doi:10.1016/j.jmb.2008.04.028 *J. Mol. Biol.* (2008) **379**, 815–829

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

# Structures and Free-Energy Landscapes of the Wild Type and Mutants of the $A\beta_{21-30}$ Peptide Are Determined by an Interplay between Intrapeptide Electrostatic and Hydrophobic Interactions

**Bogdan Tarus[1], John E. Straub[1]* and D. Thirumalai[2,3]***

[1]*Department of Chemistry, Boston University, Boston, MA 02215, USA*

[2]*Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA*

[3]*Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA*

*Received 17 December 2007; received in revised form 9 April 2008; accepted 11 April 2008*
Available online 14 May 2008

The initial events in protein aggregation involve fluctuations that populate monomer conformations, which lead to oligomerization and fibril assembly. The highly populated structures, driven by a balance between hydrophobic and electrostatic interactions in the protease-resistant wild-type $A\beta_{21-30}$ peptide and mutants E22Q (Dutch), D23N (Iowa), and K28N, are analyzed using molecular dynamics simulations. Intrapeptide electrostatic interactions were connected to calculated p$K_a$ values that compare well with the experimental estimates. The p$K_a$ values of the titratable residues show that E22 and D23 side chains form salt bridges only infrequently with the K28 side chain. Contacts between E22–K28 are more probable in "dried" salt bridges, whereas D23–K28 contacts are more probable in solvated salt bridges. The strength of the intrapeptide hydrophobic interactions increases as D23N<WT<E22Q<K28A. Free-energy profiles and disconnectivity representation of the energy landscapes show that the monomer structures partition into four distinct basins. The hydrophobic interactions cluster the $A\beta_{21-30}$ peptide into two basins, differentiated by the relative position of the DVG(23–25) and GSN(25–27) fragments about the G25 residue. The E22Q mutation increases the population with intact VGSN turn compared to the wild-type (WT) peptide. The increase in the population of the structures in the aggregation-prone Basin I in E22Q, which occurs solely due to the difference in charge states between the Dutch mutant and the WT, gives a structural explanation of the somewhat larger aggregation rate in the mutant. The D23N mutation dramatically reduces the intrapeptide interactions. The K28A mutation increases the intrapeptide hydrophobic interactions that promote population of structures in Basin I and Basin II whose structures are characterized by hydrophobic interaction between V24 and K28 side chains but with well-separated ends of the backbone atoms in the VGSN turn. The intrapeptide electrostatic interactions in the WT and E22Q peptides roughen the free-energy surface compared to the K28A peptide. The D23N mutation has a flat free-energy surface, corresponding to an increased population of random coil-like structures with weak hydrophobic and electrostatic

**Corresponding authors.* D. Thirumalai is to be contacted at Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA. E-mail addresses: straub@bu.edu; thirum@umd.edu.

Present address: B. Tarus, Computational Molecular Biophysics Group, Interdisciplinary Center for Scientific Computing (IWR), Im Neuenheimer Feld 368, 69120 Heidelberg, Germany.

Abbreviations used: Aβ, amyloid β-protein; CHARMM, Chemistry at Harvard Molecular Mechanics; DB, desolvation barrier; FE-TRDG, free-energy transition disconnectivity graph; MD, molecular dynamics; SC, side chain; TRDG, transition disconnectivity graph; WT, wild type.

0022-2836/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.

interactions. We propose that mutations or sequences that enhance the probability of occupying Basin I would promote aggregation of Aβ peptides.
© 2008 Elsevier Ltd. All rights reserved.

*Keywords:* Alzheimer's disease; amyloid β-protein; molecular dynamics simulation; transition disconectivity graphs; free-energy surface

***Edited by D. Case***

## Introduction

Amyloid β-protein (Aβ) is a major component of proteinaceous mass involved in Alzheimer's disease.[1,2] Produced by proteolytic cleavage of Aβ precursor protein, Aβ consists of 40- or 42-residue peptides *in vivo*.[1] It was initially proposed that ordered extracellular fibrillar aggregates of Aβ were pathogenic agents that led to neuronal cell death in the brains of patients with Alzheimer's disease.[3] Recently,[4–7] neurotoxicity has been associated with metastable low molecular weight oligomers, which are intermediates in the fibrillization process of Aβ.[8] It was proposed that the Aβ monomers nucleated and formed amyloid fibrils through a linear-aggregation mechanism.[9] An outstanding problem is characterization of the aggregation mechanism in terms of key intra- and interpeptide interactions in the monomer, low molecular weight oligomers, and fibrils.

