Heterogeneity in Primary Structure, Post-Translational Modifications, and Germline Gene Usage of Nine Full-Length Amyloidogenic κ1 Immunoglobulin Light Chains†

Lawreen H. Connors,*,‡,§ Yan Jiang,|| Marianna Budnik,|| Roger Théberge,|| Tatiana Prokaeva,‡,⊥ Kip L. Bodi,‡ David C. Seldin,‡,⊥ Catherine E. Costello,§,|| and Martha Skinner‡,⊥

Alan and Sandra Gerry Amyloid Research Laboratory in the Amyloid Treatment and Research Program, Mass Spectrometry Resource Center, and Departments of Biochemistry and Medicine, Boston University School of Medicine, Boston, Massachusetts 02118

Received July 12, 2007; Revised Manuscript Received September 26, 2007

ABSTRACT: Immunoglobulin light chain amyloidosis is a protein misfolding disease in which a monoclonal immunoglobulin (Ig) light chain (LC) with a critically folded β-conformation self-aggregates to form highly ordered, nonbranching amyloid fibrils. The insoluble nature of amyloid fibrils ultimately results in the extracellular deposition of the LC in tissues and organs throughout the body. Structural features that confer amyloidogenic properties on an Ig LC likely include amino acid sequence variations and post-translational modifications, but the specific natures of these changes remain to be defined. As part of an exploration of the effective range of amyloidogenic modifications, this study details the structural and genetic analyses of nine κ1 LC proteins. Urinary LCs were purified by size exclusion chromatography using FPLC, and structural analyses were performed by electrospray ionization, matrix-assisted laser desorption/ionization, and tandem mass spectrometry. RT-PCR amplification, cloning, and sequencing of the monoclonal LC genes were accomplished using bone marrow-derived mRNA. Clinical data were reviewed retrospectively. Characterization of the urinary κ1 LCs revealed extensive post-translational modification in all proteins, in addition to somatic mutations expected on the basis of results from genetic analyses. Post-translational modifications included disulfide-linked dimerization, *S*-cysteinylation, glycosylation, fragmentation, *S*-sulfonation, and 3-chlorotyrosine formation. Genetic analyses showed that several LC variable region germline gene donors were represented including O18/O8, O12/O2, L15, and L5. Clinical features included soft tissue, cardiac, renal, and hepatic involvement. This study demonstrated the extensive heterogeneity in primary structure, post-translational modifications, and germline gene usage that occurred in nine amyloidogenic κ1 LC proteins.

Amyloid deposits composed of immunoglobulin (Ig)[1] light chains (LCs) are the defining feature of immunoglobulin light chain amyloidosis (AL) or primary amyloidosis, the most common form of systemic amyloidosis in the United States (*1*). Although disease presentation usually involves a relatively minor bone marrow abnormality, AL is devastating in the usually massive extent of organ system amyloid deposits and in the short survival that accompanies the disease (*2*). The disease is associated with a plasma cell dyscrasia, and typical findings include the production of a clonal Ig LC in the bone marrow with the presence of a monoclonal LC in the serum and/or urine. Clinical manifestations of the disease become apparent when the deposits interfere with normal tissue structure and function. Primary organ involvement varies among patients, but a consistent feature of the disease is the occurrence of amyloid fibril deposition at multiple tissue locations far removed from the site of abnormal LC production (*3*, *4*).

AL is a protein misfolding disorder, and the formation of amyloid fibrils is one pathway which can result when the normal structure−function relationship of Ig LCs is disrupted (*5*, *6*). AL amyloid fibrils are constructed from aggregated subunits (or fragments thereof) of an amyloidogenic LC protein (*4*, *7*). Until recently, it was generally believed that the deposits were entirely composed of N-terminal fragments of the LC. We and others have shown that full-length LC is present in the fibrils from AL tissues, along with truncated portions of the protein (*8*−*11*). The prevailing paradigm for

† This work was supported by funding from the National Institutes of Health (Grants P01 HL68705 (to D.C.S.) and P41 RR10888, S10 RR10493, and S10 RR15942 (to C.E.C.), the Gerry Foundation, the Young Family Amyloid Research Fund, and the Amyloid Research Fund at Boston University.

* To whom correspondence should be addressed. E-mail: lconnors@bu.edu. Phone: (617) 638-4313. Fax: (617) 638-4493.

‡ Alan and Sandra Gerry Amyloid Research Laboratory in the Amyloid Treatment and Research Program.

§ Department of Biochemistry.

|| Mass Spectrometry Resource Center.

⊥ Department of Medicine.

[1] Abbreviations: Ig, immunoglobulin; LC, light chain; AL, immunoglobulin light chain amyloidosis; V_L, variable domain; C_L, constant domain; MS, mass spectrometry; MM, multiple myeloma; FPLC, fast protein liquid chromatography; RP-HPLC, reversed-phase high-pressure liquid chromatography, SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; PAS, periodic acid−Schiff base; DTT, dithiothreitol; ESI, electrospray ionization; TOF, time-of-flight; CID, collision-induced dissociation; MALDI, matrix-assisted laser desorption/ionization; M_r, relative molecular mass; PNGase F, peptide-N-glycosidase F; FR, framework region; pI, isoelectric point; V_e, elution volume; M_{calcd}, calculated mass; M_{obsd}, observed mass; D, dimer; M, monomer; CHO, covalently linked glycans.

10.1021/bi7013773 CCC: $37.00
Published on Web 11/16/2007

amyloidogenesis holds that the precursor amyloid protein becomes structurally destabilized, transitions through a series of intermediate conformations, and ultimately adopts a nonnative assembly that favors self-aggregation (*12*−*17*). Aggregation may proceed via a nucleation event, and fibril growth is thought to occur by a mechanism that involves highly ordered protofilament assembly with resultant structures that exhibit overall cross β-sheet organization. The presence of amyloid deposits can interfere with normal cellular and extracellular processes, ultimately causing organ dysfunction and damage. Recent studies have suggested that the pathological consequences in AL are not entirely due to mechanical effects, but are a result of cytotoxicity from prefibrillar aggregates, possibly through a mechanism involving the disturbance of cell−cell paracrine signaling (*18*, *19*).

The effects of variations in the LC structure on aggregation and fibril formation have been the focus of many previous investigations (*5*, *20*−*23*). Many of these studies have been limited to examination of the variable domain (V_L) of the LC protein where amino acid alterations are more prevalent than in the constant domain (C_L). While it has been postulated that key sequence alterations in the V_L are important for fibrillogenesis, other structural components, e.g., post-translational modifications, present throughout the LC, may be essential factors in the process. For this study, we have used techniques to isolate full-length LC which preserve the post-translational modifications. The purified proteins have been characterized in detail using mass spectrometry (MS), allowing us to present structural data on nine amyloidogenic κ1 LCs, which includes the identification of post-translational modifications. We have chosen to focus on κ1 proteins because this is the most common single amyloidogenic LC subfamily (*24*). LCs were purified from urine samples obtained from a group of nine patients with AL that featured the overproduction of a monoclonal κ1 LC. Molecular genetic analysis of cDNA from bone marrow clonal plasma cells and associated clinical findings in the group are also presented.

EXPERIMENTAL PROCEDURES

Patients. All nine patients were referred to the Amyloid Treatment and Research Program at Boston Medical Center for evaluation of amyloid disease and consented to participate in a research study under a protocol approved by the Institutional Review Board at Boston University Medical Center consistent with the Declaration of Helsinki. Confirmation of amyloid disease was made by observation that a tissue biopsy exhibited deposits which showed green birefringence on polarization microscopy after Congo red staining. All patients had a plasma cell dyscrasia on the basis of the finding of κ clonal plasma cells in the bone marrow and/or a monoclonal Ig κ LC in serum and/or urine on immunofixation electrophoresis. A familial form of amyloidosis was excluded in all patients presenting with cardiomyopathy, neuropathy, or renal-only disease. Patients had bone marrow aspirate samples collected for cDNA analysis of the clonal LC gene, urine samples collected for monoclonal LC protein purification and analysis, and clinical information collected and stored in a patient information data base. Previously described nonamyloidogenic, urinary κ1 LC (MM-96100) derived from a patient with multiple myeloma (MM) without amyloidosis (*20*) was used for sequence comparison to the amyloidogenic LCs.

Protein Purification and Electrophoretic Analysis (FPLC, HPLC, and PAGE). Patient urine samples were collected over a 24 h interval in containers with approximately 0.01% EDTA and 0.02% NaN_3 as preservative agents. Samples were retrieved as quickly as possible and immediately dialyzed against deionized H_2O in 6000−8000 molecular weight cutoff dialysis tubing (Spectrapor, Spectrum Laboratories, Rancho Dominguez, CA) usually through five separate changes. Following dialysis, the urine samples were lyophilized and stored at −20 °C until further use. Protocols used for processing and storage of all urine samples were strictly adhered to to avoid introducing structural changes to the LCs postcollection. An aliquot of each lyophilized urine sample was initially analyzed by SDS−PAGE (detailed below) to assess total protein content and LC abundance. Samples were then reconstituted in 0.02 M sodium phosphate, pH 7.1, and albumin-depleted using a 5 mL Hi-Trap Blue column (GE Healthcare Bio-Sciences, Piscataway, NJ). The filtrate was collected, dialyzed against deionized H_2O, and lyophilized. The purification of the LC protein involved one or two successive size-exclusion fractionation steps, as previously described (*20*, *25*). In brief, the albumin-depleted samples were chromatographed on Sephacryl S-200 (S-200, GE Healthcare Bio-Sciences, Piscataway, NJ) in 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.5 (2.6 × 100 cm column). Fractions that were immunoreactive to anti-human free κ LC polyclonal antibodies were pooled, collected, and run on a Superdex S-75 (S-75, GE Healthcare Bio-Sciences) in 50 mM Na_2HPO_4, 150 mM NaCl, pH 7.5 (1.6 × 100 cm column), if further purification was warranted. All chromatographic separations were performed in automated mode using a GE Healthcare ÄKTA FPLC system. Size-exclusion columns were calibrated using calibration kit standard proteins (GE Healthcare Bio-Sciences) which included thryoglobulin, M_r = 670 000, bovine serum albumin, M_r = 67 000, ovalbumin, M_r = 45 000, carbonic anhydrase, M_r = 29 000, and lysozyme, M_r = 13 700. Eluted proteins were monitored at 280 nm.

