Subsequently, they were diluted in denaturating and reducing sample buffer (NuPAGE, Invitrogen), heated at 95 °C for 3 min and analysed by SDS-PAGE. Protein bands corresponding to uncleaved LC monomers were quantified by densitometric analysis using Chemidoc™ MP System (Bio-Rad). In order to rule out that the very fast proteolysis of H3, H6, H7, H10 was due to partial or total unfolding several controls were performed in presence of 1 M urea: first the Far-UV spectra are superposable to the ones without urea; secondly the temperature unfolding curves monitored by Far-UV indicate that the unfolding process is starting at a temperature beyond 37 °C. For three LCs (H3, H6 and H9), which displayed very fast kinetics of proteolysis, parallel experiments were performed also in the presence of 0.5 M and 0 M urea, confirming under both conditions a trend for very fast proteolysis for these H LCs (see Figure S1). The latter observations further confirm that the fast kinetics of proteolysis were not artifacts related to excessive urea concentration.

Crystallization and X-ray structure determination. LCs were crystallized using sitting drops or hanging drops techniques. Each protein was solubilized in 50 mM sodium phosphate pH 7.4 at a concentration of 8.5–10 mg/mL at 20 °C. Crystals were obtained in: H3: 0.1 M Sodium cacodylate pH 6.5, 27% w/v PEG 2000 MME (Stura screen, Molecular dimensions); H6: 0.1 M Bicine pH 9.0, 2% w/v 1,4-Dioxane, 10% w/v PEG 20 K (Crystal Screen I/II, Hampton); H7: 0.1 M HEPES pH7.5, 10% 2-propanol, 20% w/v PEG 4 K (JBS screen, Jena Bioscience); H9: 0.1 M Sodium citrate pH 5.5, 16% w/v PEG 4000, 10% v/v 2-propanol (Stura screen, Molecular dimensions); H10: 0.05 M KBr, 30% PEG 2000 (JCSG screen, Molecular dimensions); M7: 0.1 M MMT (Malic acid, MES and Tris-base buffer) pH 4.0, 25% w/v PEG 1.5 K (microseeding with M8 crystals) (PACT screen, Molecular dimensions); M8: 0.2 M sodium acetate pH 4.6, 2.0 M NaCl (Crystal Screen I/II, Hampton).

Crystals were cryoprotected adding 33% glycerol to mother liquor and then flash frozen in liquid nitrogen. X-ray diffraction data were collected at ESRF (European Synchrotron Radiation Facility of Grenoble–France) at the beam lines: ID29, ID30, ID23-2, BM14. The diffraction data were analysed and processed using MOSFLM and XDS[41,42], the crystal symmetry was then verified by POINTLESS[43] and the intensities were merged and scaled with SCALA[44]. The crystal structures were determined by molecular replacement using PHASER, BALBES and MOLREP[45–47]. In order to perform the molecular replacement for H7, the full length LC (pdb: 1JVK)[10] was used

as search model. Then for the subsequent molecular replacements, LC structures determined in house were used. H9 initial model was generated by ARP-wARP[48]. The initial models were subjected firstly to a rigid-body refinement and then to a restrained and TLS refinement using Phenix Refine, Refmac5 and Buster[49–51]. Manual model building, water picking and structure analysis were then performed using Coot[52]. Dimer interface analysis shown in Table S3 was performed using PISA[53]. Fab elbow angles were calculated with phenix.fab_elbow_angle[49]. Figures of crystallographic structures were done using PyMOL and CCP4mg[54].

The atomic coordinates and the structure factors of the seven structures of LCs have been deposited in the Protein Data Bank with the following accession numbers: 5MTL (H3), 5MUD (H6), 5MUH (H7), 5M6A (H9), 5M76 (H10), 5MVG (M7), 5M6I (M8).