Based on general theoretical arguments, several scenarios for protein aggregation have been proposed.[10,11] In all scenarios, the key early event is the structural transitions in the monomer due to fluctuations, denaturant stress, or interactions with other monomers that populate aggregation-prone conformations. Thus, it is imperative to describe the structural characteristics that the monomers adopt and to assess whether these structures can trigger ordered oligomer formation upon interpeptide interactions. It is also important to decipher how variations in sequence shift the population of the most probable low free-energy monomer structures. In the Aβ monomers, it has long been appreciated that the VGSN region (residues 24–27) plays a key role in imparting a degree of local order, even though globally they are best treated as random coils. Kirschner *et al.*[12] used theoretical predictors of β-turn propensity[13] to show that the region of sequence centered about residue 26 (as well as residue 8) has high β-turn potential. Limited proteolysis studies[14] identified β-turn in the SNKG(26–29) region of the full-length peptide. NMR and CD spectroscopy[15] for the full-length Aβ-protein and fragments in a solution of trifluoroethanol and water show that the VGSN(24–27) region forms stable turn structures. Several other studies[16–19] also identified a helix–turn–helix structure for the Aβ-peptide in nonpolar or membrane-mimicking environments.

The aqueous solution phase NMR structure of the $A\beta_{10-35}$ congener peptide[20,21] shows that, in addition to the LVFFA(17–21) "central hydrophobic cluster" region, a turn centered at the VGSN(24–27) region of the peptide is a key structural motif. Subsequent studies of the $A\beta_{10-35}$ peptide demonstrated that the $C^{\alpha}$ proton chemical shift in the VGSNKG(24–29) region does not change over a range of temperature, which suggests that the peptide structure surrounding the VGSN(24–27) region is stable in the temperature range from 5° to 35°.[21,22]

Simulation studies[21,23,24] lent support for the stability of the proposed structure of the LVFFA (17–21) central hydrophobic cluster and VGSN(24–27) turn regions on the nanosecond time scale.[21] The conservation of the VGSN(24–27) turn region in the putative β-fibril[25] and the "collapsed coil" structures leads to the conjecture that the VGSN(24–27) may nucleate the formation of Aβ aggregates. Recent molecular dynamics (MD) simulations[26–28] and solid-state NMR experiments[29] on the $A\beta_{21-30}$ peptide revealed this decapeptide's highly fluctuating structure, with relatively low probability of forming intrapeptide E22–K28 and D23–K28 salt bridges and the V24–K28 hydrophobic interaction also supporting the role of the VGSN(24–27) in higher-ordered structures.

The bend in the VGSN(24–27) fragment also plays a crucial role in the stability of the $A\beta_{1-40}$ fibril structures.[30–32] Both intrapeptide[30,31] and interpeptide[32] D23–K28 salt bridges are key structural elements that may stabilize the bend in the VGSN (24–27) region of the peptide in the fibril state. The monomeric structural construct with the characteristic bend in the VGSN region satisfies the principle of amyloid self-assembly,[33] namely, structures that maximize both intra- and intermolecular hydrophobic interactions and the number of salt bridges are the most stable.

Here, we investigate the interplay between hydrophobic and electrostatic intrapeptide interactions in stabilizing the structure of the VGSN(24–27) region in the wild type (WT) and mutants of the $A\beta_{21-30}$ peptide. The intrapeptide electrostatic interactions were investigated through estimates of $pK_a$ values. Explicit-solvent MD simulation $pK_a$ estimates were compared with the experimental values measured through solution-state NMR chemical-shift titrations.[34] Free-energy surfaces of the $A\beta_{21-30}$(WT) peptide and three mutant peptides, E22Q ("Dutch" mutant), D23N ("Iowa" mutant), and K28A, were projected on coordinates defined by the V24–K28 and V24–N27 distances to unveil how intrapeptide hydrophobicity is modulated by electrostatic interactions. Transition disconnectivity graphs were used to uncover key details of the free-energy surfaces of the WT and the E22Q, D23N, and K28A mutant peptides. We find that mutants, such as E22Q, that lead to even modest enhancement of aggregation-prone structures (intact D23–K28 salt bridge and

stable VGSN turn), can lead to an increase in rates of fibril formation.)

