THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 41, pp. 35069–35076, October 14, 2005

Sequence Determinants of Enhanced Amyloidogenicity of Alzheimer Aβ42 Peptide Relative to Aβ40*

Received for publication, May 26, 2005, and in revised form, August 2, 2005 Published, JBC Papers in Press, August 3, 2005, DOI 10.1074/jbc.M505763200

Woojin Kim and Michael H. Hecht[1]

From the Department of Chemistry, Princeton University, Princeton, New Jersey 08544

Aggregation of proteins into insoluble deposits is associated with a variety of human diseases. In Alzheimer disease, the aggregation of amyloid β (Aβ) peptides is believed to play a key role in pathogenesis. Although the 40-mer (Aβ40) is produced *in vivo* at higher levels than the 42-mer (Aβ42), senile plaque in diseased brains is composed primarily of Aβ42. Likewise, *in vitro*, Aβ42 forms fibrils more rapidly than Aβ40. The enhanced amyloidogenicity of Aβ42 could be due simply to its greater length. Alternatively, specific properties of residues Ile[41] and Ala[42] might favor aggregation. To distinguish between these two possibilities, we constructed a library of sequences in which residues 41 and 42 were randomized. The aggregation behavior of the resulting sequences was assessed using a high throughput screen, based on the finding that fusions of Aβ42 to green fluorescence protein (GFP) prevent the folding and fluorescence of GFP, whereas mutations in Aβ42 that disrupt aggregation produce green fluorescent fusions. Correlations between the sequences of Aβ42 mutants and the fluorescence of Aβ42-GFP fusions *in vivo* were confirmed *in vitro* through biophysical studies of synthetic 42-residue peptides. The data reveal a strong correlation between aggregation propensity and the hydrophobicity and β-sheet propensities of residues at positions 41 and 42. Moreover, several mutants containing hydrophilic residues and/or β-sheet breakers at positions 41 and/or 42 were less prone to aggregate than Aβ40 wherein these two residues are deleted entirely. Thus, properties of the side chains at positions 41 and 42, rather than length *per se*, cause Aβ42 to aggregate more readily than Aβ40.

The deposition of insoluble proteins into amyloid plaque is associated with various diseases including Alzheimer disease, prion encephalopathies, Parkinson disease, and Huntington disease (1–4). In Alzheimer disease, amyloid β (Aβ)[2] peptides aggregate into fibrils, which accumulate as insoluble neuritic plaques. A variety of genetic, neuropathological, and biochemical studies suggests that either the fibrils themselves or precursors on the pathway toward these fibrils play a causative role in Alzheimer disease (1, 5–11).

Aβ peptides are produced *in vivo* by proteolytic cleavage of the amyloid precursor protein by β and γ secretases (5). Because of inhomogeneous cleavage by γ secretase, Aβ peptides range in length from 39 to 43 residues. Among these peptides, the 40-mer (Aβ40) and the 42-mer (Aβ42) are abundant in diseased brains (12–14), and these two peptides are the main components of the neuritic plaques in the parenchyma of diseased brains (5, 14).

With the exception of the C-terminal amino acids, Ile[41] and Ala[42], the sequences of Aβ40 and Aβ42 are identical. Despite their 95% sequence identity, Aβ40 and Aβ42 display dramatically different behaviors both *in vivo* and *in vitro*. Biochemical and immunocytochemical studies show that although Aβ40 is major component in cerebrospinal fluid and plasma, senile amyloid plaques formed *in vivo* are composed primarily of Aβ42 (5, 14–16). Moreover, although both peptides aggregate into fibrils *in vitro*, Aβ42 does so more rapidly than Aβ40. Concentrated solutions of Aβ40 are stable for days, whereas comparable solutions of Aβ42 aggregate almost immediately (17, 18).

Not only is Aβ42 more prone to aggregate than Aβ40, but the pathways toward aggregation are also different. Recently, Teplow and coworkers (19, 20) found that carefully prepared aggregate-free Aβ40 occurs as monomers, dimers, trimers, and tetramers in rapid equilibrium. In contrast, Aβ42 preferentially forms pentamer/hexamer units (paranuclei) that assembled further into beaded superstructures similar to early protofibrils.

The different aggregation behaviors of Aβ40 and Aβ42 led us to question the role of the two C-terminal residues, Ile[41] and Ala[42]. Is it simply the increased length of Aβ42 that causes its increased amyloidogenicity? Or are particular features of the Ile[41] and Ala[42] side chains important for amyloid formation? To address these questions, we performed random mutagenesis on positions 41 and 42. The change in aggregation behavior resulting from the mutations was monitored by using fusions of the Aβ42 variants to GFP. Fluorescence of these fusions is inversely correlated with the propensity of the fused Aβ mutant sequence to aggregate (21); GFP fused to wild type Aβ42 does not fluorescence, whereas fusions to less aggregating mutants display increased fluorescence. By correlating the observed levels of fluorescence with the identities of the side chains at positions 41 and 42, we determined the side chain properties at these positions that are responsible for the enhanced aggregation of Aβ42 relative to Aβ40.

MATERIALS AND METHODS

Construction of Libraries of Aβ42-GFP Fusions—Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), and restriction enzymes were purchased from New England Biolabs. Mutations were incorporated into a synthetic Aβ42 gene (21) by PCR using DeepVent polymerase (New England Biolabs) and an Ericomp Easycycler™ Twinblock thermocycler. After PCR, the mutagenized Aβ42 gene inserts and the pET 28 vector containing the GFP gene (21) were doubly digested with BamHI and NdeI. The digested insert and vector were then ligated together using T4 ligase. Plasmids were transformed into the XL1-Blue strain of *Escherichia coli* (Stratagene) and plated for overnight growth on plates containing 50 μg/ml kanamycin.

The primer for random mutagenesis at position 41 and 42 had the following sequence: 5′-TCTTCTGGATCCNNNNNCACCACGCCGCCCACCAT-3′. The primer for mutagenesis encoding a combinatorial mix of polar residues used the degenerate codon NAN. For a

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed. Tel.: 609-258-2901; Fax: 609-258-6746; E-mail: hecht@princeton.edu.

[2] The abbreviations used are: Aβ, amyloid β; GFP, green fluorescent protein; HFIP, hexafluoroisopropanol; HPLC, high pressure liquid chromatography; ThT, thioflavin T.

combinatorial mix of hydrophobic residues we used the degenerate codon NTN (where N denotes a mixture of A, G, C, and T).

Screening of Green/White Phenotype—DNA libraries were isolated from *E. coli* strain XL1-Blue, transformed into BL21(DE3) (23), and plated onto nitrocellulose paper (Millipore NC-HATF 83 mm) on LB plates containing 50 μg/ml kanamycin. Following overnight growth at 37 °C, the nitrocellulose papers were transferred to LB plates containing 50 μg/ml kanamycin and 1 mM isopropyl 1-thio-β-D-galactopyranoside and incubated at 30 °C for 4 h to induce the expression of the Aβ-GFP fusion protein. Colonies were counted and the green/white phenotype was recorded (21).

