# Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Abeta42 peptide

Woojin Kim, and Michael H. Hecht

*PNAS* 2006;103;15824-15829; originally published online Oct 12, 2006;
doi:10.1073/pnas.0605629103

# Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Aβ42 peptide

**Woojin Kim and Michael H. Hecht***

Department of Chemistry, Princeton University, Princeton, NJ 08544

Edited by Robert T. Sauer, Massachusetts Institute of Technology, Cambridge, MA, and approved August 28, 2006 (received for review July 7, 2006)

**One hundred years ago, Alois Alzheimer observed a relationship between cognitive impairment and the presence of plaque in the brains of patients suffering from the disease that bears his name. The plaque was subsequently shown to be composed primarily of a 42-residue peptide called amyloid β (Aβ) 42. Despite the importance of Aβ42 aggregation in the molecular etiology of Alzheimer's disease, the amino acid sequence determinants of this process have yet to be elucidated. Although stretches of hydrophobic residues in the C-terminal half of Aβ42 have been implicated, the mechanism by which these residues promote aggregation remains unclear. In particular, it is not known whether the side chains of these hydrophobic residues mediate specific interactions that direct self-assembly or, alternatively, whether hydrophobicity *per se* at these positions is sufficient to promote aggregation. To distinguish between these two possibilities, we substituted 12 hydrophobic residues in the C-terminal half of Aβ42 with random nonpolar residues. The mutant sequences were screened by using a fusion of Aβ42 to GFP. Because aggregation of Aβ42 prevents folding of the GFP reporter, this screen readily distinguishes aggregating from nonaggregating variants of Aβ42. Application of the screen demonstrated that, despite the presence of 8–12 mutations, all of the sequences aggregated. To confirm these results, several of the mutant sequences were prepared as synthetic peptides and shown to form amyloid fibrils similar to those of WT Aβ42. These findings indicate that hydrophobic stretches in the sequence of Aβ42, rather than specific side chains, are sufficient to promote aggregation.**

amyloid fibrils | GFP fusion | NTN codon | protein misfolding | binary code

Protein misfolding and aggregation are associated with a number of human diseases, including Alzheimer's disease, prion encephalopathies, Huntington's disease, and type II diabetes (1–4). In each of these diseases, a particular protein or peptide aggregates into insoluble amyloid fibrils. Although the sequences of the proteins differ dramatically from one disease to another, the structures of their aggregated forms are similar: All of them form fibrillar structures dominated by β-sheet secondary structure (5).

Although globular proteins not associated with misfolding diseases can be driven to form amyloid-like structures under appropriate experimental conditions (6, 7), it is clear that sequences associated with misfolding diseases have unusually high propensities to aggregate. Because the propensity and rate of aggregation can tip the balance between normal protein turnover and aggregation-induced pathology, there are strong motivations to understand the amino acid sequence determinants that cause aggregation.

In the case of Alzheimer's disease (AD), the aggregating sequence is the Alzheimer's peptide, amyloid β (Aβ). According to the "amyloid cascade" hypothesis, aggregated Aβ initiates a complex, multistep cascade that ultimately leads to neurodegenerative disease (8). The structure and oligomeric state of the toxic aggregate has not been fully elucidated and is a subject of intense investigation. Nonetheless, it is clear that aggregation of Aβ into some form of multimeric species (ranging from small oligomers to large fibrils) produces toxic species that lead to AD (9–12).

Our studies and those of other researchers suggest that hydrophobic stretches in the central and C-terminal parts of Aβ42 are responsible for aggregation and fibrilogenesis (13–15). In contrast, the N-terminal region (residues 1–16) of Aβ42, which is mostly polar, does not appear to promote aggregation (13, 15, 16).

Although it is clear that the central and C-terminal hydrophobic stretches of Aβ42 promote aggregation, the roles of the particular side chains in these stretches remain obscure. We consider two possibilities: (*i*) The side chains of these hydrophobic residues mediate specific interactions that direct the self-assembly of Aβ42. (*ii*) Hydrophobicity *per se* at these positions promotes aggregation, and the exact identities of the nonpolar side chains are unimportant.

Support for the first possibility (that particular nonpolar side chains promote aggregation) can be drawn from crystal structures of model amyloidogenic peptides, determined recently by Eisenberg and coworkers (17). These structures reveal highly ordered "steric zippers" comprising well packed structures with specific side-chain interactions. Support for the second possibility (that generic hydrophobicity at key positions in the sequence is sufficient to promote aggregation) comes from our work on protein design, which showed that a "binary code" specifying the patterning of polar and hydrophobic residues, but not the exact identities of these residues, is sufficient to design libraries of proteins *de novo*. By carefully controlling the binary patterning (but not the identities) of hydrophobic residues, we were able to design *de novo* proteins that favored either well folded globular structures or amyloid-like fibrils (18–21). If similar forces control the aggregation of naturally occurring amyloid proteins, then the generic hydrophobicity of the central and C-terminal stretches, rather than specific side chains in these sequences, may suffice to promote aggregation of Aβ42.

To distinguish between these two possibilities, we constructed and characterized a library of mutants in which 12 hydrophobic residues in the central and C-terminal stretches of Aβ42 were replaced by a stochastic mixture of the nonpolar residues leucine, isoleucine, valine, phenylalanine, and methionine. The phenotypes of the mutants were screened by using an artificial genetic system based on fusions of Aβ42 to GFP (14, 22, 23). This screen readily distinguishes mutants that prevent aggregation from those that aggregate like WT Aβ42. *Escherichia coli* cells expressing fusions of WT Aβ42 to GFP do not fluoresce, because rapid aggregation of Aβ42 causes the entire fusion to misfold and aggregate before GFP can form its native fluorescent structure. The appearance of green versus white colonies facilitates an unbiased assessment of the sequence determinants of Aβ42 aggregation. [The Aβ42–GFP fusion has also been used as a high-throughput screen for compounds that inhibit aggregation (24).]