References

1. Merlini, G. & Bellotti, V. Molecular mechanisms of amyloidosis. *N Engl J Med* **349**, 583–596 (2003).
2. Merlini, G. & Palladini, G. Light chain amyloidosis: the heart of the problem. *Haematologica* **98**, 1492–1495, https://doi.org/10.3324/haematol.2013.094482 (2013).
3. Merlini, G., Wechalekar, A. D. & Palladini, G. Systemic light chain amyloidosis: an update for treating physicians. *Blood* **121**, 5124–5130, https://doi.org/10.1182/blood-2013-01-453001 (2013).
4. Enqvist, S., Sletten, K. & Westermark, P. Fibril protein fragmentation pattern in systemic AL-amyloidosis. *The Journal of pathology* **219**, 473–480, https://doi.org/10.1002/path.2607 (2009).
5. Palladini, G. *et al.* New criteria for response to treatment in immunoglobulin light chain amyloidosis based on free light chain measurement and cardiac biomarkers: impact on survival outcomes. *J Clin Oncol* **30**, 4541–4549, https://doi.org/10.1200/JCO.2011.37.7614 (2012).
6. Diomede, L. *et al.* A Caenorhabditis elegans-based assay recognizes immunoglobulin light chains causing heart amyloidosis. *Blood*, https://doi.org/10.1182/blood-2013-10-525634 (2014).
7. Marin-Argany, M. *et al.* Cell Damage in Light Chain Amyloidosis: FIBRIL INTERNALIZATION, TOXICITY AND CELL-MEDIATED SEEDING. *J Biol Chem* **291**, 19813–19825, https://doi.org/10.1074/jbc.M116.736736 (2016).
8. Palladini, G. *et al.* Circulating amyloidogenic free light chains and serum N-terminal natriuretic peptide type B decrease simultaneously in association with improvement of survival in AL. *Blood* **107**, 3854–3858, https://doi.org/10.1182/blood-2005-11-4385 (2006).
9. Shi, J. *et al.* Amyloidogenic light chains induce cardiomyocyte contractile dysfunction and apoptosis via a non-canonical p38alpha MAPK pathway. *Proc Natl Acad Sci USA* **107**, 4188–4193, https://doi.org/10.1073/pnas.0912263107 (2010).
10. Bourne, P. C. *et al.* Three-dimensional structure of an immunoglobulin light-chain dimer with amyloidogenic properties. *Acta crystallographica. Section D, Biological crystallography* **58**, 815–823 (2002).
11. Baden, E. M. *et al.* Altered dimer interface decreases stability in an amyloidogenic protein. *J Biol Chem* **283**, 15853–15860, https://doi.org/10.1074/jbc.M705347200 (2008).
12. Villalba, M. I. *et al.* Site-directed mutagenesis reveals regions implicated in the stability and fiber formation of human lambda3r light chains. *J Biol Chem* **290**, 2577–2592, https://doi.org/10.1074/jbc.M114.629550 (2015).
13. DiCostanzo, A. C., Thompson, J. R., Peterson, F. C., Volkman, B. F. & Ramirez-Alvarado, M. Tyrosine residues mediate fibril formation in a dynamic light chain dimer interface. *J Biol Chem* **287**, 27997–28006, https://doi.org/10.1074/jbc.M112.362921 (2012).
14. Pelaez-Aguilar, A. E. *et al.* Inhibition of Light Chain 6aJL2-R24G Amyloid Fiber Formation Associated with Light Chain Amyloidosis. *Biochemistry* **54**, 4978–4986, https://doi.org/10.1021/acs.biochem.5b00288 (2015).
15. Peterson, F. C., Baden, E. M., Owen, B. A., Volkman, B. F. & Ramirez-Alvarado, M. A single mutation promotes amyloidogenicity through a highly promiscuous dimer interface. *Structure* **18**, 563–570, https://doi.org/10.1016/j.str.2010.02.012 (2010).
16. Blancas-Mejia, L. M. *et al.* Kinetic control in protein folding for light chain amyloidosis and the differential effects of somatic mutations. *J Mol Biol* **426**, 347–361, https://doi.org/10.1016/j.jmb.2013.10.016 (2014).
17. Feige, M. J. *et al.* The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity. *Proc Natl Acad Sci USA* **105**, 13373–13378, https://doi.org/10.1073/pnas.0802809105 (2008).
18. Hurle, M. R., Helms, L. R., Li, L., Chan, W. & Wetzel, R. A role for destabilizing amino acid replacements in light-chain amyloidosis. *Proc Natl Acad Sci USA* **91**, 5446–5450 (1994).
19. Lavatelli, F. *et al.* A novel approach for the purification and proteomic analysis of pathogenic immunoglobulin free light chains from serum. *Biochim Biophys Acta* **409-419**, 2011, https://doi.org/10.1016/j.bbapap.2010.12.012 (**1814**).
20. Blancas-Mejia, L. M. *et al.* Thermodynamic and fibril formation studies of full length immunoglobulin light chain AL-09 and its germline protein using scan rate dependent thermal unfolding. *Biophysical chemistry* **207**, 13–20, https://doi.org/10.1016/j.bpc.2015.07.005 (2015).
21. Klimtchuk, E. S. *et al.* The critical role of the constant region in thermal stability and aggregation of amyloidogenic immunoglobulin light chain. *Biochemistry* **49**, 9848–9857, https://doi.org/10.1021/bi101351c (2010).
22. Morgan, G. J. & Kelly, J. W. The kinetic stability of a full-length antibody light chain dimer determines whether endoproteolysis can release amyloidogenic variable domains. *J Mol Biol*, https://doi.org/10.1016/j.jmb.2016.08.021 (2016).
23. Andrich, K. *et al.* Aggregation of Full-length Immunoglobulin Light Chains from Systemic Light Chain Amyloidosis (AL) Patients Is Remodeled by Epigallocatechin-3-gallate. *J Biol Chem* **292**, 2328–2344, https://doi.org/10.1074/jbc.M116.750323 (2017).
24. Marin-Argany, M., Guell-Bosch, J., Blancas-Mejia, L. M., Villegas, S. & Ramirez-Alvarado, M. Mutations can cause light chains to be too stable or too unstable to form amyloid fibrils. *Protein Sci* **24**, 1829–1840, https://doi.org/10.1002/pro.2790 (2015).
25. Comenzo, R. L. *et al.* Clonal immunoglobulin light chain variable region germline gene use in AL amyloidosis: association with dominant amyloid-related organ involvement and survival after stem cell transplantation. *British journal of haematology* **106**, 744–751 (1999).
26. Dasari, S. *et al.* Proteomic detection of immunoglobulin light chain variable region peptides from amyloidosis patient biopsies. *Journal of proteome research* **14**, 1957–1967, https://doi.org/10.1021/acs.jproteome.5b00015 (2015).
27. Perfetti, V. *et al.* 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**, 948–953, https://doi.org/10.1182/blood-2002-01-0114 (2002).
28. Abraham, R. S. *et al.* Immunoglobulin light chain variable (V) region genes influence clinical presentation and outcome in light chain-associated amyloidosis (AL). *Blood* **101**, 3801–3808, https://doi.org/10.1182/blood-2002-09-2707 (2003).
29. Kourelis, T. V. *et al.* Presentation and Outcomes of Localized Immunoglobulin Light Chain Amyloidosis: The Mayo Clinic Experience. *Mayo Clinic proceedings* **92**, 908–917, https://doi.org/10.1016/j.mayocp.2017.02.016 (2017).
30. Rognoni, P. *et al.* A strategy for synthesis of pathogenic human immunoglobulin free light chains in E. coli. *PloS one* **8**, e76022, https://doi.org/10.1371/journal.pone.0076022 (2013).
31. Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890, https://doi.org/10.1038/nature02261 (2003).
32. Ramirez-Alvarado, M. Amyloid formation in light chain amyloidosis. *Current topics in medicinal chemistry* **12**, 2523–2533 (2012).

