Tertiary structure of an amyloid immunoglobulin light chain protein: A proposed model for amyloid fibril formation

(x-ray crystallography/amyloid light chain amyloidosis)

NORBERT SCHORMANN*, JILL R. MURRELL*, JURIS J. LIEPNIEKS†, AND MERRILL D. BENSON*‡‡

Departments of *Medicine and †Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202; and the ‡Richard L. Roudebush Veterans Affairs Medical Center, Indianapolis, IN 46202

Communicated by Frank W. Putnam, Indiana University, Bloomington, IN, June 22, 1995

ABSTRACT An immunoglobulin light chain protein was isolated from the urine of an individual (BRE) with systemic amyloidosis. Complete amino acid sequence of the variable region of the light chain (VL) protein established it as a κ 1, which when compared with other κ 1 amyloid associated proteins had unique residues, including Ile-34, Leu-40, and Tyr-71. To study the tertiary structure, BRE Vλ was expressed in Escherichia coli by using a PCR product amplified from the patient BRE's bone marrow DNA. The PCR product was ligated into pCZ11, a thermal-inducible replication vector. Recombinant BRE Vλ was isolated, purified to homogeneity, and crystallized by using ammonium sulfate as the precipitant. Two crystal forms were obtained. In crystal form I the BRE Vλ κ domain crystallizes as a dimer with unit cell constants isomorphous to previously published κ protein structures. Comparison with a nonamyloid Vλ κ domain from patient REI, identified significant differences in position of residues in the hypervariable segments plus variations in framework region (FR) segments 40–46 (FR2) and 66–67 (FR3). In addition, positional differences can be seen along the two types of local diads, corresponding to the monomer–monomer and dimer–dimer interfaces. From the packing diagram, a model for the amyloid light chain (AL) fibril is proposed based on a pseudohexagonal spiral structure with a rise of approximately the width of two dimers per 360° turn. This spiral structure could be consistent with the dimensions of amyloid fibrils as determined by electron microscopy.

Immunocyte-derived amyloid light chain (AL) (primary) amyloidosis is the most common form of systemic amyloidosis and, of all the amyloidoses, is associated with the shortest survival rate (1). In the majority of cases, this disease is associated with a nonmalignant expansion of a single plasma cell clone. It is the monoclonal immunoglobulin light chain product of the plasma cell clone which is the precursor of the amyloid fibril subunit protein (2).

A number of clinically observed features of AL amyloidosis suggest that structure of the immunoglobulin light chain subunit protein is an important factor in the pathogenesis of this disease. First, while the ratio of κ to λ plasma cell clones and proteins in the human is 3:2, more individuals with λ light chain-derived amyloid are identified than with κ proteins. Second, within the λ light chain subgroups, the λ VI proteins have been found to be particularly amyloidogenic (3, 4). Third, κ 1 light chain proteins represent the majority of κ amyloid proteins and primary structure analyses have suggested substitutions in key framework regions as being associated with amyloid fibril formation (5) (see Fig. 1).

To investigate structural factors which may be important for amyloid fibril formation from immunoglobulin light chain proteins, an attempt has been made to compare the structure of the variable region of an amyloid light chain protein with the structure of nonamyloid light chain proteins. The goal is to develop a structural model for the polymerization or aggregation process of immunoglobulin light chain proteins which leads to amyloid fibril formation. A κ 1 protein from an individual with amyloidosis was completely characterized at the primary structure level.

By using the determined amino acid sequence, oligonucleotide primers were constructed and used to amplify from the patient’s bone marrow DNA the sequence coding for the entire variable segment of the κ 1 light chain. The DNA construct was then used to express the recombinant variable region of the light chain (VL) of the κ 1 protein in vitro. This protein was crystallized, and its tertiary structure was determined by x-ray diffraction.