Author contributions: W.K. and M.H.H. designed research; W.K. performed research; W.K. and M.H.H. analyzed data; and W.K. and M.H.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Abbreviations: Aβ, amyloid β; GM6, green mutant 6; ThT, thioflavin T.

*To whom correspondence should be addressed. E-mail: hecht@princeton.edu.





www.pnas.org/cgi/doi/10.1073/pnas.0605629103

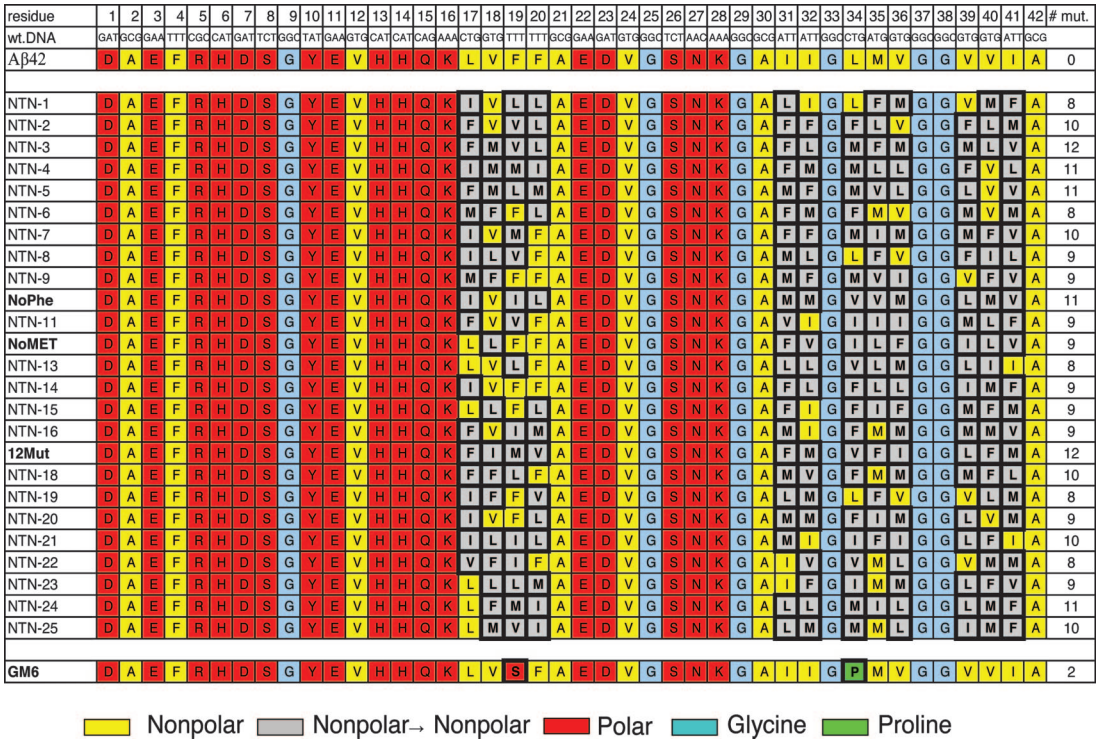


**Fig. 1.** Amino acid sequences of variants of Aβ42 containing randomized nonpolar side chains. WT Aβ42 is shown at the top, and the soluble control, GM6, is shown at the bottom. Polar and nonpolar residues are colored red and yellow, respectively. Mutated nonpolar residues are colored gray. Glycine and proline are colored blue and green, respectively. The number of amino acid substitutions is shown to the right of each sequence.

In this report, we show that none of the randomized hydrophobic mutants produce green fluorescent fusions, thereby demonstrating that, despite the presence of 8–12 nonpolar→nonpolar substitutions, all of these Aβ42 variants aggregate. Several of the mutant sequences were also studied in the context of the 42-residue synthetic peptide. Biophysical studies of these peptides showed that they aggregate into fibrils, thus confirming the results observed with the Aβ42–GFP fusions and demonstrating that hydrophobic stretches in the C-terminal half of Aβ42, rather than specific side-chain packing, are sufficient to promote aggregation.

## Results

**Screening a Library of Hydrophobic Mutants in Aβ42.** To investigate whether specific hydrophobic side chains in the C-terminal half of Aβ42 play an essential role in mediating aggregation, we constructed a library of mutant Aβ42 genes in which random hydrophobic amino acids were incorporated at positions 17–20, 31–32, 34–36, and 39–41. The mutant genes were constructed by using the degenerate NTN codon, which encodes Met, Val, Phe, Leu, and Ile. The library of semirandomized genes was fused to the N terminus of the GFP gene. In our construct, Aβ42 is separated from the N terminus of GFP by a linker encoding the sequence Gly-Ser-Ala-Gly-Ser-Ala-Ala-Gly-Ser-Gly-Glu-Phe (23). This sequence was shown previously to be effective in coupling the aggregation state of N-terminal fusions with the fluorescence of GFP (14, 22, 23). (Longer or more disordered linkers are not suitable, because they would uncouple the properties of the N-terminal peptide from those of GFP.)

The library of mutated Aβ42–GFP fusions was expressed in *E. coli* for high-throughput fluorescence screening. The fluorescence of colonies expressing these fusions was compared with two standards: fusions to WT Aβ42 and fusions to the soluble GM6 (green mutant 6) variant of Aβ42, which bears two mutations, Phe-19-Ser and Leu-34-Pro. Previous studies demonstrated that the WT Aβ42–GFP fusion aggregates and does not fluoresce, whereas the soluble control, GM6 Aβ42–GFP, produces green fluorescent colonies (14). From the library of randomized hydrophobic mutants, ≈3,000 colonies were screened. Virtually all of them were nonfluorescent. (Those few colonies that were green were shown by DNA analysis to contain aberrant sequences that had deleted all or part of the Aβ42 gene.) These findings indicate that, like WT Aβ42, the hydrophobic variants misfold and/or aggregate, thereby preventing the folding and fluorescence of GFP.