Fluorescence Measurements—To enable measurement of the fluorescence of the Aβ42-GFP fusions *in vivo*, colonies were picked, and cultures were grown in LB liquid media containing 50 μg/ml kanamycin. When cultures reached an absorbance at 600 nm of 0.8, expression was induced by addition of isopropyl 1-thio-β-D-galactopyranoside to a concentration of 1 mM, and growth was continued for an additional 5 h 30 min at 30 °C. After induction, cultures were diluted in Tris-buffered saline to an $A_{600\,nm}$ of 0.15. Fluorescence was measured using a 50B spectrofluorimeter (PerkinElmer Life Sciences) with excitation at 490 nm and emission at 510 nm. Expression of Aβ42-GFP fusions was assessed by removing 200 μl of cell culture and analyzing the whole cell content by SDS-PAGE.

Correlation of Fluorescence with Biophysical Properties—The fluorescence of the Aβ42-GFP fusions *in vivo* was plotted against the sum of hydrophobicities or against the sum of the β-sheet propensities at positions 41 and 42. Hydrophobicities were based on the scale of Kyte and Doolittle (24), and β-sheet propensities were based on the scale of Minor and Kim (25). In a further analysis, fluorescence was plotted against the predicted aggregation rates using Equation 1 as developed by Chiti *et al.* (26).

$$\ln(\nu_{mut}/\nu_{wt}) = A\Delta\mathrm{Hydr} + B(\Delta\Delta G_{coil\text{-}\alpha} + \Delta\Delta G_{\beta\text{-}coil}) + C\Delta\mathrm{Charge} \quad \text{(Eq. 1)}$$

In Equation 1, ν_{mut} and ν_{wt} are the aggregation rates of the mutant and wild type sequences, respectively; ΔHydr is the difference in hydrophobicity, and $(\Delta\Delta G_{coil\text{-}\alpha} + \Delta\Delta G_{\beta\text{-}coil})$ is the difference in the propensity to convert from α-helix to β-sheet. (Note, the two $\Delta\Delta$s are added rather than subtracted. This is because subtraction is already accomplished by the definition of the terms. In the first term α is subtracted *from coil*, and in the second term *coil is subtracted* from β.) ΔCharge is the difference in charge, and *A*, *B*, and *C* are empirically determined constants.

Synthetic 42-Residue Peptides—Crude peptides were purchased from the Keck Institute, Yale University, and were then purified using reverse phase HPLC. Their identities were confirmed by mass spectrometry, and purity was assessed by analytical reverse phase HPLC. The purities of the peptides were greater than 92%. After purification, peptides were lyophilized and dissolved in hexafluoroisopropanol (HFIP). HFIP was removed by blowing argon over the sample. Samples were then dissolved in 100 μl of Me_2SO/500 μl of 4 M NaOH. Concentrations were determined using the extinction coefficient of tyrosine.

Aggregation of Mutant Peptides in Vitro—Peptides were dissolved at a concentration of 10 μM in 50 mM NaH_2PO_4, 100 mM NaCl, 0.02% NaN_3 (pH 7.3–7.4). Following incubation at 30 °C for 1 day, samples were centrifuged for 30 min at 60,000 × *g*. Half of the supernatant was removed and loaded onto reverse phase HPLC to quantify the monomer remaining in solution. To correlate peak size with concentration, a series of standards was run using a range of concentrations of the Asp^{41}–Gln^{42} mutant. The relationship between peak size and peptide concentration was confirmed using same concentrations of wild type Aβ42, the hydrophobic mutant Leu^{41}–Leu^{42}, and the green control mutant Ser^{19}–Pro^{34}.

Thioflavin T Assay—Because the kinetics of fibril formation can be affected by small quantities of seeds, preexisting seeds were removed by the following treatment (27). Lyophilized peptides were dissolved in trifluoroacetic acid at ~1 mg/ml and sonicated for 15 min. Trifluoroacetic acid was then removed by blowing argon over the sample. The dry sample was resuspended in 2 ml of HFIP. An aliquot corresponding to 0.5 mg of peptide was dried under argon and then dissolved in 300 μl of Me_2SO and mixed with 5 ml of 8 mM NaOH. The solution was then centrifuged at 40,000 × *g* for 30 min to remove any insoluble material. The supernatant was removed, and phosphate-buffered saline was added to a final concentration of 50 mM NaH_2PO_4, 100 mM NaCl, 0.02% NaN_3. The pH was adjusted to 7.4–7.5 with 200 mM formic acid. 500-μl aliquots were removed every 30 min and mixed with 2.4 ml of a ThT solution (7 μM thioflavin T, 50 mM glycine-NaOH (pH 8.5)). Fluorescence was measured with excitation at 450 nm, and emission at 490 nm.

RESULTS

A High Throughput Screen for Aβ Aggregation—Random mutagenesis of position 41 and 42 can produce 400 possible sequences (including wild type). Synthesis, purification, and characterization of the aggregation behavior of all 400 peptides would be laborious and extremely expensive, particularly because the Aβ42 peptide is notoriously difficult to synthesize and purify (28, 29). Therefore, we developed a high throughput screen using GFP as a reporter tag (21).

The folding of GFP and the formation of its active chromaphore occur relatively slowly (30). Therefore, peptides or proteins fused to the N terminus of GFP can have a dramatic impact on fluorescence. Sequences that aggregate rapidly prevent formation of a correctly folded fluorescent GFP. In contrast, sequences that are soluble or aggregate slowly allow GFP to fold into its native fluorescent structure. In a systematic study using 20 different test proteins, Waldo *et al.* (31) demonstrated that the fluorescence of *E. coli* cells expressing fusions to the N terminus of GFP correlated with the solubility of the test protein expressed alone.

We have shown previously that the fluorescence of Aβ42-GFP fusions can be used as an unbiased screen for the sequence determinants of Aβ amyloidogenesis (21). Colonies of *E. coli* expressing fusions to wild type Aβ42 or to mutants of Aβ42 that favor rapid aggregation are white. In contrast, colonies expressing fusions to soluble (or slowly aggregating) mutants of Aβ42 are green.

Here we used Aβ42-GFP fusions to monitor the phenotypic variation resulting from amino acid substitutions at positions 41 and 42 of Aβ42. This enabled a high throughput screen in which more than 2500 colonies could be screened per plate.

To ensure that the screen had sufficient dynamic range to distinguish a range of phenotypes, we first performed pilot experiments using fusions of GFP to the wild type sequences of either Aβ42 or Aβ40. (The latter construct can be considered as the null mutant in which residues 41 and 42 are not merely substituted but deleted entirely.) At 37 °C both fusions produced white colonies. Thus, at this temperature, the screen does not have sufficient dynamic range. In contrast, at 30 °C, colonies expressing fusions to Aβ42 were white, whereas those expressing fusions to Aβ40 were green. Thus, at this lower temperature, where expression and subsequent aggregation are slower, the difference in aggregation propensity between Aβ40 and Aβ42 is easily distinguished. Further pilot experiments showed that at 37 °C, virtually all mutations at

positions 41 and 42 produced white colonies, whereas at 30 °C libraries of mutants at 41 and 42 produced both green and white colonies. Because expression at 30 °C provided excellent dynamic range, all further experiments were performed at this temperature.