The homogeneity of purified LCs was analyzed by SDS−PAGE and reversed-phase HPLC (RP-HPLC). Electrophoresis on 10−15% gradient polyacrylamide gels (*26*, *27*) was performed using an automated Phast system (GE Healthcare Bio-Sciences). The identification of LC proteins was demonstrated by immunoblotting. Proteins were initially separated by SDS−PAGE, electrophoretically transferred to nitrocellulose, and then visualized by antibody-linked staining (*28*, *29*). Low M_r standards were run as electrophoretic markers and included phosphorylase b, M_r = 97 000, albumin, M_r = 66 000, ovalbumin, M_r = 45 000, carbonic anhydrase, M_r = 30 000, trypsin inhibitor, M_r = 20 100, and α-lactalbumin, M_r = 14 400. A human κ LC (free) sample obtained commercially (Sigma, St. Louis, MO) was used as a control. Immunodetection was carried out using polyclonal rabbit IgG directed against human Ig κ LC (Dako, Carpinteria, CA). Bound antibody was visualized by incubation with alkaline phosphatase conjugated to goat anti-rabbit IgG and reacted with BCIP/NBT phosphatase substrate (Promega, Madison, WI). Purity assessment by RP-HPLC was accomplished using a Poroshell 300SB-C8 (5 μm, 2.1 × 7.5 mm) column (Agilent Technologies, Palo Alto, CA) (*25*).

Proteins were eluted over a linear gradient of 5−85% solvent B in 5 min at a flow rate of 1.0 mL/min (solvent A, 5% CH_3CN, 0.1% TFA; solvent B, 85% CH_3CN, 0.085% TFA). Spectrophotometric monitoring was at 214 and/or 280 nm. Peak fractions were collected and dried in a SPD111V SpeedVac concentrator (ThermoSavant, Holbrook, NY).

Proteins whose cDNA-deduced LC sequences predicted the presence of the *N*-glycosylation consensus sequence, N-X-T/S (where X ≠ P), were tested for glycosylation. Potential glycoproteins were checked for sugars electrophoretically by using the GelCode glycoprotein staining kit (Pierce Biotechnology, Rockford, IL). The stain reacts with periodate-oxidized carbohydrate groups, creating periodic acid−Schiff base (PAS) chemistry. Glycosylated proteins gave fuchsia-stained bands.

Mass Spectrometry. The molecular masses of the intact LCs, before and after treatment with dithiothreitol (DTT), were determined by nanospray electrospray ionization mass spectrometry (ESI-MS) using an Applied Biosystems/Sciex QStar-Pulsar I quadrupole orthogonal time-of-flight (TOF) mass spectrometer (Qstar). Proteins that stained positive for carbohydrates and did not yield molecular weights upon ESI-MS analyses were subjected to in-source collision-induced dissociation (CID) by raising the declustering potential of the mass spectrometer to 100 V. The resulting mass spectra were examined for the presence of oxonium ions diagnostic of *N*-linked glycans (*30*). The sample aliquots were proteolytically digested with trypsin, Asp-N, Lys-C, and Glu-C. Aliquots of these enzymatic digestion products were reduced with DTT. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and ESI-MS analyses of both the reduced and nonreduced digests were performed on a Bruker Reflex IV MALDI-TOF mass spectrometer and the Qstar, respectively, to generate peptide maps. Peptides of interest were sequenced using ESI-MS/MS on the Qstar. These tandem mass spectrometry experiments could be employed to acquire information on post-translational modifications. Where appropriate, *N*-linked glycans were released by treatment with peptide-N-glycosidase F (PNGase F), and molecular mass determination of the LC before and after treatment was performed using the Bruker Reflex. The released glycans were permethylated using standard protocols (*31*) and analyzed by MALDI-MS and ESI-MS/MS.

Molecular Genetic Analysis. At the time of patient evaluation, a portion of a bone marrow aspirate specimen was obtained and used for cloning and sequencing of Ig LC genes, as previously described (*24*). In brief, the aspirate cells were treated with ammonium chloride to lyse red blood cells, washed, pelleted, and frozen at −80 °C. Total RNA was extracted from 10^7 nucleated cells using Trizol, and cDNA was synthesized. The cDNA was amplified by multiplex PCR using a set of 5′ primers specific for the framework region 1 (FR1) of four V_L κ (κ1/4, κ2, κ3) families and a 3′ constant region (C_L) κ primer. Amplified products were further cloned and sequenced. Each specimen was subjected to multiple independently repeated PCR amplifications, and three clones from each reaction were sequenced in the Boston University School of Medicine Molecular Genetics Core Facility. For each AL case, the LC clonal sequence was determined from the identical matching of at least 50% of six to nine independently cloned and sequenced products. Once the LC was identified, minor nucleotide sequence errors introduced by the original FR1 primers were then corrected by additional PCR amplification using 5′ primers for the appropriate V_L leader region and the 3′ primer for the C_L region. Resequencing was carried out to obtain the correct full-length LC coding sequence. The germline donor genes were identified on the basis of maximum similarity of the nucleotide sequences using a database of rearranged Ig genes, V-BASE (http://www.mrc-cpe.cam.ac.uk), and the International Immunogenetics Information system, IMGT/V-QUEST (http://imgt.cines.fr) (*32*). For MM-96100, no nucleotide sequence was available and the germline donor gene was determined using NCBI igBLAST (http://www.ncbi.nih.nlm.gov/igblast). Amino acid positions were as designated by V-BASE numbering. The NetNGlyc server at http://www.cbs.dtu.dk/services/NetNGlyc/ was used to predict potential *N*-glycosylation sites, N-X-T/S (where X ≠ P), present in the LC sequences. All nucleotide sequences were submitted to GenBank (accession numbers EF589383, EF589396, EF589419, EF589425, EF589448, EF589452, EF589488, AF124196, AF12497). Isoelectric points (pIs) for the V_L domain and corresponding full-length protein in each of the LCs were calculated using the observed primary structure data, cDNA-deduced sequence information, and ExPASy protein analysis tools (http://ca.expasy.org/tools/pi_tool.html).

RESULTS

To study the structural features of amyloidogenic κ1 LC proteins, we purified LCs from the urine samples of nine patients with AL amyloidosis. Dialyzed and lyophilized urine samples were initially analyzed by SDS−PAGE under reducing conditions. There was considerable variation among the samples in total protein content, LC concentration, and LC electrophoretic mobility (Figure 1). Monomeric LC with a relative molecular mass (M_r) of ∼25 000 was the most abundant protein in all samples, identified by immunoblotting using a polyclonal anti-human Ig κ LC antibody (data not shown). Albumin ($M_r \approx 67\,000$) was also present in most cases and was removed using an affinity binding matrix. In addition, several samples contained proteins with M_r less than 25 000 which were prominent, but less abundant than the monomeric LC. In AL-01039 (lane 2), AL-96066 (lane 5), AL-98002 (lane 7), and AL-01102 (lane 8), these more rapidly migrating proteins were immunoreactive with the anti-κ LC antibody by immunoblot analysis (data not shown) and likely are LC fragments (Figure 1). Albumin-depleted samples were fractionated by gel filtration chromatography and assessed electrophoretically as detailed below.

Urinary LCs Are Dimers or Monomers. The chromatographic profiles of the samples could be separated into two distinct groups on the basis of the elution volume (V_e) of the LC. In three of the nine samples, AL-99067, AL-01039, and AL-02004, the LC was rapidly eluted from the column. Each had a $V_e < 300$ mL, which suggested that these proteins had relative molecular masses in the range between 67 000 and 45 000. For the other six samples, the LC in each case eluted at a slower rate and had a $V_e \approx 400$ mL, which corresponded to a protein of M_r between 29 000 and 13 700. Column-fractionated samples were assessed for homogeneity by SDS−PAGE with Coomassie Blue staining and immunoblot analysis using anti-κ LC antibody. The most abundant peak fractions in all cases yielded a single major discrete

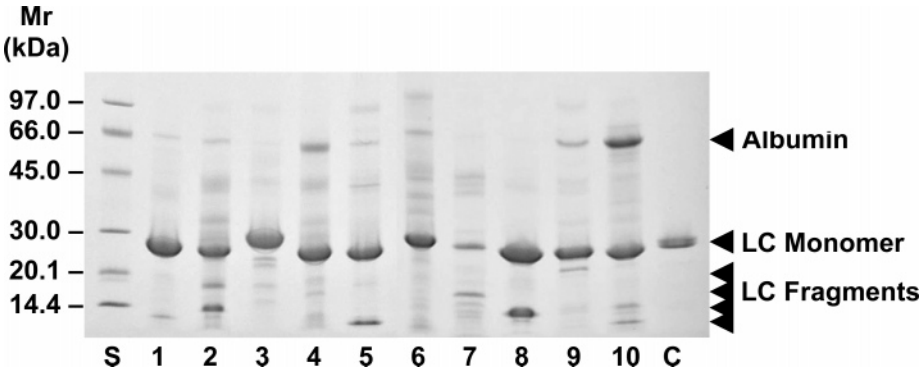

FIGURE 1: Initial SDS−PAGE analysis of urine samples that had been dialyzed exhaustively against ddH_2O, lyophilized, and resolubilized under reducing conditions. Low M_r standards (S) and a commercially obtained purified κ LC (C) were run in the far left and right lanes, respectively. Lanes numbered 1−10 contain samples from AL-99067, AL-01039, AL-02004, MM-96100, AL-96066, AL-01066, AL-98002, AL-01102, AL-01090, and AL-00131, respectively.

protein band by SDS−PAGE under reducing conditions. The electrophoretic end point of this band was always consistent with $M_r \approx 25\,000$.

An example of a faster eluting LC is shown in Figure 2; this represents the chromatographic profile of AL-99067 albumin-depleted urine fractionated on Sephacryl S-200. One major and two minor components were separated in this sample and are labeled peaks 1−3. The most abundant component (Figure 2, peak 3) had a V_e equal to 283.60 mL. The inset in Figure 2 demonstrates the purity assessment and immunoblotting results for AL-99067. The major fraction (peak 3) had $M_r \approx 25\,000$ under reducing conditions and was immunoreactive with the anti-κ LC antibody.

