Prolines and Amyloidogenicity in Fragments of the Alzheimer's Peptide β/A4

Stephen J. Wood, Ronald Wetzel,* John D. Martin,‡ and Mark R. Hurle*

Macromolecular Sciences Department, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, Pennsylvania 19406

Received September 9, 1994; Revised Manuscript Received November 10, 1994®

ABSTRACT: Although it is well accepted that the structure of amyloid fibrils is dominated by some form of antiparallel β-sheet, there are few details on the secondary structural arrangements of the constituent peptides and how these peptides pack together in the fibril. We describe here the use of scanning proline mutagenesis to map the secondary structural roles of each residue in amyloidogenic peptide fragments of the Alzheimer's amyloid peptide β/A4. In two series of fragments related to residues 15−23 and 12−26 of β/A4, we show that Pro replacement of any residue in the amyloidogenic sequence LVFFAED, corresponding to residues 17−23, leads to essentially complete loss of fibril formation and to excellent peptide solubility. Since peptidyl−prolyl bonds are incapable of forming standard extended chain conformations, the results suggest that residues 17−23 make up the β-sheet core of the fibrils formed by these fragments. In contrast to the proline replacements, alanine substitutions at residues 17, 18, and 20 have no effect on fibril formation, while replacement of Phe[19] reduces fibril formation to 15% of the level found for the wild type sequence. Scanning proline mutagenesis should play a useful role in mapping the secondary structural features of larger amyloidogenic peptide sequences, including longer, physiologically relevant forms of β/A4. In addition, these results suggest explanations for some amyloidogenic effects observed in disease-related peptides and also suggest a possible role for aggregation-inhibiting insertion of prolines in protein evolution and protein design.

Amyloid is a highly insoluble, aggregated state of certain polypeptide sequences which is associated with a number of human pathologies, including Alzheimer's disease, light chain amyloidosis, and familial amyloidosis (Cohen & Connors, 1987; Castano & Frangione, 1988; Benson et al., 1989; Stone, 1990). Although the sequences of these polypeptides are not homologous, the morphology of amyloid is similar in all diseases. The deposits consist of fibrils which are 50−100 Å in width and of indefinite length (Cohen & Calkins, 1959; Merz et al., 1983; Miyakawa et al., 1986; Roher et al., 1986). Biophysical characterization shows that the fibrils are rich in β-extended chain, most likely consisting of some form of stacked, antiparallel β-sheet called a cross-β structure (Eanes & Glenner, 1968; Bonar et al., 1969; Kirschner et al., 1986). Ordered structure is also suggested by the ability of many amyloid fibrils to display birefringence after binding the dye Congo red (Cooper, 1974). The existence of familial forms of amyloidosis (Benson & Wallace, 1989), as well as data on synthetic peptides (Hilbich et al., 1991a,b, 1992), shows that fibril formation can be very sensitive to single amino acid replacements. This argues for an underlying specific arrangement of polypeptide within the ordered fibrils. Unfortunately, the poor solubility and paracrystalline nature of amyloid fibrils have restricted knowledge of its structure to data from those few biophysical methods of modest resolution which can be used on noncrystalline solids. Although a higher resolution technique, solid state NMR, has recently been applied to fibril structure (Spencer et al., 1991), to date there is no convincing model of how a particular amino acid sequence is arranged within the cross-β structure of a fibril.

Although it remains controversial, there is increasing belief that the amyloid fibrils formed in the brain in Alzheimer's disease play an important causative link in the pathology (Selkoe, 1991; Sisodia & Price, 1992; Multhaup et al., 1993), in analogy to other amyloid diseases. These fibrils are composed primarily of a peptide called β/A4, which is derived—presumably by proteolysis—from the membrane spanning region of a receptor-like molecule called the Alzheimer's precursor protein (Kang et al., 1987). The predominant β/A4 peptides isolated from amyloid plaques are 40−42 amino acids in length (Roher et al., 1993). As is the case with most amyloid diseases investigated, purified β/A4, as well as many of its fragments, can be induced to form fibrils *in vitro* without the aid of other biological factors (Kirschner et al., 1987). Such synthetic aggregates, generated by "aging" solutions of pure β/A4 peptide, have been found to be toxic to neurons and other cells *ex vivo* (Pike et al., 1991, 1993), supporting an important role for amyloid in Alzheimer's pathology. Studies using β/A4 analogs and fragments suggest that β/A4 adopts an antiparallel β-sheet composed of strands involving regions near residues 12−25 and 30−40 (Hilbich et al., 1991a).

Systematic replacement of the amino acids in a polypeptide sequence with a probe amino acid was introduced to explore the individual contributions of amino acids to the biological activity/receptor binding of peptides (Regoli et al., 1974; Drouin et al., 1979), and with the advent of recombinant DNA methods has been extended to proteins (Cunningham & Wells, 1989). Although the most common replacing residue used is alanine, in principle such "scanning mutagenesis" can be performed with any amino acid that is diagnostic for a particular effect. For example, scanning cysteine mutagenesis has been used to identify sites compatible with disulfide formation (Kanaya et al., 1991). Since

* To whom correspondence should be addressed.
‡ Protein Biochemistry Department, SmithKline Beecham.
® Abstract published in *Advance ACS Abstracts,* December 15, 1994.

