Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases

Fabrizio Chiti*†, Martino Calamai*, Niccolò Taddei*, Massimo Stefani*, Giampietro Ramponi*, and Christopher M. Dobson†‡

*Dipartimento di Scienze Biochimiche, Università degli Studi di Firenze, Viale Morgagni 50, 50134 Florence, Italy; and †Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Protein aggregation and the formation of highly insoluble amyloid structures is associated with a range of debilitating human conditions, which include Alzheimer's disease, Parkinson's disease, and the Creutzfeldt–Jakob disease. Muscle acylphosphatase (AcP) has already provided significant insights into mutational changes that modulate amyloid formation. In the present paper, we have used this system to investigate the effects of mutations that modify the charge state of a protein without affecting significantly the hydrophobicity or secondary structural propensities of the polypeptide chain. A highly significant inverse correlation was found to exist between the rates of aggregation of the protein variants under denaturing conditions and their overall net charge. This result indicates that aggregation is generally favored by mutations that bring the net charge of the protein closer to neutrality. In light of this finding, we have analyzed natural mutations associated with familial forms of amyloid diseases that involve alteration of the net charge of the proteins or protein fragments associated with the diseases. Sixteen mutations have been identified for which the mechanism of action that causes the pathological condition is not yet known or fully understood. Remarkably, 14 of these 16 mutations cause the net charge of the corresponding peptide or protein that converts into amyloid deposits to be reduced. This result suggests that charge has been a key parameter in molecular evolution to ensure the avoidance of protein aggregation and identifies reduction of the net charge as an important determinant in at least some forms of protein deposition diseases.

A range of debilitating human diseases is known to be associated with the formation of stable highly organized protein aggregates known as amyloid fibrils. These diseases include cerebral conditions such as Alzheimer's disease, Parkinson's disease and Creutzfeldt–Jakob disease, and also a series of systemic amyloidoses in which amyloid deposition occurs in a wider variety of organs within the body (1, 2). In each of these pathological conditions, a specific peptide or protein that is normally soluble is deposited, either intact or in fragmented form, into insoluble fibrils which accumulate in one or more types of tissue. Some amyloid diseases are rare and are associated with specific mutations involving the peptide or protein associated with amyloid deposition. Examples are familial amyloidotic polyneuropathy (3), hereditary renal amyloidosis (4, 5), and the apoA1 amyloidosis (6, 7). Other diseases, such as Alzheimer's disease, frontotemporal dementia, and Parkinson's disease, are largely sporadic, with hereditary cases involving only a limited fraction of the patients suffering from the conditions in question (8–10). The identification of specific mutations associated with familial forms of a disease that is otherwise sporadic and the investigation of the mechanism by which the mutations result in pathological behavior have proved to be of fundamental importance for identifying specific genes associated with the disease and for exploring the molecular basis of the underlying pathology (8, 10, 11).

The amino acid sequences and native structures of the proteins associated with amyloid diseases have been found to be highly variable, but structural studies have revealed that amyloid fibrils from different sources share a common ultrastructure (12). Electron microscopy has shown that amyloid fibrils are typically straight and unbranched and are formed from an assembly of protofilaments 2–5 nm wide (12). X-ray fiber diffraction studies indicate a characteristic structure in which the polypeptide chains form β strands oriented perpendicular to the long axis of the fibril, resulting in β-sheets propagating in the direction of the fibril (12).

It is increasingly recognized that the ability to form amyloid fibrils is not a property restricted to the relatively few amino acid sequences associated with specific diseases but is a generic phenomenon of polypeptide chains (13). A considerable number of proteins, including several that adopt α-helical structures under native conditions such as myoglobin and cytochrome c_{552}, have been shown to form amyloid fibrils *in vitro*, provided appropriate conditions are selected (14–17). Fibril formation involving globular proteins occurs when the native structure is at least partially unfolded under conditions in which the ability to form noncovalent interactions is retained (13, 15). Importantly, aggregates formed from such nondisease-related proteins have been found, at least in some cases, to be highly toxic to both neuronal and nonneuronal cells (18). Although the ability to form amyloid structures appears generic, the propensities of different protein sequences to aggregate under given conditions can differ very substantially (19–25).

The ability of a wide range of natural proteins to form amyloid fibrils *in vitro* provides a large variety of systems with which to study the underlying nature of the conversion from the soluble to the aggregated state of proteins. This opportunity allows the fundamental principles of a process central to human disease to be investigated using a set of proteins that can be chosen to have particular advantages for their study. Human muscle acylphosphatase (AcP) is a model system that has proved to be particularly suitable for studies of misfolding and aggregation (15, 25–28). It is a relatively simple protein with 98 residues consisting of a five-stranded antiparallel β-sheet packed against two par-

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, ''Self-Perpetuating Structural States in Biology, Disease, and Genetics,'' held March 22–24, 2002, at the National Academy of Sciences in Washington, DC.

Abbreviations: Aβ, amyloid β peptide; AcP, muscle acylphosphatase; apoA1, apolipoprotein A1; TFE, 2,2,2, trifluoroethanol; ThT, thioflavine T.

‡To whom reprint requests should be addressed. E-mail: cmd44@cam.ac.uk.

Petitioner Merck, Ex. 1156, p. 16419
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

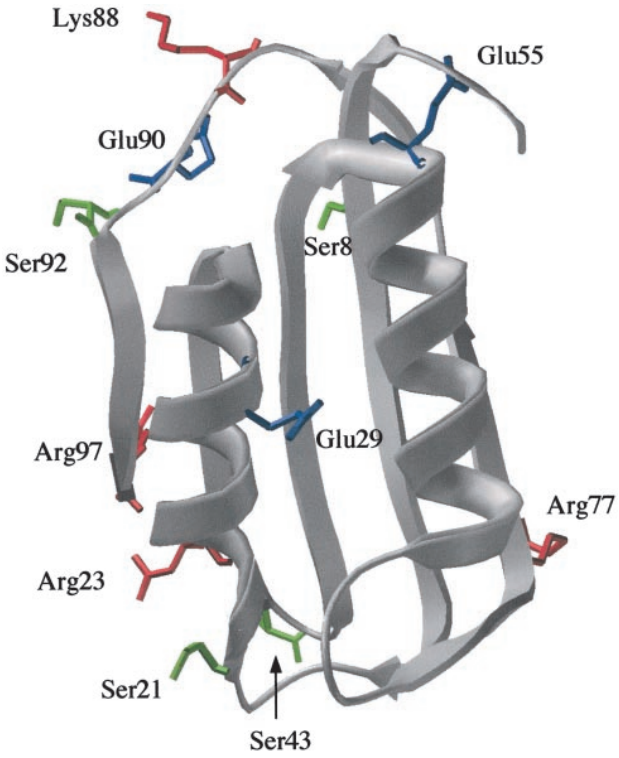


Fig. 1. Structure of AcP in its native state. Residues that have been mutated in the present study are labeled and their side chains shown. The various amino acid substitutions are listed in Table 1.

allel α-helices (Fig. 1). In addition to possessing a simple and well defined fold with no disulphide bridges or other complications such as bound cofactors, the normal folding behavior of AcP has been studied in a great detail at a residue-specific level (29, 30). This is an important issue, because folding and aggregation are potentially competitive events for a polypeptide chain in a biological environment. AcP has been shown to form readily, in the presence of moderate concentrations of denaturant such as 2,2,2, trifluoroethanol (TFE), amyloid fibrils structurally similar to those associated with disease (15). Importantly, a key step of the aggregation process, the initial formation of insoluble spherical and elongated protofibrils from soluble states, can be readily followed for AcP with a variety of biophysical techniques, and its rate is reproducible and easily measurable (25). Monitoring the first steps in the aggregation process leading to formation of prefibrillar aggregates is gaining in importance, because it is increasingly recognized that low-molecular weight oligomers that precede formation of mature amyloid fibrils, often referred to as protofibrils, represent the fundamental pathogenic species in at least many of the amyloid diseases (11, 18, 31–33).