MATERIALS AND METHODS

Isolation and Characterization of Light Chain Protein BRE. Light chain protein was isolated from the urine of patient BRE (hereafter referred to as BRE protein or BRE) by precipitation in 60% ammonium sulfate followed by chromatography on DEAE cellulose and ACA-34 (Pharmacia). Amino acid sequence analysis was performed on a Beckman 890C sequenator as described (4). For a complete structure, the BRE protein was digested with 1-1-tosamido-2-phenylthiochomethyl ketone (TPCK)-treated trypsin and resultant peptides separated on a C18 HPLC column.

DNA Isolation and Plasmid Construction for Expression of Recombinant Protein. DNA was isolated from bone marrow cells of patient BRE by using the method of Madisen et al. (6). On the basis of the amino acid sequence of BRE protein and published sequences of the signal peptide for κ 1 proteins (7), oligonucleotide primers were constructed (see Fig. 2). The resultant PCR amplification product was then ligated into pCZ11, a thermodenaturing replication vector, and used to transform competent Escherichia coli strain HB101 (8). As described for transthyretin (TTR) (8), cell pellets of positive HB101 pCZ11/BRE clones grown in culture were subjected to SDS/12.5% PAGE. Electrophoresed proteins were transferred to polyvinylidene difluoride (PVDF) membrane (Applied Biosystems) and subjected to (i) Western analysis using BRE-specific antisera and (ii) amino acid sequencing of bands of appropriate molecular weight (ABI 473A sequenator; Applied Biosystems).

Abbreviations: VL, variable region of the light chain; FR, framework region; CDR, complementarity-determining region; AL, amyloid light chain.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. U31344).
‡The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY, 11973 (references 1BRE and R1BRESF).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
Large Scale Growth and Protein Isolation. Cells containing the desired plasmid were grown in 200 ml of TY broth (0.8% tryptone/1% yeast extract/0.25% NaCl) overnight at 23°C as described (8). This culture was added to 4.5 l of TY broth in a 5-l Bioflo III Bioreactor (New Brunswick Scientific), and grown at 40°C with monitoring of O₂ saturation and pH. After 6 h, the culture was centrifuged, and the cell pellet was isolated and lysed by freeze/thaw and sonication. The supernatant was fractionated by DEAE cellulose, and fractions containing protein BRE were then subjected to further purification on ACA-34.

Crystallization and X-Ray Analysis. Purified protein BRE was concentrated to 20 mg/ml in 0.1 M Tris buffer, pH 7.5. Crystals of BRE with a size of 0.35 × 0.3 × 0.25 mm were grown by hanging-drop vapor diffusion against 2.25–2.5 M ammonium sulfate in 100 mM citrate buffer, pH 5.5, at 23°C. Two crystal forms were obtained under very similar conditions. Crystal form I belongs to the monoclinic system, space group P2₁, with unit-cell dimensions of \(a = 82.29\ \text{Å}, b = 77.73\ \text{Å}, c = 82.20\ \text{Å}\), and \(\beta = 119.95^\circ\) (\(V \approx 450,000\ \text{Å}^3\)) and six monomers (6 × 12 KDa) in the asymmetric unit. These unit-cell constants are isomorphic to those published for \(\kappa\) ROY (9); however, the \(\kappa\) BRE intensities do not exhibit hexagonal symmetry. Crystal form II, on the other hand, belongs to the orthorhombic space group C22₂₁, with unit-cell dimensions of \(a = 82.04\ \text{Å}, b = 142.09\ \text{Å},\) and \(c = 77.85\ \text{Å}\) (\(V \approx 900,000\ \text{Å}^3\)) and three monomers in the asymmetric unit.

All data were collected on an R-Axis IIIC image plate area detector (Molecular Structure, The Woodlands, TX) by using a Rigaku RU-200 x-ray source and were reduced with R-Axis processing software (10).