A Random Library of Mutants at Positions 41 and 42—A library of random mutants at positions 41 and 42 of Aβ42 was constructed using synthetic oligonucleotides to incorporate random bases (NNN) at the final 2 codons of a synthetic gene encoding Aβ42 (see "Materials and Methods"). The library of mutant genes of Aβ42 was then fused to the 5′ end of a gene encoding GFP. In this construct, the Aβ42 sequence is separated from GFP by a 12-residue linker encoding the sequence Gly-Ser-Ala-Gly-Ser-Ala-Ala-Gly-Ser-Gly-Glu-Phe (31).

The library of fusions was then transformed into XL1 Blue cells. Transformation yielded more than 5000 colonies. Because there are only 400 possible combinations of amino acids positions 41 and 42, this library is adequate to sample all (or nearly all) of the possible sequences (TABLE ONE). The library was then extracted from XL1 Blue cells and transformed into BL21(DE3) cells for protein expression and high throughput screening.

To assess the effect of amino acid substitutions at residues 41 and 42 on the aggregation of Aβ42, the library of Aβ42-GFP fusions was expressed at 30 °C, and the resulting colonies were characterized. In a typical experiment, 1520 colonies were analyzed. Of these, 140 colonies (9%) were white, similar to colonies expressing GFP fusions to wild type Aβ42. The remaining 1380 colonies (91%) showed some level of green fluorescence (TABLE TWO). 34 clones from this library were chosen arbitrarily for further characterization by sequence analysis and fluorescence measurements. The correlation between sequence and fluorescence is shown in Fig. 1.

The Effect of Hydrophobicity at Residues 41 and 42—As shown in Fig. 1A, GFP fusions to mutants with hydrophobic residues at positions 41 and 42 display low fluorescence, similar to fusions to wild type (Ile^{41}–Ala^{42}) Aβ42. In contrast, fusions to mutants with hydrophilic residues at positions 41 and 42 show high fluorescence, similar to fusions to Aβ40. These results suggest that hydrophobic side chains at positions 41 and 42 increase the propensity of Aβ42 to aggregate.

To confirm this correlation, we produced two additional libraries of mutants. In one library, residues 41 and 42 were mutated to a combinatorial mixture of hydrophobic residues. In the other library, these residues were mutated to a combinatorial mixture of hydrophilic residues. The hydrophobic library was constructed using the degenerate DNA codon NTN to encode a mixture of nonpolar residues including Ile, Leu, Val, Met, and Phe (32). The hydrophilic library was constructed using the NAN codon to encode a mixture of polar residues including Glu, Gln, Asp, Asn, Lys, His, and Tyr (where N denotes a mixture of the four DNA bases).

Both libraries were plated, and protein was expressed at 30 °C. For the 41/42-hydrophobic library, >5000 colonies were analyzed and *none were green*. Thus, randomly chosen hydrophobic residues at positions 41 and 42 support Aβ42 aggregation.

In contrast, for the 41/42-hydrophilic library, 4800 colonies were analyzed, and 87% of them (4160 colonies) were green. This result indicates that randomly chosen polar residues at positions 41 and 42 disrupt the aggregation of Aβ42.

From the hydrophilic library, only 640 colonies (13%) were white. We tested 64 of these colonies for expression of the Aβ42-GFP fusions, and we found that all but one of them did not express the fusion protein. (Because we used the NAN degenerate codon at positions 41 and 42, it was expected that the library would contain stop codons, which prevent protein expression.) The other white colony expressed normal levels of

TABLE ONE

Diversity of the libraries

The number of colonies for each library was counted after transformation into XL1-Blue cells. For each library, the number of observed colonies was much larger than number of possible amino acid combinations. Consequently, all (or nearly all) of the possible sequences are represented in the experimental libraries.

	Random at both 41, 42	Hydrophilic at both 41,42	Hydrophobic at both 41,42	Pro^{41}–$Random^{42}$	Gly^{41}–$Random^{42}$	$Random^{41}$–Pro^{42}	$Random^{41}$–Gly^{42}	$Random^{41}$–Ala^{42}	Ile^{41}–$Random^{42}$
No. colonies analyzed	>5000	>5000	>5000	>2000	>2000	>2000	>1500	>1000	>1200
No. possible amino acid combinations	400	49	25	20	20	20	20	20	20

TABLE TWO

Green /white phenotypic screening of colonies expressing Aβ(42)-GFP fusions

Colonies transformed into BL21(DE3) with plasmids encoding Aβ42-GFP fusions were screened for fluorescence after 4 h of induction at 30 °C. When positions 41 and 42 were both randomized, ~91% of colonies were green fluorescent, indicating that many mutations disrupted aggregation. When only hydrophobic residues were allowed at positions 41 and 42, all colonies were white, indicating that random hydrophobic residues support aggregation. In contrast, when only hydrophilic residues were allowed at positions 41 and 42, 87% of the colonies (4160 colonies) were green. The remaining 13% (640 colonies) were white because they did not express the fusion protein. Thus, random hydrophilic residues disrupt aggregation. The other libraries are described in the text.

	Random at both 41, 42	Hydrophobic at both 41,42	Hydrophilic at both 41,42	Pro^{41}–$Random^{42}$	Gly^{41}–$Random^{42}$	$Random^{41}$–Pro^{42}	$Random^{41}$–Gly^{42}	$Random^{41}$–Ala^{42}	Ile^{41}–$Random^{42}$
White colonies	140	5000	640	17	18	25	30	37	44
Green colonies	1380	0	4160	210	160	220	290	153	141