In our initial studies of AcP, the propensity to form fibrils was investigated for a series of mutants under conditions in which the native states of the various protein variants were substantially populated, although significantly destabilized (26). The propensity to aggregate was found to correlate inversely with the conformational stability of the native state of the protein in the different mutants (26). Consistent with this finding, stabilization of the native state of AcP by ligand binding inhibits amyloid formation (27). These results show that the stability of the native state is a major factor preventing the conversion of a globular protein into amyloid fibrils under nonpathological conditions. This conclusion is in accord with studies that have shown that destabilization of the native state is a primary mechanism by which naturally occurring mutations promote their pathogenic effect in at least some hereditary amyloid diseases (34–36). The overriding significance of the native state in preventing protein aggregation also suggests that strategies aimed at stabilizing the native states of amyloidogenic proteins could be of major value in the prevention of amyloid diseases (27, 37).

In a second study, the rates of protofibril formation of more than 30 mutants, with conservative amino acid substitutions spread throughout the sequence of the protein, were determined under conditions in which the native states of all mutants are fully destabilized (25). This approach has allowed different regions of the AcP sequence to be probed for their degree of involvement in promoting the aggregation process from an ensemble of partially denatured conformations. All mutations found to perturb significantly the rate of aggregation were found to be located in two specific regions of the protein sequence, residues 16–31 and 87–98. This result indicates that aggregation of AcP can be nucleated by specific regions of the protein sequence that are consequently directly involved in the rate-determining steps of this process (25). The measured rates of aggregation were found to correlate with changes in the hydrophobicity and in the propensity to convert from α-helical to β-sheet structure of the regions of the protein in which the mutations are located (25, 28). Interestingly, natural mutations of the prion protein that leave the conformational stability of the cellular form of the protein unaltered (38, 39) increase either the β-sheet propensity or hydrophobicity of the prion sequence.

In this paper, we have extended the protein engineering approach to investigate the role that charged residues play in the process of aggregation and amyloid formation. Electrostatic interactions have been suggested to be important in the modulation of the aggregation behavior of a number of specific disease-related proteins (14, 20, 40–42). Nevertheless, a systematic investigation of the importance in aggregation of the total charge state of a polypeptide chain or of particular electrostatic interactions involving specific residues within its sequence has not yet been reported. Furthermore, it is not yet clear whether electrostatic interactions play a key role in protein aggregation generally or whether they are particularly important for a limited number of protein systems. In addition to providing new information on the driving forces of protein aggregation, this study is also designed to give additional insights into the origin of heritable amyloid diseases, because a number of these conditions are associated with amino acid replacements that alter the charge state of the aggregating polypeptide chains.

Materials and Methods

Design, Production, and Purification of AcP Mutants. The 15 mutants were designed to perturb the α-helical and β-sheet propensities and the hydrophobicity of the protein sequence as little as possible. All replacements involving substitution of or by hydrophobic residues (Val, Ile, Leu, Ala, Gly, Tyr, Phe, Trp, Cys, Met, and Pro) were therefore discarded. To evaluate the β-sheet and α-helical propensities of the protein sequence before and after mutation, the scale of β-sheet propensities edited by Street and Mayo and the AGADIR algorithm were used, respectively (43, 44). Mutagenesis was carried out by using the QuikChange Site-Directed Mutagenesis kit (Stratagene). Protein expression and purification of wild-type and mutated AcP molecules were performed according to the procedures described previously (45). All proteins have the cysteine residue at position 21 replaced by a serine residue to avoid complexities arising from a free sulfhydryl group (29). DNA sequencing was used to ensure the presence of the desired mutation. Protein concentration was measured by UV absorption by using an ε_{280} value of 1.49 $ml \cdot mg^{-1} \cdot cm^{-1}$.

Equilibrium Unfolding Experiments. Equilibrium urea denaturation curves were obtained for each AcP variant by measuring the

Petitioner Merck, Ex. 1156, p. 16420
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

intrinsic fluorescence of 25–30 equilibrated samples containing 0.02 mg·ml^{-1} protein and urea concentrations ranging from 0 to 8.1 M, in 50 mM acetate buffer, pH 5.5, 28°C. A Perkin–Elmer LS 55 with excitation and emission wavelengths of 280 and 335 nm, respectively, was used for the measurements. The data were analyzed according to the method of Santoro and Bolen (46) to yield the free energy of unfolding in the absence of denaturant ($\Delta G^{H2O}_{U\text{-}F}$), the dependence of $\Delta G_{U\text{-}F}$ on denaturant concentration (m value) and the urea concentration at which the protein is 50% denatured (C_m). To overcome the problems associated with accurate determination of m values in individual experiments, the conformational stabilities of all protein variants are expressed as $\Delta\Delta G_{U\text{-}F} = \langle m\rangle\cdot(C_m - C'_m)$, where $\langle m\rangle$ is the average m value of all mutants; C_m and C'_m are the midpoints of denaturation for the wild-type and mutated proteins, respectively (47).

Kinetics of Aggregation. The aggregation process was monitored as described previously (25). In brief, each AcP variant was incubated at a concentration of 0.4 mg·ml^{-1} in 25% (vol/vol) TFE/50 mM acetate buffer, pH 5.5, 25°C. At regular time intervals, aliquots of 60 μl of this solution were mixed with 440 μl of 25 mM phosphate buffer, pH 6.0, containing 25 μM thioflavine T (ThT). The resulting ThT fluorescence was measured by using excitation and emission wavelengths of 440 and 485 nm, respectively. Kinetic plots were fitted to single exponential functions to determine the aggregation rate constants (k_{AGG}) for the various proteins.

Results

Selection and Characteristics of the AcP Mutations. Fifteen variants of AcP, all having single replacements of charged or hydrophilic residues located on the surface of the protein, were purified for the present study (Table 1, Fig. 1). The mutations consist of substitutions of neutral residues with residues carrying a charge under the conditions of pH investigated here (S8H, S21R, S43E, and S92H), substitutions of charged residues with uncharged ones (R23Q, E29Q, E55Q, K88N, K88Q, and R97Q), and substitutions of charged residues with others of opposite sign (E29K, E29R, R77E, E90H, and R97E). All mutations involve an increase or a decrease in the charge state of the protein by 1 or 2 units (Table 1). The rate of aggregation of AcP from a denatured ensemble of conformations was found in a previous study to be sensitive to amino acid substitutions only when these occur within the two regions of the sequence comprising residues 16–31 and 87–98 (25). The majority of the mutations were therefore designed within such regions. Four mutations at positions outside these sequence regions were, however, also produced to act as controls (S8H, S43E, E55Q, and R77E).