33. De Simone, A. *et al.* Experimental free energy surfaces reveal the mechanisms of maintenance of protein solubility. *Proc Natl Acad Sci USA* **108**, 21057–21062, https://doi.org/10.1073/pnas.1112197108 (2011).
34. Lavatelli, F. *et al. Amyloidogenic and associated proteins i*n systemic amyloidosis proteome of adipose tissue. *Molecular & cellular proteomics: MCP* **7**, 1570–1583, https://doi.org/10.1074/mcp.M700545-MCP200 (2008).
35. Bulawa, C. E. *et al.* Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proc Natl Acad Sci USA* **109**, 9629–9634, https://doi.org/10.1073/pnas.1121005109 (2012).
36. Brumshtein, B. *et al.* Inhibition by small-molecule ligands of formation of amyloid fibrils of an immunoglobulin light chain variable domain. *Elife* **4**, e10935, https://doi.org/10.7554/eLife.10935 (2015).
37. Gertz, M. A. *et al.* 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. *American journal of hematology* **79**, 319–328, https://doi.org/10.1002/ajh.20381 (2005).
38. Gertz, M. A. & Merlini, G. Definition of organ involvement and response to treatment in AL amyloidosis: an updated consensus opinion. *Amyloid: the international journal of experimental and clinical investigation: the official journal of the International Society of Amyloidosis* **17**, 48 (2010).
39. Palladini, G. *et al.* Serum N-terminal pro-brain natriuretic peptide is a sensitive marker of myocardial dysfunction in AL amyloidosis. *Circulation* **107**, 2440–2445, https://doi.org/10.1161/01.CIR.0000068314.02595.B2 (2003).
40. Perfetti, V. *et al.* Inverse polymerase chain reaction for cloning complete human immunoglobulin variable regions and leaders conserving the original sequence. *Analytical biochemistry* **239**, 107–109, https://doi.org/10.1006/abio.1996.0297 (1996).
41. Kabsch, W. XDS. *Acta crystallographica. Section D, Biological crystallography* **66**, 125–132, https://doi.org/10.1107/s0907444909047337 (2010).
42. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4+ ESF-EACMB Newsletter on Protein Crystallography* (1992).
43. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82, https://doi.org/10.1107/S0907444905036693 (2006).
44. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763 (1994).
45. Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132, https://doi.org/10.1107/S0907444907050172 (2008).
46. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
47. Vagin, A. A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *Journal of Applied Crystallography* **30**, 1022–1025 (1997).
48. Lamzin, V. S., Perrakis, A. & Wilson, K. S. The ARP/wARP software suite. *International Tables for Crystallography* **F**, 720–722 (2001).
49. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213–221, https://doi.org/10.1107/s0907444909052925 (2010).
50. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255 (1997).
51. Bricogne, G. *et al.* BUSTER version 2.11.4. *Cambridge, United Kingdom: Global Phasing Ltd* (2016).
52. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).
53. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774–797, https://doi.org/10.1016/j.jmb.2007.05.022 (2007).
54. Potterton, L. *et al.* Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr* **60**, 2288–2294 (2004).

Acknowledgements

Dr. Pietro Sormanni and Benedetta Maria Sala are kindly acknowledged for helpful discussion and technical support. PR, FL, GP and GM are supported by grants from the 'Associazione Italiana per la Ricerca sul Cancro' Special Program Molecular Clinical Oncology 5 per mille n. 9965; from the Cariplo Foundation (2015-0591, 2013-0964) and from the Italian Ministry of Health 'Ricerca Finalizzata 2013' (RF-2013-02355259). FL is also supported by 'Giovani Ricercatori 2010' (GR-2010-2317596). FL and SR are supported by Cariplo Giovani (2016-0489).

Author Contributions

L.O., P.R., A.B., R.R. and M.M. performed the experiments; S.R., G.M., G.P., F.L. and A.B. designed the study; S.R., M.B. and G.M. wrote the paper.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-16953-7.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity

Luca Oberti[1], Paola Rognoni[2], Alberto Barbiroli[3], Francesca Lavatelli[2], Rosaria Russo[1], Martina Maritan[1], Giovanni Palladini[2], Martino Bolognesi[1], Giampaolo Merlini[2] and Stefano Ricagno[1]*

Supplementary information

		Clinical features								Cardiac features					
LC code	**Germline**	**Gender, age**	**Diagnosis**	**Organs involved**	**Serum λ FLC (mg/l)**	**κ/λ FLC ratio**	**dFLC (mg/l)**	**Proteinuria (g/24h)**	**Creatinine (mg/dl)**	**Cardiac stage°**	**NT-proBNP (ng/l)BNP* (ng/L)**	**cTnI (ng/ml)**	**IVS (mm)**	**PW (mm)**	**EF (%)**
H3	1c (IGLV1-44)	M, 65	AL	H	252	0.081	231.5	0.18	1.76	III	4491	0.35	16	16	42
H6	1b (IGLV1-51)	F, 72	AL	H, K, PNS	248	0.05	235	0.7	0.81	III	18731	0.207	18.5	16	50
H7	1b (IGLV1-51)	M, 45	AL	H	477	0.01	469	0.33	0.98	III	8882	0.16	19	19	45
H9	2c (IGLV2-8)	M, 59	AL	H, ST	699	0.02	689	8	1.19	III	3530	0.152	17.5	17	50
H10	1a (IGLV1-36)	M, 73	AL	H, L	475	0.04	454.5	0.16	1.07	III	7296	0.347	15.5	15	45
H15	6a (IGLV6-57)	M, 53	AL	H, PNS, ST	839	0.002	837.5	0.98	0.74	III	1444	0.222	16.8	16	70
H16	2a2 (IGLV2-14)	M, 72	AL	H, K	383	0.05	363.6	8.1	2.73	III	1926*	1.188	14.9	14.4	61
H18	3l (IGLV3-19)	M, 69	AL	H, ST, PNS	509	0.01	500.4	0.82	0.97	III	3839	0.345	21.5	18	61
M2	2b2 (IGLV2-23)	M, 65	MM (follow-up: 1.8 y)	-	1140	0.001	1138.5	0.12 §	0.89	n.a.	201	n.a.	11	11	65
M7	3l (IGLV3-19)	F, 71	MM (follow-up: 5.2 y)	-	6130	0.001	6124	0.52 §	2.07	n.a.	42*	0.007	9	9	65
M8	2b2 (IGLV2-23)	M, 48	MM (follow-up: 5.6 y)	-	573	0.011	567	1.87§	0.84	n.a.	14.5	0.003	10	10.5	67
M9	2b2 (IGLV2-23)	M, 61	MM (follow-up: 2.7 y)	-	8510	0.00	8508	7.24 §	4.92	n.a.	74*	0.021	10.5	10.5	62
M10	2a2 (IGLV2-14)	M, 55	MM (follow-up: 1.6 y)	-	12200	0.00	12198.5	0.9 §	1.1	n.a.	88.2	0.014	11	10.5	65

Table S1 (Previous page): Main clinical and biochemical characteristics of the light chains used in this study, at the time of diagnosis and urine collection. Multiple myeloma patients did not show evidence of amyloid deposits or amyloid organ involvement throughout the follow-up time (indicated in years). § Entirely constituted by Bence-Jones proteinuria (monoclonal urinary free light chains). °According to Gertz *et al*. [1]. Abbreviations: M, male; F, female; H, Heart; K, Kidney; ST, Soft Tissues; PNS, Peripheral Nervous System; BJ, Bence Jones; FLC, Free Light Chains; BNP, Brain Natriuretic Peptide; cTnI, cardiac Troponin I; IVS, Interventricular Septum; PW, Posterior Wall; EF, Ejection Fraction; n.a., not applicable.

Reference ranges: serum λ FLC <26.3 mg/l, κ/λ ratio 0.26-1.65; serum creatinine <1.18 mg/dl in men, <1.02 mg/dl in women; NT-proBNP 5 <332 ng/l; BNP, <99 ng/l; cTnI <0.04 ng/ml.