Purified LC proteins from the rapidly eluting group, AL-99067, AL-01039, and AL-02004, were further analyzed by nonreducing SDS−PAGE to investigate the presence of disulfide-linked dimers. The chromatographic profiles of the urinary samples from these three cases suggested that little monomeric ($M_r \approx 25\,000$) LC was present. However, a major component with an M_r slightly greater than 45 000 was eluted under the nonreducing conditions in the chromatographic separation. It was hypothesized that this fractionated protein corresponded to a dimeric form of LC linked through cysteine residues on two subunits. Figure 2B shows the SDS−PAGE of purified LC from AL-99067 (peak 3, Sephacryl S-200) run in the absence (NR) and presence (R) of the reducing agent, β-mercaptoethanol. Under nonreducing (left lane) conditions, a major species was present at an electrophoretic end point consistent with dimeric LC. However, when a reductant was present (right lane), the most abundant form of the protein had $M_r \approx 25\,000$, which likely corresponds to monomeric LC. In the reduced (R) sample, a slight band in the dimer region remained visible (Figure 2B). We have observed the appearance of residual amounts of dimer in reducing SDS−PAGE analyses with other LCs and amyloidogenic proteins (including transthyretin) which are prone to subunit association. These observations are likely due to the stable noncovalent interactions between subunits and therefore attributed to SDS-resistant associations.

A second type of chromatographic elution pattern was seen in six of the urinary samples. These LCs eluted at ~400 mL corresponding to the M_r range between 29 000 and 13 700. This group of urine-derived proteins included AL-96066, AL-01066, AL-98002, AL-01090, AL-00131, and AL-01102. The gel filtration profile of AL-96066 through Sephacryl S-200 is shown in Figure 3 as a representative example of the group. One major and three minor components were partitioned and are indicated as peaks 1−4. The most abundant component (peak 3) had a V_e equal to 392.93 mL.

Column-fractionated samples were assessed for homogeneity by SDS−PAGE with Coomassie Blue staining and immunoblot analyses. The most abundant peak fraction in all cases yielded a major discrete protein band under reducing conditions. The electrophoretic end point of this band was always consistent with $M_r \approx 25\,000$. The inset in Figure 3 demonstrates the purity assessment and immunoblotting results for AL-96066. The major fraction (Figure 3, peak 3) had $M_r \approx 25\,000$ and was immunoreactive to anti-human Ig κ LC antibody. Several minor, higher M_r bands were also observed by immunoblot and likely represent LC oligomeric species. In addition, peak 4 (Figure 3) appeared to be an LC fragment with $M_r < 14\,400$ by SDS−PAGE and immunoblot analyses. Further examination of this fraction by mass spectrometric analysis determined that this protein had a molecular mass of ~12 kDa and was composed of V_L domain residues 1−108 (*25*).

Post-Translational Modifications of LCs Are Varied. Detailed and precise structural characterization studies of the purified κ1 LC proteins were conducted using MS techniques. Initially, each LC sample was examined by ESI-MS, and observed masses (M_{obsd}) were compared to those calculated (M_{calcd}) from the cDNA-deduced sequences. A summary of the mass differences between observed and calculated values, as well as the types of post-translational modifications that were identified in each of the nine κ1 LC proteins, is presented in Table 1. The wide range of mass differences in the intact proteins was due to multiple combinations of the variety of post-translational modifications that are discussed below. Results from the ESI analyses were used in combination with MALDI-MS analyses of enzymatic digestions of the LCs to verify amino acid sequences and to characterize post-translational modifications. These MS investigations revealed that there was extensive structural diversity among the urinary κ1 LC proteins, not only in amino acid sequences, but also in terms of the post-translational modifications.

(1) Dimerization. Dimerization was observed in three of the protein samples. For AL-99067, the deconvoluted ESI-MS spectra showed a major peak with an average mass of ~46 kDa. There was no evidence of this peak after treatment

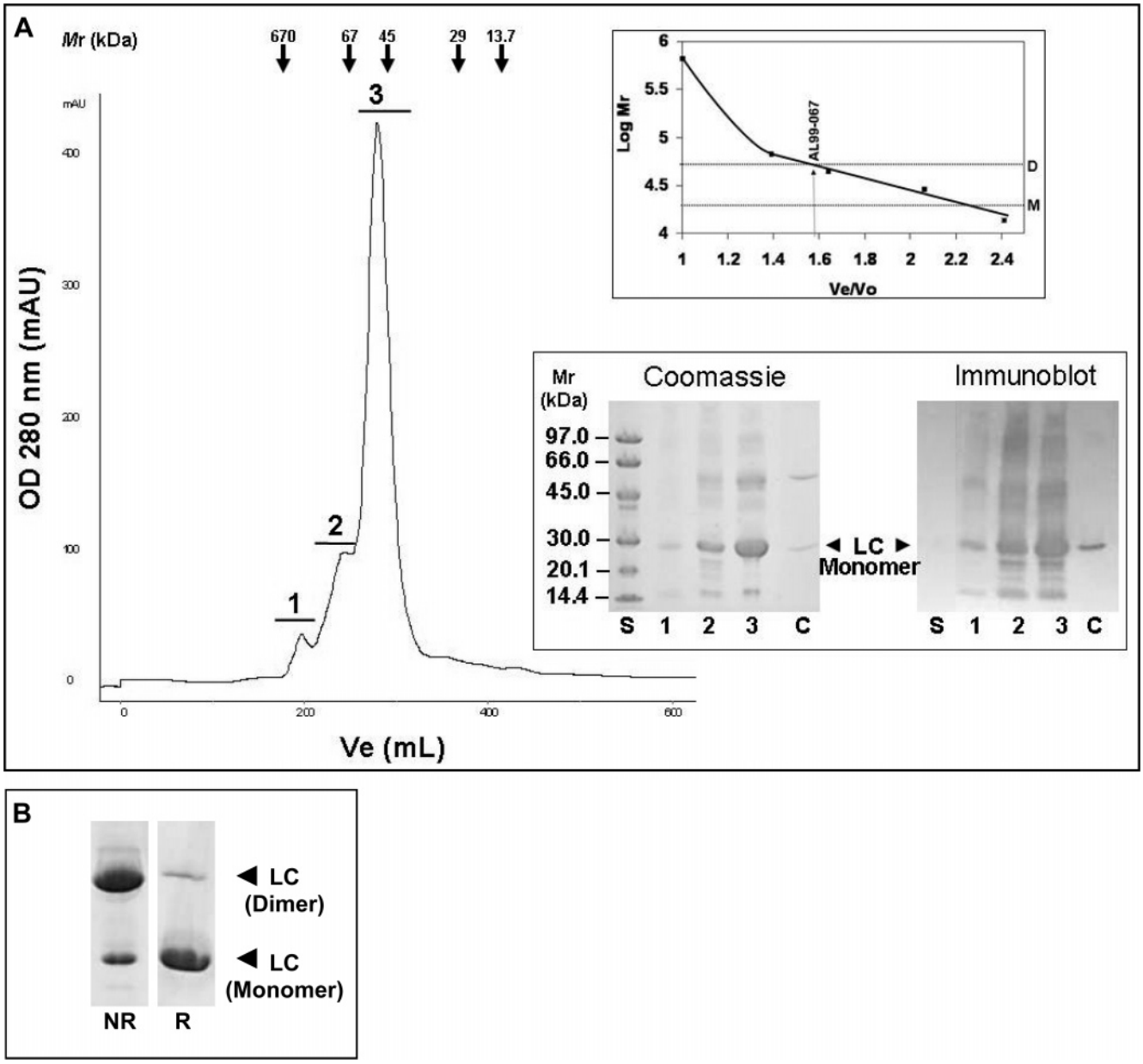

FIGURE 2: (A) Gel filtration chromatography (FPLC) of a rapidly eluting LC. An albumin-depleted urinary sample from AL-99067 was fractionated on a Sephacryl S-200 column. The elution volume (V_e) of the LC was ~300 mL, which corresponds to an M_r between 67 000 and 45 000. Arrows indicate the elution positions of thryoglobulin (M_r = 670 000), bovine serum albumin (M_r = 67 000), ovalbumin (M_r = 45 000), carbonic anhydrase (M_r = 29 000), and lysozyme (M_r = 13 700), which were standard proteins used for column calibrations. Eluted proteins were monitored at 280 nm. A representative linear plot of log M_r vs V_e/V_o for the standard proteins is shown in the inset (right, upper panel), where V_o is the void volume of the column. Horizontal lines corresponding to log M_r values for M_r = 50 000 and 25 000 represent predicted dimeric (D) and monomeric (M) LC elution volumes. The V_e/V_o for AL-99067 was ~1.6 as indicated; this value was consistent with an M_r corresponding to dimeric LC. The inset in the right, lower panel shows the Coomassie-stained SDS−PAGE performed under reducing conditions (left) and immunoblot (right) analyses of S-200 collected fractions 1−3. Immunoblotting was performed using anti-human immunoglobulin light chain κ antibodies. Low M_r standards (S) and a commercially obtained purified κ LC (C) were run in the far left and right lanes, respectively. (B) SDS−PAGE of Sephacryl S-200 peak 3 under nonreducing (left lane) and reducing (right lane) conditions.

of the LC sample with the reducing agent dithiothreitol, but a new peak at 23 233 Da was observed. This mass was consistent with the mass calculated from the cDNA-deduced sequence (M_{calcd} = 23 235 Da). LC analyzed in the absence of a reductant was shown to have a mass of 46 458 Da, which was approximately equal to the calculated mass of 46 460 Da. These results were consistent with our hypothesis that the soluble LC was purified as a covalently linked dimer. The peptide mapping provided 100% sequence coverage, as well as information on two disulfide bonds (C23−C88 and C134−C194) within one monomer and one covalent intermolecular disulfide bridge (C214−C214) between two monomers. Similar results were obtained for AL-01039 protein. AL-02004 was also shown to be a disulfide-linked dimer.