## Results

### p$K_a$ values indicate weak intrapeptide electrostatic interactions

Since Aβ$_{21–30}$ is a small peptide, its interior is mainly formed by the backbone, while the polar and charged residues are exposed to water. As a result, the most important polar or charged groups are solvated. Consequently, electrostatic interactions are unlikely to be sufficiently strong to change the protonation state of a titratable residue compared to its value in isolation. These expectations are borne out by computing the Born, background, and interaction contributions to the p$K_a$ shifts [see Eq. (2)]. We used two values (4 and 20) for the dielectric constant of the protein and found that p$K_a$ values are similar (data not shown).

#### *pK*$_a$ for WT

The computed p$K_a$ values of residues E22, D23, and K28 in the WT, presented in Table 1 and Fig. 1a, are in good accord with the experimental p$K_a$ values.[34] We used a value of 10.5 as a model p$K_a$ for K28. The p$K_a$ values of residues E22 and D23 (4.5 and 4.0, respectively) are the same as the model values, which implies that these two residues are solvated. The computationally determined p$K_a$ values of E22 and D23 (4.6 and 3.8, respectively; Table 1) suggest that there is a weak, unfavorable electrostatic interaction between E22 and the negative partial charges of the peptide. Comparison of the p$K_a$ shift of K28 (11.2) and the model value (10.5) suggests that K28 engages in favorable electrostatic interactions with the peptide.

#### *pK*$_a$ for mutants

The experimental p$K_a$ values of residues D23 and K28 in the E22Q Dutch mutant (4.0 and 11.1, respectively, in Table 1) are close to that of the WT, implying that the D23 side chain is mainly solvated and the K28 side chain interacts favorably with the polar groups of the peptide. The corresponding computationally determined p$K_a$ values of D23 and K28 in the E22Q mutant peptide are 3.7 and 10.5, respectively (Table 1), suggesting that the D23 side chain engages in a slightly favorable interaction with the peptide, while the K28 side chain is more tightly solvated.

In the D23N Iowa mutant (Table 1 and Fig. 1c), the E22 side chain is isolated, with both experimental and calculated p$K_a$ (4.5 and 4.3, respectively) being close to the model value. A decrease in the p$K_a$ for K28 compared to the WT and E22Q indicates that the favorable intrapeptide electrostatic interactions of the K28 side chain in the WT and E22Q peptides are diminished in the D23N mutant. Taken together, these results show that the D23 side chain in the WT and the E22Q can infrequently form a salt bridge with K28.

In the K28A mutant, the favorable electrostatic influences on the E22 and D23 residues from the rest of the peptide are abolished. The negatively charged side chains of the E22 and D23 residues interact unfavorably, increasing their p$K_a$ values relative to the model values (Table 1 and Fig. 1d). Thus, substitution of K28 by a hydrophobic residue drastically alters the nature of the electrostatic interactions.

A couple of comments about the accuracy of the p$K_a$ calculations and their implications are worth making. (i) One way to test the accuracy of the method used to calculate the p$K_a$ values is to verify the linearity of the equation:[35]

$$\mathrm{p}K_{\mathrm{a,i}}^{\mathrm{exp}} = \mathrm{p}K_{\mathrm{a,i}}^{\mathrm{model}} - \frac{1}{2.303 k_{\mathrm{B}} T}\Delta G_i,$$

where $i$ is the titratable residue, E22, D23, or K28. $\mathrm{p}K_{\mathrm{a},i}^{\mathrm{exp}}$ and $\mathrm{p}K_{\mathrm{a},i}^{\mathrm{model}}$ are the experimental p$K_a$ and the one for the isolated residue, respectively. The electrostatic free energy $\Delta G_i$ is obtained by solving the linearized Poisson–Boltzmann equation and includes the Born free energy, the background free energy, and the electrostatic free energy of interaction [Eq. (2)]. A good correlation between calculated and measured p$K_a$ values should give intercepts at p$K_a$ of the isolated residue with a slope of $-1/(2.303 k_{\mathrm{B}}T)$ or $-0.73$ kcal/mol. A plot of the experimental p$K_a$ values *versus* the $\Delta G_i$ yields intercepts close to the p$K_a$ values of the isolated E22, D23, and K28 residues (Fig. 2). The decreased correlation between the experimental and computed p$K_a$ values of K28 (Fig. 2) suggests interactions of K28