Twenty-five of the nonfluorescent clones were chosen arbitrarily for sequence analysis and further characterization. The sequences of these Aβ42 variants are shown in Fig. 1. They contain numerous mutations, with 8–12 amino acid substitutions in the 42-residue Aβ42 sequence. Val-24 was not mutated, because it falls in a region used for DNA annealing and gene assembly. Also, alanine was excluded from the mutations, because it cannot be encoded by the NTN codon. With the exception of these two cases, none of the hydrophobic residues in residues 17–42 was conserved. Although the randomized hydrophobic variants have 8–12 mutations in the 26 residues composing the C-terminal half of Aβ42 (Fig. 1), all of the variant Aβ42–GFP fusions produce white colonies.

**Randomized Hydrophobic Mutants of Aβ42 Cause GFP Fusions to Misfold into Insoluble Aggregates.** To assess the quantity of misfolded Aβ42–GFP fusions, we measured the fluorescence of liquid cultures expressing these sequences. Fig. 2 compares the fluorescence of cultures expressing the mutant Aβ42–GFP fusions with those expressing either WT or GM6 GFP fusions. All of the nonpolar→nonpolar mutants showed low levels of fluorescence. Thus, the mutant forms of Aβ42, like WT Aβ42, cause the Aβ42–GFP fusion to misfold. These results indicate that hydrophobic amino acids at the mutated positions support aggregation of Aβ42, irrespective of the identity of the nonpolar side chain.

To ensure that the nonfluorescent phenotypes of mutant Aβ42–GFP fusions were due to misfolding rather than poor expression, we tested the expression levels of the fusions. SDS/PAGE analysis of

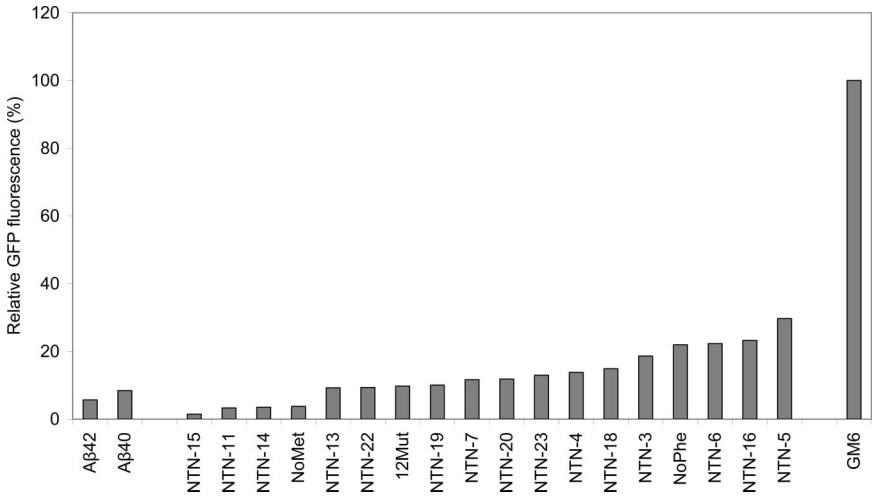


**Fig. 2.** *In vivo* fluorescence of GFP fusions of WT Aβ42, WT Aβ40, and various mutants of Aβ42. Low fluorescence indicates that misfolding and aggregation of the Aβ42 sequence causes the entire fusion to misfold. Fusions to WT Aβ40 and Aβ42 are shown on the left side; fusion to the soluble control, GM6 (Phe-19-Ser and Leu-34-Pro), is shown on the right. Fusions to randomized hydrophobic variants of Aβ42 display low fluorescence, indicating that random mutations of hydrophobic residues to other hydrophobic residues support aggregation.

whole-cell lysates confirmed that mutant versions of the Aβ42–GFP fusions expressed at the same high levels as WT Aβ42–GFP fusions. Several examples are shown in Fig. 3*A*. Despite these high levels of expression, GFP fusions to the randomized hydrophobic variants of Aβ42 do not fluoresce.

These findings suggest that the variants of Aβ42, like WT Aβ42, cause the entire fusion protein to misfold into insoluble aggregates. To confirm this interpretation, we assayed the amount of protein in the soluble fractions of the cell lysates. As shown in Fig. 3*B*, GFP fusions to randomized hydrophobic variants or to WT Aβ42 are insoluble and not detected in the soluble fraction of lysates of the overexpressing cells. Only the soluble control GM6 (rightmost lane in Fig. 6*B*) yielded a soluble GFP fusion protein.

**Hydrophobic Variants of Synthetic Aβ42 Peptide Form Insoluble Aggregates.** The experiments described above were performed on Aβ42–GFP fusions expressed in *E. coli*. It seems reasonable to expect that the effects of mutations in Aβ42 on the solubility of the fusion protein would mirror the effects of these same mutations on the biologically relevant 42-residue peptide in isolation. However, because the Aβ42 sequence (4 kDa) is only a small portion of the entire Aβ42–GFP fusion protein (28 kDa), this expectation must be verified explicitly. To confirm that our findings with Aβ42–GFP fusions apply to isolated Aβ42 peptides, we studied the aggregation properties of 42-residue peptides prepared by solid-phase synthesis.

Five sequences were chosen for peptide synthesis and detailed characterization: NoPhe, NoMet, 12Mut, GM6, and WT Aβ42. NoPhe and NoMet are sequences that have no phenylalanines or methionines, respectively, in the central and C-terminal hydrophobic regions. These sequences were chosen because several studies have focused on the importance of Phe or Met residues in Aβ aggregation and amyloidogenesis (25, 26). 12Mut was chosen because it contains amino acid substitutions at all 12 of the targeted residues. GM6 and WT are controls for soluble and aggregating sequences, respectively.