(2) S-Cysteinylation. The addition of disulfide-linked cysteine at the C-terminal C214 residue was identified in three of the amyloidogenic proteins, AL-96066, AL-01066, and AL-00131. The ESI mass spectrum of AL-96066 is shown in Figure 4. The most abundant component had a mass of 23 394 Da. Less abundant species with masses between

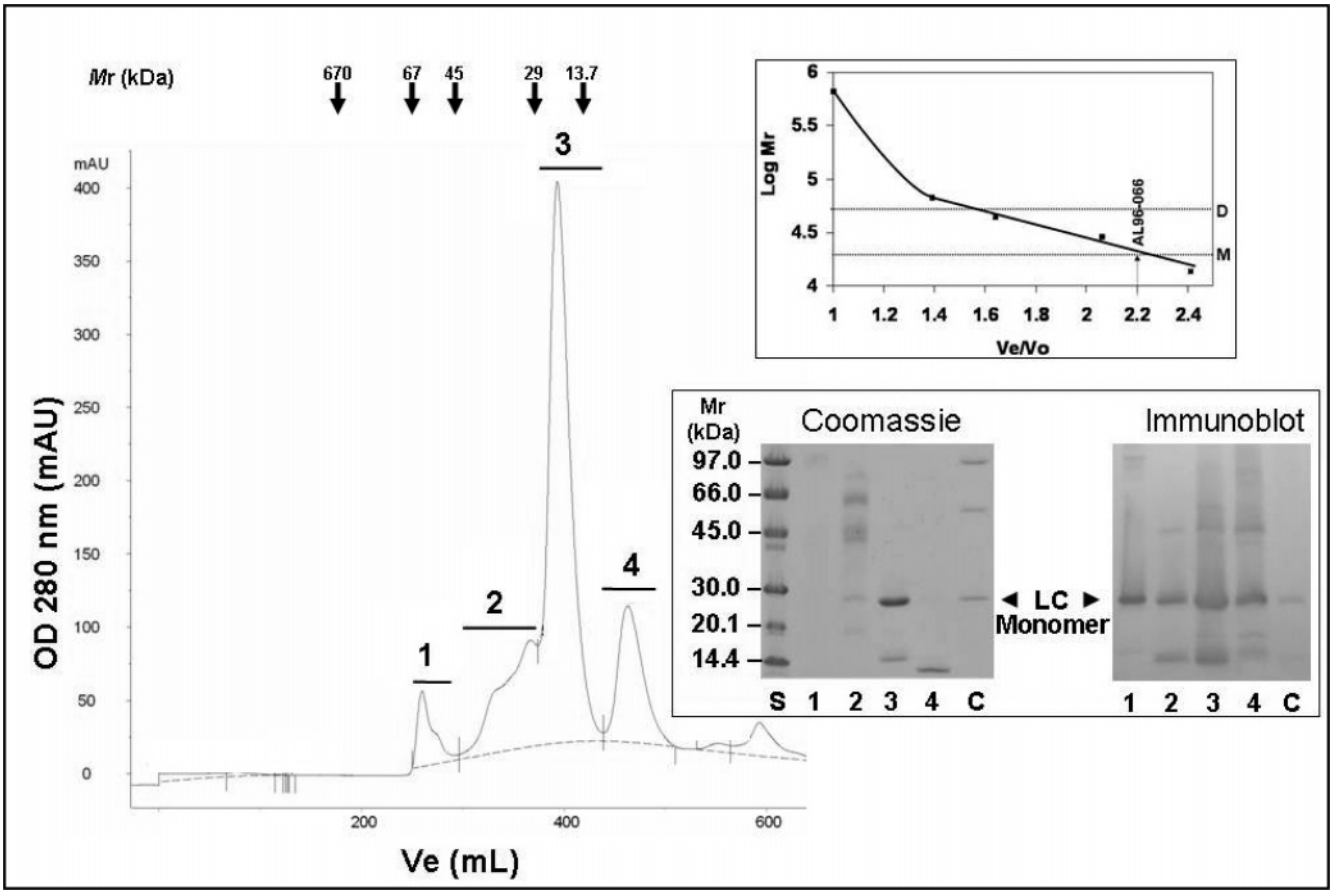

FIGURE 3: Gel filtration chromatography (FPLC) of a slowly eluting LC. An albumin-depleted urinary sample from AL-96066 was fractionated on a Sephacryl S-200 column. The elution volume (V_e) of the LC was ~400 mL, which corresponds to an M_r between 29 000 and 13 700. Arrows indicate the elution positions of thryoglobulin (M_r = 670 000), bovine serum albumin (M_r = 67 000), ovalbumin (M_r = 45 000), carbonic anhydrase (M_r = 29 000), and lysozyme (M_r = 13 700), which were standard proteins used for column calibrations. Eluted proteins were monitored at 280 nm. A representative linear plot of log M_r vs V_e/V_o for the standard proteins is shown in the inset (right, upper panel), where V_o is the void volume of the column. Horizontal lines corresponding to log M_r values for M_r = 50 000 and 25 000 represent predicted dimeric (D) and monomeric (M) LC elution volumes. The V_e/V_o for AL-96066 was ~2.2 as indicated; this value was consistent with an M_r corresponding to monomeric LC. The inset in the right, lower panel shows the Coomassie-stained SDS−PAGE performed under reducing conditions (left) and immunoblot (right) analyses of S-200 collected fractions 1−4. Immunoblotting was performed using anti-human immunoglobulin light chain κ antibodies. Low M_r standards (S) and a commercially obtained purified κ LC (C) were run in the far left and right lanes, respectively.

23 274 and 23 504 Da were also noted. The predicted mass for this LC was 23 277 Da. This was approximately 117 mass units lower than the observed mass. Subsequent analysis of the most abundant species showed that this was an intact κ1 LC containing the *S*-cysteinyl post-translational modification at C214 and internal disulfide bonds at C23−C88 and C134−C194 (*21*).

For AL-01066, the deconvoluted ESI mass spectrum showed a major peak at 23 690 Da. After treatment of the LC with the reducing agent, dithiothreitol, a new peak at 23 573 Da was observed and there was no evidence of the peak at 23 690 Da. This observed mass of 23 573 Da was in agreement with the theoretical value of 23 570 Da calculated from the cDNA-deduced amino acid sequence. The 120 Da mass decrease after reduction corresponded to a loss of a disulfide-linked cysteine and the reduction of two to three disulfide bonds in the native protein. The cDNA-deduced amino acid sequence was verified by peptide mapping using MALDI-TOF MS, and three disulfide bonds (C23−C88, C134−C194, and C214−cysteine) were identified.

Results from the structural analysis of AL-00131 protein are discussed in further detail below.

(3) N-Glycosylation. Two cDNA-deduced LC sequences, AL-98002 and AL-02004, were predicted to have potential *N*-glycosylation sites (N-X-T/S, where X ≠ P) at N70 and N20, respectively. PAS treatment of the purified urinary LCs from these two cases confirmed that these were glycoproteins. The results for PAS staining of AL-98002 are shown in Figure 5 (lane 3). For reference, nonglycosylated (lane 1) and glycosylated (lane 2) proteins were electrophoresed in adjacent lanes.

Due to the complexity of the AL-98002 sample, the molecular mass of the LC could not be determined by ESI-MS. However, using MALDI-MS, the average mass of 22 196 Da was obtained for the intact protein (Figure 6A, top panel). The *N*-glycosylation site at N70 was confirmed by mass spectrometry. After treatment with PNGase F to remove *N*-glycans, MALDI-MS of the sample produced a peak at *m*/*z* 17 625 Da (Figure 6A, lower panel), suggesting that some of the heterogeneity was due to *N*-linked glycosylation. The PNGase F-released glycans were permethylated using a standard protocol (*31*, *33*). MALDI-MS of the permethylated glycans released from the LC indicated that the glycans contained bi- and triantennary structures (Figure

Table 1: Structural Variations in Nine Amyloidogenic Urinary κ1 LC Proteins

κ1 LC	M_{calcd}[a]	M_{obsd}[b]	ΔM[c]	modification
		Dimerization		
AL-99067	46 460	46 458 (D)[d]	−2	C214−C214
	23 235	23 233 (M)[e]	−2	
AL-01039	46 488	46 504 (D)	+16	C214−C214
	23 253	23 238 (M)	−15	
AL-02004[f]	46 880	46 770 (D)	−110	N20-CHO[g]
	23 449	23 382 (M)	−67	C214−C214
		S-Cysteinylation		
AL-96066	23 277	23 394	+117	C214-cysteine
AL-01066	23 570	23 690	+120	C214-cysteine
		N-Glycosylation		
AL-98002	23 674	22 196	−1478	N70-CHO
AL-02004	46 880	46 770 (D)	−110	N20-CHO
	23 449	25 726 (M)	+2277	C214−C214
		Fragmentation		
AL-01102	23 351	12 047 and 12 136		N-terminal truncation
	11 607 (V_L) 11 762 (C_L)			
		Multiple Other		
AL-01090	23 383	23 386	+3	C214-*S*-SO_3 N-terminal pyruvate M4-O Y36-Cl
AL-00131	23 248	23 279	+31	C214-cysteine C214-*S*-SO_3 C214-SO_3 N-terminal pyruvate Y49-Cl

[a] M_{calcd} = calculated mass (Da). [b] M_{obsd} = observed mass (Da). [c] $\Delta M = M_{obsd} - M_{calcd}$ (Da). [d] D = dimer. [e] M = monomer. [f] Listed twice. [g] CHO = covalently linked glycans.

6B). These structures were similar to those reported for other LC proteins associated with AL (*22*).

In the case of AL-02004, the MALDI-TOF mass spectrum of the reduced LC yielded an average mass of 25 726 Da with a peak width at half-height of over 800 Da (data not shown). The calculated mass (obtained from the cDNA sequence) was 23 449 Da. The considerably higher observed mass, coupled with the large peak width, strongly suggested the presence of glycosylation. ESI-MS analysis did not provide molecular weight information. However, increasing the cone voltage of the instrument to cause in-source fragmentation generated oxonium ions typical of the fragmentation of carbohydrates. The masses of oxonium ions observed were consistent with those of sialated *N*-linked glycans (*m*/*z* 657 Da). An *N*-linked glycan consensus site was present at N20 in AL-02004.