The 15 mutations used here were chosen for their ability to leave the α-helical and β-sheet propensities of the protein sequence unchanged, because secondary structure-forming propensities have been shown to be major determinants of the aggregation behavior of AcP (see *Materials and Methods* for details) (25, 28). For the same reason, mutations in which hydrophilic residues are replaced by hydrophobic ones were not considered, because changes of hydrophobicity have been shown to modify considerably the aggregation rate of AcP (25). Through analysis of the 15 mutational variants selected here, we have therefore endeavored to relate effects on the aggregation process resulting from mutations simply to the changes in the charge state of the protein at the mutated position by minimizing mutational changes of hydrophobicity and secondary structural propensity of the protein.

The conformational stabilities of all of the 15 AcP variants were evaluated by means of equilibrium urea denaturation measurements (Fig. 2). The resulting parameters show that the mutations induce a destabilization of the native state of the protein ranging from 0 to 12 kJ mol^{-1} (Table 1). The two mutations S8H and R77E resulted in the largest values of the free energy of destabilization ($\Delta\Delta G_{U\text{-}F}$) of 11.9 and 11.7

Table 1. Parameters describing the conformational stabilities and aggregation rates of AcP mutants

	Net charge of the mutant*	C_m, M†	$\Delta\Delta G_{U\text{-}F}$, kJ mol^{-1}‡	ln k_{AGG}§
Wild-type	+5	4.0 ± 0.1	–	−6.98 ± 0.04
S8H	+6	1.8 ± 0.1	11.9 ± 0.6	−6.89 ± 0.12
S21R	+6	3.8 ± 0.1	1.4 ± 0.6	−9.04 ± 0.14
R23Q	+4	3.5 ± 0.1	3.0 ± 0.6	−6.31 ± 0.13
E29K	+7	3.3 ± 0.1	4.0 ± 0.6	−7.60 ± 0.13
E29Q	+6	4.1 ± 0.1	−0.6 ± 0.6	−7.39 ± 0.12
E29R	+7	3.3 ± 0.1	4.0 ± 0.6	−8.52 ± 0.14
S43E	+4	4.0 ± 0.1	−0.1 ± 0.6	−6.33 ± 0.14
E55Q	+6	3.3 ± 0.1	4.1 ± 0.6	−6.85 ± 0.12
R77E	+3	1.8 ± 0.1	11.7 ± 0.6	−6.16 ± 0.14
K88N	+4	3.7 ± 0.1	1.6 ± 0.6	−7.22 ± 0.12
K88Q	+4	3.9 ± 0.1	0.7 ± 0.6	−7.20 ± 0.12
E90H	+7	3.0 ± 0.1	5.5 ± 0.6	−9.50 ± 0.14
S92H	+6	3.8 ± 0.1	1.4 ± 0.6	−7.32 ± 0.12
R97E	+3	3.7 ± 0.1	1.8 ± 0.6	−6.57 ± 0.13
R97Q	+4	3.6 ± 0.1	2.2 ± 0.6	−6.98 ± 0.12

All experimental errors reported in the table are standard deviations unless stated otherwise.

*Calculated at pH 5.5 by using standard pKa values for protein side chains. Under the denaturing conditions used to study aggregation of AcP and its mutants, the pKa values of the various residues of AcP are assumed not to deviate significantly from normal ranges, because these residues are likely to be highly solvent exposed. At pH 5.5, residues of Asp, Glu, Arg, Lys, and His are therefore assumed to be charged.

†Concentration of urea at which the mutant is 50% denatured.

‡$\Delta\Delta G_{U\text{-}F}$ values were obtained by using $\Delta\Delta G_{U\text{-}F} = \langle m\rangle\cdot(C_m - C'_m)$, where $\langle m\rangle$ is the average m value of all mutants ($\langle m\rangle$ = 5.40 ± 0.15 kJ·mol^{-1}·M^{-1}); C_m and C'_m are the midpoint of denaturation for the wild-type and mutated protein, respectively. The $\Delta\Delta G_{U\text{-}F}$ values correspond to those obtained by subtracting the $\Delta G_{U\text{-}F}$ value of the mutant from that of the wild-type protein ($\Delta G_{U\text{-}F}$ = 21.7 ± 0.8 kJ·mol^{-1} for the wild type).

§The best estimate and experimental error reported for the wild-type protein are, respectively, the average value and standard error obtained from nine independent measurements.

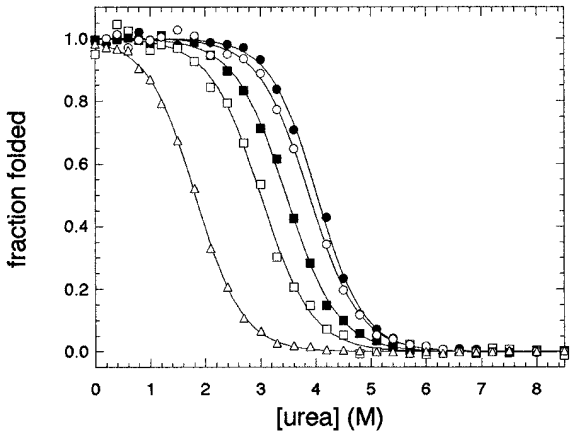


Fig. 2. Urea denaturation curves of representative AcP variants in 50 mM acetate buffer, pH 5.5, 28°C. Curves are normalized to the fraction of folded protein and correspond to those of wild-type AcP (filled circles), K88Q (open circles), R23Q (filled squares), E90H (open squares), and S8H (open triangles) mutants. The solid lines through the data represent the best fits of the data points to the equation given by Santoro and Bolen (46). The resulting thermodynamic parameters for all protein variants are listed in Table 1.

Petitioner Merck, Ex. 1156, p. 16421
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

kJ·mol^{-1}, respectively. Apart from these mutations, however, none of the remaining 13 substitutions resulted in a loss of stability of >6 kJ mol^{-1}. A previous mutational study of AcP, in which residues in the hydrophobic core of the protein were replaced by other hydrophobic residues with a smaller size, indicated far larger $\Delta\Delta G_{U\text{-}F}$ values than those observed here (30). This difference is illustrated by the average values of $\Delta\Delta G_{U\text{-}F}$ determined for the substitutions at hydrophobic core positions and surface hydrophilic residues, which are 11.5 and 3.5 kJ·mol^{-1}, respectively. Charged or hydrophilic residues do not, therefore, play as significant a role in AcP conformational stability as residues in the hydrophobic core of the protein.