0006-2960/95/0434-724$09.00/0

the occurrence of the amino acid proline in an extended chain is energetically highly unfavorable (Kim & Berg, 1993; Minor & Kim, 1994; Smith et al., 1994) due to constraints on its ability to support the required peptide backbone conformation (Richardson & Richardson, 1989), we considered that a corresponding technique using prolines, "scanning proline mutagenesis", might be useful for determining the secondary structure features adopted by different portions of a polypeptide in an amyloid fibril. We describe here the use of this technique to explore the ability of peptides consisting of the 12−26 region of β/A4 to support fibril formation when containing Pro at various positions. The results support models in which the 16−24 sequence KLVFFAEDV forms the core of one strand of β-sheet in the β/A4 fibril and suggest that scanning proline mutagenesis will prove a useful technique for mapping secondary structure in fibrils of β/A4 itself as well as other, longer amyloidogenic sequences.

MATERIALS AND METHODS

Materials. The peptides corresponding to β/A4 (12−26) with various proline substitutions were purchased from Chiron Mimetopes. The peptides corresponding to β/A4-(15−23;Q15K) along with the various proline-substituted analogs were synthesized using standard Merrifield solid phase methods (Tam et al., 1983). All synthetic peptides were HPLC purified on a Vydac C18 reversed phase column using linear gradients of buffer B (0.1% TFA in acetonitrile) in buffer A (0.1% TFA in water). Purity was confirmed by analytical HPLC. Correct amino acid compositions of purified peptides were confirmed by laser desorption mass spectrometry (Table 1).

Fibrillization of β/A4 Analogs. Peptides were solubilized in 0.1% acetic acid at concentrations ranging from 1 to 2.5 mg/mL. For fibrillization studies peptides were incubated for 2 h in 25 μL of 50 mM 2-(*N*-morpholine)ethanesulfonic acid (MES), pH 5.8 at room temperature (RT), and at peptide concentrations ranging from 0.26 to 0.49 mg/mL. The mildly acidic pH of 5.8 was used since optimum fibril formation of both β/A4(1−40) and β/A4(12−28) occurs in the pH 5.5−6.5 range (Wood et al., 1994). In this peptide concentration range and under these conditions, both fibril formation and the binding of Congo red to fibril are essentially complete: thus, the Congo red signal is linear with peptide concentration for fibril formation reactions of β/A4(1−40) when peptide concentrations between 0.2 and 1.0 mg/mL are allowed to make fibrils (data not shown). All fibrillization incubations were set up in duplicate with parallel incubations in a microtiter plate well (for Congo red analysis) and in a 1.5 mL Eppendorf tube (for RP-HPLC analysis).

Congo Red Binding Assay. Fibrils were detected using the Congo red binding method (Klunk et al., 1989). After the fibrillization step, 225 μL of a 22.2 μM Congo red solution in 5 mM phosphate buffer with 0.15 M NaCl, pH 7.4, was added to the fibrillization incubation and allowed to bind at room temperature. After 30 min, the absorbance of the Congo red/fibril mixture was read at A^{540} and A^{480}. The concentration of Congo red bound to fibril, or C_b, was calculated using the equation:

$$C_b\ (\mu M) = A^{540}/25295 - A^{480}/46306$$

HPLC Assay. Fibril formation was induced, as described above, in a 1.5 mL Eppendorf tube. After 2 h, the tubes were centrifuged at high speed for 5 min in a tabletop centrifuge. The supernatant was then analyzed by RP-HPLC on a C18 column. The area of the A^{215} peptide peak was integrated, and the amount of peptide represented by the peak area was estimated from a standard curve generated using related peptides of known concentration (based on amino acid composition analysis).

Circular Dichroism Spectroscopy. Samples for CD were mixed at RT in 10 mM phosphate buffer, pH 7.0, to a final concentration of 1 mg/mL. Non-proline-containing peptides were filtered in a Millipore 0.45 μM filter unit to remove some aggregated material. (However, the peptide concentration of 1 mg/mL was confirmed on the soluble filtrate by HPLC analysis.) Spectra within the range of 180−260 nm were read in a Jasco-700 spectropolarimeter. Each spectrum represents the average of four repeats run in a 20 mm path length cell at a sensitivity of 20 mdeg/cm.

Electron Microscopy. Peptides were applied to carbon-stabilized collodion-coated copper mesh grids and allowed to stand for 5 min. Excess material was wicked off and the residue negatively stained with 2% aqueous uranyl acetate. After 30 s, excess stain was wicked off and the samples were allowed to dry. Grids were examined with a JEOL 100CX transmission electron microscope operating at 80 kV.

Mass Spectrometry. Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) data for the synthetic peptides were obtained on a Versec "Voyager RP" linear time-of-flight instrument. Samples were suspended at about 0.3 mg/mL in an α-cyano-4-hydroxycinnamic acid/acetonitrile/trifluoroacetic acid mixture along with the 8−32 fragment of salmon calcitonin (MW = 2708.5) as internal standard. Spectra were averaged over about 50 scans, with data collected and analyzed using a customized version of IGOR Pro (WaveMetrics).

RESULTS

Many fragments of β/A4 have been described which are themselves capable of aggregating into fibrils *in vitro* (Castano et al., 1986; Gorevic et al., 1987; Kirschner et al., 1987; Halverson et al., 1990; Fraser et al., 1991; Hilbich et al., 1991a,b, 1992; Barrow et al., 1992; Caputo et al., 1992; Inouye et al., 1993; Jarrett et al., 1993; Pike et al., 1993). Since we were already exploring structure−function relationships in the sequence region 12−26 (Wood et al., 1994), which corresponds to one of two extended chain regions in models of β/A4 fibril structure (Hilbich et al., 1991a), we chose to explore this same region for the effect of introduced prolines. As expected from previous work (Fraser et al., 1991; Hilbich et al., 1991a), we found that the wild type 12−26 region aggregated to form a precipitate which binds Congo red (Table 1). We also designed a smaller peptide, β/A4(15−23), a sequence which other studies have suggested forms the core, β-sheet forming region of the 12−26 extended chain in β/A4(1−40) (Hilbich et al., 1992). Unfortunately, the 15−23 peptide composed of the wild type sequence of this peptide, QKLVFFAED, was highly insoluble and difficult to work with (data not shown) and, therefore, was not pursued. However, we determined that the single amino acid replacement peptide, β/A4(15−23;Q15K), is soluble in aqueous acetic acid. This Gln^{15}→Lys mutant of the 15−23 peptide forms Congo red binding aggregate when exposed to appropriate conditions (see