(4) Fragmentation. Evidence of extensive fragmentation was found in one LC protein sample. ESI-MS analysis of AL-01102 showed that there was no full-length LC in this sample. Two fragments were detected with molecular masses of 12 047 and 12 136 Da. These masses were consistent with the calculated mass for residues 105−214 of the C_L domain. MALDI-MS analyses of the enzymatic digests of the whole sample gave 79% sequence coverage (residues 17−27, 46−92, and 105−214). Two disulfide bonds were identified as C23−C88 and C134−C194.

(5) Multiple and Other Post-Translational Modifications. Two amyloidogenic LCs contained multiple types of modifications. In AL-01090, modifications at both the amino and carboxy termini were noted. Aspartic acid was processed to pyruvate at the first position in the sequence, and C214 was *S*-sulfonated at the C-terminus. In addition, M4 was found in the oxidized form, methionine sulfoxide, and a low abundance of chlorinated Y36 was also detected.

The most extensively modified protein in this group of nine amyloidogenic κ1 LCs was AL-00131 (Figure 7). C214 was found to be *S*-cysteinylated and *S*-sulfonated, as well as oxidized to cysteic acid. Evidence of partial C-terminal truncation was also noted. At the N-terminus, aspartic acid was modified to pyruvate (as demonstrated by the 45 Da shift from 23 276 Da to the lower mass of 23 232 Da) and M4 was oxidized to methionine sulfoxide. It was also determined that Y49 was chlorinated.

V_L Gene Usage and Somatic Mutations Analysis. A variety of germline genes were used in the rearrangement of V_L κ1 gene sequences in this study (Table 2). Six of the nine amyloidogenic LCs, AL-96066, AL-98002, AL-00131, AL-01066, AL-01102, and AL-02004, were found to use the O18/O8 V_L κ1 germline gene donor. Three other LCs, AL-99067, AL-01039, and AL-01090, were derived from L15, L5, and O12/O2 V_L κ1 germline genes, respectively. Without exception, V_L sequences were somatically mutated throughout all framework (FR) and complementarity-determining (CDR) regions, with the number of replacement mutations varying from 2 to 18 per sequence. Specifically for FR1−3, amino acid substitutions were located at multiple residues throughout these regions in all but one (AL-96066) of the κ1 sequences (Table 2). Furthermore, the number and nature of these mutations were unique to each LC. Certain positions appeared to be commonly targeted for mutational activity. Thus, serine codons, AGC and AGT, known to be intrinsic mutational hotspots (*34*) were targeted in three sequences, e.g., S77G (AL-99067), S77N (AL-98002), and S77G (AL-01090). Overall, the majority of somatic mutations were the result of neutral amino acid replacements. However, mutations that resulted in the gain or loss of a charged residue in at least one FR position occurred in all but one of the group (AL-00131). These included R18T in AL-99067, Y49D in AL-96066, Q3E and S76R in AL-01066, K42T and D70N in AL-98002, R61S in AL-02004, and K45N in AL-01090 (Table 2). In two κ1 proteins, the identical charge change from basic to acidic residue was noted (K45E in AL-01039 and K42E in AL-01102), although the substitution positions were different. It is also of interest that, in two of the O18/O8-derived cases, the gain of an asparagine residue, D70N (AL-98002) and T20N (AL-02004), created potential *N*-glycosylation sites whose occupancy was substantiated by PAS staining and MS analyses.

Amyloidogenic vs Nonamyloidogenic LC Comparison. The V_L sequences of the nine AL proteins were compared to MM-96100, a nonamyloidogenic κ1 LC previously characterized as a monomeric LC that was *S*-cysteinylated at C214 and derived from the O18/O8 germline gene (*20*). In FR1−3, amino acid differences were noted among the AL and MM proteins. The total number of residue differences in the AL proteins ranged from 3 to 10. While the majority of these amino acid variations represented neutral changes, 7/9 of the amyloidogenic LCs showed at least one position where a net gain or loss of charge occurred. These non-neutral changes (with the corresponding MM sequence shown in

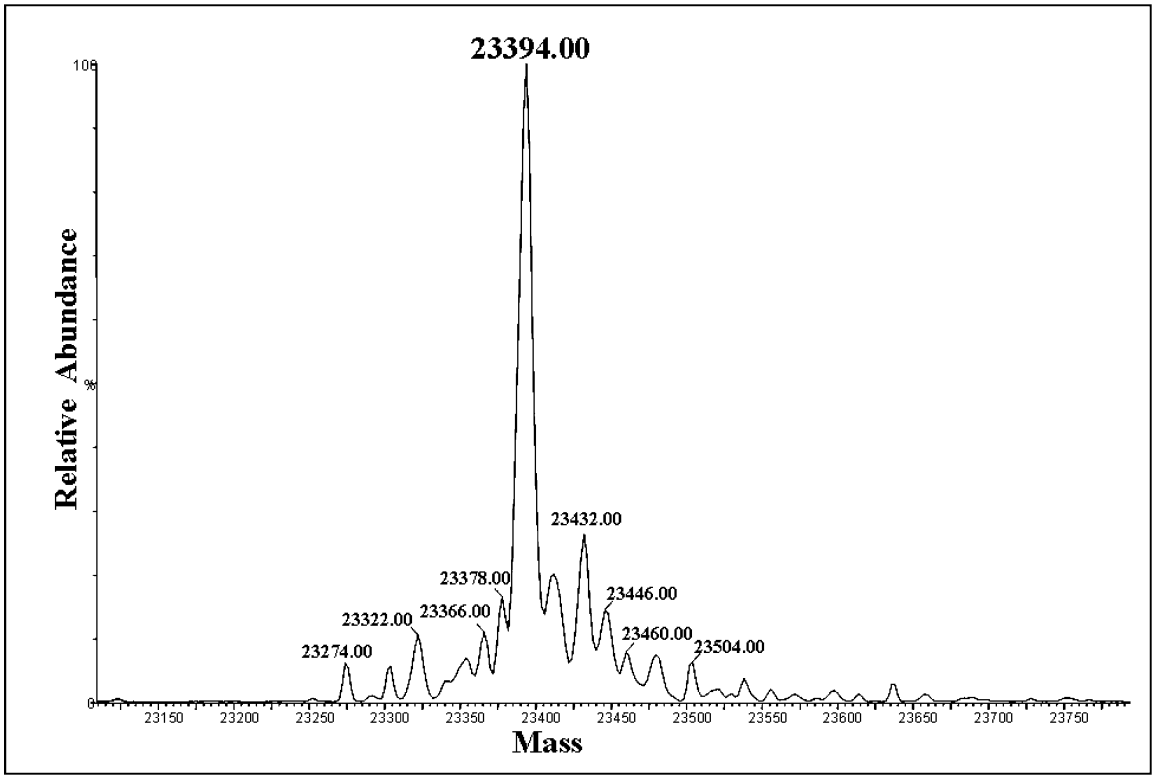

FIGURE 4: Molecular ion region of the ESI mass spectrum obtained for purified urinary LC from AL-96066.

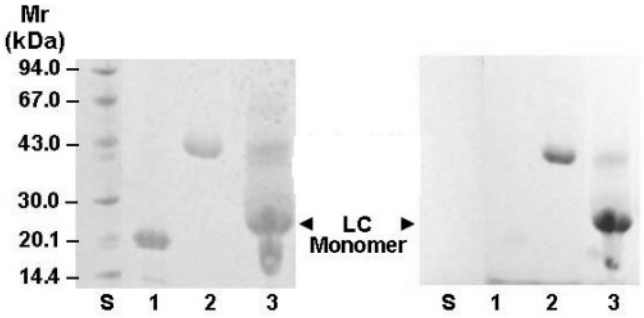

FIGURE 5: Duplicate SDS−PAGE (10−15%) of purified κ1 LC protein (AL-98002) stained with Coomassie (left) and PAS (right). Lane S contains reference M_r marker proteins. Nonglycosylated (soybean trypsin inhibitor) and glycosylated (α1-acid glycoprotein) proteins were run in lanes 1 and 2, respectively, as controls. Lane 3 contains a glycosylated κ1 LC (AL-98002).

parentheses) included T18 (R18) and G41 (E41) in AL-99067, E45 (K45) in AL-01039, D49 (Y49) in AL-96066, E3 (Q3) and R76 (S76) in AL-01066, T42 (K42) in AL-98002, S61 (R61) in AL-02004, and E42 (K42) in AL-01102.

For the AL and MM proteins, pI values were calculated for the V_L domains and corresponding full-length LCs using the observed primary structure data and cDNA-deduced sequence information. Results are shown in Table 3. Without exception, the V_L domain was more acidic than the full-length LC in each case. Values for the V_L domain of the amyloidogenic proteins ranged from 4.07 (AL-01102) to 6.74 (AL-01090). The pI_{calcd} value of 4.53 was determined for the V_L of nonamyloidogenic MM-96100 protein. Calculated pI values for the V_L regions of each LC were compared to those predicted for the appropriate germline gene donor. The pI_{calcd} values for the V_L germline gene sequences O18/O8, L15, L5, and O12/O2 were 4.27, 6.25 7.99, and 7.97, respectively. For the full-length LCs, pI values ranged from 4.95 (AL-01102) to 6.99 (AL-01090). The pI_{calcd} of 5.44 was predicted for full-length MM-96100. This value was comparable to those of the AL proteins. In fact, nonamyloidogenic MM-96100 shared identical V_L and full-length pI_{calcd} values with amyloidogenic LC AL-01066.

Demographic, Clinical, and Laboratory Data. The demographic, clinical, and laboratory features for the nine cases of AL are summarized in Table 4. The group included seven male and two female patients ranging in age at the time of diagnosis from 42 to 74 years. Soft tissue or cardiac disease was the primary clinical feature in this AL group. Other commonly involved organs were kidney and liver. Serum free LCs were present in seven cases. The urinary excretion of monoclonal κ LC ranged from 20% to 90% of the total urinary protein excretion. No association of clinical features with post-translational modifications of κ1 LCs was found. The survival from diagnosis was short, less than 2 years in all but one patient.