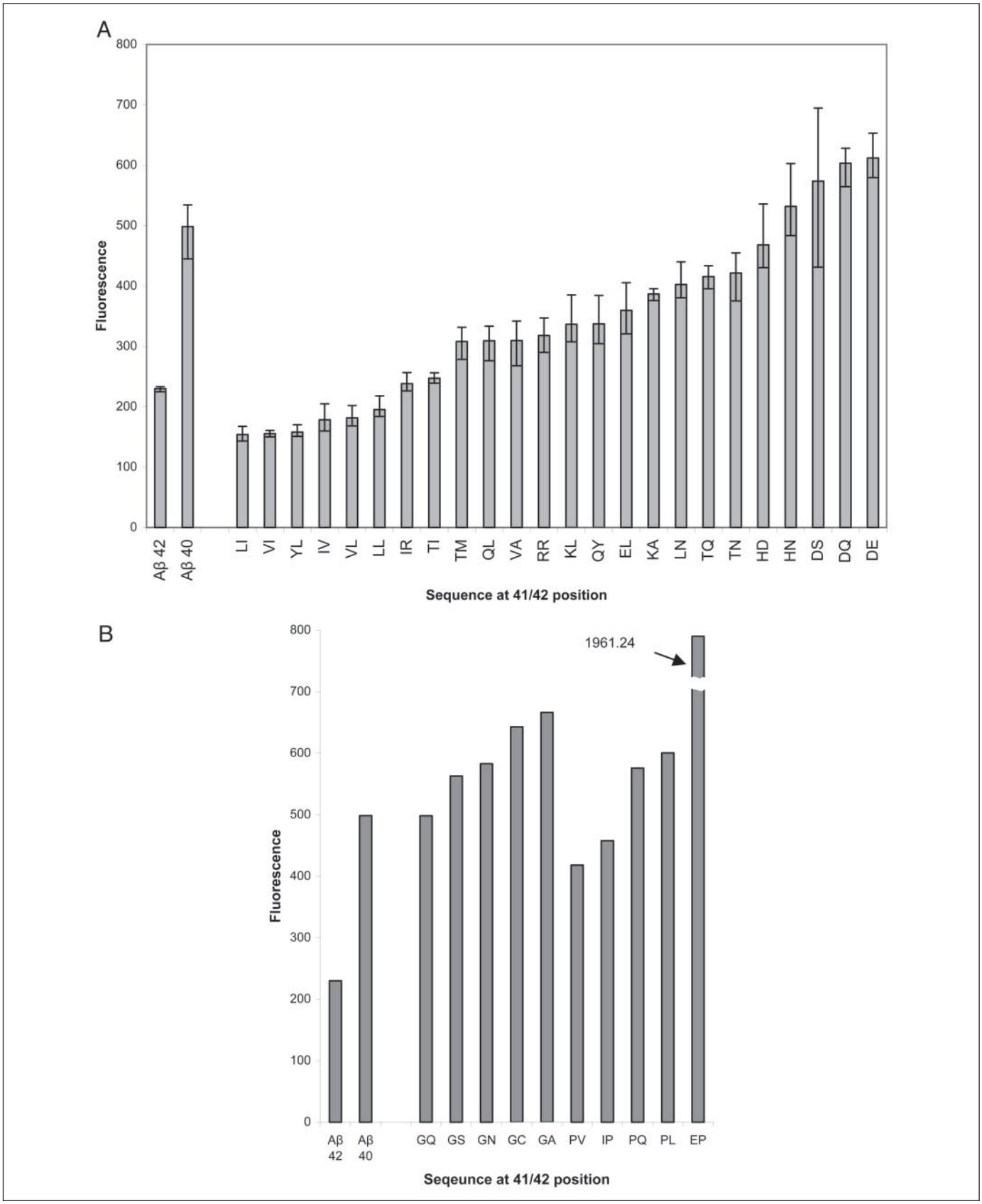

FIGURE 1. **Fluorescence of Aβ42-GFP fusions.** Cells were cultured in liquid media, and protein expression was induced at 30 °C. After induction, cells were diluted in buffer, and fluorescence was measured at 510 nm. *A,* mutants with hydrophobic residues at positions 41 and 42 show lower fluorescence, consistent with higher levels of aggregation preventing the formation of the GFP chromaphore. *B,* mutants with β-sheet breakers (Pro or Gly) show a high level of fluorescence. The Glu^{41}–Pro^{42} mutant shows an exceptional level of fluorescence, indicative of a very low propensity to aggregate, resulting from both charged and proline residues.

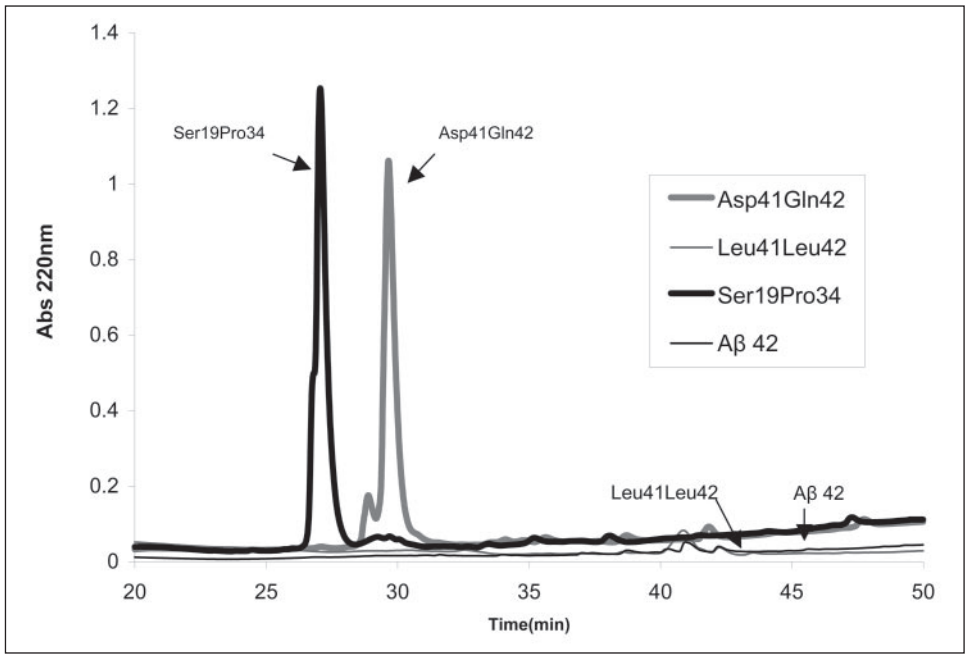

FIGURE 2. **Aggregation of synthetic 42-residue peptides.** Peptides (10 μM) were incubated at 30 °C for 1 day. Samples were then centrifuged to remove precipitated aggregates, and the supernatant was run on reverse phase HPLC. The size of the peak indicates the amount of monomer that remained in solution. The hydrophilic mutant, Asp^{41}–Gln^{42}, and the non-C-terminal green control, Ser^{19}–Pro^{34}, yielded peaks at the elution time expected for monomers. In contrast, wild type Aβ42 and the hydrophobic mutant, Leu^{41}–Leu^{42}, showed almost no soluble peptide.

the fusion protein. Sequence analysis of this clone revealed that it had arginine at both positions 41 and 42. Thus, although most randomly chosen polar residues at positions 41 and 42 interfere with Aβ aggregation, arginine is somehow special and allows aggregation (see below).

Effect of the β-Sheet Propensity of Residues 41 and 42—Amyloid structures are dominated by β-sheet secondary structure (33, 34), and previous studies suggested a strong correlation between the tendency of a sequence to form amyloid and the β-sheet propensities of its component amino acids (26, 35–38). Of the 34 sequences analyzed above, 10 sequences contained proline or glycine at positions 41 or 42. These residues have very low intrinsic propensities for β-sheet structure (39). All 10 of these mutants displayed high levels of fluorescence (Fig. 1B).