Structure Determination. Crystal form I (BRE I). A native data set was collected on a single crystal to a nominal resolution of 1.75 Å (see Table 1). Molecular replacement (rotation and translation searches together with Patterson correlation refinement) was performed with X-PLOR (11, 12) by using data in the 15- to 3.5-Å shell and the refined REI structure as the search model (13). Each independent solution from the cross-rotation function after PC refinement (11) was used in a translation search in the \(a\) and \(c\) directions since this space group requires only a search in \(x\) and \(z\). The best solution for the first translation function gave the position of dimer 1 in the asymmetric unit. To determine the relative position of dimer 2 and 3 with respect to dimer 1, we used combined translation functions (11). A rough overall model with appropriate amino acid substitutions made by using QUANTA (QUANTA 4.0, 1994); Molecular Modeling Software Package, Polygen Corp., Waltham, MA) was then submitted to rigid body and Powell minimization refinement with X-PLOR (12). After subsequent refinement using a simulated annealing protocol in

<table>
<thead>
<tr>
<th>Table 1. Statistics of x-ray data collection for BRE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BRE I</strong></td>
</tr>
<tr>
<td>Resolution range, Å</td>
</tr>
<tr>
<td>Number of unique reflections</td>
</tr>
<tr>
<td>% completeness of data</td>
</tr>
<tr>
<td>(resolution limit, Å)</td>
</tr>
<tr>
<td>(R)merge, %*</td>
</tr>
</tbody>
</table>

*\(R\)merge = \(\Sigma_i|I_i - \bar{I}_i| / \Sigma_i\bar{I}_i\); \(R\)merge (all reflections) was calculated for the observations with \(I > \sigma\) (4).
X-PLOR (14) with intermittent cycles of model rebuilding using O (15) by inspection of 2Fo − Fo maps, all residues except Arg-108 in each monomer were clearly observed. Refinement was carried out to a resolution of 2.0 Å, and in later stages water molecules found by a peak search routine in CCP4 (16) were gradually added. No noncrystallographic symmetry restraints were used at this point. After positional refinement, as implemented in X-PLOR (12), the crystallographic R factor dropped to 21.8% for all data in the resolution range 5.0–2.0 Å. To assess the quality of the refinement, we used the free R-factor test in X-PLOR (17). The high value (Rfree ~ 33.7%) indicated possible overfitting of our model. To improve the data-to-parameter ratio, we included noncrystallographic symmetry restraints and repeated the refinement with X-PLOR (12, 14). The crystallographic R factor converged at 22.9% (Rfree ~ 26%). Finally, we subjected the model to 10 cycles of restrained least-squares refinement by using PROLSQ (18). The current model consists of 4974 protein atoms (107 of the 108 residues modeled) and 40 water molecules (see Table 2). The present R-factor is 19.8% (data in resolution range of 5.0–2.0 Å) with an Rfree of 20.4% (5% of data). The mean coordinate error is estimated to be 0.33 Å by using the SIGMAA method (19).

### RESULTS AND DISCUSSION

Sequence analysis of intact protein BRE isolated from the urine of a patient with systemic amyloidosis identified it as belonging to the κ I subgroup. The complete sequence of its Vκ region was determined by analysis of peptides obtained after digestion with trypsin. Comparison of BRE Vκ sequence to other κ I amyloid associated proteins showed a few unique residues, including Ile-34, Leu-40, and Tyr-71 (Fig. 1). By using oligonucleotide primers based on the amino acid sequence of BRE Vκ and published κ I signal peptides, PCR amplification of bone marrow DNA from patient BRE gave a product of expected size which by direct sequencing was shown to have nucleotide sequence in agreement with the protein structure of BRE Vκ (Fig. 2). Ligation of this DNA construct into pCZ11 and expression in E. coli produced a protein of approximately 12,000 Da. This protein reacted specifically with an antiserum raised against the BRE urine light chain. Recombinant BRE protein was purified to homogeneity and crystallized by using ammonium sulfate as precipitant. Diffraction data to 1.75-Å resolution were collected in two crystal forms (Table 1). The structure of the BRE Vκ κ domain in crystal form I was determined by x-ray diffraction and refined to 2 Å (Table 2).

