

A Strategy for Synthesis of Pathogenic Human Immunoglobulin Free Light Chains in *E. coli*

Paola Rognoni¹, Francesca Lavatelli¹, Simona Casarini¹, Giovanni Palladini¹, Laura Verga², Paolo Pedrazzoli³, Giovanna Valentini⁴, Giampaolo Merlini¹, Vittorio Perfetti^{3*}

1 Amyloidosis Research and Treatment Center and Biotechnology Research Laboratories, Fondazione IRCCS Policlinico S. Matteo and Department of Molecular Medicine, University of Pavia, Pavia, Italy, **2** Pathologic Unit, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy, **3** Department of Hematology and Oncology, Section of Oncology, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy, **4** Department of Biology and Biotechnology "Lazzaro Spallanzani", University of Pavia, Pavia, Italy

Abstract

Monoclonal immunoglobulin light chains are normally synthesized in excess compared to the heavy chain partners and can be detected in serum and urine ("free" LC). Occasionally free LC are *per se* cause of organ toxicity, as in free LC-related disorders. In AL amyloidosis, the most common of these conditions, free LC with peculiar biophysical properties related to their primary structure damage target organs and organize in amyloid fibrils. Unlimited availability of well-characterized free LC is instrumental to investigate the toxic effect of these proteins and to study their interactions with targets. We present a straightforward strategy to obtain recombinant monoclonal free LC by using a bacterial system. These proteins, expressed as inclusion bodies, were subjected to solubilization and refolding procedures to recover them in native form. To minimize differences from the circulating natural LC, full-length recombinant LC were expressed, i.e. complete of variable and constant regions, with the original amino acid sequence along the entire protein, and with no purification tags. The strategy was exploited to generate free LC from three AL amyloidosis patients. After purification, recombinant proteins were biochemically characterized and compared to the natural Bence Jones protein isolated from one of the patients. Results showed that the recombinant free LC were properly folded and formed homodimers in solution, similar to the natural Bence Jones protein used for comparison. Furthermore, as proof of pathogenicity, recombinant proteins formed amyloid fibrils *in vitro*. We believe that the present strategy represents a valuable tool to speed research in free LC-related disorders.

Citation: Rognoni P, Lavatelli F, Casarini S, Palladini G, Verga L, et al. (2013) A Strategy for Synthesis of Pathogenic Human Immunoglobulin Free Light Chains in *E. coli*. PLoS ONE 8(9): e76022. doi:10.1371/journal.pone.0076022

Editor: Ellen R. Goldman, Naval Research Laboratory, United States of America

Received: June 12, 2013; **Accepted:** August 18, 2013; **Published:** September 27, 2013

Copyright: © 2013 Rognoni et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by Fondazione IRCCS Policlinico San Matteo, Grant 2009-2012 by AIRC (Associazione Italiana per la Ricerca sul Cancro) Special Program Molecular Clinical Oncology 5 per mille n. 9965, Italian Ministry of Health (GR-2010-2317596), Amyloid Foundation and Fondazione Mintas, Ghislieri College, Pavia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: v.perfetti@smatteo.pv.it

Introduction

Human immunoglobulin light chains (LC) consist of two domains, each of 100-110 amino acids (12 kDa, approximately), that fold independently of each other: the variable region domain (V_L), generated via rearrangement of a V to a J gene segment, and the constant region domain (C_L), of minimal variation. LC can occur in two isotypes, namely κ and λ [1]. The two LC domains have similar compact tertiary structure with a hydrophobic core composed of two twisted β -sheets, which are stabilized by a single intrachain disulfide bond. On the other hand, a cysteine (for a total of 5 cysteines in a single LC molecule) located at the carboxyterminal end, forms an interchain disulfide bond with the immunoglobulin heavy chain. LC are usually synthesized in large excess over heavy chains, and this results in production of "free" LC that can be detected in serum and urine [2]. Kappa free LC are normally monomeric and can be found with various degrees of glycosylation, while lambda free-light chains tend to be dimeric and unglycosylated [3-5].

Monoclonal free LC are not only a hallmark of clonal B-cell expansions; these molecules can be responsible of a variety of pathological conditions, including LC deposition disease, acquired Fanconi syndrome, LC cast nephropathy and AL amyloidosis, the

most common of systemic amyloidoses and of the free LC-related disorders [6].

In AL amyloidosis, bone marrow plasma cells secrete monoclonal free LC, typically of the λ isotype, that undergo extracellular deposition as fibrils (amyloid) at distant sites [7]. Since amyloid deposition occurs in just a fraction of subjects with monoclonal gammopathy and free LC production, and it recurs in transplanted organs if LC synthesis is not halted by chemotherapy, peculiar molecular features of free LC must be involved in determining protein "amyloidogenicity". Indeed, the V_L gene repertoire of amyloid LC is substantially skewed [8-10] and different from the normal one [8], with significant preferential expression of two V_L segments, namely *IGLV6-57* [8-10] and *IGLV3-1* [8,10], indicating the existence of amyloid-associated germline gene segments. Furthermore, by the analysis of the relationship between LC gene use and the two most frequent organ involved at diagnosis, kidney and heart, a close association was established between expression of *IGLV6-57* and renal involvement [8-10], and between *IGLV1-44* and dominant heart involvement [11]. This evidence provided important insights to understand the bases of amyloid organ tropism, a phenomenon

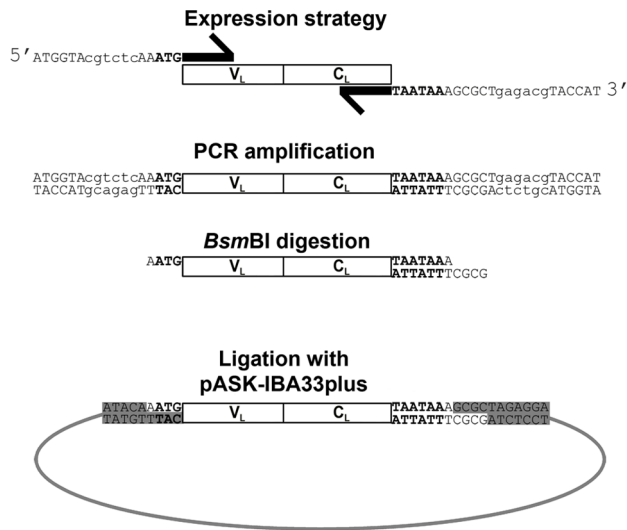


Figure 1. Schematic representation of the expression strategy. Each patient's full-length monoclonal LC ($V_L + C_L$) obtained by RT-PCR was re-amplified using primers designed according to the pASK-IBA cloning strategy. Start Met codon in 5' primer and the two stop codons in 3' primer were introduced and shown in bold, *BsmBI* recognition sites are indicated in lower case in both primers. After *BsmBI* digestion, PCR product was ligated into *BsaI/BsmI* sites of pASK-IBA33plus vector. doi:10.1371/journal.pone.0076022.g001

that influences clinical presentation, prognosis and therapeutic options.