To study further the role of β-sheet propensity at residues 41 and 42, we constructed additional libraries in which β-sheet breakers were incorporated intentionally at position 41 or 42. Four libraries were constructed as follows: Pro^{41}–Random^{42}, Gly^{41}–Random^{42}, Random^{41}–Pro^{42}, and Random^{41}–Gly^{42}. For each library, 20 different sequences are possible. To ensure extensive sampling, 178–320 colonies were analyzed for each library (TABLE TWO). For all four libraries, ~90% of colonies were green. The non-green colonies were assayed for protein expression, and none expressed the fusions. (As described above, non-expressers arise from stop codons encoded by the NNN degenerate codon.) These results indicate that the ability of residues 41 and 42 to form β-sheet structure is important for Aβ aggregation.

Effect of the Site of Mutation—To assess which of the two positions affects aggregation more significantly, we constructed two additional libraries. In each of these libraries one of the two C-terminal residues was held constant, whereas the other was randomized. These libraries were called Ile^{41}–Random^{42} and Random^{41}–Ala^{42}. (The wild type sequence is Ile^{41}–Ala^{42}.) For each library, nearly 200 colonies were analyzed. The Random^{41}–Ala^{42} library yielded 37 white colonies (19% white), and 153 green colonies. The Ile^{41}–Random^{42} library yielded 44 white colonies (24% white), and 141 green colonies (TABLE TWO). Because the fraction of white colonies in the two libraries is not significantly different, we conclude that the wild type residues at positions 41 and 42 contribute to aggregation to a similar extent.

Aggregation of Mutant Peptides in Vitro—Waldo *et al.* (31) and Wurth *et al.* (21) demonstrated the validity of GFP fusions as a screen for the aggregation behavior of proteins and peptides. In particular, Wurth *et al.* (21) showed that the fluorescence *in vivo* of a mutant of Aβ42 fused to GFP correlated well with the solubility/aggregation behavior of the same mutant in the context of the 42-residue synthetic peptide. In the current study, however, because we have focused on residues at the C terminus of Aβ42, it is reasonable to question whether GFP fusions can produce meaningful readouts for perturbations so close to the fusion linker and GFP.

To address this question, we studied four sequences from our library as synthetic 42-residue peptides: Wild type Aβ42, Asp^{41}–Gln^{42}, Leu^{41}–Leu^{42}, and Ser^{19}–Pro^{34}. The mutants Asp^{41}–Gln^{42} and Leu^{41}–Leu^{42} were chosen as representatives of the hydrophilic green mutants and the hydrophobic white mutants, respectively. The Ser^{19}–Pro^{34} peptide is a control. This non-C-terminal mutant was isolated previously by Wurth *et al.* (21), who found the following: (i) as an Aβ42-GFP fusion, Ser^{19}–Pro^{34} was the most fluorescent mutant in their collection; and (ii) as a synthetic 42-residue peptide, it was far more soluble and less prone to aggregate than wild type Aβ42.

The extent of aggregation was compared for the four peptides. Peptide samples were prepared at 10 μM concentration in 50 mM NaH_2PO_4, 100 mM NaCl, and 0.02% NaN_3 (pH 7.3–7.4) and were incubated for 1 day at 30 °C. After incubation, samples were centrifuged at 60,000 × *g* for 30 min to precipitate the aggregated material. The amount of unaggregated peptide in the supernatant was then quantified by reverse phase HPLC (see "Materials and Methods").

As shown in Fig. 2, the hydrophilic mutant, Asp^{41}–Gln^{42}, and the non-C-terminal green control, Ser^{19}–Pro^{34}, yielded substantial peaks at the elution time expected for monomers of each peptide. Integration of the peak areas indicated that for these two peptides almost the entire sample remained in solution and unaggregated. In contrast, wild type Aβ42 and the hydrophobic mutant, Leu^{41}–Leu^{42}, showed almost no peaks corresponding to monomeric soluble peptide (Fig. 2). These results confirm that the fluorescent phenotype of the GFP fusions *in vivo* is a reasonable predictor of the aggregation behavior of the isolated 42-residue peptides *in vitro*, even for mutants at the C terminus of the peptide.

Rate of Amyloidogenesis—Although the experiments described above demonstrate that aggregation of Aβ42 peptides *in vitro* mimics the phe-

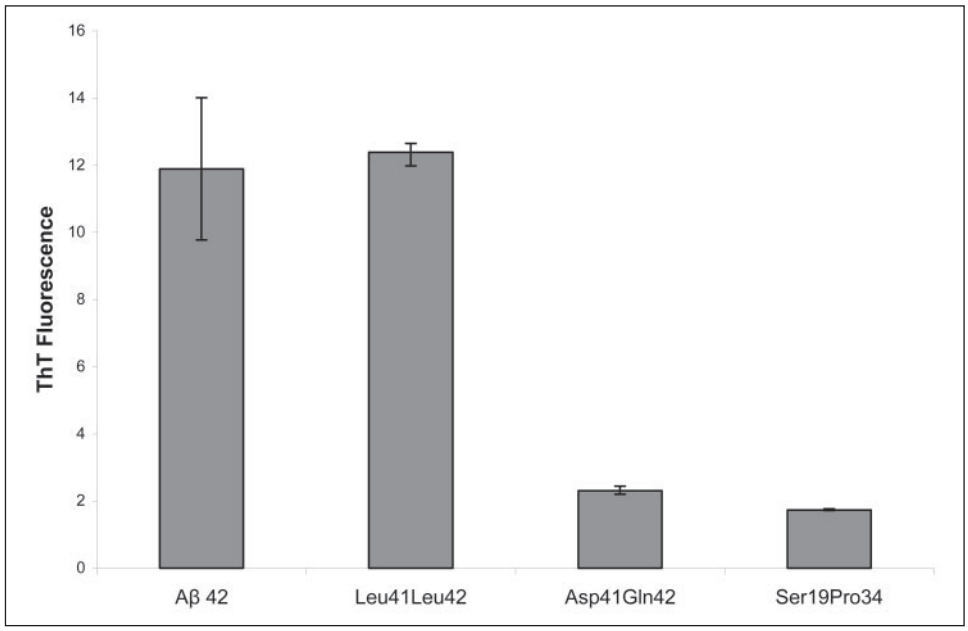

FIGURE 3. **Thioflavin T fluorescence of wild type and mutant 42-residue peptides.** Synthetic peptides (20 μM) were incubated in quiescent condition (not agitated). After 1 day of incubation samples were mixed with thioflavin T, and fluorescence was measured at 490 nm. Wild type Aβ42 and Leu^{41}–Leu^{42} showed thioflavin T fluorescence indicative of amyloid structure.