![Fig. 3. Ribbon diagram of Vκ dimer based on the refined BRE structure. The figure was prepared with MOLSCRIPT (21) and rendered with RASTER3D (22, 23) on the basis of secondary structure assignments made by DSSP (24).](image-url)
The BRE V\textsubscript{L} \(\kappa\) domain crystallizes as a dimer, as found for proteins ROY and REI (9, 13). The tertiary structure of BRE is similar to other \(\lambda\) and \(\kappa\) light chain fragments in that the monomer is composed of a \(\beta\)-barrel formed from nine antiparallel \(\beta\)-strands arranged in two sheets packed against each other, with two monomers making a Bence Jones-type dimer (Fig. 3). Two forms of interdomain contacts can be seen along the two types of local diads. Amino acid residues Tyr-36, Glu-38, Pro-44, Asn-45, Tyr-87, and Phe-98 of both monomers participate in the monomer–monomer contact region, as described for REI (13). The contact region between the non-covalent dimers is formed by amino acid residues Ser-9, Ser-10, and Ser-12 extending the hydrogen bonding network and by Pro-8 and Leu-11 making hydrophobic contacts.

Sequence comparison of BRE with REI shows the altered residues and their assignment to FRs and complementarity-determining regions (CDRs) (Fig. 4). Comparison with the published tertiary structure of REI (13) was accomplished by superimposition of the structures by using a least-squares procedure. Significant differences in the position of residues in the hypervariable segments (residues 28–32, 50–53, and 91–96), as well as variations in FR segments 40–46 (FR2) and 66–67 (FR3), are seen. The largest difference between BRE and REI monomers is in the 40–44 loop which forms part of the dimer interface, partly due to a substitution of Leu for Pro at residue 40. This substitution also introduces a hydrophobic residue into a hydrophilic environment since residue 40 is exposed to the solvent. Important positional differences of residues involved in monomer–monomer interactions at the dimer interface and in monomer–monomer interactions along the dimer–dimer contact region are seen (Fig. 5 A and B).

While topology and hydrogen-bonding pattern of the monomers and interactions overall between monomers at the dimer interface are preserved, there are perturbations in contacts along the dimer–dimer interface. Due to moving apart from each other, the hydrogen bond between the hydroxyl of Ser-10 and the nitrogen of Ser-212 is deleted in BRE, which should result in a destabilization of dimer–dimer contacts.

**Proposed Model for the AL Amyloid Fibril.** The study of tertiary structure of \(\kappa\) BRE, an amyloid-forming light chain protein, has resulted in the formation of a model of V\textsubscript{L} dimers in a pseudohexagonal arrangement with a diameter of 115 Å and a rise of 38.8 Å per turn of 180° (Fig. 6). This would allow a structure that can be visualized as a spiral with six light chain dimers and a rise of approximately the diameter of two dimers per turn of 360° (Fig. 7). This model fits well with the dimensions noted by electron microscopy for fibrils of approximately 10 nm (10 Å) in diameter and indeterminate length.

As for TTR, structural alterations are found in immunoglobulin protein(s) associated with amyloidosis when compared with their nonamyloid counterparts. These changes usually involve residues at the interface (monomer–monomer and monomer–dimer interactions). Both light chain proteins and TTR variants show an extensive \(\beta\)-sheet framework believed to be a prerequisite for fibril formation. In TTR variants, single-site substitutions lead to structural perturbations, but so far all variants studied crystallographically crystallize isomorphously (same space group, very similar unit-cell parameters)—i.e., their packing mode is preserved (25). To date, no unifying model regarding initiation of polymerization and subsequent fibril formation has been established; although, partial denaturation of TTR in an acidic environment and polymerization
light chain proteins of both subtypes (k and a). If this model holds, it may then be possible to identify key residues responsible for the transformation.

We thank Dr. Marianne Schiffer for helpful discussions and encouragement in the early phases of these studies. This work was supported by Veterans Affairs Medical Research (MRIS 583-0888), and from U.S. Public Health Service Grants RR0750, DK34881, DK42111, AG10668, AR20582, the Arthritis Foundation, and the Marion E. Jacobson Fund.