It is now widely recognized that the presence of amyloid fibrils alone is not sufficient to explain the cellular and organ damage observed in patients. This aspect is thought to be a common element in all amyloid diseases [12,13]. In AL amyloidosis, a major clue towards this conclusion came from clinical evidence; variations in the concentration of circulating amyloidogenic free LC are immediately paralleled by concordant and proportional variations of NT-proBNP concentration, a marker of cardiac damage; changes in NT-proBNP concentrations, in turn, reflect rapid improvement or worsening of heart dysfunction, even in the absence of a reduction of cardiac amyloid load, as estimated by echocardiography [14]. These data indicated that cardiac dysfunction is primarily affected by circulating free LC, exerting direct cardiac toxicity. Indeed, *in vitro* studies on animal cultured cells showed alterations upon exposure to LC purified from urine [Bence Jones (BJ) proteins] of AL patients with heart involvement. In particular, a rapid induction of cell toxicity was reported in cultured rat cardiomyocytes and in isolated mouse hearts [15–17].

For the above reasons, it is predicted that further insights into the pathogenesis of these disorders may come from studies dealing with the soluble pathogenic molecules; however, discrete amounts of free LC are needed for these investigations, whilst these proteins are frequently found in low concentrations in human fluids [18] and are not always available. Recombinant (r) free LC production in *E. coli* may be a way to provide well-characterized proteins from archive cDNA. However, proteins containing multiple disulfide bonds, such as complete free LC ($V_L + C_L$), are difficult to produce in bacteria, and mainly accumulate in the cytoplasm as inclusion bodies, particles formed by densely packed denatured protein molecules [19]. For this reason, engineering the relevant cDNA with bacteria-derived peptide leader signals that allow for protein translocation into the more favorable, oxidizing environment of the bacteria periplasm is considered the method of choice for LC production [19]. However, potential drawbacks are suboptimal

yield, complex isolation and incomplete removal of the peptide leader.

In the present paper we devised a novel, straightforward strategy comprising inclusion body formation as a natural way to obtain consistent amounts of recombinant proteins virtually free of contaminants, with carefully optimized conditions for proper refolding and disulfide bond formation *in vitro*. Given the paramount relevance of the LC structure in the pathogenic process, the strategy was designed to obtain proteins as similar as possible to natural circulating LC: original primary sequence, full-length protein ($V_L + C_L$), no purification tags (usually consisting of a polyhistidine stretch added to one of the two amino acid terminals of the protein). We here describe free rLC production from two patients with predominant heart amyloidosis and one with exclusive kidney involvement.

Materials and Methods

Cloning of complete monoclonal free LC cDNA

Total RNA was extracted from 10^7 bone marrow mononuclear cells from three AL amyloidosis patients using TRIzol reagent (Life Technologies, Paisley, United Kingdom). Samples were obtained during clinical evaluation, after acquisition of the written informed consent for genetic investigations on patient's material. The study was approved by the IRCCS Policlinico San Matteo Ethics Committee. Monoclonal V_L region nucleotide sequences were cloned by an inverse PCR strategy that preserves the original sequence at 5' and 3' ends [20]. After sequencing, to obtain the original full-length monoclonal LC ($V_L + C_L$, approximately 650 bp, from codons +1 to +216), 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_L sequence) and a universal reverse C_L carboxyterminal cloning primer, corresponding to the last amino acids of C_L (codons +208 to +215, 5'-TGAACATTCTGTAGGGGCCACTGT-3'). In case of κ LC expression, the universal 3' C_L primer must be used: 5'-

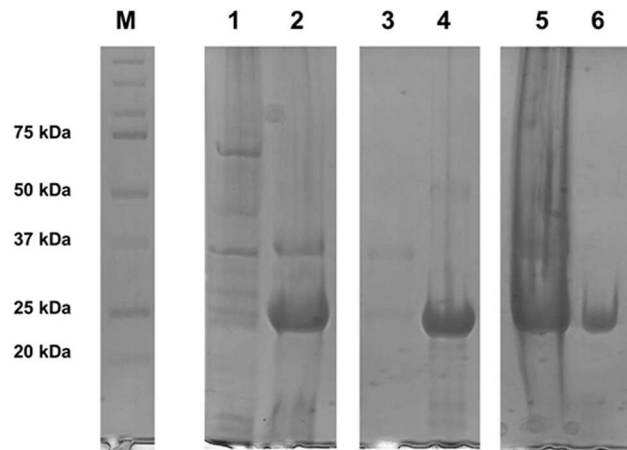


Figure 2. Different steps of rAL-1 renaturation from inclusion bodies. Protein samples were subjected to reducing 12% SDS-PAGE and stained with Coomassie Brilliant Blue. M: molecular mass markers; lane 1, supernatant of bacteria sonication; lane 2, pellet of cell sonication, containing inclusion bodies; lane 3, pellet of inclusion bodies solubilization; lane 4, supernatant of inclusion bodies solubilization; lane 5, pellet of recombinant protein refolding; lane 6, supernatant of recombinant protein refolding. A significant proportion of the expressed protein was soluble after refolding from inclusion bodies. doi:10.1371/journal.pone.0076022.g002