Table S2 (next page): Data collection and refinement statistics.

[a] $R_{merge} = \Sigma_{hkl} \Sigma_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \Sigma_j I_{hkl,j}$, where I_{hkl} is the observed intensity and $\langle I_{hkl} \rangle$ is the average intensity for the hkl reflection.

[b] $R_{work} = \Sigma_{hkl} |F_o - F_c| / \Sigma_{hkl} F_o$ for all data except 5–10%, which were used for the R_{free} calculation.

Values given in parenthesis refer to the high-resolution shell.

Structure	H3	H6	H7	H9	H10	M8	M7
Beam Line	ID23-2 (ESRF)	ID30 (ESRF)	ID29 (ESRF)	ID29 (ESRF)	ID29 (ESRF)	ID 23-1 (ESRF)	ID29 (ESRF)
Space group	P1	I 1 2 1	C2	P 1	P 21 21 21	I 4 2 2	P 31 2 1
Unit cell constants (Å)	a = 46.48, b = 65.08 c =88.39, α= 84.96° β = 86.66° γ=79.93°	a = 80.06 , b = 72.81 c = 84.99, α= 90° β = 105.29° γ= 90°	a = 100.01, b = 72.82 c = 73.72, α= 90° β = 121.19° γ= 90°	a = 46.49 , b = 63.43 c = 86.64, α= 95.32° β = 85.84° γ= 99.58°	a = 55.49 , b = 71.42 c = 103.93, α= 90° β = 90° γ= 90°	a = 229.38 , b = 229.38 c = 64.43, α= 90° β =90 ° γ= 90°	a = 70.72 , b = 70.72 c = 171.95, α= 90° β =90 ° γ= 120°
Resolution (Å)	28.57 – 2.45 (2.54 – 2.45)	52.98 – 2.34 (2.42 – 2.34)	55.45 – 2.70 (2.74 – 2.70)	45.77 – 1.64 (1.70 – 1.64)	48.95 - 2.50 (2.60 - 2.50)	57.3 – 2.20 (2.32 – 2.20)	57.70 – 2.20 (2.27 – 2.20)
R_{merge} (%)	12.6 (61.0)	4.5 (51.2)	7.7 (57.4)	8.7 (54.0)	10.8 (51.0)	27.1 (254.8)	7.70 (53.6)
I/σI	8.8 (2.2)	16.6 (2.5)	7.9 (1.6)	7.3 (1.6)	8.6 (2.3)	12.4 (2.3)	26.0 (4.6)
Completeness (%)	98.2 (97.9)	96.0 (95.8)	97.0 (98.2)	95.7 (94.2)	99.2 (97.9)	99.7 (99.6)	100.0 (100.0)
Multiplicity	3.3 (3.2)	3.5 (3.5)	3.5 (3.2)	3.6 (3.7)	4.3 (4.1)	14.2 (14.6)	10.9 (11.1)
Unique reflections	36581 (3606)	19281 (1908)	12871 (1137)	113746 (11177)	14723 (1407)	43679 (6279)	26134 (2224)
Refinement							
R_{work} (%)	23.42	20.70	26.92	18.01	23.87	19.50	22.15
R_{free} (%)	27.68	25.00	32.10	20.07	28.24	22.65	25.14
Average B factor, all atoms (Å²)	38.5	59.7	48.5	25.7	38.2	30.6	47.7
Number of atoms	5940	3143	2718	7156	3169	3419	3134
Protein	5896	3121	2692	6235	3110	3158	2870
Waters	44	22	26	882	49	259	258
Heteroatoms	-	-	-	39	1	2	6
Ramachandran plot, *n* (%)							
Most favoured region	94.97	95.95	91.0	97.51	97.13	97.86	94.92
Allowed region	4.51	3.57	8.0	2.49	2.87	2.14	5.08
Outliers	0.52	0.48	1.0	0	0	0	0
Elbow Angle (°)	116.7	108.0	124.0	116.2	112.7	157.4	127.7

RMSD (Å/Cα)	H3	H6	H7	H9	H10	M8	M7
H3	———						
H6	0.50/196	———					
H7	0.87/188	0.61/188	———				
H9	0.56/196	0.58/198	1.05/188	———			
H10	0.70/194	0.63/196	1.60/189	0.61/196	———		
M8	0.53/196	0.56/198	0.61/189	0.50/198	0.59/196	———	
M7	0.72/192	0.62/194	0.78/189	0.62/194	0.48/194	0.56/194	———

Table S3A: Root mean square C deviations (r.m.s.d.) calculated for the superposition of C_l dimers.

RMSD (Å/Cα)	H3	H6	H7	H9	H10	M7	M8
H3	———						
H6	0.79/92	———					
H7	0.67/87	0.87/92	———				
H9	0.78/93	0.77/107	0.81/92	———			
H10	0.83/92	0.75/108	0.90/92	0.75/105	———		
M7	1.23/70	0.97/105	0.107/91	0.83/103	0.98/104	———	
M8	0.74/93	0.84/105	0.77/92	0.62/107	0.80/103	0.90/104	—---

Table S3B: R. m. s. d. calculated for the superposition of V_l monomers.

RMSD (Å/Cα)	H3	H6	H7	H9	H10	M8
H3	———					
H6	3.06/171	———				
H7	1.06/168	2.95/161	———			
H9	1.28/186	2.58/174	1.10/171	———		
H10	1.41/188	3.55/183	1.44/172	1.68/210	———	
M8	3.06/140	2.65/129	3.46/122	2.80/151	2.72/152	———

Table S3C: R.m.s.d. calculated for the superposition of V_l dimers. The M7 structure has not been used in this calculation given the partial traceability of one of the V_l domains.

LC	Solvation energy, chain A* (kcal/mol)	Solvation energy, chain B* (kcal/mol)	Dimer interface chain A (Å²)	Dimer interface chain B (Å²)
H3	-9.2	-9.3	1504	1486
H6	-7.4	-8.2	1564	1568
H7	-8.7	-7.3	1444	1347
H9	-8.6	10.8	1577	1569
H10	-12.1	-11.3	1670	1671
M7	-10.2	-9.1	1336	1568
M8	-6.8	-8.7	1137	1125

Table S4: dimer interface analysis. (*): Energy gain on complex formation.