**Table 1.** Computational p$K_a$ values are compared with the experimental[34] p$K_a$ values (in parenthesis) for titratable residues E22, D23, and K28 in the Aβ$_{21–30}$ peptide

| | E22 | D23 | K28 |
|---|---|---|---|
| Model's p$K_a$ | 4.5 | 4.0 | 10.5 |
| Ac-Aβ$_{21–30}$-NH$_2$(WT) | 4.6±0.3 (4.5±0.02) | 3.8±0.7 (4.0±0.02) | 10.7±0.2 (11.2±0.01) |
| Ac-Aβ$_{21–30}$-NH$_2$(E22Q) | — | 3.7±0.3 (4.0±0.02) | 10.5±0.2 (11.1±0.01) |
| Ac-Aβ$_{21–30}$-NH$_2$(D23N) | 4.3±0.2 (4.5±0.01) | — | 10.5±0.2 (10.4±0.01) |
| Ac-Aβ$_{21–30}$-NH$_2$(K28A) | 4.9±0.2 (4.9±0.04) | 3.9±0.3 (4.7±0.03) | — |

The p$K_a$ values of the isolated residues and the residues in the peptide structure are similar, indicating that the p$K_a$ shifts result from weak intrapeptide electrostatic interactions.

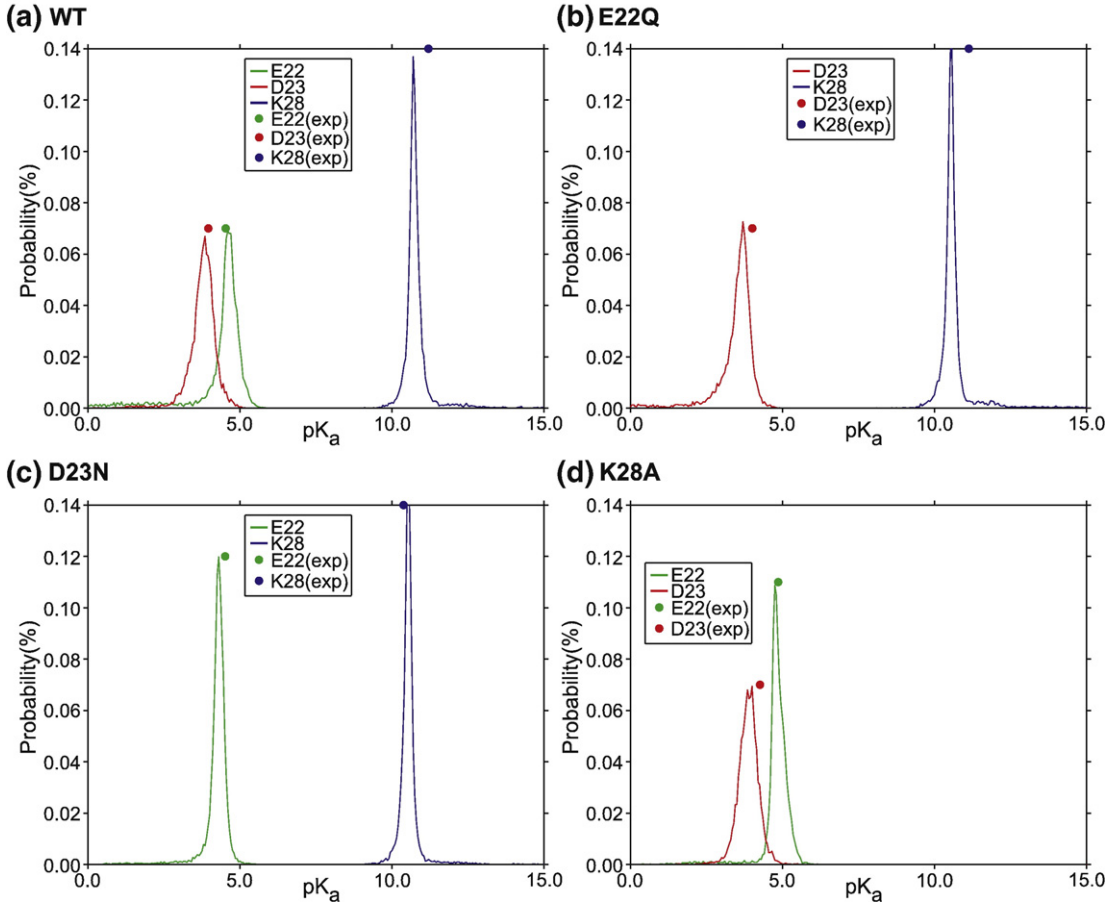


**Fig. 1.** Comparison of calculated and experimental p$K_a$ values of WT and mutants of the Aβ$_{21–30}$ peptide. The computed p$K_a$ values of the titratable residues are in good agreement with the experimental[34] p$K_a$ values. The reported p$K_a$ values are computed with a dielectric coefficient of 4.

with the peptide are not accurately described by the simulations. However, the difference between the experimental (11.1±0.01) and computational (10.5±0.2) p$K_a$ for the K28 side chain in the E22Q mutant is equivalent only to 0.8 kcal/mol favorable electrostatic interaction between the K28 side chain and the rest of the peptide. This value is comparable with the thermal fluctuations, 0.6 kcal/mol at 300 K, and hence is relatively small.