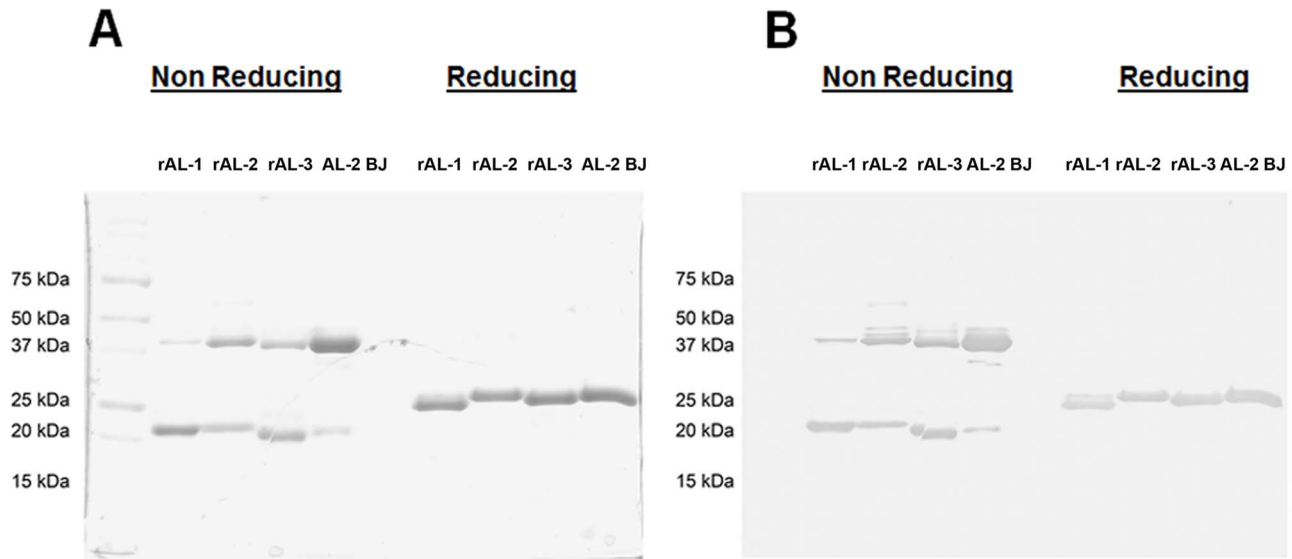


Figure 3. SDS-PAGE analysis of the purified rLC. Coomassie Brilliant Blue staining (A) and western blotting (B) analysis of purified rLC compared to AL-2 Bence Jones protein (AL-2 BJ), purified from urine. Protein samples (5 μ g loaded) were resolved on 12% SDS-PAGE run under non reducing and reducing conditions. Western blot analysis was performed using a rabbit anti-human λ LC antiserum as primary antibody (Dako Cytomation, Glostrup, Denmark), revealed by a goat anti-rabbit conjugated with horseradish peroxidase (Dako Cytomation), probed with diaminobenzidine (DAB) substrate. Molecular mass markers are shown in the left lane. The proteins presented the typical dimeric-monomeric species of free LC. Different degrees of protein denaturation are responsible of the minor differences in electrophoretic migration of monoclonal free LC (A), a common finding in SDS PAGE running under denaturing conditions of these proteins.
doi:10.1371/journal.pone.0076022.g003

ACACTCTCCCCTGTTGAAGCT-3'. The PCR fragment was sequenced from both sides, and sequences were submitted to GenBank (GenBank accession nos.: *KC433670*, *KC433671*, *KC433672*).

Construction of the expression vector encoding the recombinant free LC cDNA

The vector used to express monoclonal complete LC was the pASK-IBA33plus (IBA, Göttingen, Germany), carrying the inducible tetracycline promoter/operator and the ampicillin

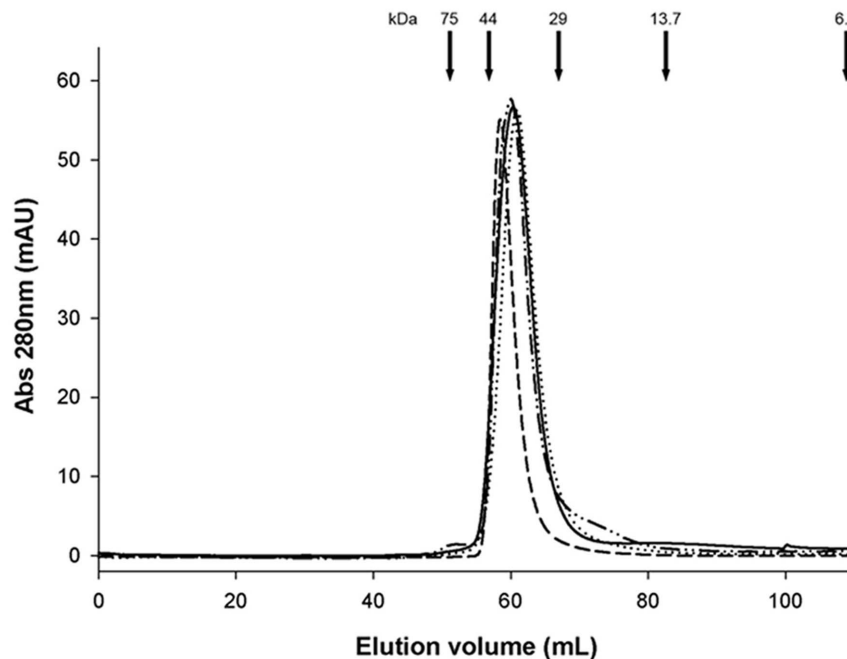


Figure 4. Size exclusion chromatography of rLC. rAL-1 (dotted), rAL-2 (short dashes), rAL-3 (dash-dot-dot) compared with AL-2 BJ protein (solid). Conalbumin (75.0 kDa), ovalbumin (44.0 kDa), carbonic anhydrase (29.0 kDa), ribonuclease A (13.5 kDa), aprotinin (6.5 kDa) were used as molecular mass standards. Virtually identical elution profiles were observed.
doi:10.1371/journal.pone.0076022.g004