Effect of the Mutations on the Aggregation Process of AcP. Some of the mutations analyzed here cause a significant, albeit small, destabilization of the native state as described in the previous section. Amino acid replacements that destabilize the native state of a protein are known to favor the process of aggregation by populating unfolded or partially folded states that are more prone to aggregation than the fully native state; such effects have been observed for AcP as well as for other protein systems (26, 34–36, 48). To overcome this problem, we probed the rate of aggregation of the polypeptide chain under conditions where the native fold of the protein is substantially disrupted, i.e., in aqueous solutions containing 25% (vol/vol) TFE. These conditions were found to denature even the most stable AcP mutants within a few seconds but still to allow aggregation to occur (15, 25). This procedure permits any change in aggregation rate to be attributed entirely to the intrinsic effect of the amino acid substitutions on the aggregation process, without the complications of additional contributions arising from the destabilization of the native state.

Fig. 3*a* describes the increase of ThT fluorescence resulting from aggregation of wild-type AcP and some representative mutants under these conditions. Such increases in ThT fluorescence reflect the formation of small aggregates, revealed by electron microscopy to be spherical or elongated 4-nm-wide protofibrils (15, 25). The rate constants obtained by fitting the data points to single exponential functions are reported in Table 1 for all of the protein variants studied here. The majority of the mutations involving residues within the two regions of the sequence 16–31 and 87–98 change the rate of aggregation to a significant extent (Table 1). In addition to having a relatively high hydrophobicity and a considerable propensity to form β structure, these two regions are also characterized by a net charge of zero. The 16–31 region contains six charged residues, three with a positive charge and the other three with a negative charge. The 87–98 region is also neutral, because it contains two positively charged residues, one negatively charged residue, and the negatively charged C terminus of the protein. If the neutralities of these two regions were directly responsible for their critical role in the process of aggregation, all substitutions of charged residues located within these two regions would be expected to disfavor the process of aggregation. This does not appear to be the case, however, as two mutations (R23Q and R97E), both of which generate a local negative charge in these regions, accelerate the process of aggregation significantly (Table 1).

A trend of a different type, however, can be observed when the aggregation rates reported in Table 1 are examined from a different perspective. Mutations decreasing the aggregation rate invariably involve the addition to the protein of an extra positive charge and/or deletion of a negative one. Similarly, mutations that increase the aggregation rate involve deletion of positively charged residues and/or insertion of negatively charged groups. This behavior can be accounted for by considering the overall charge of AcP. Under these conditions of pH, the wild-type protein has a net charge of +5 (AcP contains 9 lysines, 6 arginines, 6 glutamates, 4 aspartates, and no histidines). Mutations that disfavor the process of aggregation therefore increase the overall net charge, whereas those that favor aggregation reduce the net charge. This finding can be rationalized on a simple electrostatic argument, that self association will tend to be disfavored if the electrostatic repulsion between distinct molecules is increased, provided all other factors remain constant. For example, the mutation of the glutamate residue at position 29 to glutamine increases the net charge of the protein by 1 unit, from 5 to 6. This change results in an increase of the overall aggregation rate (Table 1). When Glu-29 is mutated to positively charged residues such as arginine or lysine, a more dramatic increase of the aggregation rate is observed (Table 1). The greater effect of the mutation to arginine relative to that to lysine can be attributed to the higher hydrophilicity of arginine as compared with lysine.

Unlike the hydrophobic interactions and secondary structural preferences, the effects of the net charge of the protein are not confined to local regions, again consistent with the simple

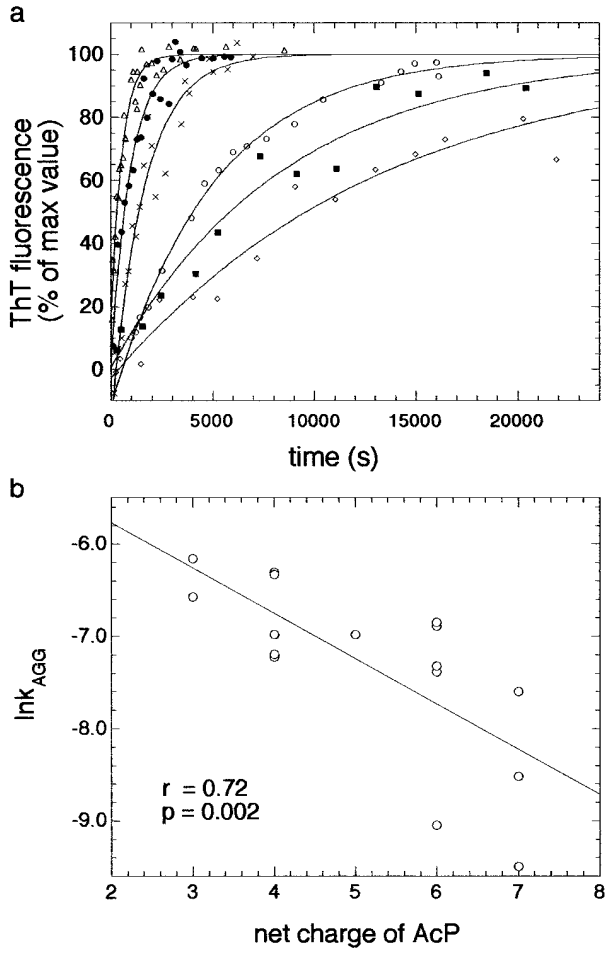


Fig. 3. (*a*) Aggregation of six representative AcP variants followed by ThT fluorescence. Aggregation was initiated in each case in 25% TFE/50 mM acetate buffer, pH 5.5, 25°C. Aliquots were withdrawn at regular time intervals for the ThT assay. The AcP variants shown are: wild-type (filled circles), R23Q (open triangles), E29Q (crosses), E29R (open circles), S21R (filled squares), and E90H (diamonds). The solid lines through the data points represent the best fits to single exponential functions. The resulting rate constant values are reported for all variants in Table 1. (*b*) Aggregation rate versus net charge constructed with the data points of the wild-type protein and the 15 mutants. Changes of net charge on mutation are calculated at pH 5.5 assuming standard pKa values for amino acid residues.

Petitioner Merck, Ex. 1156, p. 16422
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

Table 2. Mutations associated with hereditary forms of amyloid disease involving changes in the net charge of the peptide or protein associated with each disease

Mutation*	Amyloid disease	Protein/peptide involved	Effect of mutation on net charge†	Reference
E693Q	Hereditary cerebral hemorrhage with amyloidosis	Aβ	From −3 to −2‡	57
E693K	Hereditary cerebral hemorrhage with amyloidosis	Aβ	From −3 to −1‡	58
E693G	Early-onset Alzheimer's disease	Aβ	From −3 to −2‡	24, 59
D694N	Dementia with cerebral amyloid angiopathy	Aβ	From −3 to −2‡	23, 60
G26R	apoA1 amyloidosis	apoA1	From −9 to −8§	61, 62
W50R	apoA1 amyloidosis	apoA1	From −9 to −8§	63
L60R	apoA1 amyloidosis	apoA1	From −9 to −8§	64
Δ70–72	apoA1 amyloidosis	apoA1	From −9 to −8§	40
Δ60–71¶	apoA1 amyloidosis	apoA1	From −9 to −8§	7
K257T	Frontotemporal dementia with Parkinsonism	τ	From +11 to +10‖	65
ΔK280	Frontotemporal dementia with Parkinsonism	τ	From +11 to +10‖	66
K369I	Frontotemporal dementia with Parkinsonism	τ	From +11 to +10§	67
G389R	Frontotemporal dementia with Parkinsonism	τ	From +11 to +12‖	68
R406W	Frontotemporal dementia with Parkinsonism	τ	From +11 to +10‖	69
E526V	Hereditary renal amyloidosis	Fibrinogen α chain	From −3 to −2**	4
R554L	Hereditary renal amyloidosis	Fibrinogen α chain	From −3 to −4**	5

This list does not include mutations for which a causative link with pathogenesis has been proposed (see text). The 17 mutations listed include only cases for which this link is not yet established or is still under debate. A recently compiled database (70) has been utilized to identify some of the mutations listed in this table.