(ii) Both the experimental and computed small values of the p$K_a$ shift values of the titratable residues within the Aβ$_{21–30}$(WT) and its variants, E22Q, D23N, and K28A, suggest that the titratable residues are largely exposed to solvent. Our simulations, which show relatively small p$K_a$ shifts, suggest that the probability of forming structures with intact salt bridge with a stable bend (aggregation-prone structures) is small.

## Strength of salt bridges varies greatly from WT to mutants

The experimentally measured p$K_a$ values do not allow us to directly interpret the strength of electrostatic interaction between the E22–K28 and D23–K28 charged side-chain pairs. The distribution of the electrostatic energies for the E22–K28 and D23–K28 pairs shows that $\Delta G_{inter}$ [Eq. (2)] for most of the simulated structures (between 78% and 95%) is centered around zero (Fig. 3). This implies that the E22, D23, and K28

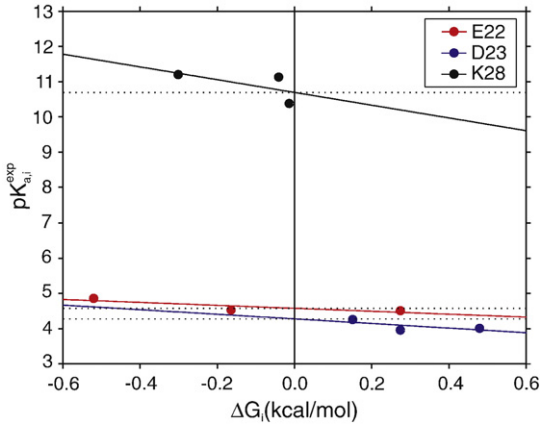


**Fig. 2.** Variation of the experimental p$K_a$, $pK_{a,i}^{exp}$, as a function of the electrostatic free energy, $\Delta G_i$. The linear dependence validates the computational method used to calculate the p$K_a$s. The intercepts for the E22, D23, and K28 residues are 4.6, 4.3, and 10.7, respectively, with the slopes of −0.4, −0.7, and −1.8, respectively. The root-mean-squared errors are 0.1, 0.2, and 0.5, respectively.