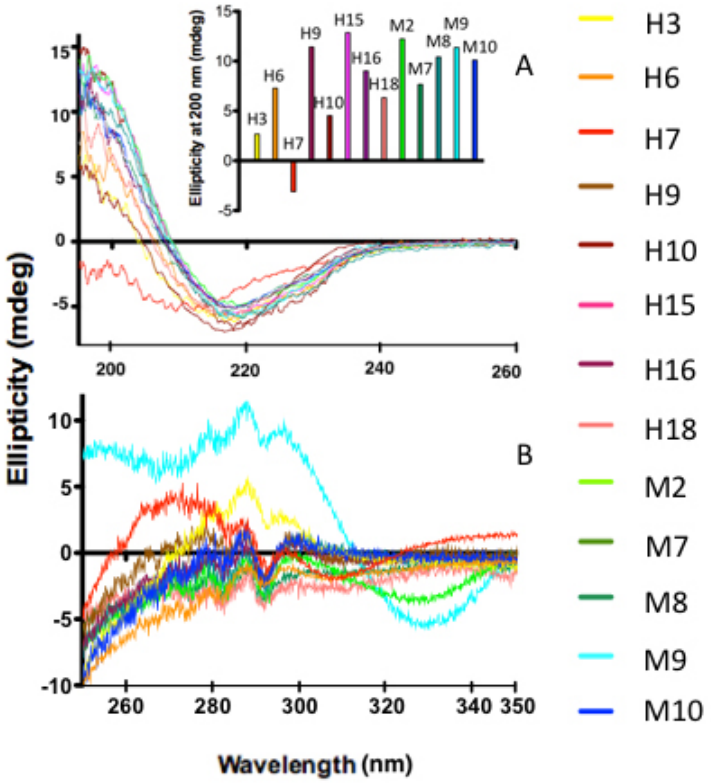


Figure S1: Inspection of LC secondary and tertiary structures. (A) panel A shows the Far-UV CD spectra: all LCs display spectra with quite comparable shapes, typical of proteins with high β-structure content: unexpectedly, the H7 spectrum, which however is well reproducible, is atypical and not superposable to all other spectra. Since H7 was successfully crystallized (see below) and showed cooperative unfolding curves (Figure 2), we can exclude that this spectrum indicates an aggregated or an unfolded species. Closer inspection of all the LC spectra shows that M LCs cluster together, displaying, on average, a slightly stronger CD signal (inset), as if the M LCs were typically characterized by an overall slightly more regular β-structure in solution. (B) Near-UV CD spectra are considered a tertiary structure fingerprint. The well conserved shape shared by the recorded protein spectra, suggests that LCs maintain a comparable structure in solution. However, both the number of near-UV chromophores (mainly Trp and Tyr residues) and their chemical environments are not strongly conserved in different LCs due to sequence variability in the V_L domains. Therefore, the resulting near-UV CD spectra show expected, although unpredictable, LC-specific differences, *e.g.* the M9 spectrum is more intense than all the other LCs.

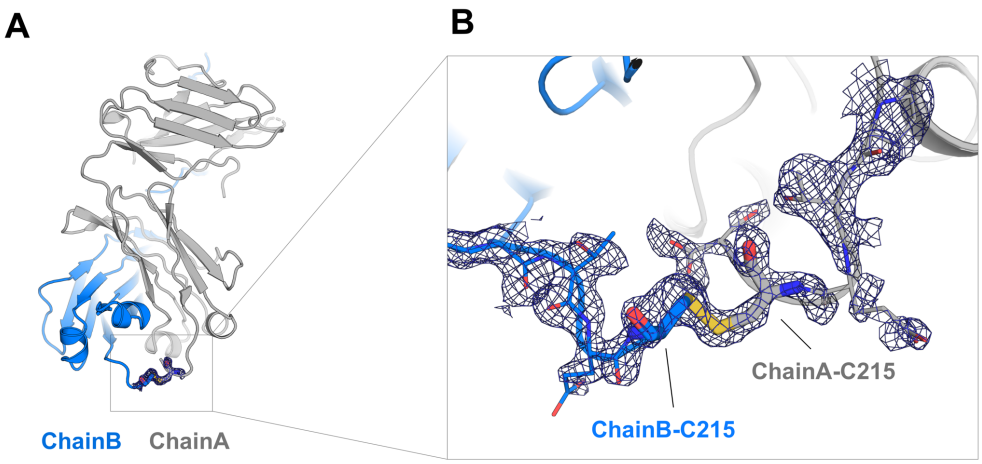


Figure S2: Intermolecular disulphide bond in the H9 dimer. (A) Cartoon representation of the overall fold of the H9 homodimer, having at the C-terminal a disulphide bond linking C215 the two LC monomers (ChainA and ChainB, colored in grey and blue, respectively). (B) Zoom into the well-defined electron density of C215-C215 disulphide bond (sticks representation) and the surrounding C-terminal residues (lines representation). 2*F*o-*F*c electron density map is contoured at 1σ.

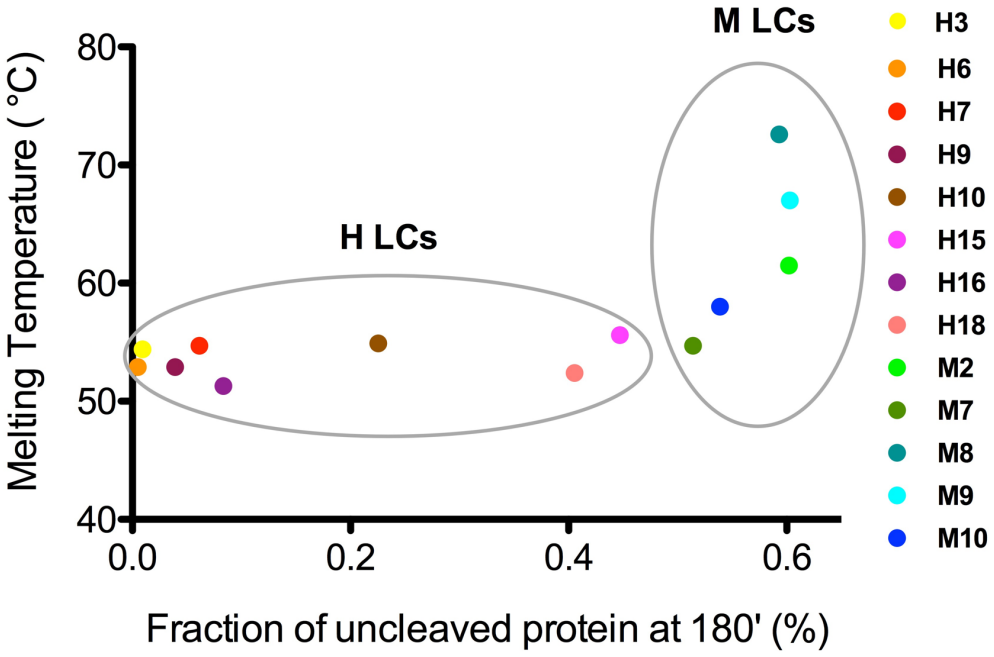


Figure S3: Plot of uncleaved LC at the end of the proteolysis experiments as shown in Figure 3A (180 minutes) against melting temperatures (Tm) monitored by Far-UV (see Table 2). H and M-LCs are colour coded as in Figures 2 and 3. The H and M LC sets are circled highlighting the clustering between H and M LCs.