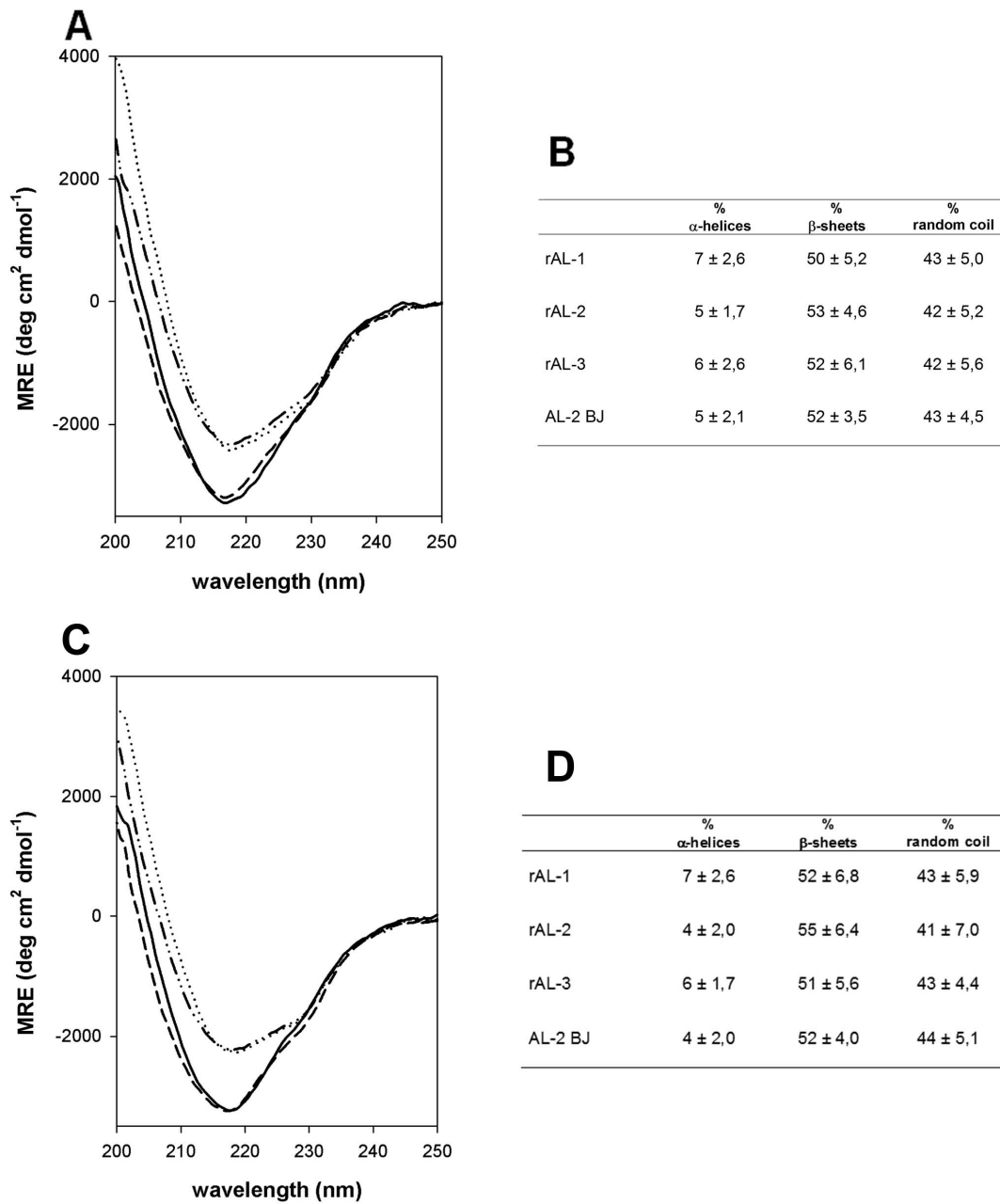
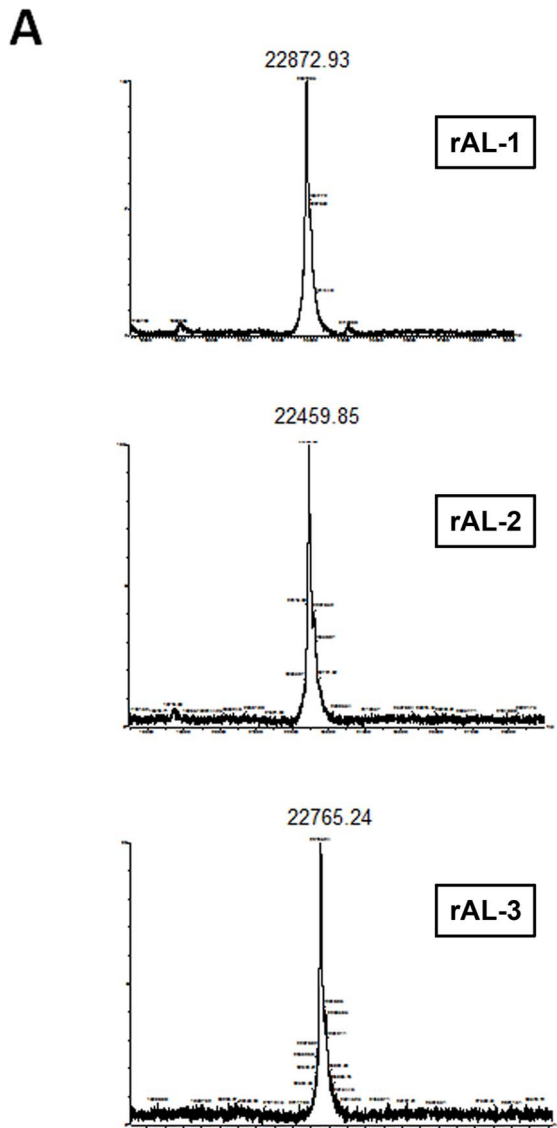


Figure 5. Circular dichroism analyses. Far-ultraviolet CD spectra (panel A, pH 7.4; panel C, pH 5.0) of purified rAL-1 (short dashes), rAL-2 (dotted), rAL-3 (solid) compared with the spectrum of patient AL-2 BJ protein (dash-dot-dot). Each CD spectrum represents the average of ten scans. Data are shown as mean residue ellipticity (MRE) as a function of wavelength. Panels B and D report the secondary structure content of the analyzed proteins (at pH 7.4 and 5.0, respectively), calculated with K2D, CDSSTR and CONTIN software applications. Data are expressed as mean \pm sd. Spectra were virtually identical for rAL-2 and the corresponding natural BJ (panels A and C), overlapping for the other two rLC (94% amino acid identity), and display the minimum at around 218 nm, characteristic of β -sheets. Proteins showed very similar percentage of secondary structures. doi:10.1371/journal.pone.0076022.g005

resistance cassette, originally designed for cytoplasmic protein expression with C-terminal His-tag. The recombinant monoclonal LC (Figure 1) was obtained via PCR reamplification (94°C, 5 sec; 65°C, 45 sec; 72°C, 45 sec; 30 cycles) of the V_L+C_L cDNA, using primers designed according to the pASK-IBA protein cloning strategy, which requires internal *Bsm*BI recognition sites both at the 5' and 3' primers (Figure 1). The patient-specific forward primer was: 5'-ATGGTAcgtctcAAATG-patient-specific V_L nucleotides-3', which included *Bsm*BI recognition site (lower case), initiation codon (bold) and patient-specific V_L +1 to +22

nucleotides (Figure 1). The universal λ reverse primer was: 5'-ATGGTAcgtctcAGCGCTTTATTATGAACATTCTGTAGGGGCCACTG-3', which included *Bsm*BI recognition site (lower case), and two stop codons (bold italic) after the last monoclonal LC constant region nucleotide triplet (Figure 1). For κ LC expression the following universal reverse primer is adopted: 5'-ATGGTAcgtctcAGCGCTTTATTAACTCTCCCCTGTTGAAGCT-3'. After *Bsm*BI digestion, each PCR product was ligated into *Bsa*I/*Bsa*I sites of the expression vector pASK-IBA33plus (Figure 1). Plasmid amplifica-



B

	MW (theoretical)	MW (experimental)	Δ ppm
rAL-1	22869	22873	174
rAL-2	22460	22459	45
rAL-3	22769	22765	175

Figure 6. Intact mass determination of purified rLC by MALDI-TOF mass spectrometry. Panel A shows the peaks corresponding to the single charged ions. Panel B shows the calculated (including the presence of an N-terminal methionine) and the observed intact masses of each LC. The error (ppm) from the calculated mass is reported. The results are consistent with the predicted molecular masses.
doi:10.1371/journal.pone.0076022.g006

tion was carried out into the *E. coli* DH5 α cells; insert sequence and reading frame were checked.