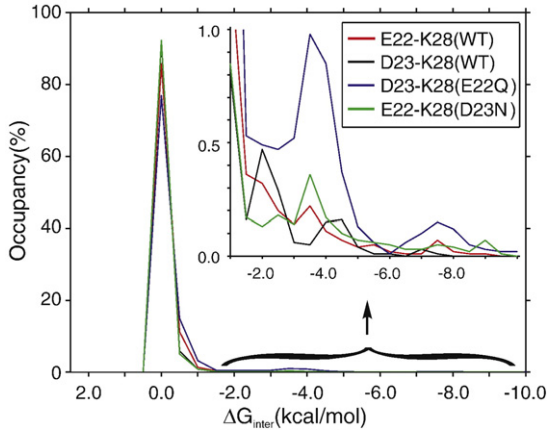


**Fig. 3.** The electrostatic interaction energy distributions between the residues E22–K28 and D23–K28 indicate that the titratable residues in the Aβ$_{21–30}$ peptide are highly solvated during the simulation. Hydrated electrostatic contacts associated with weakly favorable interactions are of low favorable energies, explaining their instability. In the WT peptide, the D23–K28 electrostatic interaction is more frequent in a low favorable energy domain (around −2.0 kcal/mol), while the E22–K28 is more frequent in a larger favorable energy domain (around −3.5 kcal/mol). The E22Q mutation increases the probability of electrostatic interaction between the D23 and K28 residues. The D23N mutation decreases the probability of electrostatic interaction between the E22 and K28 residues in the low-value domain (−2.0 kcal/mol), while it increases the probability in the larger-value domain (around −3.5 kcal/mol).

side chains are highly solvated, in accordance with the experimental and calculated p$K_a$ values (see Table 1) and the computed p$K_a$-shift values (see Fig. 1). The strength of the electrostatic interactions between the E22–K28 and D23–K28 side-chain pairs appears to be similar in the WT peptide (Fig. 3).

The E22Q Dutch mutation increases the importance of electrostatic interaction between the D23 and K28 side chains, around −2.0 and −3.5 kcal/mol, and introduces stronger electrostatic interactions, around −7.5 kcal/mol (Fig. 3). A value of −2.0 kcal/mol corresponds to a solvated electrostatic interaction between two unlike unit charges. A lower value, such as −7.5 kcal/mol, implies exclusion of the solvating water molecules through burial of the salt bridge in the peptide interior. The somewhat more favorable interaction between D23 and K28 results in a modest increase in the probability of salt bridge formation in the Dutch mutant compared to the WT. The tendency to form this stable salt bridge with intact VGSN turn provides a structural explanation for the E22Q to aggregate more rapidly than the WT. The D23N mutation (see Fig. 3) decreases the frequency of the weaker electrostatic interactions (−2.0 kcal/mol) and only slightly increases the moderate electrostatic interactions (−3.5 kcal/mol) between the E22 and K28 residues.

## Nature of salt bridge depends on the sequence

The strength of the electrostatic interactions within the E22–K28 and D23–K28 pairs depends on both the distance between the charged side chains and the local dielectric coefficient, which is related to the solvent exposure. The distributions of distances between the atoms C$^{\delta}$(E22)–N$^{\zeta}$(K28) and C$^{\gamma}$(D23)–N$^{\zeta}$(K28) in the WT, E22Q, and D23N are shown in Fig. 4a. A basin centered at 4.3 Å between the N$^{\zeta}$(K28) and C$^{\delta}$(E22) or C$^{\gamma}$(D23) atoms, respectively, accommodates one hydrogen bond between the donor N$^{\zeta}$(K28) atom and the acceptor atoms of either E22 or D23 side chains, O$^{\epsilon}$(E22) or O$^{\delta}$(D23), respectively, corresponding to hydrogen-bonded