To assess aggregation, each peptide was prepared as a 10 $\mu$M solution in 50 mM $NaH_2PO_4$ and 100 mM NaCl (pH 7.3–7.4) and incubated at 37°C under gently agitated conditions. After overnight incubation, insoluble aggregates were removed by centrifugation at 100,000 $\times$ *g*, and soluble material in the supernatant was quantified by RP-HPLC. As shown in Fig. 4, only the control peptide GM6 (Ser-19/Pro-34) yielded a significant peak after 10 h of incubation. This peak, at 28 min, corresponds to soluble monomeric GM6 peptide. In contrast, WT Aβ42 and the hydrophobic mutants showed small or negligible peaks. Thus, in agreement with the GFP fusion results, the synthetic peptides of NoPhe, NoMet, and 12Mut resemble WT Aβ42 in their tendency to form insoluble aggregates.

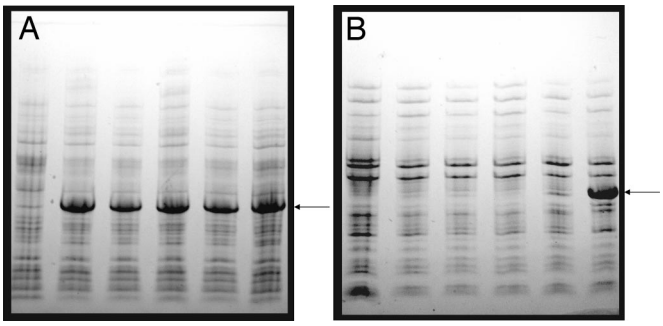


**Fig. 3.** Expression and solubility of Aβ42–GFP fusions. (*A*) Expression levels of the Aβ42–GFP fusions. SDS/PAGE of whole cells demonstrates that mutant versions of the Aβ42–GFP fusions express at the same high levels as the WT Aβ42–GFP fusion. (*B*) Soluble fraction of *E. coli* lysates. GFP fusions to randomized hydrophobic variants or to WT Aβ42 are insoluble and not detected in the soluble fraction of the cell lysates. Only the control, GM6, yielded a soluble GFP fusion protein. The samples in both gels, from left to right, are as follows: lysate from uninduced *E. coli*; lysates from *E. coli* expressing the GFP fusion of WT Aβ42, 12Mut, NoMet, and NoPhe; and the soluble control, GM6. The arrow indicates the position of the Aβ42–GFP fusion.

**Hydrophobic Variants of Synthetic Aβ42 Peptide Form Amyloid.** As shown above, the randomized hydrophobic variants cause aggregation, both in the context of the GFP fusion expressed in *E. coli* and in the 42-residue peptide prepared synthetically. Aggregative behavior in the GFP fluorescence assay does not prove that the material forms amyloid. Indeed, the type of aggregate formed by the Aβ42–GFP fusions expressed in *E. coli* is not known. Likewise, the insolubility of the synthetic peptides (Fig. 4) does not necessarily indicate amyloid formation. To assess whether this insoluble material is amyloid rather than amorphous aggregates, we monitored binding of thioflavin T (ThT). This dye is known to bind to a variety of amyloids, including those formed by insulin, transthyretin, and Aβ42; yet it does not bind to monomeric or small oligomeric complexes of these peptides and proteins (27).

Binding of ThT was monitored for each of the five synthetic peptides. Each peptide was dissolved to a final concentration of 20 $\mu$M in a buffer containing 50 mM $NaH_2PO_4$ and 100 mM NaCl (pH 7.3–7.4). Samples were gently agitated at 37°C for 5 h. ThT was then added, and binding was assayed by measuring the fluorescence of the ThT complex at 490 nm. As shown in Fig. 5, NoPhe, NoMet, and 12Mut form ThT-staining amyloid to an extent similar to WT Aβ42. These peptides assemble into amyloid despite the presence of 9–12

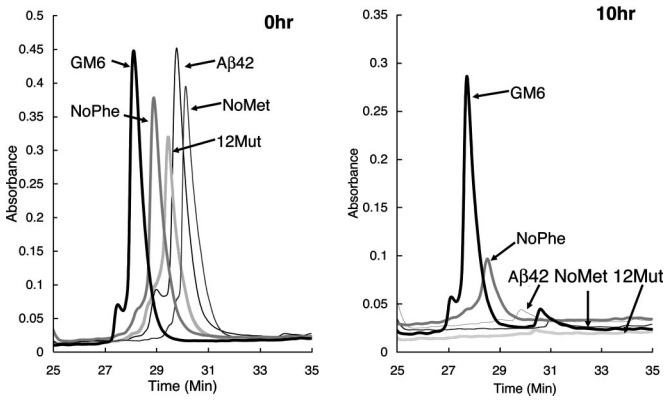


**Fig. 4.** HPLC quantification of peptide solubility. Synthetic 42-residue peptides were dissolved at a concentration of 10 $\mu$M. Samples were incubated with gentle agitation at 37°C for either 0 or 10 h. After the incubation, insoluble material was removed by centrifugation. The amount of peptide remaining in solution was quantified by RP-HPLC. Only the control peptide, GM6, yielded a significant peak at 28 min, corresponding to soluble monomeric peptide. WT A$\beta$42 and the three hydrophobic variants were mostly insoluble and yielded very little HPLC-quantifiable monomer.

amino acid substitutions in the 42-residue sequence. In contrast, the soluble control, GM6, which contains only two substitutions, hardly forms any ThT-staining material.