Table 1: Amyloid Formation/Solubilities of β/A4 Fragment Analogs

Peptide	m/z[1] calculated	m/z[1] observed	Fibril Formation[2]	% Maximal Fibril Formation[3]	% Insoluble[4]
V**P**HQKLVFFAEDVGS	1673.33	1672.9	0.33	81.6	98.8
VH**P**QKLVFFAEDVGS	1673.33	1672.9	0.29	71.0	61.7
VHH**P**KLVFFAEDVGS	1682.27	1681.9	0.19	45.9	43.9
VHHQ**P**LVFFAEDVGS	1682.87	1681.8	0.06	14.7	55.0
KK**P**VFFAED	1080.82	1080.6	0	0	n.d.
KKL**P**FFAED	1094.84	1094.6	0	0	n.d.
KKLV**P**FAED	1046.84	1046.6	0	0	n.d.
VHHQKLV**P**FAEDVGS	1663.33	1662.9	0	0	0
KKLVF**P**AED	1046.84	1046.6	0	0	n.d.
KKLVFF**P**ED	1122.86	1122.6	0	0	4.3
VHHQKLVFFA**P**DVGS	1681.34	1680.9	0	0	0
VHHQKLVFFAE**P**VGS	1695.36	1694.9	0	0	7.4
VHHQKLVFFAED**P**GS	1711.33	1710.9	0.05	12.3	28.1
VHHQKLVFFAEDV**P**S	1753.37	1752.9	0.14	33.9	46.9
KKLVFFAED	1096.85	1096.8	0.34	100	96.5
VHHQKLVFFAEDVGS	1713.34	1712.9	0.41	100	87.1

[1] Mass/charge ratio (molecular weight of protonated parent ion) in laser desorption mass spectrometry. [2] Ratio of concentration of Congo red bound to concentration of peptide in assay (see Materials and Methods). [3] % calculated using the appropriate control, Pro-free peptide for each sequence. [4] % insoluble is calculated from the RP-HPLC assay described in Materials and Methods.

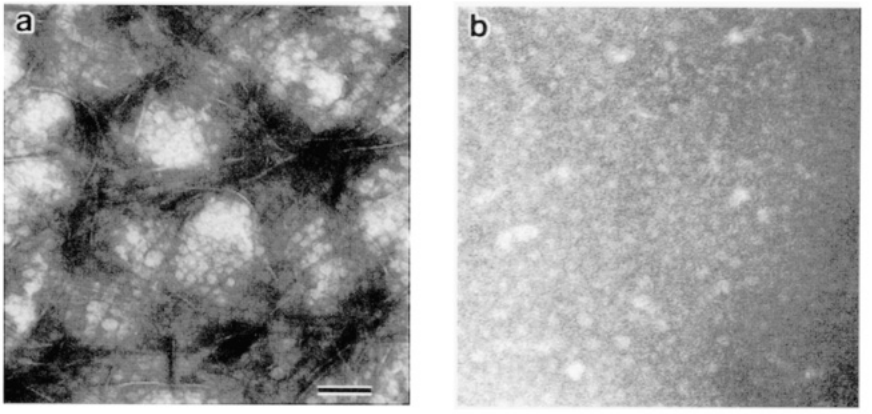

FIGURE 1: Electron micrograph of fibril formation reactions. Bar inset = 50 nm. (a) Grid containing reaction of the sequence KKLVFFAED incubated at 2 mg/mL in 0.25 M potassium acetate, pH 5.2. Mean fibril width, from 50 measurements, is 45.9 ± 6.7 Å using Optimas image analysis software (Edmonds, WA). (b) Grid containing reaction of a solution of 4 mg/mL KKLV**P**FAED incubated in the same buffer.

Materials and Methods), and this aggregate displays typical fibril structure in electron micrographs (Figure 1a). We therefore used this Q15K mutant of β/A4(15−23) to explore the effect of Pro replacements in a smaller peptide series.

The degree of fibril formation was quantified by two methods: (a) the binding of Congo red, a diazo dye which binds amyloid structures (Puchtler et al., 1962), and (b) the depletion of soluble β/A4 monomer as monitored by HPLC. The quantitation of Congo red binding to fibril, by monitoring the shift in the Congo red spectrum upon binding, was used previously by Klunk to determine binding constants and stoichiometry of binding of Congo red to preformed fibril (Klunk et al., 1989). By assuming unchanged binding stoichiometry and strong affinity for Congo red binding to fibrils regardless of peptide sequence, we adapt the method here to quantify the degree of fibril formation. Table 1 displays the relative fibril formation for the series of peptides examined in this work as determined both by the Congo red binding method and by estimation of remaining soluble peptide using HPLC. In general, the results agree very well. Only one peptide, in which a Lys is replaced with Pro, gives significant disagreement between the two methods. In this case, less soluble peptide is obtained (by HPLC) than is expected on the basis of the amount of Congo red binding