DISCUSSION

This is the first detailed structural report of both the V_L and C_L domains of a group of κ1 LC proteins from patients with AL. Mass spectrometric analyses provided precise characterization of the amino acid sequence, in addition to, identification of the location and chemical nature of post-translational modifications in the κ1 LCs.

In our MS analyses of the κ1 LCs presented, there were no cases of unmodified proteins detected in any of the molecular weight profiles. We have provided qualitative structural data rather than the extent of each LC post-translational modification as ionization efficiencies vary widely among peptides depending on their chemistry. All LCs were isolated in either dimeric or monomeric form under nondenaturing and nonreducing conditions. Three proteins

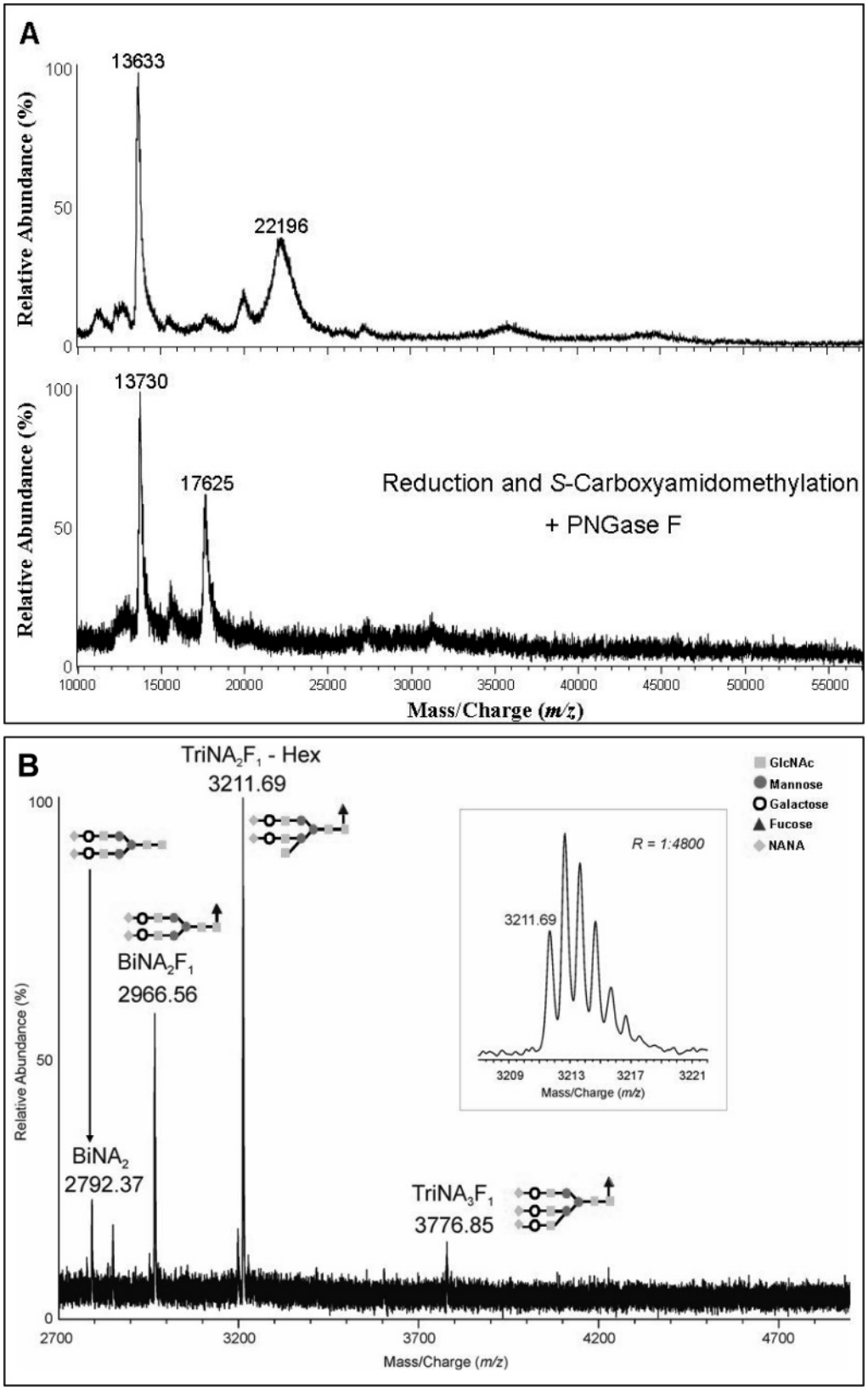

FIGURE 6: (A) MALDI-TOF mass spectra of the purified LC from AL-98002 before (upper panel) and after (lower panel) treatment with dithioreitol, iodoacetamide, and PNGase F. (B) MALDI-TOF mass spectrum of the released permethylated glycans from the glycosylated light chain.

were purified from urine as covalently linked dimers. Mass spectrometric analyses demonstrated that these three LCs were bonded via C214 disulfide bridging between two monomers. While disulfide-linked dimers are known to be a common structural feature of free λ LCs (*35*−*37*), κ LCs are usually found in monomeric form (*36*, *38*). In fact, previous studies have suggested that the predominance of λ to κ found in AL amyloid deposits may be because dimerization of LCs is an initial step in the process of fibril formation (*7*, *39*). It is possible that previous studies have

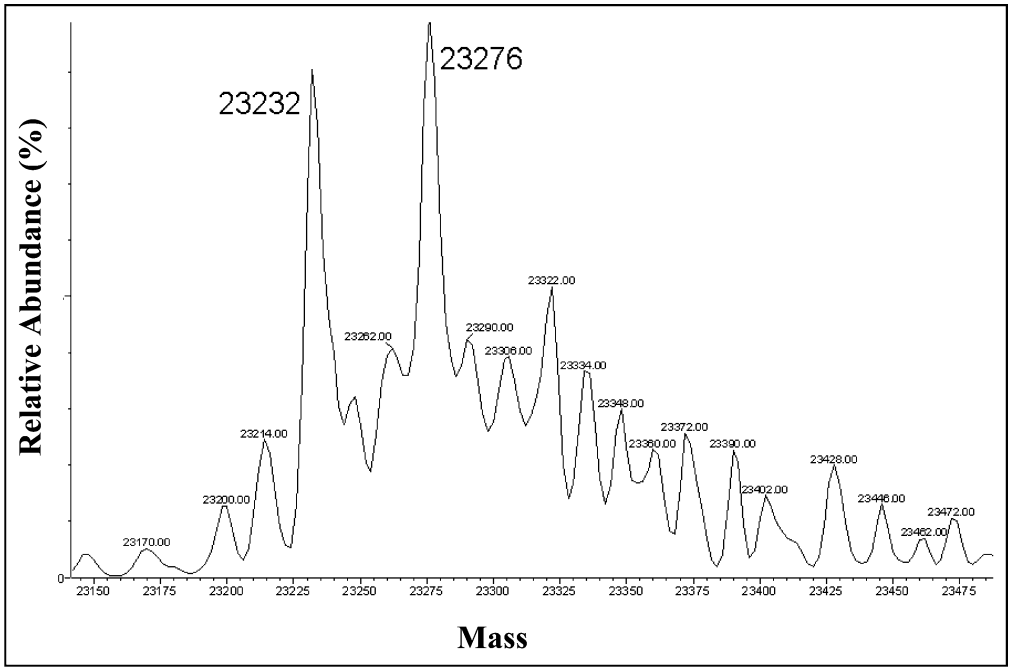

FIGURE 7: Deconvoluted ESI mass spectrum of purified LC from AL-00131.

Table 2: Somatic Mutations in κ1 V_L Domain Framework Regions of Amyloidogenic LCs Grouped by Post-Translational Modifications

	dimerization		*S*-cysteinylation		*N*-glycosylation		other		
V_L domain	AL-99067	AL-01039	AL-96066	AL-01066	AL-98002	AL-02004	AL-01102	AL-00131	AL-01090
FR1 (1−23)					I2V				
				Q3E					
							V15I		
	R18T								
					V19I				T19V
						T20N[a]			
					T22A				
FR2 (35−49)									Q38H
		K39R							
					K42T		K42E		
		K45E							K45N
				L46V					
							I48V		
			Y49D						Y49F
FR3 (57−88)						R61S			
								S63T	
									G66A
					D70N[a]				
						T72I			
								T74S	
	S76T			S76R					
	S77G				S77N				S77N
		P80A							
						A84G			
					Y87H		Y87F		
no. of replacement mutations, *n*	3	3	1	3	7	4	4	2	6
V_L germline gene	L15	L5	O18/O8	O18/O8	O18/O8	O18/O8	O18/O8	O18/O8	O12/O2

[a] Gain of an asparagine residue that created an *N*-glycosylation site.

failed to recognize the dimeric nature of κ amyloidogenic LCs because sample preparation employed reducing agents. Alternatively, structural modifications of the urinary LCs could have been affected by the length and conditions of sample storage. For example, the low levels of albumin observed in the three urine samples containing dimeric LCs (Figure 1, lanes 1−3) might suggest that the dimers were not found in the circulation but formed in the urine. While this possibility cannot be excluded, it is unclear why only three of the LCs were found to be disulfide-linked dimers as consistency in handling and storage protocols were strictly maintained and LC purifications at multiple time points on individual urines rendered similar results.

Three LCs were found to be *S*-cysteinylated at C214, the site which normally forms an interchain disulfide bond with an available Ig heavy chain. Cysteinylation has been previously reported in a nonamyloidogenic LC (*20*) purified from a patient with multiple myeloma without amyloidosis (MM-96100). *S*-cysteinylation of other proteins has been linked to elevated levels of homocysteine, a biomolecule which is

Table 3: Comparison of pI Values for V_L and Full-Length κ1 LC Proteins

κ1 LC	Germline Gene	pI_{calcd}[a]	
		V_L Domain	LC Monomer
–	*O18/O8*	*4.27*	–
MM-96100		4.52	5.44
AL-96066		4.27	5.21
AL-01066		4.52	5.44
AL-98002		5.43	6.10
AL-02004		4.59	5.61
AL-01102		4.07	4.95
AL-00131		4.60	5.72
–	*L15*	*6.25*	–
AL-99067		4.60	5.72
–	*L5*	*7.99*	–
AL-01039		4.93	6.16
–	*O12/O2*	*7.97*	–
AL-01090		6.74	6.99

[a] pI_{calcd} = isoelectric point calculated from observed primary structure data and cDNA-deduced sequences.

converted to cysteine in oxidizing environments (*40*−*42*). Several reports have suggested that cysteinylation of proteins can modulate bioactivity through consequent conformational alterations (*42*, *43*). Interestingly, *S*-cysteinylation has not been described in λ LCs.