Recombinant free LC expression

E. coli BL21(DE3) cells (Merck Chemicals, Nottingham, UK) were transformed with the recombinant LC expression vectors and grown at 37°C in Luria-Bertani medium containing 100 μ g/mL ampicillin. When the culture optical density at 600nm reached a value of 0.5–0.6, expression was induced by anhydrotetracycline addition to a final concentration of 0.43 μ M. The protein expression was obtained incubating the induced culture at 37°C for 4 h. rLC were found as inclusion body proteins.

Isolation of inclusion bodies from *E. coli* cells

Typically, *E. coli* cells obtained from 1 L culture were collected by centrifugation (7000 g, 10 min) and resuspended in 70 mL ice-cold buffer (Buffer A: 50 mM TRIS-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid, 1 mM phenylmethylsulfonyl-fluoride). The cell suspension was sonicated at 200 W for 6 min with UP200S sonicator (Hielscher, Teltow, Germany), and centrifuged at 30000 g for 15 min at 4 °C. The pellet, containing inclusion bodies, was washed twice with ice-cold Buffer A.

Solubilization and refolding of monoclonal LC from inclusion bodies

The inclusion bodies were resuspended in 15 mL ice-cold denaturing buffer (Buffer A containing 6 M guanidinium chloride and 5 mM β -mercaptoethanol), incubated under mild agitation at 4 °C for 5 h and centrifuged at 30000 g for 15 min. Refolding of solubilized rLC into their bioactive form was obtained by a dilution method in the presence of reduced and oxidized glutathione [21,22]. To this end, protein solution containing the solubilized rLC was 20-fold diluted by dripping into an ice-cold refolding-redox shuffle buffer (Buffer A containing 5 mM reduced L-glutathione and 0.5 mM oxidized L-glutathione), under mild stirring. After overnight incubation at 4°C under mild agitation, the protein solution was centrifuged (30000 g, 15 min), and the supernatant was dialyzed overnight against 50 mM TRIS-HCl pH 8.0.

Recombinant LC purification

The recombinant proteins were purified by a single step procedure, consisting in an anion exchange chromatography on a HiPrep 16/10 Q FF pre-packed column (GE Healthcare, Piscataway, NJ, USA) equilibrated in 50 mM TRIS-HCl pH 8.0. Refolded rLC obtained from inclusion bodies were applied to the column, on an AKTAPurifier® FPLC system (GE Healthcare), and eluted by the use of a 0 to 1 M NaCl linear gradient. The homogeneity of the protein preparation was assessed by 12% SDS-PAGE and by immunoblotting, following standard procedures. The protein concentration was evaluated using Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, IL, USA); bovine serum albumin was used as standard.

Size exclusion chromatography

The purified rLC were subjected to size exclusion chromatography on an HiLoad 16/60 Superdex 75 pre-packed column (GE Healthcare) equilibrated in 50 mM sodium phosphate pH 7.4. Samples (50 μ g) were applied to the column on an AKTAPurifier® FPLC system and eluted with the equilibration buffer at 0.6 mL/min, monitoring at 280 nm. Column was calibrated with the following proteins: conalbumin (75.0 kDa), ovalbumin (44.0 kDa),

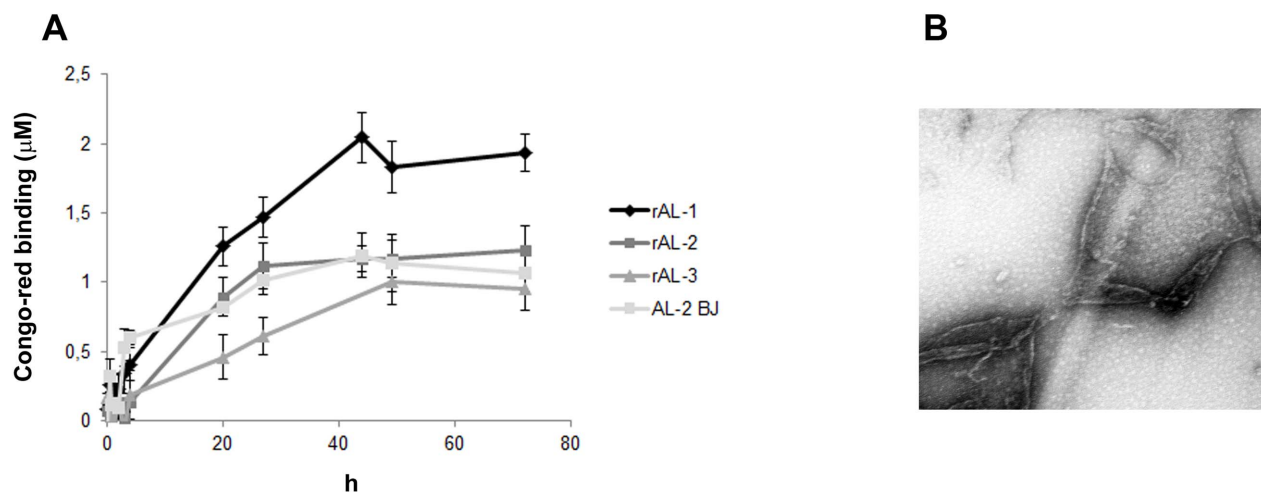


Figure 7. rLC *in vitro* fibrillogenesis. Congo-red binding profiles and electron microscopy. Panel A reports Congo-red binding patterns during fibrillogenesis. rAL-2 and AL-2 BJ revealed similar binding profiles. Panel B shows electron microscopy of typical amyloid fibrils generated *in vitro* by the rLC (fibrils from rAL-2 are shown). Magnification is 60000 x. doi:10.1371/journal.pone.0076022.g007

carbonic anhydrase (29.0 kDa), ribonuclease A (13.5 kDa), apro-tinin (6.5 kDa) (GE Healthcare).

Quantification of thiols in free LC

Free thiol content of purified proteins was quantified by the Measure-iT™ Thiol Assay kit (Invitrogen). The fluorescence of samples (tested from 5 to 30 pmol) was determined at $\lambda_{ex} = 490$ nm and $\lambda_{em} = 520$ nm in clear 96-well plates on an Infinite F200Pro plate reader (Tecan, Männedorf, Switzerland).