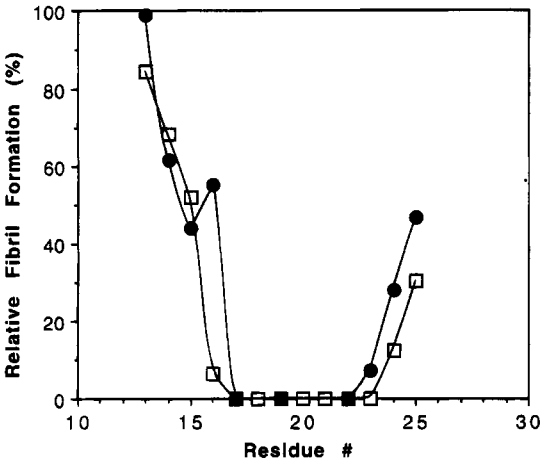

FIGURE 2: Extent of fibril formation of a series of proline-containing peptides relative to the fibril formation in corresponding proline-minus sequences VHHQKLVFFAEDVGS and KKLVFFAED. The x-axis is the residue number of the amino acid replaced by proline corresponding to the data point. The y-axis is the percent of fibril formation compared to control peptide as deduced by quantitative Congo red binding (□) or the percent insolubilization compared to control as determined by quantifying soluble peptide in the supernate by reversed phase HPLC (●). See Materials and Methods.

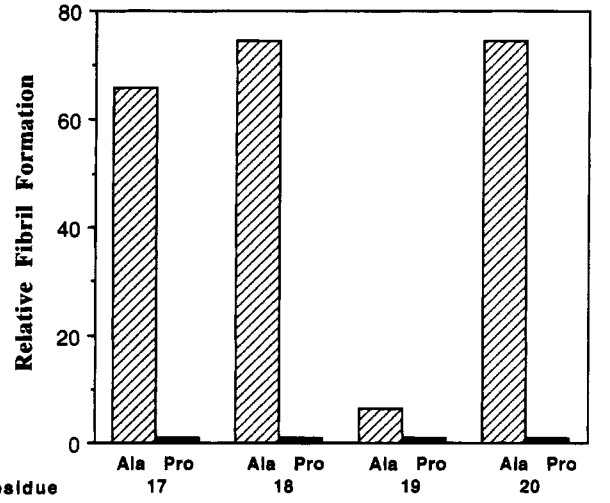

FIGURE 3: Relative fibril formation by analogs of β/A4(15−23;Q15K), or KKLVFFAED, containing either proline or alanine replacements. Data are expressed as a percentage of the fibril formation by KKLVFFAED itself, as determined by the Congo red binding assay. Structures of the alanine replacement peptides were confirmed by mass spectrometry or amino acid composition analysis.

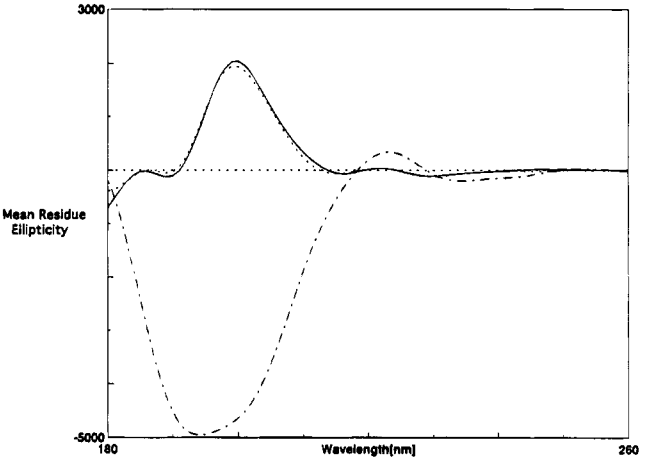

FIGURE 4: Circular dichroism spectra of KKLVFFAED (—), KKLAFFAED (- - -), and KKLPFFAED (— • —) determined as described in Materials and Methods.

aggregate. This suggests a significant amount of nonfibrillar aggregate, which for this peptide may be attributed to the effect of the removal of a charged residue by the Pro replacement. The good agreement between these methods shown in Figure 2 suggests that the Congo red assay can be legitimately used to quantify fibril formation in such experiments.

The results in Table 1 and Figures 1 and 2 show dramatic effects of Pro replacements on the extent of fibril formation. A proline substitution anywhere in the region of residues 17−23 renders the peptide incapable of binding Congo red. Prolines at the neighboring positions, 16 or 24, diminish the fibril forming capacity of a peptide by approximately 80%. Prolines at position 15 or 25 reduce the capacity by about 50% whereas prolines at positions 13 or 14 only slightly inhibit fibrillization. Since it could be argued that the diminution of fibril formation after Pro replacement might reflect a specific positive role of the WT amino acid, rather than a negative role for the Pro residue, we evaluated a series of alanine replacements as controls. Figure 3 shows that, of the four residue positions explored by alanine replacements, three had no effect on fibril formation. Only the replacement of Phe^{19} significantly reduced fibril formation. In contrast, proline replacement at all four positions completely eliminated fibril formation. The differing extents to which the replacement of adjacent phenylalanines with alanine affect fibril formation are indicative of the specific packing interactions which must be at work in making up the fibril structure.

Importantly, Pro replacement at a particular residue influences fibril formation in a similar fashion in both peptide series. Replacement of Leu^{17} or Ala^{21} with Pro completely eliminates fibril formation in the 15−23 peptide, even though the replaced residue is only two residues away from the terminus of the peptide. In contrast, in the 12−26 series, Pro replacement of the amino acids His^{14} and Gln^{15}, which are 2−3 residues removed from the N-terminus, continues to result in good fibril formation, suggesting that residues 12−15 are not intimately involved in fibril structure. (We have not investigated the possibility that the histidines may be more important for fibril formation at higher pH values where they are expected to be deprotonated.) A similar argument can be made for the lack of importance of residues 25−26 in fibril structure.