*The numbering refers to the sequence of the intact proteins (τ) or of the precursors (Aβ, apolipoprotein A1, fibrinogen α chain). The numbering of Aβ and τ refers to the longest isoform in each case.

†Calculated at physiological pH 7.4, when only Lys, Arg, Glu, and Asp are assumed to be charged (71).

‡The net charge of −3 refers to the 40- or 42-residue form of the Aβ peptide.

§The net charge of −9 refers to the 93-residue form of apoA1, which is found in amyloid deposits. The charge is −6 if the alternative form of 82 residue is considered.

¶Consisting of the deletion of residues 60–71 and insertion of Val-Thr at the same position.

‖The net charge of +11 refers to the four-repeat domain that forms the core of the amyloid fibrils by τ (paired helical filaments) (72). The net charge becomes +2 if the whole sequence of the longest τ isoform is considered.

**The net charge of −3 refers to the fragment of the fibrinogen α chain (residues 500–580) that has been extracted from *ex vivo* fibrils (5).

electrostatic argument. Thus, two of the four substitutions that involve residues outside the two regions 16–31 and 87–98 result in aggregation rates significantly different from that of wild-type AcP (Table 1). The most marked rate change is for the R77E mutation, a substitution that results in a decrease of the net charge by 2 units. When all 15 variants are considered, a highly significant negative correlation exists between the net charge and the aggregation rate, with a linear correlation coefficient (r) and P values of 0.72 and 0.002, respectively (Fig. 3*b*).

Despite showing a highly significant correlation, the plot shown in Fig. 3*b* also indicates that mutations that cause the same change in net charge can exhibit considerably different aggregation rates. This variation may reflect a number of factors that influence this analysis such as the fact that the hydrophobicities and secondary structure preferences of the protein are not completely unchanged by any of the mutations. Such effects are expected to be particularly pronounced when the mutations are located in the two regions previously identified as key nucleation sites for aggregation. In accord with this expectation, the S21R and E90H mutations, both of which are located well inside the two key regions, decelerate the aggregation process more than do analogous replacements outside or at the edges of these regions. In addition, specific electrostatic effects may contribute to the aggregation rates, introducing further complications into the analysis. It is notable, however, that the presence of charged residues as such does not by itself act to inhibit aggregation. If this were the case, replacements of negatively charged residues by positive ones (E29K, E29R, E90H) and of positively charged residues by negative ones (R97E, R77E) would not be expected to produce significant effects on the aggregation kinetics. By contrast, these mutations result in substantial decelerations and accelerations of aggregation, respectively (Table 1).

Discussion

Net Charge, Hydrophobicity, and Secondary Structure Preferences in Protein Aggregation. The mutational study described here indicates that the total charge of the aggregation-prone state of a protein strongly influences its propensity to aggregate. The relevance of charged residues as "structural gatekeepers" against aggregation does not, however, appear to be based only on the ability of these residues to interrupt contiguous stretches of hydrophobic residues, as suggested (21), but also on their ability to generate electrostatic repulsions between protein molecules. This idea is also supported by separate observations that shielding of positively charged groups at low pH accelerates aggregation (14) and that aggregation of different proteins induced at neutral pH by preformed fibrils of a positively charged peptide correlates inversely with the isoionic point of the various proteins tested (41). Consistent with this view, proteins that are unfolded under physiological conditions ("natively unfolded proteins") generally have a total net charge that is significantly higher than proteins that fold into globular structures (49). This mechanism is likely to be a strategy through which proteins that do not fold into globular structures avoid aggregation and remain soluble in the crowded environment of the cell.

Unlike the amino acid replacements that alter the hydrophobicity, α-helical or β-sheet propensities of the AcP sequence, mutations that modify the charge of the protein are able to alter the aggregation rate even when these modifications occur outside the regions of sequence 16–31 and 87–98 that appear to be primarily responsible for regulating the aggregation process. Different regions may play a critical role at different stages of the aggregation process and therefore produce a change in the aggregation rate when mutated. However, the change of aggregation rate resulting from such mutations correlates well with the

Petitioner Merck, Ex. 1156, p. 16423
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

overall change in the charge of the protein, rather than with that of specific regions of the sequence. The lack of sequence specificity of the charge effects suggests that the formation of specific electrostatic interactions is not an important determinant of the rate of aggregation of AcP, at least under the conditions examined in the present study. These data do not, of course, rule out the possibility that formation of salt bridges stabilizes the resulting aggregates; rather, they suggest that their formation plays a minor role in determining the kinetics of aggregation. The common ability of globular proteins to form amyloid fibrils at acidic pH values is not inconsistent with these general findings, because the primary effect of lowering the pH is the destabilization of the protein native fold, a process necessary for initiating aggregation (14, 50, 51). Indeed, aggregation of AcP is substantially slower at low pH than at neutral pH values if TFE is added to both solutions such that the protein remains partially denatured through this pH range (data not shown).

As mentioned above, the aggregation rate is highly sensitive to conservative mutations that alter the hydrophobicity or β-sheet propensity only when these occur within two relatively narrow regions of the AcP sequence, both of which have intrinsically high hydrophobicities and high propensities to form β-sheet structure (25). Because hydrogen bonding within β-structure and hydrophobic interactions between side-chains are likely to be the major stabilizing interactions within aggregates, increases in the propensities for such interactions are likely to enhance the rate at which aggregation occurs. As with the process of protein folding, however, formation of such interactions appear to be rather specific; indeed, both processes appear to involve well defined groups of residues that have the capability of nucleating the formation of either the intramolecular or intermolecular interactions (25). Interestingly, the residues that nucleate folding and aggregation are located in different regions of the polypeptide sequence, indicating that the two processes are highly distinct in natural proteins (25).

Origins of Familial Amyloid Diseases. Nearly 20 human diseases so far have been associated with amyloid deposition in either the brain or other organs in the body (1, 2). Hereditary forms have been described for many of these diseases, with single-point mutations often being the genetic changes that are responsible for their onset. The mechanisms through which the pathogenic effects of such naturally occurring mutations are mediated have been well established in a number of cases. Most of the mutations associated with early-onset Alzheimer's disease and related pathologies, for example, have been demonstrated to alter the efficiency or specificity of β- and γ-secretases, the two proteases that generate the amyloid β (Aβ) peptide from the precursor protein, APP (52–54). Such alterations are responsible for either an overproduction of the Aβ peptide or an increased release of the more amyloidogenic form of the peptide consisting of 42 rather than 40 residues. Alteration of proteolytic processing is also thought to be crucial in the pathogenesis of apolipoprotein A1 amyloidosis induced by the L174S mutation of this protein (6).