The rate of amyloid formation was also monitored. We used two different experimental protocols that are known to produce amyloid with very different kinetic rates (28). Under agitated conditions, A$\beta$42 is known to form amyloid rapidly, whereas under quiescent conditions, amyloidogenesis is much slower. As shown in Fig. 6*A*, under agitated conditions, amyloid formation by WT A$\beta$42 increases rapidly and reaches a plateau after ≈70 min. NoMet and 12Mut form amyloid at approximately one-third to one-half the rate of WT. NoPhe has a somewhat longer lag time and appears to have an overall rate approximately one-fourth that of WT A$\beta$42. The soluble control, GM6, shows dramatically slower kinetics: Even at long time points, GM6 forms very little ThT-staining material.

Under quiescent conditions, amyloid formation is significantly slower and occurs only after several days of incubation. As shown in Fig. 6*B*, WT A$\beta$42, 12Mut, NoMet, and NoPhe all form amyloid over the course of several days. 12Mut is slightly faster than WT,

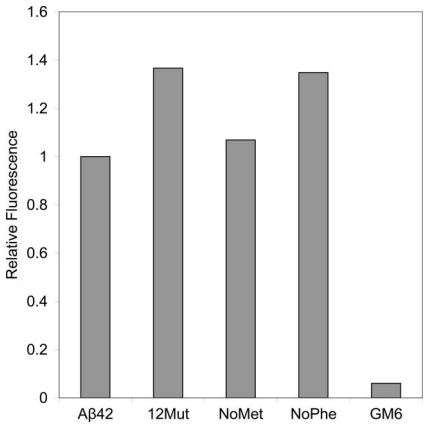


**Fig. 5.** Formation of amyloid assayed by the binding and fluorescence of ThT. Synthetic peptides were incubated at 37°C with agitation. After 5 h, samples were mixed with ThT, and fluorescence was measured at 490 nm. WT A$\beta$42 and the three hydrophobic variants yielded significant ThT-staining amyloid, whereas the soluble control, GM6 (Phe-19-Ser and Leu-34-Pro), did not.

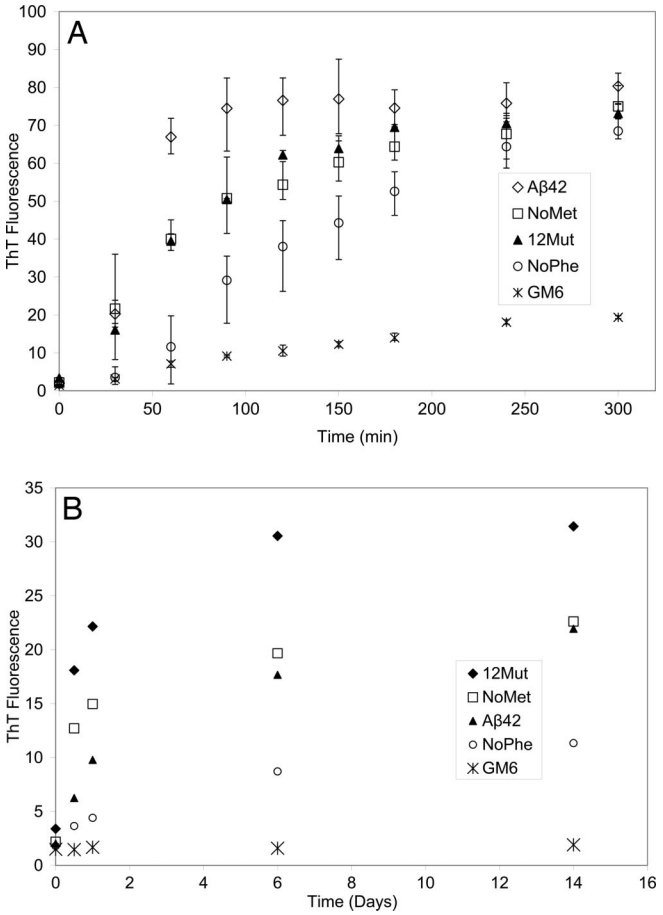


**Fig. 6.** Time course of amyloid formation. Synthetic peptides were incubated at 37°C. Samples were removed at various time points and assayed by addition of ThT and subsequent measurement of fluorescence at 490 nm. (*A*) Under agitated conditions, WT A$\beta$42, NoMet, and 12Mut form amyloid rapidly; NoPhe is slightly slower; and the soluble control, GM6, barely forms any amyloid, even after long incubations. (*B*) Under quiescent conditions, WT A$\beta$42, NoMet, and 12Mut form fibrils over the course of several days. NoPhe shows somewhat slower kinetics. GM6 does not form fibrils, even after 2-week incubation.

whereas NoPhe is slightly slower (see *Discussion*). The soluble control, GM6, does not form amyloid at all, even after 2 weeks.

**Visualization of Amyloid Fibrils by EM.** To determine whether the hydrophobic mutants form fibrils similar to those formed by WT A$\beta$42, we allowed the 42-residue synthetic peptides to aggregate under quiescent conditions and visualized the resulting structures by EM. Samples were imaged after 1, 3, 7, 14, 21, and 28 days of incubation. Fig. 7 shows the fibril morphology of 20 $\mu$M samples incubated for 21 days. All of the hydrophobic mutants formed abundant fibrils, whereas GM6 produced few fibrils, if any. Among the hydrophobic mutants, some differences in morphology can be observed: WT A$\beta$42 and NoPhe form long fibrils, whereas NoMet and 12Mut form shorter fibrils. The different lengths presumably result from slightly different kinetics of nucleation and elongation (see *Discussion*). Nonetheless, it is clear from these images that, despite the presence of 9–12 side-chain replacements, the hydrophobic mutants readily assemble into amyloid fibrils.