Circular dichroism spectroscopy

Far-UV circular dichroism (CD) measurements were performed with a Jasco J-700 spectropolarimeter (Jasco Europe, Cremella, Italy) using a quartz cuvette with a path length of 1 mm. Scans were conducted from 250 to 200 nm at a speed of 20 nm/min with a spectral band width of 2 nm, a sensitivity of 20 millidegrees and response time of 1 sec. CD spectra measurements were performed at 25°C in 50 mM sodium phosphate pH 7.4 and pH 5.0, and represent the average of 10 scans. The protein concentration was 0.4 mg/mL. The α -helical and β -sheet content was calculated with K2D, CDSSTR and CONTIN software applications [23–25].

Analysis of rLC by MALDI-TOF mass spectrometry

rLC were dialyzed overnight against 50 mM ammonium bicarbonate and reduced by addition of dithiothreitol (final concentration 10 mM, 5 min at 95°C, 45 min at 25°C). Intact mass determination was performed immediately after reduction by MALDI-TOF mass spectrometry, on a Micromass Micro MX instrument (Waters, Milford, MA, USA), operated in positive-ion, linear mode, over the 7500–30000 m/z range. Samples (protein concentration: 0.25–0.5 mg/mL) were co-crystallized onto target plate with the matrix α -cyano-4-hydroxycinnamic acid (CHCA, 1:1 v/v). External calibration was performed using a mixture of protein standards (cytochrome C, trypsinogen, myoglobin, all from Sigma Aldrich, Steinheim, DE). Spectra were processed and analyzed using MassLynx software (Waters). Theoretical masses of the intact proteins (from bone marrow-derived LC sequences, and from the same sequences upon addition of a methionine residue at the N-terminus) were calculated with the on-line

software tool MS-Digest (Protein Prospector, University of California, San Francisco, CA, USA).

AL-2 patient's BJ protein purification

BJ protein from AL-2 patient (AL-2 BJ) was purified from 24 h urine collection. The urine, immediately made 0.1% w/v NaN_3 to reduce degradation phenomena and bacterial proliferation and centrifuged at 3000 g for 30 min to remove cells and insoluble material, was subjected to ammonium sulfate precipitation (65% of saturation) and incubated overnight at 4 °C. The precipitate was collected by centrifugation (3000 g, 30 min) and dialyzed overnight against 10 mM sodium phosphate, pH 7.0. All steps were performed at 4°C.

Dialyzed proteins were applied to the pre-packed column HiPrep 16/10 Q FF (GE Healthcare), equilibrated in 10 mM sodium phosphate, pH 7.0 and eluted using a 0 to 1 M NaCl linear gradient. The homogeneity of the isolated species was assessed by 12% SDS-PAGE and immunoblotting. The protein concentration was evaluated using Pierce BCA Protein Assay Kit; bovine serum albumin was used as standard.

In Vitro Fibril Formation of rLC and AL-2 BJ

For fibril formation assays, 1 mg/mL LC samples were prepared in 10 mM TRIS-HCl pH 7.0, filtered (0.22 μm) and incubated at 60°C rocking at 250 rpm on a Innova 43R shaker (New Brunswick Scientific, Edison, NJ) for up to 72 h [26]. Aliquots (10 μL) were removed at various time points, added to a blank solution of 20 μM Congo-red (Fisher Scientific, Loughborough, UK) in 25 mM PBS buffer at pH 7.4 and left incubated for 30 min at room temperature. Fibrils formation was monitored by recording the UV spectrum between 400 and 700 nm on an Infinite M200Pro spectrophotometer (Tecan, Männedorf, Switzerland). Congo-red binding was calculated according to Klunk and coworkers [27].

Electron microscopy of amyloid fibrils

After 72 h [26], 10 μL samples from *in vitro* fibrillogenesis (see above), were applied to carbon-coated Formvar nickel grid (200 mesh) (Electron Microscopy Sciences, Washington, PA, USA). Samples were allowed to sediment on the carbon film for 15 min

and the negative staining was performed with 10 μL of 2% w/v uranyl acetate solution (Electron Microscopy Sciences). After draining off the excess of staining solution by means of a filter paper, the specimens were transferred to the electron microscope for examination, using a Philips CM12 transmission electron microscope, operating at 80 kV [28]. Electron micrographs of negatively stained samples were photographed on Kodak film. Magnification 60000 x.

Results and Discussion

In this paper we describe a strategy to obtain significant amounts of human immunoglobulin free LC by the use of a bacterial cell factory. By means of this novel procedure, it is possible to obtain recombinant proteins that closely reproduce the natural pathogenic LC, in this case from AL amyloidosis, a disorder characterized by altered, unstable, toxic free LC whose primary structure is recognized to be a determinant in the pathogenic process.

Monoclonal LC cloning and sequencing

In order to clone the V_L , we adopted a consolidated inverse PCR procedure to preserve the original nucleotide sequence at the extremities of the V_L region [8,11]. We consider this point of great relevance given the role in protein folding, and consequently in the function of the LC, of the first amino acids of V_L [29]. Somatic mutations leading to amino acid changes in these areas of the molecule are also quite common in amyloidosis [8,11]. Although we consider the use of inverse PCR preferable, alternative V_L amplification methods could be used without introducing modifications to the expression strategy.

Based on amino terminal sequence information, clone-specific primers were designed for standard RT-PCR to obtain the complete full-length LC sequence (V_L+C_L). Sequencing of the cloned cDNA fragments showed that LC were potentially functional, being devoid of stop codons or frameshift mutations. LC derived from somatic mutation of different germline gene segments, *IGLV1-44*

(AL-1 and AL-3) and *IGLV1-51* (AL-2). No changes from the reported λ C_L sequences were observed. In order to recreate proteins that were most similar to the natural circulating LC, and at variance from previous studies focused on truncated proteins [30–32], we synthesized the complete LC. Indeed, the relevance of working with the whole LC is becoming to be appreciated [33,34], given the important influence of the constant region in the thermodynamic properties of the LC and in the process of amyloid formation [33]. Indeed, it has been shown that full-length LC are more likely to initiate aggregation during unfolding and provide a template for the V_L deposition [33].

Expression vector construction and rLC production

The PCR fragment encoding the full-length cDNA of the amyloid LC was cloned into the expression vector pASK-IBA33plus (Figure 1), stop codons were introduced at 3' via PCR to express recombinant proteins free of the C-terminal His-tag. By means of this strategy (Figure 1), the rLC were predicted to be identical to the original one, with the last amino acid consisting of the invariant Ser216, which follows the Cys215 involved in the interchain disulfide bond, with just a single amino acid addition consisting of the 5' Met encoded by the initiating codon ATG. This extra codon was included into the expression primer to allow transcription (Figure 1). However, Met is expected not to be detrimental for the biochemical and biophysical properties of the protein, since methionine is a small and uncharged amino acid

residue, in a position that in LC is typically exposed to solvent and not involved in interactions with other amino acids [1]. The preservation of Cys215, at the end of C_L , is relevant for the process of dimerization [1], a natural tendency of the free LC (see below).