In other cases, destabilization of the native state of the globular protein responsible for amyloid deposition has been identified as the primary mechanism by which amino acid substitutions give rise to disease; this mechanism is particularly important in a number of systemic noncerebral amyloidoses such as familial amyloid polyneuropathy, lysozyme amyloidosis, and light chain amyloidosis (34–36). The two mutations of gelsolin associated with familial Finnish type amyloidosis (D187N and D187Y) have also been found to destabilize the native state of the protein, facilitating proteolytic attack that generates the highly amyloidogenic fragment (55, 56). An increased rate of oligomerization and an increased ability to permeate the cytoplasmic membrane have been suggested to be possible mechanisms by which the two natural mutations of α-synuclein, the A30P and A53T substitutions, cause familial Parkinson's disease (9, 11).

Despite the progress made recently in understanding the origins of a number of familial amyloid diseases, the mechanism of action through which the pathogenic effects of many amino acid replacements are mediated remains elusive. The conclusion from the present study raises the possibility that some disease-associated mutations act primarily through a reduction of the net charge of the corresponding peptides or proteins. To explore this possibility, we have examined pathogenic mutations involving substitution, or introduction, of charged residues within the proteins or peptides associated with amyloid diseases, and for which the causative link with the pathogenic effects is not yet established. The mutations of this type that we have been able to identify are listed in Table 2. Remarkably, 14 of 16 mutations listed in the table do indeed reduce the net charge of the polypeptides involved.

Unlike other mutations of APP, and despite the common assumption that all substitutions within the APP protein alter the processing of the protein, none of the four mutations of the Aβ peptide listed in the table have been found to act in this way (23, 24, 53). By contrast, all four mutations have been shown to increase the intrinsic propensity of Aβ to form fibrils or protofibrils *in vitro* (20, 22–24). That these mutations involve a reduction of the charge of Aβ suggests this perturbation could be the primary molecular basis for their pathogenic effect. In this regard, it is interesting to observe that the charge-preserving E693D substitution of APP is not pathogenic (20). Similarly, the six mutations of the apolipoprotein A1 listed in the table are not thought to stimulate the release of the N-terminal 82 residue fragment of the protein that is found in the fibrils, because they are too distant from the site of proteolysis (6). All six mutations, however, have the effect of reducing the negative charge within the amyloidogenic fragment (Table 2). It has, in fact, previously been suggested that the pathogenic action of these mutations could be related to alterations of the charge state (7, 40). This conclusion is well supported by the overall analysis of charge mutations described in the present work.

As far as the mutations associated with τ pathologies are concerned, a conflicting picture has emerged as to whether their pathogenicity results primarily from an impairment of the interactions of τ with microtubules or an accelerated rate of aggregation (19, 73, 74). An exception to this situation is the N279K variant of the protein, as this has been shown not to perturb the interaction with microtubules (75). The nucleotide substitution that leads to this particular mutation, however, has been shown to be pathogenic because of an alteration of the RNA splicing, which results in the production of a more aggregating variant (76, 77). Among the remaining five mutations that involve charge alterations within the τ protein, four of these reduce the net positive charge of the protein (K257T, ΔK280, K369I, and R406W). A mechanism unrelated to the reduction of charge must account for the pathogenicity of the mutation that increases the charge (G389R). Finally, the two amino acid substitutions of the fibrinogen α-chain associated with hereditary renal amyloidosis cause opposite effects on the net charge (Table 2); the E526V mutation reduces the charge of the protein, but the R554L increases it. In both cases, however, highly hydrophobic residues replace charged residues. The hydrophobic effect of the R554L substitution may outweigh its effect on the charge of the protein.

Overall, therefore, on the assumption that the pKa values are comparable to those of model compounds, 14 of the 16 mutations listed in Table 2 result in a reduction of the net charge of the corresponding peptide or protein under physiological conditions, whereas only two mutations increase it. This analysis

Petitioner Merck, Ex. 1156, p. 16424
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

does not consider mutations for which an alternative causative link with pathogenesis, such as reduction of protein stability, alterations of splicing, or proteolytic processes, has been proposed; the 16 mutations discussed here include only cases for which this link is not established or is still under debate. For example, disease-related mutations involving globular proteins, such as the prion protein or transthyretin, are not considered, because in at least some cases they are likely to be pathogenic as a consequence of their destabilizing effect on the native fold of the precursor protein (34–36). A preliminary analysis, however, indicates that 12 of 13 and 19 of 25 mutations involving alterations of charge do, in fact, cause a reduction in the total charge for the prion protein and transthyretin, respectively (for the prion protein, the 125–231 domain is considered). The present analysis suggests that the mechanism of action through which the deposition of amyloid fibrils and the onset of the disease are mediated could be rather straightforward for the familial diseases associated with these particular mutations. Indeed, these mutations may stimulate pathological effects simply by increasing the intrinsic propensity of the aggregating polypeptide chains to self-assemble through a reduction in the electrostatic repulsion between the molecules.

Conclusion

Investigating the fundamentals of protein aggregation by using a model system not linked to amyloid disease and then using any principles that emerge to interpret the behavior of peptides and proteins that form amyloid deposits *in vivo* is a powerful approach to gaining an understanding of the general principles involved in the development of pathological conditions of this type. Such an approach may reveal common features in the process of protein aggregation in a particularly straightforward manner. The studies carried out on AcP show that mutations can favor aggregation either by destabilizing the native state, hence allowing unfolded or partially unfolded species prone to aggregation to be significantly populated, or by directly favoring the process of self-assembly through an increase in hydrophobicity, β-sheet propensity, or a decrease of the net charge of the protein sequence. Examples of all these mechanisms can be found in the pathogenesis of the different familial amyloid diseases, indicating the value of studying such model systems.

Biological cells and extracellular spaces are highly crowded environments that enable the huge variety of interactions between molecules that are essential for life to take place efficiently and productively. Nevertheless, problems of self-association, or inappropriate interactions between different molecular species, will also be favored within such highly concentrated environments. The evolutionary selection of molecules able to interact specifically with their targets, but not with other species, is likely to have relied heavily on the values and distributions of charges on the surface of the molecules involved. For example, a very high net charge on a protein is likely to inhibit its interactions with other molecules of the same charge but might cause it to be unstable in a compact folded state as a result of large intramolecular repulsions (49, 78, 79). In addition, such high charges would presumably generate nonspecific association between proteins with opposite charges. It is therefore evident that protein sequences have evolved to optimize a balance between the attractions and repulsions between molecules within both the cellular and extracellular environments. The present findings on the critical role that charge plays in determining the onset of some forms of amyloid diseases, and presumably in modulating the aggregation process of proteins more generally, are a testament not only to the efficiency of the evolutionary selection of protein sequences but also to the fragility of biological organisms when such sequences are perturbed by mechanisms such as genetic mutations or misprocessing.