## Discussion

An extensive range of genetic and biochemical studies support the amyloid cascade hypothesis, which posits that accumulation of aggregated A$\beta$ peptide initiates a multistep cascade that ultimately leads to Alzheimer's disease. The aggregation of A$\beta$42, like the

BIOPHYSICS

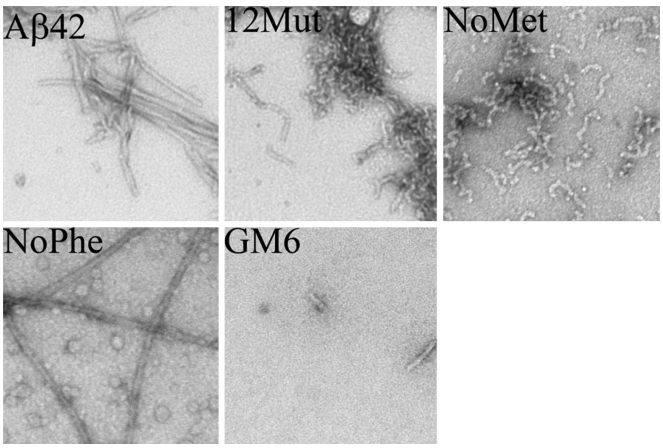


**Fig. 7.** EM of amyloid fibrils. Peptides at a concentration 20 $\mu$M were incubated at 37°C under quiescent conditions. After 21 days, samples were stained with uranyl acetate and imaged. WT A$\beta$42 and NoPhe produce long fibrils, whereas NoMet and 12Mut produce an abundance of short fibrils. Under these conditions, the soluble control, GM6 (Phe-19-Ser and Leu-34-Pro), rarely produces any fibrillar material (only two short fibrils were observed in this 3-week image).

folding of globular proteins, is thought to be driven primarily by the hydrophobic effect. For globular proteins, removal of hydrophobic side chains from contact with aqueous solvent produces folded structures with the nonpolar residues buried in the interior (29, 30). For A$\beta$42, the sequestration of hydrophobic side chains promotes aggregation into oligomers, and ultimately into amyloid fibrils, with the nonpolar side chains buried in intra- and intermolecular interfaces (31, 32).

For globular proteins, the role of the hydrophobic effect has been tested by numerous experimental and theoretical studies (18, 20, 21, 30, 33). A very stringent test of the importance of the hydrophobic effect in globular proteins was its use as the sole feature for the design of proteins *de novo*. We reported previously the design of combinatorial libraries of proteins based on the premise that folded structures could be achieved by requiring only that surface positions be occupied by polar side chains and buried positions be occupied by nonpolar side chains (18, 20). We used a binary code to construct collections of proteins wherein all surface positions were encoded by the degenerate DNA codon VAN, which encodes Lys, His, Glu, Gln, Asp, and Asn, and all buried positions were encoded by the degenerate DNA codon NTN, which encodes Met, Leu, Ile, Val, and Phe (V represents A, G, or C; N represents A, G, C, or T). The binary patterned proteins were expressed, purified, and characterized. Two high-resolution structures were determined, and the proteins were shown to form well ordered structures consistent with the design (20) (A. Go, S. Kim, J. Baum, and M.H.H., unpublished data). Our finding that well folded globular proteins can be obtained from a design strategy that uses a stochastic process to specify arbitrary nonpolar residues in the hydrophobic core demonstrated that hydrophobicity *per se*, rather than particular side chains, at specified locations in a sequence can suffice (at least in some cases) to encode native-like globular proteins.

In this work, we used a similar approach to test whether the same principles apply to the aggregation of A$\beta$42. As was the case in our work on globular protein design, we used the NTN degenerate codon to randomize the hydrophobic residues that are presumed to direct the assembly of the structure. Our results in the current study with A$\beta$42 mirror our results with the design of globular proteins. In both cases, the expected structure was obtained despite randomization of most of the nonpolar side chains.

Both for the designed globular proteins described previously and for the mutants of A$\beta$42 described here, the positioning of the hydrophobic residues is more important than the exact identities of the hydrophobic side chains. For the designed proteins, we demonstrated the importance of positioning by designing sets of sequences that had similar compositions of polar (P) and nonpolar (N) residues but differed in the positioning of these residues. Sequences with the pattern PNPPNNP, which approximates the structural repeat of an amphiphilic $\alpha$-helix, formed $\alpha$-helical structures, whereas those with the pattern PNPNPNP, which matches the repeat of amphiphilic $\beta$-strands, formed $\beta$-sheet structures (18, 19). Thus, isocompositional sequences with different polar/nonpolar periodicities formed dramatically different structures. Similarly, for A$\beta$42, the positioning of hydrophobic residues is the major determinant of structure: Generic hydrophobic side chains at the positions marked in yellow (or gray) in Fig. 1 promote fibrilogenesis; however, the WT sequence synthesized backwards (A$\beta$42-1), which has the same composition but different hydrophobic positioning, does not form fibrils (34).

It is important to emphasize that our results, for both designed globular proteins and mutants of A$\beta$42, do not imply that side-chain packing is unimportant. Quite the contrary, the well ordered structures (as compared with molten globules) that we observed for our designed four-helix bundles require that side chains adapt unique rotomers and form well defined interresidue contacts (20) (A. Go, S. Kim, J. Baum, and M.H.H, unpublished data). Similarly, in the current work on A$\beta$42, the fact that each of the nonpolar→nonpolar mutant peptides can assemble into a fibril (Fig. 7) indicates that, although each peptide uses different nonpolar residues, all of them can form ordered structures. Our results demonstrate that there are many ways to achieve ordered structures and that many different amino acid sequences are compatible with a particular folded or aggregated structure.

Although our results demonstrate that particular hydrophobic side chains are not required to promote A$\beta$42 aggregation and amyloidogenesis, some of the mutated sequences aggregate with slightly altered kinetics (Fig. 6). In particular, the NoPhe mutant, which was slightly more soluble than the other variants (Fig. 4), aggregated somewhat slower than either WT A$\beta$42 or the other nonpolar→nonpolar variants. Gazit and Azriel (25, 35) have proposed that $\pi$ stacking interactions by aromatic residues promote fibrilogenesis. They showed that although sequences devoid of aromatics can form amyloid structures, they do so more slowly. Consistent with these findings, we propose that the slower aggregation rate of the NoPhe variant results from the lack of aromatic residues in its C-terminal half. This proposal is supported by EM studies of the peptides at low concentrations: Incubation of 10 $\mu$M solutions of peptide for 3 days under quiescent conditions produced fibrils for the NoMet and 12Mut peptides, which contain four Phe residues in their C-terminal halves, but not for WT A$\beta$42 or NoPhe, which contain two or zero Phe residues, respectively, in their C-terminal halves (data not shown).