The rLC were expressed in BL21(DE3) cells. Protein production was monitored by SDS-PAGE analysis after induction. At the best time point (4 h of induction at 37 °C), after bacteria collection and sonication, the homogenate was centrifuged, supernatant (containing soluble proteins) and pellet were analyzed for recombinant protein content. As expected, recombinant products were recovered in the pellet fraction as inclusion bodies (Figure 2, lane 2). This feature was advantageous for purification (see below).

Recombinant LC solubilization and refolding

rLC were solubilized from inclusion bodies using reducing and denaturing conditions, working at low temperature, to prevent bacterial proteolysis. As shown in Figure 2, lane 4, the recombinant protein was found in the supernatant, indicating that the denaturing conditions allowed the total solubilization of the rLC.

To obtain rLC refolding with concomitant disulfide bond formation, a key step in the recovery of bioactive proteins from inclusion bodies, we set up a glutathione redox-shuffle [21,22].

condition, which involved an appropriate mixture of oxidized and reduced thiol reagents (reduced/oxidized L-glutathione, ratio 10/1 mM). A similar approach was successfully used for immunoglobulin Fab [35,36]. After refolding and centrifugation, supernatant was subjected to 12% SDS-PAGE (Figure 2, lane 6). Nearly 40% of total rLC refolded, as estimated by visual inspection (Figure 2, lanes 5 and 6), a typical result for the glutathione redox-shuffle [21,22,35–37]. Data obtained from 3 different preparations indicated that the final yield of refolded rLC was approximately 35 mg of protein per liter of bacterial culture for rAL-2, whereas approximately 24 mg/liter and 20 mg/liter were the values for rAL-1 and rAL-3, respectively.

Recombinant LC purification

Refolded rLC were purified to homogeneity by a single step anion-exchange chromatography. Figure 3 shows 12% SDS-PAGE of the 3 purified rLC, with the AL-2 BJ for comparison, run under non-reducing and reducing conditions and stained with Coomassie Brilliant Blue (Figure 3A) or analyzed by western blotting with anti-human λ LC antiserum (Figure 3B). LC migrated as single bands of apparent molecular mass of approximately 25 kDa in the presence of reducing agents, whereas they showed two major bands of apparent molecular mass of approximately 40 and 20 kDa under non-reducing conditions. These data were in agreement with the dimeric and monomeric species of rLC, and were supported by the similar electrophoretic profile observed for AL-2 BJ purified protein run in parallel (Figure 3).

Recombinant LC biochemical and functional characterization

The assessment of the oligomeric state of the rLC was performed by size exclusion chromatography. The natural AL-2 BJ protein was also run as control. Each protein eluted from the column as a single main peak, symmetric in shape (Figure 4). Peaks were at a position of approximately 40–44 kDa, in agreement with the expected value for the dimeric form of LC.

In order to test the redox state of cysteines in refolded rLC and isolated AL-2 BJ, free thiol content was determined by means of a highly sensitive fluorimetric assay. Results showed no detectable

free thiols in the three rLC and in the natural BJ protein (<0.004 pmol per pmol LC), as expected for LC found in a dimeric state, and in agreement with the size exclusion chromatographic analyses (Figure 4).

Figure 5 reports the free LC secondary structure patterns investigated by CD spectroscopy at the pH of serum (pH 7.4, panels A and B) and urine (pH 5.0, panels C and D). Spectra profiles typical of the immunoglobulin LC fold were observed (panels A and C), characterized by the high content of β -sheets of the two structural domains (panels B and D). In analogy to natural free LC, that are stable in serum and urine, the synthetic free LC were capable to preserve their secondary structures at different pH (panels A and C, respectively). In the whole, these data documented comparable and proper folding, and indicated that the expression strategy led to the bioactive native conformation of the rLC.

The molecular masses of purified rLC were assessed by linear MALDI-TOF mass spectrometry. Figure 6A shows the singly charged forms of the full-length rLC; the observed masses were consistent with complete reduction of the disulfide bonds. A mass shift of approximately 131 Da from the theoretical mass was observed in all rLC (Figure 6B), compatible with the presence of the additional methionine residue derived from the transcription codon (see above).

Then, biochemical characterization indicated that the recombinant proteins exhibited the typical properties of immunoglobulin LC, and that the rAL-2 was virtually indistinguishable from the relative BJ protein (Figure 4 and 5). Moreover, rLC proteins were arranged as dimers in solution (Figure 4). This latter feature added further evidence of proper folding, since it is known that the

capability of immunoglobulin LC of forming dimers depends essentially on the correct formation of protein's secondary structure [1].

As proof of functionality and conserved pathogenic property, we therefore tested the ability of rLC to generate amyloid fibrils *in vitro*. Again, the natural AL-2 BJ was used as control. Congo-red binding assays (Figure 7A) showed typical patterns consistent with fibril formation. rAL-2 and AL-2 BJ showed very similar Congo-red binding profiles (Figure 7A). Electron microscopy of materials from the 72 h end point [26] demonstrated the presence of amyloid fibrils in all cases (Figure 7B, rAL-2).

In conclusion, we developed a procedure that allowed to obtain, through a bacterial system and in a short period of time (about 3 weeks from RNA extraction), significant quantities of free LC (20–35 mg/liter of culture, depending on the protein) from three patients with AL amyloidosis, a disease typically characterized by LC instability and toxicity. The protein expression in the form of inclusion bodies made purification easy, reducing the number of chromatographic steps generally required to obtain homogeneous protein preparations. Refolding yield from inclusion bodies was noteworthy, providing pathogenic free-LC virtually identical to the natural BJ proteins, capable of amyloid fibril formation *in vitro*, and in quantities suitable for most *in vitro* applications.

Author Contributions

Conceived and designed the experiments: PR VP. Performed the experiments: PR FL SC LV. Analyzed the data: PR FL GV GP GM VP PP. Wrote the paper: PR VP. Reviewed the paper and participated to research supervision: GV. Critically reviewed the manuscript: GP GM PP.