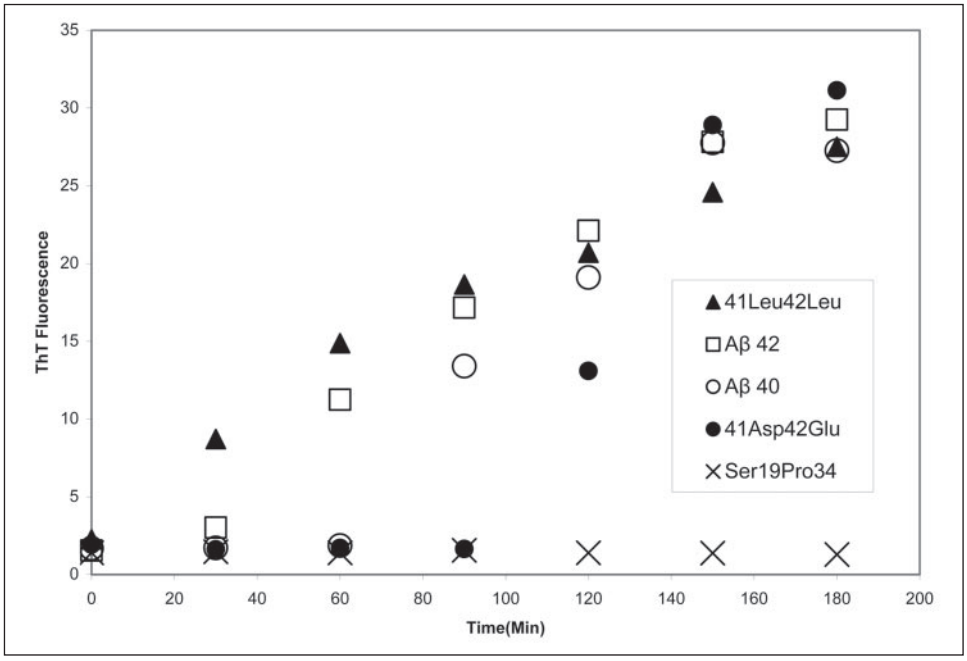

FIGURE 4. **Kinetics of amyloid formation followed by thioflavin T fluorescence.** Purified peptides (20 μM) were agitated at 30 °C. Aliquots were removed every 30 min and mixed with thioflavin T, and fluorescence was measured at 490 nm. Wild type Aβ42 and the hydrophobic mutant, Leu^{41}–Leu^{42} form amyloid from the outset, whereas the hydrophilic mutant, Asp^{41}–Gln^{42}, like Aβ40, aggregates only after a 90–120-min lag. The control peptide, Ser^{19}–Pro^{34}, showed little or no ThT binding even after 5 h.

notypes of the GFP fusions *in vivo*, it is important (i) to establish that the aggregated states of these peptides are indeed amyloid, and (ii) to compare the rates of amyloid formation for the different peptides.

The presence of amyloid fibrils can be detected by monitoring the binding of dyes, such as Congo Red or thioflavin T (ThT). These dyes are known to bind specifically to the cross-β-structure in a variety of amyloids including those formed by insulin, transthyretin, polyglycine, apoAII, and Aβ; yet they do not bind to monomeric or small oligomeric complexes of these peptides and proteins (40–43). Thioflavin T binding is easily monitored by the shift in fluorescence from 430 to 493 nm.

To assess binding to thioflavin T, each of the synthetic peptides (20 μM) was incubated for 1 day without agitation, and samples were then mixed with ThT, and fluorescence was measured at 490 nm. As shown in Fig. 3, the hydrophobic mutant Leu^{41}–Leu^{42} bound ThT to an extent similar to wild type Aβ42. In contrast, the hydrophilic mutant, Asp^{41}–Gln^{42}, and the soluble control, Ser^{19}–Pro^{34}, showed poor ThT binding. Thus, wild type Aβ42 and Leu^{41}–Leu^{42} readily form amyloid fibrils, whereas Asp^{41}–Gln^{42} and Ser^{19}–Pro^{34} do so only poorly.

Next, we followed the rate of fibril formation under conditions that favor aggregation: agitation, rather than quiescent incubation. Time course measurements of ThT binding showed that wild type Aβ42 and the hydrophobic mutant, Leu^{41}–Leu^{42}, began to form amyloid within a few minutes. In contrast, the hydrophilic mutant, Asp^{41}–Gln^{42}, like Aβ40, began to form amyloid only during the 2nd h of incubation (Fig. 4). (The control peptide, Ser^{19}–Pro^{34}, showed little or no ThT binding even after 5 h.)

Thus, the behavior of the synthetic 42-residue peptides *in vitro* validates the results observed for the GFP fusions *in vivo*. Hydrophobic

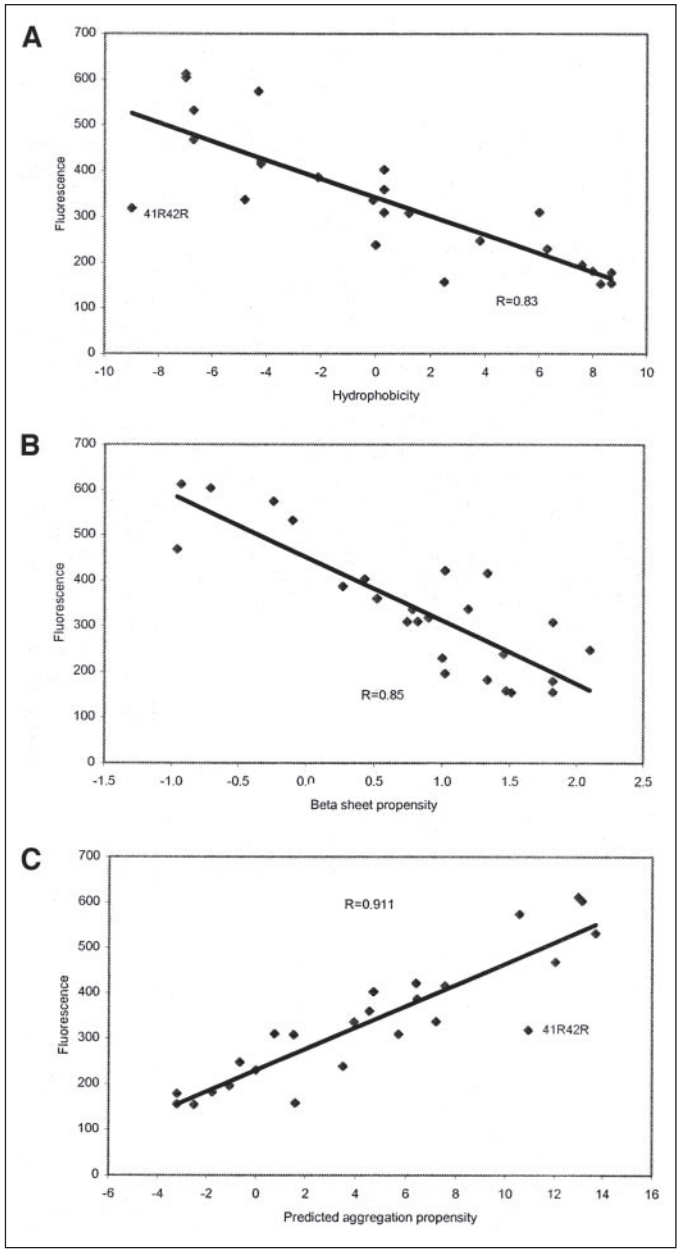

FIGURE 5. **Correlation of the fluorescence *in vivo* of Aβ42-GFP fusions (except Gly and Pro) with biophysical properties of the residues at positions 41 and 42.** *A,* correlation with the sum of hydrophobicities at positions 41 and 42. *B,* correlation with the sum of β-sheet propensities at positions 41 and 42. *C,* correlation with the aggregation propensity predicted by the equation of Chiti *et al. (26).*

mutants at positions 41 and 42 mimic wild type behavior, whereas hydrophilic mutants retard amyloidogenesis.