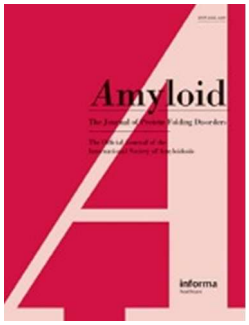
References:

1 Gertz, M. A. & Merlini, G. Definition of organ involvement and response to treatment in AL amyloidosis: an updated consensus opinion *Amyloid : the international journal of experimental and clinical investigation : the official journal of the International Society of Amyloidosis* **17**, 48 (2010).

The concurrency of several biophysical traits links immunoglobulin light chains with toxicity in AL amyloidosis

My contribution in this work consisted in summarizing the data obtained in order to provide a short overview on our experiments.

This work has been submitted to *Amyloid* on July 2018 and it is in review.



The concurrency of several biophysical traits links immunoglobulin light chains with toxicity in AL amyloidosis

Journal:	*Amyloid*
Manuscript ID	DAMY-2018-0173
Manuscript Type:	Letter to the Editor
Date Submitted by the Author:	30-Jul-2018
Complete List of Authors:	Oberti, Luca; Università degli studi di Milano, Dipartimento di bioscienze Maritan, Martina; Università degli studi di Milano, Dipatimento di bioscienze Rognoni, Paola; Amyloidosis Research and Treatment Center, Foundation Scientific Institute San Matteo, Department of Molecular Medicine, University of Pavia Barbiroli, Alberto; Universita degli Studi di Milano, DeFens Lavatelli, Francesca; Fondazione IRCCS Policlinico San Matteo, Amyloid Research and Treatment center Russo, Rosaria; Università degli studi di Milano, Dipartimento di bioscienze Palladini, Giovanni; University of Pavia and Fondazione IRCCS Policlinico San Matteo, Amyloidosis Research and Treatment Center Bolognesi, Martino; University of Milan, Dept. of Biosciences Merlini, Giampaolo; Amyloidosis Research and Treatment Center, Foundation Scientific Institute San Matteo, Department of Molecular Medicine, University of Pavia Ricagno, Stefano; Università degli studi di Milano, Dipartimento di bioscienze
Keywords:	Immunoglobulin light chain, Amyloidosis, Biophysics, protein flexibility, fold stability

The concurrency of several biophysical traits links immunoglobulin light chains with toxicity in AL amyloidosis

Luca Oberti[1], Martina Maritan[1], Paola Rognoni[2], Alberto Barbiroli[3], Francesca Lavatelli[2], Rosaria Russo[4], Giovanni Palladini[2], Martino Bolognesi[1], Giampaolo Merlini[2] and Stefano Ricagno[1]

[1] *Department of Bioscience, University of Milan, Italy,*
[2] *Amyloidosis Research and Treatment Center, Fondazione IRCCS Policlinico San Matteo, and Department of Molecular Medicine, University of Pavia, 27100, Pavia, Italy,*
[3] *DeFENS - Dep. of Food, Environmental and Nutritional Sciences, University of Milan, Italy,*
[4] *Department of Physiopathology and Transplantation, University of Milan, Italy*

Address for correspondence: Stefano Ricagno, Department of Bioscience, University of Milan, Via Giovanni Celoria 26, 20133, Milan, Italy. E-mail: stefano.ricagno@unimi.it

Background. Light chain amyloidosis (LC-AL) is the most common systemic amyloidosis. It is caused by the overproduction and the aggregation of toxic and monoclonal immunoglobulin light chains (LC) in target organs. Among all the organs injured by the pathology, the heart is the most affected one. In particular, the ventricular compliance is reduced, resulting in a symptomatic congestive heart failure [1]. It is extremely relevant that due to the genetic rearrangement and somatic hypermutation, a high variability among LCs' sequences is generated. This means, virtually, that each AL patients present different amyloidogenic LCs [2], stressing in the necessity to investigate a large set of cases. To date, it is particularly interesting the observation described in Milani *et al.,* 2018 [1], where it was observed that the severity of heart symptoms is linked with the blood concentration of full-length LCs, which are the major circulating species. Specifically, these facts highlight a crucial role for full-length LC in the pathology beyond the amyloid deposits. However, up to now, despite the several works conducted, the molecular bases of proteotoxicity and the aggregation mechanism(s) are still unclear.

Material and methods. Our approach consisted in a biophysical and structural characterization. LCs were purified from patients' urine, or by the recombinant expression in *E.coli* [3]. Protein stability was evaluated by circular dichroism (CD) and by anilinonaphthalene-1-sulfonic acid (ANS) fluorescence, while flexibility and dynamics was studied by limited proteolysis [3]. The X-ray diffraction experiments, on LCs crystals, were carried out at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France.

Results. In order to achieve as generalizable as possible data, our study is based on a large pool of thirteen λ full-length LCs divided in two groups. In particular, eight LCs are amyloidogenic and responsible of severe cardiac symptoms in AL patients (H LCs); five are non-amyloidogenic LCs and selected from patients affected by multiple myeloma, and used as control (M LCs). In order to highlight potential determinants of LCs proteotoxicity and aggregation, all proteins were extensively structurally and biophysically characterized. Firstly, X-ray diffraction of seven LCs crystals (five H LCs and two M LCs) reveals that H and M LCs structures match very closely. From the biophysical point of view, the melting temperature measurements assessed by three independent spectroscopic techniques demonstrate that H LCs tend to be less stable than M LCs (fig. 1A). Moreover, limited proteolysis by trypsin and protease K strongly indicate that a more pronounced flexibility for H LCs compared to M LCs (fig. 1B). These findings suggest thermodynamic properties and protein dynamics to have a role in the molecular proteotoxicity mechanism(s). In particular, since the crystal structures are conserved, it is very likely that dynamics plays a critical role in defining the propensity of a LC to be toxic or not.