The fundamental β-extended chain structure of the aggregated states formed by these peptides, and the ability of Pro replacement to inhibit β-extended chain formation, is confirmed by examination of spectral properties of three β/A4(15−23;Q15K) peptides. KKLVFFAED, KKLAFFAED, and KKLPFFAED were examined by circular dichroism (Figure 4), which has been used previously to observe the transition of β/A4 peptides from random coil to a β-sheet structure (Hilbich et al., 1991a,b; Barrow et al., 1992). The spectra of the two non-proline-containing peptides are superimposable, both exhibiting a positive maximum at 200 nm and minima at 218 and 228 nm. The 200 nm maximum and the 218 nm minimum are typical for antiparallel β-sheet proteins (Yang et al., 1986). The spectrum of proline-containing KKLPFFAED has neither of

Table 2: Human Amyloid Proteins[a]

protein	longest sequence lacking Pro	length of peptide found in amyloid	ref
atrial natriuretic factor	28	~26	Westermark, 1994
amylin[b]	37	37	Westermark et al., 1986
β/A4	42	42	Castano & Frangione, 1988
gelsolin	62	70	Maury, 1990
cystatin C[c]	64	110	Ghiso et al., 1986
serum amyloid A	76	74−87	Prelli et al., 1991
transthyretin[d]	42	≥81	Westermark et al., 1990b
apolipoprotein A1[e]	58	~85	Nichols et al., 1987
Ig light chain[f]	30−50	~110	Kabat et al., 1991
β_2 microglobulin	39	99	Connors et al., 1985
lysozyme	70	130	Pepys et al., 1993

[a] Amino acid sequences were accessed from the Swiss Protein Data Base. [b] Amylin is also known as islet amyloid polypeptide (IAPP).[c] The large fragment of cystatin C found in amyloid begins at amino acid residue number 11, just before the long Pro-minus sequence 14−77; the excised N-terminal decapeptide contains three prolines. [d] The smallest transthyretin fragment found in amyloid begins at residue 46, at the beginning of the long Pro-minus sequence 44−85. [e] Native apoA1 has 243 residues. The fragment found in amyloid is approximately the N-terminal 85 residues, and this fragment includes the longest Pro-minus sequence in the entire 243 amino acid protein, from residues 8 to 65. [f] Immunoglobulin light chains vary considerably in sequence, so it is practical to only list the range of Pro-minus sequences within which the majority of light chains fall.

these features; its predominant feature is a minimum near 200 nm, normally assigned to random coil structure (Yang et al., 1986). A similar random coil spectrum was obtained for the proline analog KKLVFFPED (not shown).

DISCUSSION

It has been well established for over twenty years (Chou & Fasman, 1973) that proline residues are found only infrequently in β-sheets. There are three characteristics of the proline residue which can account for this behavior (Richardson & Richardson, 1989): (a) the conformation of a peptidyl−prolyl bond ($\phi = -60°$) imposed by the proline ring is incompatible with peptide bond geometries found in typical β-sheets ($\phi = -120°$, $\psi = 140°$), (b) the ring of the proline cannot otherwise sterically fit into the β-sheet H-bonding network, and (c) the peptide bond nitrogen of a peptidyl−prolyl group is not available to contribute to the β-sheet H-bonding network. In fact, a recent analysis of β-sheet structures shows that prolines *never* occur in the interior of antiparallel β-sheets (Wouters & Curmi, 1994). Such structural analyses are further supported by recent studies quantifying the effect on protein stability by substituting a surface-exposed β-sheet position with various amino acids (Kim & Berg, 1993; Minor & Kim, 1994; Minor et al., 1994; Smith et al., 1994). If the fundamental structural unit of amyloid fibrils is the β-extended chain, then amyloidogenic sequences should likewise be devoid of an amino acid which is incapable of supporting β-sheet formation. (It is possible that Pro residues might be accommodated within amyloidogenic sequences, if they are located in turns of loops responsible for chain reversal.) In fact, in most amyloidosis-associated proteins, proline either does not occur at all or is found only rarely, leaving long stretches of proline-free sequence (Table 2). These observations suggested to us that amyloidogenic sequences may be intolerant of Pro replacement and, in addition, suggested that Pro replacement might be a useful diagnostic tool in identifying amyloidogenic portions.

Due to the synthetic difficulties involved in producing a series of "full-length" β/A4(1−40) analogs, we decided to test these hypotheses first on smaller amyloidogenic fragments of β/A4. The results presented here on the 12−26 and 15−23 fragments support these concepts and suggest that residues 17−23 compose the core of the β-sheet formed when these peptides make amyloid fibrils. Since it is conceivable that β/A4(1−40) may undertake fibril formation in such a way as to utilize the 17−23 region differently, these experiments would have to be repeated on full-length peptides to confirm the results for physiologically relevant sequences. It is, nonetheless, interesting that other workers have also identified the 17−20 region of full-length β/A4 peptides as being important for fibril formation (Hilbich et al., 1992). Further, in agreement with data presented here, other work with different amino acid replacements in β/A4 fragments also suggests that the His-His sequence at residue positions 13 and 14 is not involved in cross-β sheet (Fraser et al., 1994). The results presented here neither support nor contradict models of fibril formation for the full-length peptide that suggest an additional important role for the 29−42 sequence (Hilbich et al., 1992; Jarrett et al., 1993).