One of the dimeric and one of the monomeric LCs were found to be glycosylated. Glycosylation of LCs has been described for both κ and λ proteins (*22*, *44*−*48*). In addition, several reports have identified a subset of κ1 LCs as glycoproteins. In fact, it has been suggested that glycosylation may be important in the mechanism of amyloid fibril formation (*4*, *22*, *23*). In a study that compared the V_L domain sequences of 37 amyloidogenic and 63 nonamyloidogenic κ1 LCs, it was proposed that the gain of an *N*-glycosylation site was a risk factor for amyloid fibrillization (*23*).

Mass spectrometric analysis of one κ1 protein (AL-01102) demonstrated only polypeptide fragments of $M_r \approx 12\,000$, although initial electrophoretic and chromatographic evaluations of the urine sample had indicated the presence of a full-length monomeric LC. Extensive N-terminal truncation was identified from the mass spectrometric data and may have been a result of substantial structural instability. Two other amyloidogenic LCs were each unique in the nature and multiplicity of their post-translational modifications. Modifications were identified at the N- and C-termini of both AL-01090 and AL-00131. The processing of N-terminal aspartic acid to pyruvate and oxidation of Met4 were identified in both LCs. We have recently observed the conversion of N-terminal aspartic acid to pyruvate in λ LCs as well. This transformation represents a new post-translational modification, probably stemming from oxidative decarboxylation followed by oxidative deamination (*49*−*51*). It is interesting that chlorotyrosine residues (position 36 in AL-01090 and position 49 in AL-00131) were identified in these two proteins. Chlorotyrosine residues have been found to be elevated in atherosclerosis, lung disease, sepsis, vasculitis, and other inflammatory diseases and are regarded as an indicator for oxidative damage (*52*).

Although λ AL amyloidosis is more common than κ, the most frequent LC subfamily in our patient population is κ1, representing 22% of LCs (*24*). In the κ1 LC group, about 40% are derived from the V_L germline gene O18/O8, and about 20% are derived from O12/O2 (*53*). Seven of the κ1 LC proteins in this study were derived from the O18/08 V_L germline gene. The pattern of somatic mutations in the V_L domains of the LCs was unique to each. It has been previously suggested that specific residues in the V_L domain can play a key role in promoting amyloid fibril formation (*7*, *11*). On the basis of the analysis of 37 amyloidogenic and 84 nonamyloidogenic V_L κ1 sequences, Stevens suggested that certain amino acids at approximately 24 positions were either over- or underrepresented in the AL proteins (*23*). Four features were proposed to be risk factors for fibrillogenesis and included loss of I27b (an important packing side chain), replacement of K31 with D (destabilizing at this position), loss of R61 (involved in a critical buried salt bridge to D82), and the gain of an *N*-linked glycosylation site

Table 4: Demographic, Clinical, and Laboratory Data on AL κ1 Cases

AL κ1 LC	sex	age[a]	serum free κ LC	monoclonal LC/urine protein (mg/24 h)	primary organ involvement[b]	other organ involvement[b]	survival (months)
Dimerization							
AL-99067	F	59	yes	1500/4836	ST	L, K	16
AL-01039	M	68	yes	290/648	ST	H	10
AL-02004[c]	M	42	yes	2500/2756	H	L, K	6
S-Cysteinylation							
AL-96066	M	68	yes	636/1060	ST	K	144
AL-01066	M	56	yes	172/766	ST	L, K, H	20
N-Glycosylation							
AL-98002	M	60	no	nd/1000	H	ANS	3
AL-02004[c]	M	42	yes	2500/2756	H	L, K	6
Fragmentation							
AL-01102	M	49	yes	2660/3700	ST	GI, K, H	2
Multiple Other							
AL-01090	M	59	yes	766/1036	H	ST, L	18
AL-00131	F	74	no	nd/1454.0	H	L, K	6

[a] Age at time of diagnosis. [b] Abbreviations for organ involvement: ST, soft tissue; H, heart; K, kidney; ANS, autonomic nervous system; L, liver; GI, gastrointestinal; nd, not determined. [c] Listed twice.

through mutations. Overall, 80% of the analyzed amyloidogenic V_L sequences were characterized by the presence of at least one of three specific substitutions or the acquisition of an *N*-linked glycosylation site (*23*). In the present study, only two of nine (22.2%) κ1 proteins actually met the criteria of this predictive strategy. Thus, replacement of D70 in AL-98002 and T20 in AL-02004 with asparagine resulted in the introduction of an *N*-glycosylation site, and in fact, AL-02004 also featured R61 changed to serine.

We also investigated the κ1 LC group for net charge differences in both the V_L domains and the full-length proteins. Previous studies have demonstrated that LCs from cases of AL were acidic, with pI values ranging from 3.8 to 5.2 (*54*−*56*). In contrast, nonamyloidogenic proteins have been shown to exhibit more basic or near neutral pIs (*21*). We calculated pI values for O18/O8, O12/O2, L15, and L5 germline gene sequences and compared these to values obtained from the observed κ1 primary structure data. Predicted values for O18/O8 germline gene and the group encoded by this donor were all in the acidic range. Of note, the acidic pI value of 4.52 for the nonamyloidogenic V_L LC MM-96100 fell within the range of the pIs of the amyloidogenic proteins. Furthermore, while the O12/O2, L15, and L5 germline gene sequences were predicted to have near neutral pIs, the amyloidogenic V_L sequences were acidic in two of the three cases. When the C_L domain was included in pI calculations for the 9 κ1 LC sequences, the predicted values overall were less acidic. However, it is interesting to note that the nonamyloidogenic LC had the near lowest pI at 5.44, with only one amyloidogenic protein (AL-0102) predicted to be more acidic (pI 4.95).

We have previously reported the association of κ1 LC sequences with soft tissue manifestations of AL (*24*). In this study, soft tissue or cardiac disease was the dominant clinical feature, with other organs including kidney and liver showing more minor involvement. No association of clinical features with post-translational modifications of κ1 LCs was found.

In summary, this molecular characterization of nine amyloidogenic κ1 LCs demonstrated that these proteins were derived from multiple germline genes which have undergone extensive somatic mutations. Without exception, the expressed proteins were subsequently modified in a variety of ways. All nine amyloidogenic LCs of the κ1 subtype examined in this study showed extensive structural diversity in both amino acid sequence and post-translational modifications. Modifications included disulfide-linked dimerization, *S*-cysteinylation, *N*-glycosylation, fragmentation, *S*-sulfonation, oxidation of methionine, chlorination of tyrosine, the conversion of aspartic acid to pyruvate, and the presence of cysteic acid conjugated to cysteine. While the individual contributions of these modifications to pathology are unknown, several of these structural features have been linked to oxidative stress and may be markers of the process of amyloidogenesis or cell or tissue damage. To further understand the role of protein structure in AL, it will be essential to obtain detailed molecular information on a set of full-length, nonamyloidogenic κ1 LCs for comparison with the data presented herein.

ACKNOWLEDGMENT

We thank Gregory Karamitis, Jeremy Eberhard, and Kate Laporte for their technical expertise in purification and analysis of LC proteins and the Boston University School of Medicine Molecular Genetics Core Facility for providing nucleotide sequencing services.

REFERENCES

1. Falk, R. H., Comenzo, R. L., and Skinner, M. (1997) The systemic amyloidoses, *N. Engl. J. Med. 337*, 898−909.
2. Skinner, M., Sanchorawala, V., Seldin, D. C., Dember, L. M., Falk, R. H., Berk, J. L., Anderson, J. J., O'Hara, C., Finn, K. T., Libbey, C. A., Wiesman, J., Quillen, K., Swan, N., and Wright, D. G. (2004) High-dose melphalan and autologous stem-cell transplantation in patients with AL amyloidosis: an 8-year study, *Ann. Intern. Med. 140*, 85−93.
3. Falk, R. H., and Skinner, M. (2000) The systemic amyloidoses: an overview, *Adv. Intern. Med. 45*, 107−137.
4. Merlini, G., and Bellotti, V. (2003) Molecular mechanisms of amyloidosis, *N. Engl. J. Med. 349*, 583−596.
5. Bellotti, V., Mangione, P., and Merlini, G. (2000) Review: Immunoglogulin light chain amyloidosis—the archetype of structural and pathologic variability, *J. Struct. Biol. 130*, 280−289.
6. Lansbury, P. T., Jr. (1999) Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease, *Proc. Natl. Acad. Sci. U.S.A. 96*, 3342−3344.
7. Stevens, F. J., Myatt, E. A., Chang, C. H., Westholm, F. A., Eulitz, M., Weiss, D. T., Murphy, C., Solomon, A., and Schiffer, M. (1995) A molecular model for self-assembly of amyloid fibrils: immunoglobulin light chains, *Biochemistry 34*, 10697−10702.
8. Gertz, M. A., Skinner, M., Cohen, A. S., Connors, L. H., and Kyle, R. A. (1985) Isolation and characterization of a kappa amyloid fibril protein, *Scand. J. Immunol. 22*, 245−250.
9. Liepnieks, J. J., Dwulet, F. E., and Benson, M. D. (1990) Amino acid sequence of a kappa I primary (AL) amyloid protein (AND), *Mol. Immunol. 27*, 481−485.
10. Myran, T., Husby, G., Kyle, R. A., and Sletten, K. (2004) The amino acid sequence of a glycosylated AL-chain from a patient with primary amyloidosis, *Amyloid 11*, 109−112.
11. Solomon, A., Weiss, D. T., Murphy, C. L., Hrncic, R., Wall, J. S., and Schell, M. (1998) Light chain-associated amyloid deposits comprised of a novel kappa constant domain, *Proc. Natl. Acad. Sci. U.S.A. 95*, 9547−9551.
12. Hurshman, A. R., White, J. T., Powers, E. T., and Kelly, J. W. (2004) Transthyretin aggregation under partially denaturing conditions is a downhill polymerization, *Biochemistry 43*, 7365−7381.
13. Kelly, J. W. (2000) Mechanisms of amyloidogenesis, *Nat. Struct. Biol. 7*, 824−826.
14. Konno, T. (2001) Multistep nucleus formation and a separate subunit contribution of the amyloidgenesis of heat-denatured monellin, *Protein Sci. 10*, 2093−2101.
15. Lee, C. C., Nayak, A., Sethuraman, A., Belfort, G., and McRae, G. J. (2007) A three-stage kinetic model of amyloid fibrillation, *Biophys. J. 92*, 3448−3458.
16. Thirumalai, D., Klimov, D. K., and Dima, R. I. (2003) Emerging ideas on the molecular basis of protein and peptide aggregation, *Curr. Opin. Struct. Biol. 13*, 146−159.
17. Wogulis, M., Wright, S., Cunningham, D., Chilcote, T., Powell, K., and Rydel, R. E. (2005) Nucleation-dependent polymerization is an essential component of amyloid-mediated neuronal cell death, *J. Neurosci. 25*, 1071−1080.
18. Liao, R., Jain, M., Teller, P., Connors, L. H., Ngoy, S., Skinner, M., Falk, R. H., and Apstein, C. S. (2001) Infusion of light chains from patients with cardiac amyloidosis causes diastolic dysfunction in isolated mouse hearts, *Circulation 104*, 1594−1597.
19. Brenner, D. A., Jain, M., Pimentel, D. R., Wang, B., Connors, L. H., Skinner, M., Apstein, C. S., and Liao, R. (2004) Human amyloidogenic light chains directly impair cardiomyocyte function through an increase in cellular oxidant stress, *Circ. Res. 94*, 1008−1010.
20. Chung, C. M., Chiu, J. D., Connors, L. H., Gursky, O., Lim, A., Dykstra, A. B., Liepnieks, J., Benson, M. D., Costello, C. E., Skinner, M., and Walsh, M. T. (2005) Thermodynamic stability of a kappaI immunoglobulin light chain: relevance to multiple myeloma, *Biophys. J. 88*, 4232−4242.
21. Kaplan, B., Livneh, A., and Gallo, G. (2007) Charge differences between in vivo deposits in immunoglobulin light chain amyloidosis and non-amyloid light chain deposition disease, *Br. J. Haematol. 136*, 723−728.