DISCUSSION

Aβ40 is the primary product of the proteolytic cleavage of the Alzheimer amyloid precursor protein. Yet senile plaques are composed primarily of Aβ42 (5, 14–16). Moreover, studies *in vitro* have demonstrated that Aβ42 has a far higher propensity to aggregate than Aβ40 (5, 15, 17, 44). In the current studies we sought to elucidate which features of the C-terminal residues are responsible for the greater amyloidogenicity of the longer peptide.

The Role of Hydrophobicity—Fig. 1A shows that mutants with hydrophobic residues at both positions 41 and 42 have high propensities to aggregate. Some of these hydrophobic mutants, such as Val^{41}–Ile^{42} and Leu^{41}–Ile^{42}, are even more prone to aggregate than the wild type (Ile^{41}–Ala^{42}) sequence. The greater hydrophobicity of Ile relative to Ala at position 42 presumably accounts for this enhanced aggregation. Likewise, for the Val^{41}–Ala^{42} mutant, the lower hydrophobicity of Val relative to Ile at position 41 may account for the slightly decreased aggregation of this mutant relative to wild type.

Mutants that have one hydrophilic and one hydrophobic residue at the C terminus (*e.g.* Lys^{41}–Leu^{42}, Gln^{41}–Leu^{42}, and Glu^{41}–Leu^{42}) are slightly less prone to aggregate than the double hydrophobic sequences, whereas mutants that have two polar residues at the C terminus (*e.g.* Asp^{41}–Glu^{42}, Asp^{41}–Gln^{42}, and His^{41}–Asn^{42}) are the least aggregating sequences in the library.

These results indicate that the hydrophobicity of residues at positions 41 and 42 is a major contributor to the enhanced amyloidogenicity of Aβ42 relative to Aβ40, and would be consistent with residues Ile^{41} and Ala^{42} occurring in the buried core of Aβ fibrils.[3]

β-Sheet Propensity—Although the hydrophobicity of residues at positions 41 and 42 clearly plays a dominant role in promoting aggregation, hydrophobicity alone does not explain the properties of all mutants in our collections. For example, four libraries of mutants were constructed in which only position 41 or 42 was randomized, whereas the other position was constrained to be either proline or glycine (TABLE TWO). For all four of these libraries, the vast majority (~90%) of mutants yielded green fluorescent (i.e. less aggregating) GFP fusions (TABLE TWO), and the few white colonies that were observed did not express the fusions protein (presumably because of the presence of stop codons). These results suggest that all (or nearly all) sequences with proline or glycine at positions 41 or 42 have significantly reduced propensities to aggregate. This behavior was observed despite the fact that proline is a hydrophobic residue, and glycine is neither polar nor nonpolar. Proline and glycine are known to be β-sheet breakers (39), and their occurrence at positions 41 or 42 presumably interferes with the formation of cross-β amyloid structure.

The wild type residues, Ile^{41} and Ala^{42}, are both hydrophobic and good β-sheet formers. In Aβ42, these residues, together with residues Val^{39} and Val^{40} (which are also hydrophobic β-sheet formers), may form a β-strand at an early stage of aggregation, and thereby lower the kinetic barrier for the aggregation of Aβ42 relative to Aβ40.

Tertiary Interactions—Although hydrophobicity and β-sheet propensity account for most of the observations reported here, two of the mutants shown in Fig. 1, Ile^{41}–Arg^{42} and Arg^{41}–Arg^{42}, yield less fluorescent (*i.e.* more aggregating) GFP fusions than would be expected for sequences containing arginine, which is an extremely polar residue (24). The behavior of these mutants might be explained by the structure of Aβ fibrils modeled by Tycko and co-workers (45) from solid-state NMR experiments. In that structure, Aβ forms a hairpin, which would place the C-terminal amino acids in close proximity to glutamic acid at position 11. Interaction between the positively charged arginine and the negatively charged glutamic acid might compensate for burying the charged arginine side chain. This interpretation would suggest that a sequence with lysines at positions 41 and 42 might also yield less fluorescent (more aggregating) GFP fusions than expected for sequences containing charged residues at both positions. To test this expectation, we used site-directed mutagenesis to construct the Lys^{41}–Lys^{42} mutant. Fluorescence of the GFP fusion of this mutant, although not as diminished as the Arg^{41}–Arg^{42} mutant, was indeed lower than expected for a doubly charged dipeptide. (Its fluorescence was approximately midway between those of Asp^{41}–Glu^{42} and Arg^{41}–Arg^{42}; data not shown).

[3] R. Tykco, personal communication.

Correlation of Aggregation with Hydrophobicity, β-Sheet Propensity, and Previous Models—To verify the importance of hydrophobicity and β-sheet propensity at positions 41 and 42, we plotted the fluorescence of Aβ42-GFP fusions against each of these parameters. As shown in Fig. 5A, the fluorescence of our mutants and the sum of hydrophobicities at positions 41 and 42 exhibit a strong negative correlation ($R = 0.83$). Thus, sequences having hydrophobic residues at these positions have higher propensities to aggregate.

Fluorescence of the Aβ42-GFP fusions was also plotted against the sum of the β-sheet propensities (25) at positions 41 and 42. As shown in Fig. 5*B*, this plot also showed a strong negative correlation ($R = 0.85$), thereby supporting the hypothesis that β-sheet propensity at these positions favors aggregation.

Chiti *et al.* (26) showed that rates of aggregation could be predicted from intrinsic factors including hydrophobicity, β-sheet propensity, α-helical propensity, and charge. They derived an equation to quantify this prediction (see Equation 1 under "Materials and Methods").

For our collections of mutants in residues 41 and 42, we plotted the observed fluorescence of the Aβ42-GFP fusions against the aggregation rate predicted by Equation 1 (Fig. 5*C*). The observed fluorescence and the aggregation predicted by this equation, which includes both hydrophobicity and β-sheet propensity, showed a better correlation ($R = 0.911$) than either hydrophobicity alone (Fig. 5*A*) or β-sheet propensity alone (Fig. 5*B*).

The Arg^{41}–Arg^{42} mutant deviated significantly from the trend (Fig. 5*C*). As described above, the behavior of this sequence may be explained by attractive electrostatic interactions between the arginine side chain and glutamic acid at position 11.

Sequence Determinants of the Enhanced Amyloidogenicity of Aβ42 Relative to Aβ40—Although both Aβ40 and Aβ42 assemble into amyloid fibrils, the longer peptide aggregates more readily both *in vitro* and *in vivo* (17, 29, 44). Although it might have been hypothesized that length *per se* is sufficient to increase the amyloidogenicity of Aβ, this hypothesis is disproved by our results. In particular, fluorescence studies of Aβ42-GFP fusions containing polar amino acids at positions 41 and 42 demonstrate that these sequences are in fact less prone to aggregate than the null mutant, Aβ40, which has both residues deleted entirely (Fig. 1).