The abundance of short fibrils formed by NoMet and 12Mut (Fig. 7) might also be explained by enhanced opportunities for aromatic $\pi$ stacking resulting from the additional Phe residues in these sequences. Because of their faster aggregation kinetics at early time points (Fig. 6*B*), these sequences would nucleate more short fibrils than the NoPhe or WT A$\beta$42 peptides (Fig. 7). To test this interpretation, we recently constructed a variant of A$\beta$42 that has Phe residues at the same four positions as NoMet but is WT at all other positions. Consistent with the above hypothesis, this sequence also formed short fibrils (unpublished data).

The high-resolution structure of A$\beta$42 amyloid is not known. However, a number of structural models have been proposed (13, 31, 32). All of the models place the central and C-terminal hydrophobic stretches in $\beta$-strands. Although the specific side-chain contacts differ from one model to the next, all of them presume extensive interactions between the hydrophobic side chains in the $\beta$-strands. Our results are globally consistent with all of these models and do not distinguish between them.

In conclusion, the properties of randomized nonpolar→nonpolar mutants in A$\beta$42, characterized in the context of GFP fusions and as synthetic 42-residue peptides, demonstrate that particular nonpolar side chains in the C-terminal half of A$\beta$42 are not required for aggregation and amyloidogenesis. Randomized nonpolar side chains at all 12 positions also promote aggregation. Although some minor differences in kinetics and fibril morphology were observed, we nonetheless conclude that hydrophobicity *per se* at these positions, rather than side-chain identity, is sufficient to promote A$\beta$42 aggregation.

## Materials and Methods

**Synthesis of Gene Libraries.** Four oligonucleotides (IDT, Coralville, IA) were annealed to synthesize a library of genes encoding random hydrophobic amino acids in positions 17–20, 31–32, 34–36, and 39–41 of A$\beta$42. The sequences of these oligonucleotides are as follows: oligo-1, 5′-biotin-AGT CGT CAT ATG GAT GCG GAA TTT CGC CAT GAT TCT GGC TAT *GAA GTG CAT CAT CAG*; oligo-2, 5′-*CGC GCC TTT GTT AGA* GCC CAC ATC TTC CGC **SAH SAH SAH SAH** TTT *CTG ATG ATG CAC TTC*; oligo-3, 5′-*TCT AAC AAA GGC GCG* **DTS DTS** GGC **DTS DTS DTS** GGC GGC **DTS DTS DTS** *GCG GGA TCC TGA GTA*; oligo-4, 5′-biotin-*TAC TCA GGA TCC CGC*.

Italicized bases are complementary between oligo-1 and oligo-2, oligo-2 and oligo-3, and oligo-3 and oligo-4. **H** represents a mixture of A, C, and T in the ratio of 2:1:2. **D** represents a mixture of A, G, and T in the ratio of 2:1:2. **S** represents a mixture of G and C in the ratio of 1:1. The ratios of nucleotides were chosen to favor an equal representation of the five nonpolar residues leucine, isoleucine, valine, phenylalanine, and methionine.

Oligonucleotides 1 and 2 were annealed, and second-strand synthesis was accomplished by using Klenow enzyme (Promega, Madison, WI) at 37°C for 14 h. In a separate test tube, oligonucleotides 3 and 4 were annealed, and second-strand synthesis was accomplished similarly. Finally, the complete genes were synthesized by using a thermocycler (Easycycler; Ericomp, San Diego, CA). The genes were double-digested by using BamHI and NdeI (New England Biolabs, Ipswich, MA) and purified away from single-digested or undigested genes by using a biotin–streptavidin interaction (Pierce, Rockford, IL). Purified double-digested inserts were fused to the N terminus of the GFP gene [separated by a 12-residue linker (14)] in the pET28 vector (Novagen, San Diego, CA) and expressed in *E. coli* as described in refs. 14 and 22.

**Screening the Library.** BL21(DE3) cells (Stratagene, La Jolla, CA) containing A$\beta$42–GFP fusion constructs were plated onto nitrocellulose paper (Millipore, Billerica, MA) on LB plates supplemented with kanamycin (50 $\mu$g/ml). After overnight incubation, nitrocellulose papers were transferred to LB plates supplemented with kanamycin (50 $\mu$g/ml) and isopropyl $\beta$-D-thiogalactoside (1 mM) to induce protein expression. Quantitative fluorescence of each mutant culture was measured as described by Wurth *et al.* (14).

**Peptide Purification.** Crude peptides were purchased from the Keck Institute at Yale University (New Haven, CT) and purified by using a C4 reverse-phase column. Solvent gradients were run at 65°C using solvent A (95% water/5% acetonitrile/0.1% TFA) and solvent B (50% acetonitrile/50% water/0.1% TFA). Molecular weights of the purified peptides were confirmed by using mass spectrometry, and purity was checked by using an analytical RP-HPLC C4 column (Vydac, Columbia, MD). Purified peptides were treated with TFA to remove preexisting aggregates (36).

**Assessment of Peptide Aggregation.** Peptides were dissolved to give a final concentration of 10 $\mu$M in 50 mM $NaH_2PO_4$/100 mM NaCl/0.02% $NaN_3$, pH 7.3–7.4. Each sample was incubated at 37°C under gently agitated conditions. After incubation, samples were centrifuged at 100,000 $\times$ *g* for 30 min to remove insoluble materials, and soluble peptides in the supernatant were quantified by analytical RP-HPLC.