The work presented here suggests that scanning proline mutagenesis carried out on longer sequences, of β/A4 and other amyloidogenic polypeptides, should be an effective tool for identifying the segments that are intimately involved in amyloid fibril structure—important information which is currently difficult to obtain. We have also recently used this technique to block amyloid formation in another β/A4 sequence fragment known to be both amyloidogenic and toxic, in order to confirm the relationship between toxicity and aggregation (Wetzel et al., 1994).

In solvents which generally favor α-helix formation, β/A4-(1−28) exists as an α-helix (Talafous et al., 1994). At the same time, the 12−28 sequence and its fragments strongly favor β-sheet structure and amyloid formation when incubated in native buffer around pH 6. Yet again, the insertion of a single Pro residue within one of these fragments leads to a random coil conformation (Figure 4). These results can be understood in terms of the low energy barriers to conformational interchange of most peptides in solution. The β/A4(1−28) sequence and its fragments presumably possess such flexibility and are thus capable of being held into α-helix by appropriate solvent additives, or into β-sheet by the thermodynamic driving forces associated with β-sheet-mediated aggregation and precipitation. A single Pro replacement only modestly alters the spectrum of conformations accessible to monomeric peptide in native buffers, but it significantly changes peptide thermodynamics by eliminating the possibility of β-sheet formation. The intrinsic ability of many peptide sequences to adapt either α-helix or β-sheet, and the ability of the packing forces associated with the formation of both globular and amyloid protein structure to influence and control conformation, is presumably the basis for the α-helix to β-sheet transitions which must occur in the formation of amyloid by predominantly α-helical proteins like apolipoprotein A1 (Nichols et al., 1987) and lysozyme (Pepys et al., 1993), and which are now thought to be important in aggregation and infectivity of prions (Pan et al., 1993).

Knowledge of the proline effect may contribute to the understanding of particular amyloidogenic sequence alterations which have been described. (1) In human amylin, the GAILS sequence at residues 24−28 has been implicated by other experiments to be the amyloidogenic core; in animals which do not exhibit this form of amyloidosis, this region contains 1−2 proline residues, whereas in animals which do, this sequence region lacks prolines (Westermark et al., 1990a). (2) In the senescence accelerated mouse (SAM), the 79 amino acid long polypeptide apoA-II is deposited into amyloid (Kunisada et al., 1986). Wild type murine apoA-II contains a 45 amino acid long sequence free of proline. In the SAM, apoA-II has undergone two mutations rendering it amyloidogenic: $Pro^{5}\rightarrow$Gln (extending the Pro-minus sequence to 50 residues) and $Val^{38}\rightarrow$Ala (within this putative amyloidogenic, Pro-minus region). (3) The sequence 100−130 is a region of particular interest in the PrP protein associated with scrapie and human spongiform encephalopathies (Prusiner, 1992). This region contains four of the mutations associated with genetic forms of the disease (Prusiner, 1992; Kitamoto et al., 1993), and a peptide portion of this region is itself capable of forming fibrils *in vitro* (Gasset et al., 1992; Forloni et al., 1993) and is neurotoxic on cultured cells (Forloni et al., 1993). Two of the four genetic lesions identified in this region are Pro→Leu mutations, at positions 102 and 105 (Doh et al., 1989; Hsiao et al., 1989; Kitamoto et al., 1993). (4) The placement of Pro residues may contribute to the amyloidogenicity of certain immunoglobulin variable domain sequences (R. Wetzel, unpublished observations). Such proline effects would also be consistent with the view that sequence effects on amyloidogenicity of globular proteins can be of two types, replacements which destabilize tertiary structure to allow the formation of an aggregation-prone folding intermediate (Hurle et al., 1994), and replacements which alter the ability of exposed segments of the intermediate to undergo effective cross-β formation (Wetzel, 1994).

Given its unusual structure and properties, it might be expected that proline would generally play an important and unique role in the structure and properties of globular proteins. For the most part, however, this does not seem to be the case. Occassionally, prolines appear to make unique structural contributions, such as in the case of some *cis* peptidyl−prolyl bonds, but in many proteins prolines can be replaced by other amino acids without greatly influencing the folding pattern and/or stability of the product (Alber et al., 1988; Kiefhaber et al., 1990; Herning et al., 1992; Texter et al., 1992; Tweedy et al., 1993). On the basis of the results described here, we speculate that prolines are sometimes utilized in protein evolution not for any significant, unique contribution they make to the structure and function of the *native* state, but rather for their role in effectively blocking potentially devastating off-pathway aggregation reactions, such as amyloid and inclusion body formation, during folding *in vivo* (Wetzel, 1994). Formation of "amorphous" protein aggregates both *in vitro* and *in vivo* is associated with high levels of β-sheet (Clark et al., 1981; Oberg et al., 1994; Przybycien et al., 1994), suggesting that proline insertion might block other modes of the abnormal assembly of proteins besides amyloid formation. Since proline residues can be accommodated in most elements of regular and irregular secondary structure, including in some helices (Alber et al., 1988) and on the edge strands of β-sheets (Wouters & Curmi, 1994), there would seem to be many sites in the polypeptide sequence where aggregation-blocking prolines might be inserted without affecting the structure and function of the native protein. Whether or not protein insertion plays such a role in protein evolution, the above reasoning suggests it should be useful in protein design, since aggregation *in vivo* and *in vitro* is a major barrier to the production of novel designed proteins.

ACKNOWLEDGMENT

We are grateful to Kyle Mansfield, who assisted in the purification and characterization of many of these peptides, and Dean McNulty for amino acid composition analysis. We also thank Anne Klinkner, Beverly Maleeff, and Timothy Hart of SB for the electron micrographs.