We thank Cristina Capanni, Vittorio Bellotti, Maria Grazia Spillantini, Jesus Zurdo, and Jeff Kelly for valuable discussions of the ideas contained within this paper. We are very grateful for support from the Fondazione Telethon-Italia (to F.C.) and the Wellcome Trust (to C.M.D.). The Dipartimento di Scienze Biochimiche in Florence is supported by the Italian Ministero dell'Istruzione dell'Università e della Ricerca (L. 449/97 Progetto "Genetica Molecolare") and the Fondazione Telethon-Italia (Project 453/bi).

1. Kelly, J. W. (1996) *Curr. Opin. Struct. Biol.* **6,** 11–17.
2. Serpell, L. C., Sunde, M. & Blake, C. C. (1997) *Cell. Mol. Life Sci.* **53,** 871–887.
3. Plante-Bordeneuve, V. & Said, G. (2000) *Curr. Opin. Neurol.* **13,** 569–573.
4. Uemichi, T., Liepnieks, J. J., Alexander, F. & Benson, M. D. (1996) *Q. J. Med.* **89,** 745–750.
5. Benson, M. D., Liepnieks, J., Uemichi, T., Wheeler, G. & Correa, R. (1993) *Nat. Genet.* **3,** 252–255.
6. Obici, L., Bellotti, V., Mangione, P., Stoppini, M., Arbustini, E., Verga, L., Zorzoli, I., Anesi, E., Zanotti, G., Campana, C., *et al.* (1999) *Am. J. Pathol.* **155,** 695–702.
7. Booth, D. R., Tan, S. Y., Booth, S. E., Tennent, G. A., Hutchinson, W. L., Hsuan, J. J., Totty, N. F., Truong, O., Soutar, A. K., Hawkins, P. N., *et al.* (1996) *J. Clin. Invest.* **97,** 2714–2721.
8. Selkoe, D. J. (2001) *Physiol. Rev.* **81,** 741–766.
9. Conway, K. A., Lee, S. J., Rochet, J. C., Ding, T. T., Williamson, R. E. & Lansbury, P. T., Jr. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 571–576.
10. Goedert, M., Ghetti, B. & Spillantini, M. G. (2000) *Ann. N.Y. Acad. Sci.* **920,** 74–83.
11. Volles, M. J. & Lansbury, P. T., Jr. (2002) *Biochemistry* **41,** 4595–4602.
12. Sunde, M. & Blake, C. (1997) *Adv. Protein Chem.* **50,** 123–159.
13. Dobson, C. M. (1999) *Trends Biochem. Sci.* **24,** 329–332.
14. Zurdo, J., Guijarro, J. I., Jimenez, J. L., Saibil, H. R. & Dobson, C. M. (2001) *J. Mol. Biol.* **311,** 325–340.
15. Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G. & Dobson, C. M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3590–3594.
16. Fandrich, M., Fletcher, M. A. & Dobson, C. M. (2001) *Nature (London)* **410,** 165–166.
17. Pertinhez, T. A., Bouchard, M., Tomlinson, E. J., Wain, R., Ferguson, S. J., Dobson, C. M. & Smith, L. J. (2001) *FEBS Lett.* **495,** 184–186.
18. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M. & Stefani, M. (2002) *Nature (London)* **416,** 507–511.
19. Barghorn, S., Zheng-Fischhofer, Q., Ackmann, M., Biernat, J., von Bergen, M., Mandelkow, E. M. & Mandelkow, E. (2000) *Biochemistry* **39,** 11714–11721.
20. Melchor, J. P., McVoy, L. & Van Nostrand, W. E. (2000) *J. Neurochem.* **74,** 2209–2212.
21. Otzen, D. E., Kristensen, O. & Oliveberg, M. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 9907–9912.
22. Sian, A. K., Frears, E. R., El-Agnaf, O. M., Patel, B. P., Manca, M. F., Siligardi, G., Hussain, R. & Austen, B. M. (2000) *Biochem. J.* **349,** 299–308.
23. Van Nostrand, W. E., Melchor, J. P., Cho, H. S., Greenberg, S. M. & Rebeck, G. W. (2001) *J. Biol. Chem.* **276,** 32860–32866.
24. Nilsberth, C., Westlind-Danielsson, A., Eckman, C. B., Condron, M. M., Axelman, K., Forsell, C., Stenh, C., Luthman, J., Teplow, D. B., Younkin, S. G., *et al.* (2001) *Nat. Neurosci.* **4,** 887–893.
25. Chiti, F., Taddei, N., Baroni, F., Capanni, C., Stefani, M., Ramponi, G. & Dobson, C. M. (2002) *Nat. Struct. Biol.* **9,** 137–143.
26. Chiti, F., Taddei, N., Bucciantini, M., White, P., Ramponi, G. & Dobson, C. M. (2000) *EMBO J.* **19,** 1441–1449.
27. Chiti, F., Taddei, N., Stefani, M., Dobson, C. M. & Ramponi, G. (2001) *Protein Sci.* **10,** 879–886.
28. Taddei, N., Capanni, C., Chiti, F., Stefani, M., Dobson, C. M. & Ramponi, G. (2001) *J. Biol. Chem.* **276,** 37149–37154.
29. van Nuland, N. A., Chiti, F., Taddei, N., Raugei, G., Ramponi, G. & Dobson, C. M. (1998) *J. Mol. Biol.* **283,** 883–891.
30. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C. M. (1999) *Nat. Struct. Biol.* **6,** 1005–1010.
31. Lin, H., Bhatia, R. & Lal, R. (2001) *FASEB J.* **15,** 2433–2444.
32. Sousa, M. M., Cardoso, I., Fernandes, R., Guimaraes, A. & Saraiva, M. (2001) *Am. J. Pathol.* **159,** 1993–2000.
33. Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J. & Selkoe, D. J. (2002) *Nature (London)* **416,** 535–539.
34. Hurle, M. R., Helms, L. R., Li, L., Chan, W. & Wetzel, R. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5446–5450.
35. McCutchen, S. L., Lai, Z. H., Miroy, G. J., Kelly, J. W. & Colon, W. (1995) *Biochemistry* **34,** 13527–13536.

Petitioner Merck, Ex. 1156, p. 16425
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033

36. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C., *et al.* (1997) *Nature (London)* **385,** 787–793.
37. Baures, P. W., Oza, V. B., Peterson, S. A. & Kelly, J. W. (1999) *Bioorg. Med. Chem.* **7,** 1339–1347.
38. Swietnicki, W., Petersen, R. B., Gambetti, P. & Surewicz, W. K. (1998) *J. Biol. Chem.* **273,** 31048–31052.
39. Liemann, S. & Glockshuber, R. (1999) *Biochemistry* **38,** 3258–3267.
40. Persey, M. R., Booth, D. R., Booth, S. E., van Zyl-Smit, R., Adams, B. K., Fattaar, A. B., Tennent, G. A., Hawkins, P. N. & Pepys, M. B. (1998) *Kidney Int.* **53,** 276–281.
41. Konno, T. (2001) *Biochemistry* **40,** 2148–2154.
42. Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S., Uversky, V. N. & Fink, A. L. (2001) *Biochemistry* **40,** 6036–6046.
43. Street, A. G. & Mayo, S. L. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9074–9076.
44. Lacroix, E., Viguera, A. R. & Serrano, L. (1998) *J. Mol. Biol.* **284,** 173–191.
45. Taddei, N., Stefani, M., Magherini, F., Chiti, F., Modesti, A., Raugei, G. & Ramponi, G. (1996) *Biochemistry* **35,** 7077–7083.
46. Santoro, M. M. & Bolen, D. W. (1988) *Biochemistry* **27,** 8063–8068.
47. Matouschek, A. & Fersht, A. R. (1991) *Methods Enzymol.* **202,** 82–112.
48. Ramirez-Alvarado, M., Merkel, J. S. & Regan, L. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 8979–8984.
49. Uversky, V. N., Gillespie, J. R. & Fink, A. L. (2000) *Proteins* **41,** 415–427.
50. McParland, V. J., Kad, N. M., Kalverda, A. P., Brown, A., Kirwin-Jones, P., Hunter, M. G., Sunde, M. & Radford, S. E. (2000) *Biochemistry* **39,** 8735–8746.
51. Lai, Z., Colon, W. & Kelly, J. W. (1996) *Biochemistry* **35,** 6470–6482.
52. Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T. D., Hardy, J., Hutton, M., Kukull, W., *et al.* (1996) *Nat. Med.* **2,** 864–870.
53. Farzan, M., Schnitzler, C. E., Vasilieva, N., Leung, D. & Choe, H. (1999) *Proc. Natl. Acad. Sci. USA* **97,** 9712–9717.
54. De Jonghe, C., Esselens, C., Kumar-Singh, S., Craessaerts, K., Serneels, S., Checler, F., Annaert, W., Van Broeckhoven, C. & De Strooper, B. (2001) *Hum. Mol. Genet.* **10,** 1665–1671.
55. Isaacson, R. L., Weeds, A. G. & Fersht, A. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11247–11252.
56. Kazmirski, S. L., Isaacson, R. L., An, C., Buckle, A., Johnson, C. M., Daggett, V. & Fersht, A. R. (2002) *Nat. Struct. Biol.* **9,** 112–116.
57. Levy, E., Carman, M. D., Fernandez-Madrid, I. J., Power, M. D., Lieberburg, I., van Duinen, S. G., Bots, G. T., Luyendijk, W. & Frangione, B. (1990) *Science* **248,** 1124–1126.
58. Miravalle, L., Tokuda, T., Chiarle, R., Giaccone, G., Bugiani, O., Tagliavini, F., Frangione, B. & Ghiso, J. (2000) *J. Biol. Chem.* **275,** 27110–27116.
59. Kamino, K., Orr, H. T., Payami, H., Wijsman, E. M., Alonso, M. E., Pulst, S. M., Anderson, L., O'dahl, S., Nemens, E., White, J. A., *et al.* (1992) *Am. J. Hum. Genet.* **51,** 998–1014.
60. Grabowski, T. J., Cho, H. S., Vonsattel, J. P., Rebeck, G. W. & Greenberg, S. M. (2001) *Ann. Neurol.* **49,** 697–705.
61. Nichols, W. C., Gregg, R. E., Brewer, H. B., Jr., & Benson, M. D. (1990) *Genomics* **8,** 318–323.
62. Vigushin, D. M., Gough, J., Allan, D., Alguacil, A., Penner, B., Pettigrew, N. M., Quinonez, G., Bernstein, K., Booth, S. E., Booth, D. R., *et al.* (1994) *Q. J. Med.* **87,** 149–154.
63. Booth, D. R., Tan, S. Y., Booth, S. E., Hsuan, J. J., Totty, N. F., Nguyen, O., Hutton, T., Vigushin, D. M., Tennent, G. A., Hutchinson, W. L., *et al.* (1995) *Q. J. Med.* **88,** 695–702.
64. Soutar, A. K., Hawkins, P. N., Vigushin, D. M., Tennent, G. A., Booth, S. E., Hutton, T., Nguyen, O., Totty, N. F., Feest, T. G., Hsuan, J. J., *et al.* (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7389–7393.
65. Rizzini, C., Goedert, M., Hodges, J. R., Smith, M. J., Jakes, R., Hills, R., Xuereb, J. H., Crowther, R. A. & Spillantini, M. G. (2000) *J. Neuropathol. Exp. Neurol.* **59,** 990–1001.
66. Rizzu, P., Van Swieten, J. C., Joosse, M., Hasegawa, M., Stevens, M., Tibben, A., Niermeijer, M. F., Hillebrand, M., Ravid, R., Oostra, B. A., *et al.* (1999) *Am. J. Hum. Genet.* **64,** 414–421.
67. Neumann, M., Schulz-Schaeffer, W., Crowther, R. A., Smith, M. J., Spillantini, M. G., Goedert, M. & Kretzschmar, H. A. (2001) *Ann. Neurol.* **50,** 503–513.
68. Murrell, J. R., Spillantini, M. G., Zolo, P., Guazzelli, M., Smith, M. J., Hasegawa, M., Redi, F., Crowther, R. A., Pietrini, P., Ghetti, B., *et al.* (1999) *J. Neuropathol. Exp. Neurol.* **58,** 1207–1226.
69. Hutton, M., Lendon, C. L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., Pickering-Brown, S., Chakraverty, S., Isaacs, A., *et al.* (1998) *Nature (London)* **393,** 702–705.
70. Siepen, J. A. & Westhead, D. R. (2002) *Protein Sci.* **11,** 1862–1866.
71. Creighton, T. E. (1993) in *Proteins* (Freeman, New York), 2nd Ed., pp. 6–7.
72. Wischik, C. M., Novak, M., Thogersen, H. C., Edwards, P. C., Runswick, M. J., Jakes, R., Walker, J. E., Milstein, C., Roth, M. & Klug, A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4506–4510.
73. Goedert, M., Jakes, R. & Crowther, R. A. (1999) *FEBS Lett.* **450,** 306–311.
74. Friedhoff, P., von Bergen, M., Mandelkow, E. M. & Mandelkow, E. (2000) *Biochim. Biophys. Acta* **1502,** 122–132.
75. Hong, M., Zhukareva, V., Vogelsberg-Ragaglia, V., Wszolek, Z., Reed, L., Miller, B. I., Geschwind, D. H., Bird, T. D., McKeel, D., Goate, A., *et al.* (1998) *Science* **282,** 1914–1917.
76. Delisle, M. B., Murrell, J. R., Richardson, R., Trofatter, J. A., Rascol, O., Soulages, X., Mohr, M., Calvas, P. & Ghetti, B. (1999) *Acta Neuropathol.* **98,** 62–77.
77. Hasegawa, M., Smith, M. J., Iijima, M., Tabira, T. & Goedert, M. (1999) *FEBS Lett.* **443,** 93–96.
78. Goto, Y., Calciano, L. J. & Fink, A. L. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 573–577.
79. Luisi, D. L. & Raleigh, D. P. (2000) *J. Mol. Biol.* **299,** 1091–1100.

Petitioner Merck, Ex. 1156, p. 16426
Merck Sharp & Dohme LLC v. Halozyme Inc.
PGR2025-00033