Recently, Diomede *et al.,* 2017 [4] described H LCs toxicity *in vivo* to be increased by the presence of copper ions. Along this line, we are investigating potential interactions between LCs and bivalent cationic metal ions, including copper. In particular, we observe a direct and specific binding between H LCs and Cu^{2+}. Our findings show that the LC-copper interaction destabilizes the proteins and it seems to turn H LCs in a more flexible state, underling the importance of dynamics in determining LC toxicity.

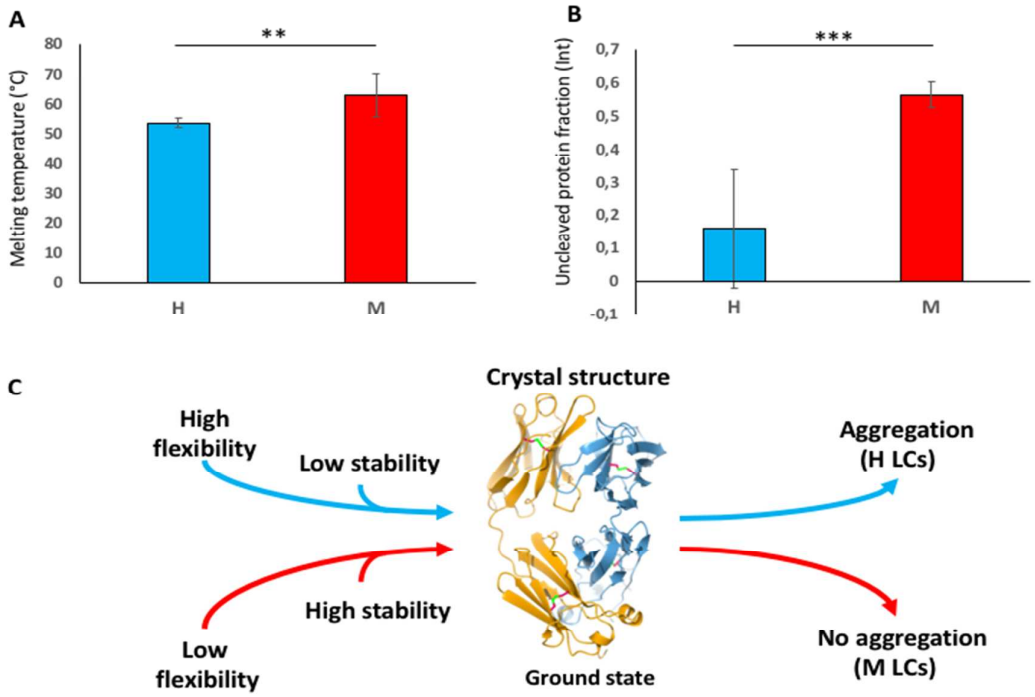


Figure 1. A. Average of H LC and M LC melting temperatures. **B.** Average of the H and M uncleaved protein fraction at 180 min after limited proteolysis; *** *p-value* < 0.0005; ** *p-value* < 0.005. **C.** Graphical representation of the biophysical properties contributing in the amyloidogenic properties.

Discussion and conclusions. The new contribute of this work highlights the concurrency of different biophysical traits to be linked with LCs amyloid propensity. Our data suggest that thermal stability and flexibility/dynamics correlate with the proteotoxicity LCs tendency, whereas the overall structural determinants are conserved between H and M LCs (fig. 1C).

Declaration of interest. The authors report no conflict of interest.

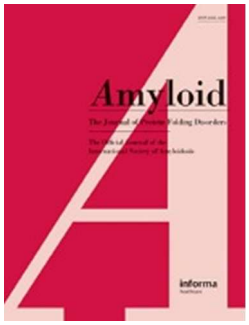
References

1. Milani, P. et al. Light Chain Amyloidosis. Mediterranean Journal of Hematology and Infectious Diseases. 2018;10:e2018022.
2. Merlini G, Palladini G. Light chain amyloidosis: the heart of the problem. Haematologica. 2013; 98:1492-5.
3. Oberti, L. et al. Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity. Scientific reports. 2017; 7.1.
4. Diomede, L. et al. Cardiac light chain amyloidosis: the role of metal ions in oxidative stress and mitochondrial damage. Antioxidants & redox signalling. 2017; 27.9: 567-582.

Modulating the cardiotoxic behaviour of immunoglobulin light chain dimers through point mutations

My contribute in this work consisted in the collaboration for the LCs production and in the evaluation of protein flexibility.

This work has been submitted to *Amyloid* on July, 2018 and now it's in review.



Modulating the cardiotoxic behaviour of immunoglobulin light chain dimers through point mutations

Journal:	*Amyloid*
Manuscript ID	DAMY-2018-0171
Manuscript Type:	Letter to the Editor
Date Submitted by the Author:	30-Jul-2018
Complete List of Authors:	Maritan, Martina; Università degli studi di Milano, Dipatimento di bioscienze Ambrosetti, Arianna; Università degli studi di Milano, Dipartimento di bioscienze Oberti, Luca; Università degli studi di Milano, Dipartimento di bioscienze Barbiroli, Alberto; Universita degli Studi di Milano, DeFens Diomede, Luisa; Istituto Di Ricerche Farmacologiche Mario Negri, Molecular Biochemistry and Pharmacology Romeo, Margherita; Istituto Di Ricerche Farmacologiche Mario Negri, Molecular Biochemistry and Pharmacology Lavatelli, Francesca; Fondazione IRCCS Policlinico San Matteo, Amyloid Research and Treatment center Sormanni, Pietro; University of Cambridge, Chemistry Palladini, Giovanni; University of Pavia and Fondazione IRCCS Policlinico San Matteo, Amyloidosis Research and Treatment Center Bolognesi, Martino; University of Milan, Dept. of Biosciences Merlini, Giampaolo; Amyloidosis Research and Treatment Center, Foundation Scientific Institute San Matteo, Department of Molecular Medicine, University of Pavia Ricagno, Stefano; University of Milan, Dept. of Bioscience
Keywords:	Amyloidosis, Immunoglobulin light chain, Biophysics, Point mutations, protein toxicity

Modulating the cardiotoxic behaviour of immunoglobulin light chain dimers through point mutations

Martina Maritan[1], Arianna Ambrosetti[1], Luca Oberti[1], Alberto Barbiroli[2], Luisa Diomede[3], Margherita Romeo[3], Francesca Lavatelli[4], Pietro Sormanni[5], Giovanni Palladini[4], Martino Bolognesi[1], Giampaolo Merlini[4] and Stefano Ricagno[1].