REFERENCES

Alber, T., Bell, J. A., Sun, D.-P., Nicholson, H., Wozniak, J. A., Cook, S. & Matthews, B. W. (1988) *Science 239*, 631−635.

Barrow, C. J., Yasuda, A., Kenny, P. T., & Zagorski, M. G. (1992) *J. Mol. Biol. 225*, 1075−1093.

Benson, M. D., & Wallace, M. R. (1989) in *The Metabolic Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D., Ed.) pp 2439−2460, McGraw-Hill, New York.

Benson, M. D., Dwulet, F. E., Madura, D., & Wheeler, G. (1989) *Scand. J. Immunol. 29*, 175−179.

Bonar, L., Cohen, A. S., & Skinner, M. M. (1969) *Proc. Soc. Exp. Biol. Med. 131*, 1373−1375.

Caputo, C. B., Fraser, P. E., Sobel, I. E., & Kirschner, D. A. (1992) *Arch. Biochem. Biophys. 292*, 199−205.

Castano, E. M., & Frangione, B. (1988) *Lab. Invest. 58*, 122−132.

Castano, E. M., Ghiso, J., Prelli, F., Gorevic, P. D., Migheli, A., & Frangione, B. (1986) *Biochem. Biophys. Res. Commun. 141*, 782−789.

Chou, P. Y., & Fasman, G. D. (1973) *Biochemistry 13*, 211−222.

Clark, A. H., Saunderson, D. H., & Suggett, A. (1981) *Int. J. Pept. Protein Res. 17*, 353−364.

Cohen, A. S., & Calkins, E. (1959) *Nature 183*, 1202−1203.

Cohen, A. S., & Connors, L. H. (1987) *J. Pathol. 151*, 1−10.

Connors, L. H., Shirahama, T., Skinner, M., Fenves, A., & Cohen, A. S. (1985) *Biochem. Biophys. Res. Commun. 131*, 1063−1068.

Cooper, J. H. (1974) *Lab. Invest. 31*, 232−238.

Cunningham, B. C., & Wells, J. A. (1989) *Science 244*, 1081−1085.

Doh, K., Tateishi, J., Sasaki, H., Kitamoto, T., & Sakaki, Y. (1989) *Biochem. Biophys. Res. Commun. 163*, 974−979.

Drouin, J. N., Gaudreau, P., St, P. S. A., & Regoli, D. (1979) *Can. J. Physiol. Pharmacol. 57*, 562−566.

Eanes, E. D., & Glenner, G. G. (1968) *J. Histochem. Cytochem. 16*, 673−677.

Forloni, G., Angeretti, N., Chiesa, R., Monzani, E., Salmona, M., Bugiani, O., & Tagliavini, F. (1993) *Nature 362*, 543−546.

Fraser, P. E., Nguyen, J. T., Surewicz, W. K., & Kirschner, D. A. (1991) *Biophys. J. 60*, 1190−1201.

Fraser, P. E., McLachlan, D. R., Surewicz, W. K., Mizzen, C. A., Snow, A. D., Nguyen, J. T., & Kirschner, D. A. (1994) *J. Mol. Biol.* (in press).

Gasset, M., Baldwin, M. A., Lloyd, D. H., Gabriel, J.-M., Holtzman, D. M., Cohen, F., Fletterick, R., & Prusiner, S. B. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10940−10944.

Ghiso, J., Jensson, O., & Frangione, B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2974−2978.

Gorevic, P. D., Castano, E. M., Sarma, R., & Frangione, B. (1987) *Biochem. Biophys. Res. Commun. 147*, 854−862.

Halverson, K., Fraser, P. E., Kirschner, D. A., & Lansbury, P. T. J. (1990) *Biochemistry 29*, 2639−2644.

Herning, T., Yutani, K., Inaka, K., Kuroki, R., Matsushima, M., & Kikuchi, M. (1992) *Biochemistry 31*, 7077−7085.

Hilbich, C., Kisters, W. B., Reed, J., Masters, C. L., & Beyreuther, K. (1991a) *J. Mol. Biol. 218*, 149−163.

Hilbich, C., Kisters, W. B., Reed, J., Masters, C. L., & Beyreuther, K. (1991b) *Eur. J. Biochem. 201*, 61−69.

Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C. L., & Beyreuther, K. (1992) *J. Mol. Biol. 228*, 460−473.

Hsiao, K., Baker, H. F., Crow, T. J., Poulter, M., Owen, F., Terwilliger, J. D., Westaway, D., Ott, J., & Prusiner, S. B. (1989) *Nature 338*, 342−345.

Hurle, M. R., Helms, L. R., Li, L., Chan, W., & Wetzel, R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5446−5450.

Inouye, H., Fraser, P. E., & Kirschner, D. A. (1993) *Biophys. J. 64*, 502−519.

Jarrett, J. T., Berger, E. P.; & Lansbury, P. T., Jr. (1993) *Biochemistry 32*, 4693−4697.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., & Foeller, C. (1991) *Sequences of Proteins of Immunological Interest*, U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Washington, DC.

Kanaya, E., Kanaya, S., & Kikuchi, M. (1991) *Biochem. Biophys. Res. Commun. 173*, 1194−1199.

Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K.-H., Multhaup, G., Beyreuther, K., & Muller-Hill, B. (1987) *Nature 325*, 733−736.

Kiefhaber, T., Grunert, H.-P., Hahn, U., & Schmid, F. X. (1990) *Biochemistry 29*, 6475−6480.

Kim, C. A., & Berg, J. M. (1993) *Nature 362*, 267−270.