**ThT Assay.** Peptides were dissolved in 300 $\mu$l of DMSO and diluted with 6 ml of 8 mM NaOH (20 $\mu$M final peptide concentration). After centrifugation at 40,000 $\times$ *g* for 30 min, 5.0 ml of supernatant was taken, and 300 $\mu$l of concentrated PBS buffer was added to the supernatant (final concentration: 50 mM $NaH_2PO_4$/100 mM NaCl/0.02% $NaN_3$, pH 7.3–7.4). Samples were incubated at 37°C either with agitation or under quiescent conditions. At various time points, 500 $\mu$l of sample was mixed with 2.4 ml of a solution of ThT (7 $\mu$M ThT/50 mM glycine-NaOH, pH 8.5), and fluorescence was measured at 490 nm (excitation at 450 nm).

**EM.** Solutions were prepared as described above at peptide concentrations of 10, 20, or 50 $\mu$M. Samples were incubated at 37°C under quiescent conditions for 1, 3, 7, 14, 21, and 28 days. After the incubation, Formvar carbon-coated grids were floated on a drop of each sample for 2 min, washed twice with distilled water, and then stained for 2 min with 1% uranyl acetate. Samples were imaged by using a 912ab electron microscope (Zeiss, Thornwood, NY).

We thank Margaret Bisher for assistance with EM and Janet Crawford for peptide synthesis.

1. Alzheimer A (1906) *Neurol Centralbl* 23:1129–1136.
2. Carrell RW, Lomas DA (1997) *Lancet* 350:134–138.
3. Thomas PJ, Qu BH, Pedersen PL (1995) *Trends Biochem Sci* 20:456–459.
4. Sipe JD (1992) *Annu Rev Biochem* 61:947–975.
5. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CC (1997) *J Mol Biol* 273:729–739.
6. Goda S, Takano K, Yamagata Y, Nagata R, Akutsu H, Maki S, Namba K, Yutani K (2000) *Protein Sci* 9:369–375.
7. Fändrich M, Fletcher MA, Dobson CM (2001) *Nature* 410:165–166.
8. Hardy JA, Higgins GA (1992) *Science* 256:184–185.
9. Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ, Selkoe DJ (2002) *Nature* 416:535–539.
10. Lee EB, Leng LZ, Zhang B, Kwong L, Trojanowski JQ, Abel T, Lee VM (2006) *J Biol Chem* 281:4292–4299.
11. Murakami K, Irie K, Ohigashi H, Hara H, Nagao M, Shimizu T, Shirasawa T (2005) *J Am Chem Soc* 127:15168–15174.
12. Lesne S, Koh MT, Kotilinek L, Kayed R, Glabe CG, Yang A, Gallagher M, Ashe KH (2006) *Nature* 440:352–357.
13. Williams AD, Portelius E, Kheterpal I, Guo JT, Cook KD, Xu Y, Wetzel R (2004) *J Mol Biol* 335:833–842.
14. Wurth C, Guimard NK, Hecht MH (2002) *J Mol Biol* 319:1279–1290.
15. Kheterpal I, Williams A, Murphy C, Bledsoe B, Wetzel R (2001) *Biochemistry* 40:11757–11767.
16. Morimoto A, Irie K, Murakami K, Masuda Y, Ohigashi H, Nagao M, Fukuda H, Shimizu T, Shirasawa T (2004) *J Biol Chem* 279:52781–52788.
17. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D (2005) *Nature* 435:773–778.
18. Kamtekar S, Schiffer JM, Xiong H, Babik JM, Hecht MH (1993) *Science* 262:1680–1685.
19. West MW, Wang W, Patterson J, Mancias JD, Beasley JR, Hecht MH (1999) *Proc Natl Acad Sci USA* 96:11211–11216.
20. Wei Y, Kim S, Fela D, Baum J, Hecht MH (2003) *Proc Natl Acad Sci USA* 100:13270–13273.
21. Hecht MH, Das A, Go A, Bradley LH, Wei Y (2004) *Protein Sci* 13:1711–1723.
22. Kim W, Hecht MH (2005) *J Biol Chem* 280:35069–35076.
23. Waldo GS, Standish BM, Berendzen J, Terwilliger TC (1999) *Nat Biotechnol* 17:691–695.
24. Kim W, Kim Y, Min J, Kim DJ, Chang Y-T, Hecht MH (2006) *ACS Chem Biol* 1:461–469.
25. Gazit E (2002) *FASEB J* 16:77–83.
26. Bitan G, Tarus B, Vollers SS, Lashuel HA, Condron MM, Straub JE, Teplow DB (2003) *J Am Chem Soc* 125:15359–15365.
27. Levine H, III (1993) *Protein Sci* 2:404–410.
28. Petkova AT, Leapman RD, Guo Z, Yau WM, Mattson MP, Tycko R (2005) *Science* 307:262–265.
29. Kauzmann W (1959) *Adv Protein Chem* 14:1–63.
30. Dill KA (1990) *Biochemistry* 29:7133–7155.
31. Lührs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, Döbeli H, Schubert D, Riek R (2005) *Proc Natl Acad Sci USA* 102:17342–17347.
32. Petkova AT, Yau WM, Tycko R (2006) *Biochemistry* 45:498–512.
33. Lim WA, Sauer RT (1989) *Nature* 339:31–36.
34. Coraci IS, Husemann J, Berman JW, Hulette C, Dufour JH, Campanella GK, Luster AD, Silverstein SC, El Khoury JB (2002) *Am J Pathol* 160:101–112.
35. Azriel R, Gazit E (2001) *J Biol Chem* 276:34156–34161.
36. Jao S, Ma K, Talafous J, Orlando R, Zagorski MG (1997) *Amyloid Int J Exp Clin Invest* 4:240–252.