[1]*Department of Bioscience, University of Milan, Milan, Italy;*
[2]*Department of Nutritional Science, University of Milan, Milan, Italy;*
[3]*IRCCS-Istituto di Ricerche Farmacologiche "Mario Negri", Milano, Italy;*
[4]*Amyloidosis Research and Treatment Center, Policlinico San Matteo, Italy;*
[5]*Department of Chemistry, University of Cambridge, Cambridge, UK*

Address for correspondence: Stefano Ricagno, Department of Biosciences, University of Milan, 20133 Milano, Italy. E-mail: stefano.ricagno@unimi.it

Background. Light chain amyloidosis (AL) is caused by the overproduction, misfolding and aggregation of immunoglobulin light chains (LCs) that tend to deposit as amyloid fibrils in the cardiac tissue[1]. Nowadays, the molecular details underling LCs soluble cardiotoxicity and fibril formation remain to be fully elucidated. It has been suggested a relationship between conformational flexibility and amyloidogenicity, indicating protein thermal stability and dynamics as factors able to influence the complex processes of misfolding and aggregation[2,3]. To date, no compelling correlations have been identified between LCs cardiotoxicity and primary sequence. Here we present a pivotal mutagenesis study in which we produce and characterise a mutated variant of the toxic H6 LC (mH6). Our purpose is to abrogate the toxicity typical of H6 yielding to a non-cardiotoxic LC and to understand the biophysical and biochemical changes underlying such loss of toxicity.

Material and Methods. H6 mutant design was carried out combining the results of bioinformatics tools for protein stability predictions and sequence multi-alignment. H6 sequence was mutated into mH6 using side directed mutagenesis. *E. coli* was used to express LCs, which are then refolded and purified by chromatographic methods as previously reported[3]. For protein characterization both spectroscopic (Circular Dichroism, Fluorescence) and structural (X-ray crystallography) techniques were applied. *In-vivo* toxicity assessment was carried out measuring *C. elegans'* pharynx pumping rate[4].

Results. The design of mutated H6 was performed applying an innovative method that correlates structural data with toxicity (**Figure 1**). A set of thermal stabilizing mutation candidates were found through an hybrid approach involving: i) multi-sequence analysis and ii) energetic calculations. H6 amino acid sequence was aligned against a multiple myeloma sequence database, in order to identify the conserved amino acids spots associated with non-toxic LC (**Figure 1A**). The crystal structure of H6 was analysed by bioinformatics tools that calculated the most favourable mutations to achieve protein stabilisation (**Figure 1B**). Final mutations selection was obtained by merging and ranking the best outcomes of the two methods (**Figure 1C**), which led to the identification of three point mutations located in the variable domain.
Preliminary data on mH6 *in vivo* toxicity seem to indicate a significant decrease in the mutant toxic profile compared to the wild type cardiotoxic H6. This points out the critical role of specific amino acid positions in determining the toxic behaviour.
Interestingly, the melting temperatures of H6 and mH6 monitored by Far-UV circular dichroism are comparable (T_m H6= 43.9 – 54.5 °C, T_m mH6=53.8 °C). However unfolding profiles are not perfectly superposable suggesting different unfolding transitions: mH6

appear to have a more defined transition, which possibly reflect a more compact domain organization than H6 three step unfolding[3]. Limited proteolysis data revealed slower proteolytic kinetics for the mutant, hence reduced fold stability.

The crystal structure of mH6 has been determined at 2.1 Å resolution, displaying a homodimeric organization common to all LC dimers. Secondary structure superposition does not show evident differences between the H6 structure and the mH6 one, giving an r.m.s.d. value of 0.3 Å for the overall dimers superposition, indicating that the mutagenesis has no impact on the native protein organization and dimerization.

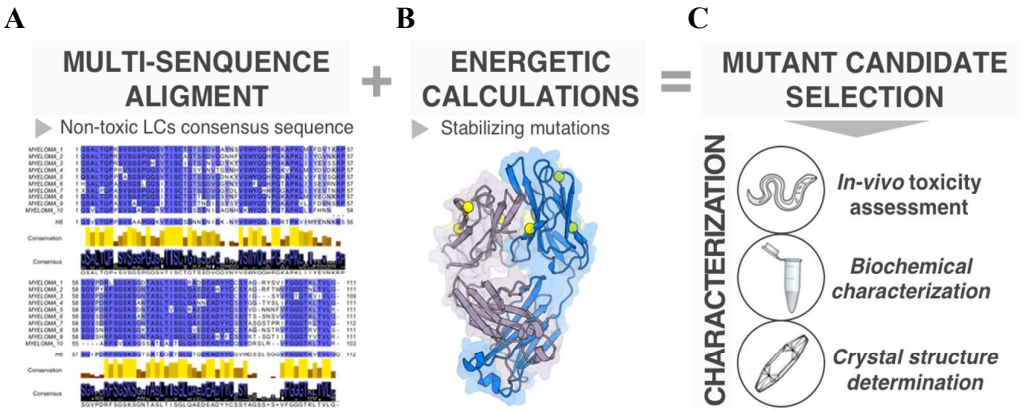


Figure 1. (A) Representative sequence multi-alignment of myeloma LCs against H6. Sequence conservation and consensus sequence are depicted by histograms. (B) Bioinformatics analysis using H6 structure as input run to identify possible mutation spots. (C) The combination of outputs from (A) and (B) lead the mutation spot selection and subsequent protein production and characterization.

Discussion and conclusions. Previous biophysical findings indicate that low fold stability and high protein dynamics correlate with LC amyloidogenicity[3]. Our data showed that the toxicity of a soluble LC could be tuned by introducing specific point mutations. This evidence points out the importance of the primary sequence for cardiotoxic LCs. Moreover, our data strengthen the hypothesis that links LCs toxic behaviour *in vivo* to thermal stability and dynamics.

Declaration of interest. The authors report no conflict of interest.

References

1 Ramirez-Alvarado, M. Amyloid formation in light chain amyloidosis. Current topics in medicinal chemistry. 2012; 2523-2533.

2 Poshusta, T. L., Katoh, N., Gertz, M. A., Dispenzieri, A. & Ramirez-Alvarado, M. Thermal stability threshold for amyloid formation in light chain amyloidosis. International journal of molecular sciences. 2013; 22604-22617.

3 Oberti, L. *et al.* Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity. Scientific reports. 2017; 7.1.

4 Diomede, L. *et al.* A Caenorhabditis elegans-based assay recognizes immunoglobulin light chains causing heart amyloidosis. Blood. 2014; 123:3543-3552.