Kirschner, D. A., Abraham, C., & Selkoe, D. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 503−507.

Kirschner, D. A., Inouye, H., Duffy, L. K., Sinclair, A., Lind, M., & Selkoe, D. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6953−6957.

Kitamoto, T., Amano, N., Terao, Y., Nakazato, Y., Isshiki, T., Mizutani, T., & Tateishi, J. (1993) *Ann. Neurol. 34*, 808−813.

Klunk, W. E., Pettegrew, J. W., & Abraham, D. J. (1989) *J. Histochem. Cytochem. 37*, 1273−1281.

Kunisada, T., Higuchi, K., Aota, S., Takeda, T., & Yamagishi, H. (1986) *Nucleic Acids Res. 14*, 5729−5740.

Maury, C. P. J. (1990) *Biochim. Biophys. Acta 1096*, 84−86.

Merz, P. A., Wisniewski, H. M., Somerville, R. A., Bobin, S. A., Masters, C. L., & Iqbal, K. (1983) *Acta Neuropathol. 60*, 113−124.

Minor, D. L., Jr., & Kim, P. S. (1994) *Nature 367*, 660−663.

Minor, J., Daniel, L., & Kim, P. S. 91994) *Nature 371*, 264−267.

Miyakawa, T., Watanabe, K., & Katsuragi, S. (1986) *Virchows Arch., B 52*, 99−106.

Multhaup, G., Masters, C. L., & Beyreuther, K. (1993) *Biol. Chem. Hoppe-Seyler 374*, 1−8.

Nichols, W. C., Dwulet, F. E., & Benson, M. D. (1987) *Clin. Res. 35*, 595A.

Oberg, K., Chrunyk, B. A., Wetzel, R., & Fink, A. (1994) *Biochemistry 33*, 2628−2634.

Pan, K.-M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., & Prusiner, S. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10962−10966.

Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M., & Hsuan, J. J. (1993) *Nature 362*, 553−557.

Pike, C. J., Walencewicz, A. J., Glabe, C. G., & Cotman, C. W. (1991) *Brain Res. 563*, 311−314.

Pike, C. J., Burdick, D., Walencewicz, A. J., Glabe, C. G., & Cotman, C. W. (1993) *J. Neurosci. 13*, 1676−1687.

Prelli, F., Pras, M., Shtrasburg, S., & Frangione, B. (1991) *Scand. J. Immunol. 33*, 783−786.

Prusiner, S. B. (1992) *Biochemistry 31*, 12277−12288.

Przybycien, T. M., Dunn, J. P., Valax, P., & Georgiou, G. (1994) *Protein Eng. 7*, 131−136.

Puchtler, H., Sweat, F., & Levine, M. (1962) *J. Histochem. 10*, 355−364.

Regoli, D., Park, W. K., & Rioux, F. (1974) *Pharmacol. Rev. 26*, 69−123.

Richardson, J. S., & Richardson, D. C. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., Ed.) pp 1−98, Plenum, New York.

Roher, A., Wolfe, D., Palutke, M., & KuKuruga, D. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2662−2666.

Roher, A. E., Lowenson, J. D., Clarke, S., Woods, A. S., Cotter, R. J., Gowing, E., & Ball, M. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10836−10840.

Selkoe, D. J. (1991) *Neuron 6*, 487−498.

Sisodia, S. S., & Price, D. L. (1992) *Curr. Opin. Neurobiol. 2*, 648−652.

Smith, C. K., Withka, J. M., & Regan, L. (1994) *Biochemistry 33*, 5510−5517.

Spencer, R. G. S., Halverson, K. J., Auger, M., McDermott, A. E., Griffin, R. G., & Lansbury, P. T., Jr. (1991) *Biochemistry 30*, 10382−10387.

Stone, M. J. (1990) *Blood 75*, 531−545.

Talafous, J., Marcinowski, K. J., Klopman, G., & Zagorski, M. G. (1994) *Biochemistry 33*, 7788−7796.

Tam, J. P., Heath, W. F., & Merrifield, R. B. (1983) *J. Am. Chem. Soc. 105*, 6442.

Texter, F. L., Spencer, D. B., Rosenstein, R., & Matthews, C. R. (1992) *Biochemistry 31*, 5687−5691.

Tweedy, N. B., Nair, S. K., Paterno, S. A., Fierke, C. A., & Christianson, D. W. (1993) *Biochemistry 32*, 10944−10949.

Westermark, P. (1994) *Amyloid 1*, 47−60.

Westermark, P., Westermark, C., Wilander, E., & Sletten, K. (1986) *Biochem. Biophys. Res. Commun. 140*, 827−831.

Westermark, P., Engstrom, U., Johnson, K. H., Westermark, G. T., & Betsholtz, C. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5036−5040.

Westermark, P., Sletten, K., Johansson, B., & Cornwell, G. G. 3. (1990b) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2843−2845.

Wetzel, R. (1994) *Trends Biotechnol. 12*, 193−198.

Wetzel, R., Wood, S. J., Davis, J. B., & Hurle, M. R. (1994) *Neurobiol. Aging 15* (Suppl. 1), S50.

Wood, S. J., Wetzel, R., & Hurle, M. (1994) in *Proceedings of the 8th International Symposium on Amyloidosis*, pp 523−525, Parthenon, New York.

Wouters, M. A., & Curmi, P. M. G. (1994) *Protein Sci. 3* (Suppl. 1), 60, Abstr. 43S.

Yang, J. T., Wu, C.-S. C., & Martinez, H. M. (1986) *Methods Enzymol. 130*, 208−266.

BI942141H