References

- Padlan EA (1994) Anatomy of the antibody molecule. *Mol Immunol* 31: 169–217.
- Bradwell AR, Carr-Smith HD, Mead GP, Tang LX, Showell PJ, et al. (2001) Highly sensitive, automated immunoassay for immunoglobulin free light chains in serum and urine. *Clin Chem* 47: 673–680.
- Solomon A (1985) Light chains of human immunoglobulins. *Methods Enzymol* 116: 101–121.
- Connors LH, Jiang Y, Budnik M, Théberge R, Prokaeva T, et al. (2007) Heterogeneity in primary structure, post-translational modifications, and germline gene usage of nine full-length amyloidogenic kappa1 immunoglobulin light chains. *Biochemistry* 46: 14259–14271.
- Omtvedt LA, Bailey D, Renouf DV, Davies MJ, Paramonov NA, et al. (2000) Glycosylation of immunoglobulin light chains associated with amyloidosis. *Amyloid* 7: 227–244.
- Merlini G, Stone MJ (2006) Dangerous small B-cell clones. *Blood* 108: 2520–2530.
- Merlini G, Bellotti V (2003) Molecular mechanisms of amyloidosis. *N Engl J Med* 349: 583–596.
- Perfetti V, Casarini S, Palladini G, Vignarelli MC, Klersy C, et al. (2002) Analysis of V(lambda)-J(lambda) expression in plasma cells from primary (AL) amyloidosis and normal bone marrow identifies 3r (lambdaIII) as a new amyloid-associated germline gene segment. *Blood* 100: 948–953.
- Comenzo RL, Zhang Y, Martinez C, Osman K, Herrera GA (2001) The tropism of organ involvement in primary systemic amyloidosis: contributions of Ig V(L) germ line gene use and clonal plasma cell burden. *Blood* 98: 714–720.
- Abraham RS, Geyer SM, Price-Troska TL, Allmer C, Kyle RA, et al. (2003) Immunoglobulin light chain variable (V) region genes influence clinical presentation and outcome in light chain-associated amyloidosis (AL). *Blood* 101: 3801–3808.
- Perfetti V, Palladini G, Casarini S, Navazza V, Rognoni P, et al. (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.
- Baglioni S, Casamenti F, Bucciantini M, Lusheski LM, Taddei N, et al. (2006) Prefibrillar amyloid aggregates could be generic toxins in higher organisms. *J Neurosci* 26: 8160–8167.
- Uversky VN (2010) Mysterious oligomerization of the amyloidogenic proteins. *FEBS J* 277: 2940–2953.
- Palladini G, Lavatelli F, Russo P, Perlini S, Perfetti V, et al. (2006) 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.
- Liao R, Jain M, Teller P, Connors LH, Ngoy S, et al. (2001) Infusion of light chains from patients with cardiac amyloidosis causes diastolic dysfunction in isolated mouse hearts. *Circulation* 104: 1594–1597.
- Brenner DA, Jain M, Pimentel DR, Wang B, Connors LH, et al. (2004) Human amyloidogenic light chains directly impair cardiomyocyte function through an increase in cellular oxidant stress. *Circ Res* 94: 1008–1010.
- Shi J, Guan J, Jiang B, Brenner DA, Del Monte F, et al. (2010) Amyloidogenic light chains induce cardiomyocyte contractile dysfunction and apoptosis via a non-canonical p38alpha MAPK pathway. *Proc Natl Acad Sci U S A* 107: 4188–4193.
- Lavatelli F, Brambilla F, Valentini V, Rognoni P, Casarini S, 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.
- Skerra A (1993) Bacterial expression of immunoglobulin fragments. *Curr Opin Immunol* 5: 256–262.
- Perfetti V, Sassano M, Ubbiali P, Vignarelli MC, Arbustini E, et al. (1996) Inverse polymerase chain reaction for cloning complete human immunoglobulin variable regions and leaders conserving the original sequence. *Anal Biochem* 239: 107–109.
- Lilie H, Schwarz E, Rudolph R (1998) Advances in refolding of proteins produced in *E. coli*. *Curr Opin Biotechnol* 9: 497–501.
- De Bernardez Clark E, Schwarz E, Rudolph R (1999) Inhibition of aggregation side reactions during *in vitro* protein folding. *Methods Enzymol* 309: 217–236.
- Andrade MA, Chacón P, Merelo JJ, Morán F (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng* 6: 383–390.
- Sreerama N, Venyaminov SY, Woody RW (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. *Protein Sci* 8: 370–380.
- Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287: 252–260.
- Ren R, Hong Z, Gong H, Laporte K, Skinner M, et al. (2010) Role of glycosaminoglycan sulfation in the formation of immunoglobulin light chain amyloid oligomers and fibrils. *J Biol Chem* 285: 37672–37682.
- Klunk WE, Pettegrew JW, Abraham DJ (1989) Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. *J Histochem Cytochem* 37: 1273–1281.

28. Colombo R, Carotti A, Catto M, Racchi M, Lanni C, et al. (2009) CE can identify small molecules that selectively target soluble oligomers of amyloid beta protein and display antifibrillogenic activity. *Electrophoresis* 30: 1418–1429.
29. Honegger A, Plückthun A (2001) The influence of the buried glutamine or glutamate residue in position 6 on the structure of immunoglobulin variable domains. *J Mol Biol* 309: 687–699.
30. Hurler MR, Helms LR, Li L, Chan W, Wetzel R (1994) A role for destabilizing amino acid replacements in light-chain amyloidosis. *Proc Natl Acad Sci U S A* 91: 5446–5450.
31. Wall J, Schell M, Murphy C, Hrcic R, Stevens FJ, et al. (1999) Thermodynamic instability of human lambda 6 light chains: correlation with fibrillogenicity. *Biochemistry* 38: 14101–14108.
32. Raffin R, Dieckman IJ, Szpunar M, Wunsch C, Pokkuluri PR, et al. (1999) Physicochemical consequences of amino acid variations that contribute to fibril formation by immunoglobulin light chains. *Protein Sci* 8: 509–517.
33. Klimtchuk ES, Gursky O, Patel RS, Laporte KL, Connors LH, et al. (2010) The critical role of the constant region in thermal stability and aggregation of amyloidogenic immunoglobulin light chain. *Biochemistry* 49: 9848–9857.
34. Levinson RT, Olatoye OO, Randles EG, Howell KG, DiCostanzo AC, et al. (2013) Role of mutations in the cellular internalization of amyloidogenic light chains into cardiomyocytes. *Sci Rep* 3: 1278.
35. Buchner J, Rudolph R (1991) Renaturation, purification and characterization of recombinant Fab-fragments produced in *Escherichia coli*. *Biotechnology (N Y)* 9: 157–162.
36. Saxena VP, Wetlaufer DB (1970) Formation of three-dimensional structure in proteins. I. Rapid nonenzymic reactivation of reduced lysozyme. *Biochemistry* 9: 5015–5023.
37. Mo H, Moore RC, Cohen FE, Westaway D, Prusiner SB, et al. (2001) Two different neurodegenerative diseases caused by proteins with similar structures. *Proc Natl Acad Sci U S A* 98: 2352–2357.