Our experiments using several libraries of Aβ42-GFP fusions as well as synthetic 42 residue peptides demonstrate that rather than length, it is the hydrophobicity and β-sheet propensity of residues 41 and 42 that cause the enhanced amyloidogenicity of Aβ42 relative to Aβ40.

REFERENCES

1. Koo, E. H., Lansbury, P. T., Jr., and Kelly, J. W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 9989–9990
2. Carrell, R. W., and Lomas, D. A. (1997) *Lancet* **350,** 134–138
3. Thomas, P. J., Qu, B., and Pedersen, P. L. (1995) *Trends Biochem. Sci.* **20,** 456–459
4. Slipe, J. D. (1992) *Annu. Rev. Biochem.* **61,** 947–975
5. Selkoe, D. (2001) *Physiol. Rev.* **81,** 741–766
6. Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T. D., Hardy, J., Hutton, M., Kukull, W., Larson, E., Levy-Lahad, E., Viitanen, M., Peskind, E., Poorkaj, P., Schellenberg, G., Tanzi, R., Wasco, W., Lannfelt, L., Selkoe, D., and Younkin, S. (1996) *Nat. Med.* **2,** 864–870
7. Weggen, S., Eriksen, J. L., Das, P., Sagi, S. A., Wang, R., Pietrzik, C. U., Findlay, K. A., Smith, T. E., Murphy, M. P., Bulter, T., Kang, D. E., Marquez-Sterling, N., Golde, T. E., and Koo, E. H. (2001) *Nature* **414,** 212–216
8. Geula, C., Wu, C., Saroff, D., Lorenzo, A., Yuan, M., and Yankner, B. (1998) *Nat. Med.* **4,** 827–831
9. Schenk, D., Barbour, R., Dunn, W., Gordon, G., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Liao, Z., Lieberburg, I., Motter, R., Mutter, L., Soriano, F., Shopp, G., Vasquez, N., Vandevert, C., Walker, S., Wogulis, M., Yednock, T., Games, D., and Seubert, P. (1999) *Nature* **400,** 173–177
10. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002) *Nature* **416,** 507–511
11. Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J., and Selkoe, D. J. (2002) *Nature* **416,** 535–539
12. Kuo, Y., Emmerling, M. R., Vigo-Pelfrey, C., Kasunic, T. C., Kirkpatrick, J. B., Murdoch, G. H., Ball, M. J., and Roher, A. E. (1996) *J. Biol. Chem.* **271,** 4077–4081
13. Franz, G., Beer, R., Kampfl, A., Engelhardt, K., Schmutzhard, E., Ulmer, H., and Deisenhammer, F. (2003) *Neurology* **60,** 1457–1461
14. Roher, A. E., Lowenson, J. D., Clarke, S., Woods, A. S., Cotter, R. J., Gowing, E., and Ball, M. J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 10836–10840
15. Selkoe, D. J. (1991) *Neuron* **6,** 487–498
16. Gravina, S. A., Ho, L., Eckman, C. B., Long, K. E., Otvos, L., Jr., Younkin, L. H., Suzuki, N., and Younkin, S. G. (1995) *J. Biol. Chem.* **270,** 7013–7016
17. Jarrett, J. T., and Lansbury, P. T., Jr. (1993) *Cell* **73,** 1055–1058
18. Harper, J. D., and Lansbury, P. T., Jr. (1997) *Annu. Rev. Biochem.* **66,** 385–407
19. Bitan, G., Kirkitadze, M., Lomakin, A., Vollers, S., Benedek, G., and Teplow, D. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 330–335
20. Bitan, G., Vollers, S., and Teplow, D. (2003) *J. Biol. Chem.* **278,** 34882–34889
21. Wurth, C., Guimard, N. K., and Hecht, M. H. (2002) *J. Mol. Biol.* **319,** 1279–1290
22. Deleted in proof
23. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89
24. Kyte, K., and Doolittle, R. (1982) *J. Mol. Biol.* **157,** 105–132
25. Minor, D., Jr., and Kim, P. (1994) *Nature* **367,** 660–663
26. Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003) *Nature* **424,** 805–808
27. Jao, S., Ma, K., Talafous, J., Orlando, R., and Zagorski, M. G. (1997) *Amyloid: Int. J. Exp. Clin. Investig.* **4,** 240–252
28. Murakami, K., Irie, K., Morimoto, A., Ohigashi, H., Shindo, M., Nagao, M., Shimizu, T., and Shirasawa, T. (2002) *Biochem. Biophys. Res. Commun.* **294,** 5–10
29. Burdick, D., Soreghan, B., Kwon, M., Kosmoski, J., Knauer, M., Henschen, A., Yates, J., Cotman, C., and Glabe, C. (1992) *J. Biol. Chem.* **267,** 546–554
30. Cubitt, A., Heim, R., Adams, S., Boyd, A., Gross, L., and Tsien, R. (1995) *Trends Biochem. Sci.* **20,** 448–455
31. Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) *Nat. Biotechnol.* **17,** 691–695
32. Hecht, M., Das, A., Go, A., Bradley, L., and Wei, Y. (2004) *Protein Sci.* **13,** 1711–1723
33. Serpell, L. C. (2000) *Biochim. Biophys. Acta* **1502,** 16–30
34. Makin, O. S., Atkins, E., Sikorski, P., Johansson, J., and Serpell, L. C. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 315–320
35. Johansson, J., Weaver, T. E., and Tjernberg, L. O. (2004) *Cell. Mol. Life Sci.* **61,** 326–335
36. Päiviö, A., Nordling, E., Kallberg, Y., Thyberg, J., and Johansson, J. (2004) *Protein Sci.* **13,** 1251–1259
37. Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J., and Johansson, J. (2001) *J. Biol. Chem.* **276,** 12945–12950
38. Pawar, A. P., Dubay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M., and Dobson, C. M. (2005) *J. Mol. Biol.* **350,** 379–392
39. Creighton, T. (1992) *Proteins,* 2nd Ed., pp. 171–200, W. H. Freeman & Co., New York
40. Bonifacio, M., Sakaki, Y., and Saraiva, M. (1996) *Biochim. Biophys. Acta* **1316,** 35–42
41. LeVine, H., III (1993) *Protein Sci.* **2,** 404–410
42. LeVine, H., III (1995) *Amyloid: Int. J. Exp. Clin. Investig.* **2,** 1–6
43. Naiki, H., Higuchi, K., Shimada, T., Takeda, T., and Nakakuki, K. (1993) *Lab. Investig.* **68,** 332–337
44. Jarrett, J., Berger, E., and Lansbury, P. T., Jr. (1993) *Biochemistry* **32,** 4693–4697
45. Petkova, A., Ishii, Y., Balbach, J., Antzutkin, O., Leapman, R., Delaglio, F., and Tycko, R. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 16742–16747