22. Omtvedt, L. A., Bailey, D., Renouf, D. V., Davies, M. J., Paramonov, N. A., Haavik, S., Husby, G., Sletten, K., and Hounsell, E. F. (2000) Glycosylation of immunoglobulin light chains associated with amyloidosis, *Amyloid 7*, 227−244.
23. Stevens, F. J. (2000) Four structural risk factors identify most fibril-forming kappa light chains, *Amyloid 7*, 200−211.
24. Prokaeva, T., Spencer, B., Kaut, M., Ozonoff, A., Doros, G., Connors, L., Skinner, M., and Seldin, D. C. (2007) Soft tissue, joint, and bone manifestations of AL amyloidosis, *Arthritis Rheum.* (in press).
25. Lim, A., Wally, J., Walsh, M. T., Skinner, M., and Costello, C. E. (2001) Identification and location of a cysteinyl posttranslational modification in an amyloidogenic kappa1 light chain protein by electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry, *Anal. Biochem. 295*, 45−56.
26. Weber, K., and Osborn, M. (1969) The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis, *J. Biol. Chem. 244*, 4406−4412.
27. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
28. Towbin, H., Staehelin, T., and Gordon, J. (1992) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979, *Biotechnology Ser. 24*, 145−149.
29. Jagersten, C., Edstrom, A., Olsson, B., and Jacobson, G. (1988) Blotting from PhastGel media after horizontal sodium dodecyl sulfate-polyacrylamide gel electrophoresis, *Electrophoresis 9*, 662−665.
30. Conboy, J. J., and Henion, J. (1992) High-performance anion-exchange chromatography coupled with mass spectrometry for the determination of carbohydrates, *Biol. Mass Spectrom. 21*, 397−407.
31. Ciucanu, I., and Costello, C. E. (2003) Elimination of oxidative degradation during the per-O-methylation of carbohydrates, *J. Am. Chem. Soc. 125*, 16213−16219.
32. Giudicelli, V., Chaume, D., and Lefranc, M. P. (2004) IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis, *Nucleic Acids Res. 32*, W435−440.
33. Reinhold, V. N., Reinhold, B. B., and Costello, C. E. (1995) Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID, *Anal. Chem. 67*, 1772−1784.
34. Rogozin, I. B., and Kolchanov, N. A. (1992) Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis, *Biochim. Biophys. Acta 1171*, 11−18.
35. Abraham, R. S., Charlesworth, M. C., Owen, B. A., Benson, L. M., Katzmann, J. A., Reeder, C. B., and Kyle, R. A. (2002) Trimolecular complexes of lambda light chain dimers in serum of a patient with multiple myeloma, *Clin. Chem. 48*, 1805−1811.
36. Berggard, I., and Peterson, P. A. (1969) Polymeric forms of free normal kappa and lambda chains of human immunoglobulin, *J. Biol. Chem. 244*, 4299−4307.
37. Grey, H. M., and Kohler, P. F. (1968) A case of tetramer Bence Jones proteinaemia, *Clin. Exp. Immunol. 3*, 277−285.
38. Peterson, P. A., and Berggard, I. (1971) Urinary immunoglobulin components in normal, tubular, and glomerular proteinuria: quantities and characteristics of free light chains, IgG, igA, and Fc-gamma fragment, *Eur. J. Clin. Invest. 1*, 255−264.
39. Schormann, N., Murrell, J. R., Liepnieks, J. J., and Benson, M. D. (1995) Tertiary structure of an amyloid immunoglobulin light chain protein: a proposed model for amyloid fibril formation, *Proc. Natl. Acad. Sci. U.S.A. 92*, 9490−9494.
40. Bar-Or, D., Heyborne, K. D., Bar-Or, R., Rael, L. T., Winkler, J. V., and Navot, D. (2005) Cysteinylation of maternal plasma albumin and its association with intrauterine growth restriction, *Prenatal Diagn. 25*, 245−249.
41. Bizzozero, O. A., DeJesus, G., Callahan, K., and Pastuszyn, A. (2005) Elevated protein carbonylation in the brain white matter and gray matter of patients with multiple sclerosis, *J. Neurosci. Res. 81*, 687−695.
42. Chen, W., Yewdell, J. W., Levine, R. L., and Bennink, J. R. (1999) Modification of cysteine residues in vitro and in vivo affects the immunogenicity and antigenicity of major histocompatibility complex class I-restricted viral determinants, *J. Exp. Med. 189*, 1757−1764.
43. Watarai, H., Nozawa, R., Tokunaga, A., Yuyama, N., Tomas, M., Hinohara, A., Ishizaka, K., and Ishii, Y. (2000) Posttranslational modification of the glycosylation inhibiting factor (GIF) gene product generates bioactive GIF, *Proc. Natl. Acad. Sci. U.S.A. 97*, 13251−13256.
44. Dwulet, F. E., O'Connor, T. P., and Benson, M. D. (1986) Polymorphism in a kappa I primary (AL) amyloid protein (BAN), *Mol. Immunol. 23*, 73−78.
45. Engvig, J. P., Olsen, K. E., Gislefoss, R. E., Sletten, K., Wahlstrom, O., and Westermark, P. (1998) Constant region of a kappa III immunoglobulin light chain as a major AL-amyloid protein, *Scand. J. Immunol. 48*, 92−98.
46. Foss, G. S., Nilsen, R., Cornwell, G. C., 3rd, Husby, G., and Sletten, K. (1998) A glycosylated Bence Jones protein and its autologous amyloid light chain containing potentially amyloidogenic residues, *Scand. J. Immunol. 47*, 348−354.
47. Milstein, C. P., and Milstein, C. (1971) Glycopeptides from human kappa-chains, *Biochem. J. 121*, 211−215.
48. Toft, K. G., Sletten, K., and Husby, G. (1985) The amino-acid sequence of the variable region of a carbohydrate-containing amyloid fibril protein EPS (immunoglobulin light chain, type lambda), *Biol. Chem. Hoppe-Seyler 366*, 617−625.
49. Dimroth, P. (1982) Decarboxylation and transport, *Biosci. Rep. 2*, 849−860.
50. Kowalik-Jankowska, T., Ruta, M., Wisniewska, K., Lankiewicz, L., and Dyba, M. (2004) Products of Cu(II)-catalyzed oxidation in the presence of hydrogen peroxide of the 1−10, 1−16 fragments of human and mouse beta-amyloid peptide, *J. Inorg. Biochem. 98*, 940−950.
51. Okamura, T., Noda, H., Fukuda, S., and Ohsugi, M. (1998) Aspartate dehydrogenase in vitamin B12-producing Klebsiella pneumoniae IFO 13541, *J. Nutr. Sci. Vitaminol. 44*, 483−490.
52. Mohiuddin, I., Chai, H., Lin, P. H., Lumsden, A. B., Yao, Q., and Chen, C. (2006) Nitrotyrosine and chlorotyrosine: clinical significance and biological functions in the vascular system, *J. Surg. Res. 133*, 143−149.
53. 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*, 3801−3808.
54. Bellotti, V., Merlini, G., Bucciarelli, E., Perfetti, V., Quaglini, S., and Ascari, E. (1990) Relevance of class, molecular weight and isoelectric point in predicting human light chain amyloidogenicity, *Br. J. Haematol. 74*, 65−69.
55. Gallo, G., Goni, F., Boctor, F., Vidal, R., Kumar, A., Stevens, F. J., Frangione, B., and Ghiso, J. (1996) Light chain cardiomyopathy. Structural analysis of the light chain tissue deposits, *Am. J. Pathol. 148*, 1397−1406.
56. Schiffer, M. (1996) Molecular anatomy and the pathological expression of antibody light chains, *Am. J. Pathol. 148*, 1339−1344.

BI7013773