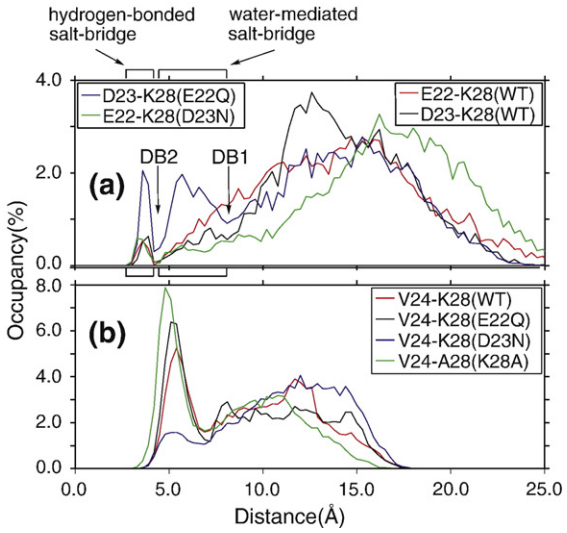


**Fig. 4.** Distribution of distances characterizing the intrapeptide electrostatic (a) and hydrophobic (b) interactions. (a) The distributions of the distances between C$^{\delta}$(E22)–N$^{\zeta}$(K28) and C$^{\gamma}$(D23)–N$^{\zeta}$(K28) indicate the weak electrostatic interaction between these titratable residue pairs. The hydrogen-bonded salt bridge of the E22–K28 and D23–K28 pairs are equally populated in the Aβ$_{21–30}$(WT) peptide, while the water-mediated salt bridge is more populated for the E22–K28 pair. The E22Q mutation increases the D23–K28 close interaction probability. The D23N mutation does not change the probability of the hydrogen-bonded E22–K28 salt bridge compared to the WT, while the water-mediated E22–K28 salt bridge is less probable. DB1 indicates the first desolvation barrier, while DB2 indicates the second desolvation barrier. (b) The distance between the centers of mass of the V24 side chain and the hydrophobic portion of residue 28 reveals the distribution of the hydrophobic intrapeptide interactions. Position 28 is occupied by K in the WT and E22Q and D23N variants, and by A in the K28A mutant. Two distinct basins characterize the distributions. The distances between the side chain V24 and the residue 28 in contact are centered around 5.0 Å. The solvated side chains 24 and 28 have their centers of mass separated by at least one solvation shell with a minimum distance of 7.5 Å. The E22Q mutation increases the intrapeptide hydrophobic interaction, while the D23N decreases it. The K28A mutation maximizes the population of the 24–28 hydrophobic contacts.

salt bridges. A second basin, centered on 5.5 Å (Fig. 4), accommodates one water molecule between the $N^{\zeta}$(K28) and $O^{\varepsilon}$(E22) or $O^{\delta}$(D23), and corresponds to water-mediated salt bridges. The border between the water-mediated salt bridges and the highly solvated side chains corresponds to the first desolvation barrier, DB1, at 7.5 Å. A second desolvation barrier, DB2, separates the water-mediated and hydrogen-bonded salt bridges at 4.5 Å. The E22–K28 and D23–K28 pairs form hydrogen-bonded salt bridges with low probability, 1.3%, in the WT (Fig. 4a). The interactions between the negatively charged E22 and D23 side chains and positively charged K28 are somewhat more probable, 6.4% and 4.4%, respectively, when mediated by a water molecule. Figure 4a suggests that the formation of the water-mediated E22–K28 contact is not an activated process but the D23–K28 water-mediated contact must overcome a small barrier at 7.5 Å, DB1 in Fig. 4a. Removal of the last solvation shell between the E22–K28 or D23–K28 side chains in the WT is an activated process characterized by poorly populated basins separated by the solvated charged residues with a barrier at 4.5 Å, DB2 in Fig. 4a.

The E22Q mutation increases the stability of the D23–K28 hydrogen-bonded and water-mediated salt bridges, with formation probabilities of 5% and 14%, respectively (see Fig. 4a). The increase in the population (cumulatively by a factor of 3) of structures with the D23–K28 salt bridge is sufficient to explain the modest enhancement in the rate of fibril formation in E22Q. The D23N mutation generates structures with a hydrogen-bonded E22–K28 salt bridge similarly populated (1.6%) as in the WT (1.3%) and is also separated by an activation barrier at 4.5 Å (see DB2 in Fig. 4a). The probability of observing E22–K28 water-mediated salt bridges is reduced to 3.2% compared to the similar contact in the WT peptide (6.4%). The water-mediated salt-bridge formation does not involve crossing of activation barrier in order to remove the solvation water in the D23N peptide.

## Intrapeptide hydrophobic contacts are similar in the WT and E22Q

We focused on the interaction between the side chain of V24 and the aliphatic portion of the K28 side chain as the main contribution to intrapeptide hydrophobicity. Recent NMR experimental measurements[29] and coarse-grained MD simulations[26] suggest that the V24 propyl side chain and the butyl portion of K28 side chain are in contact. We have examined the correlation between the electrostatic and hydrophobic intrapeptide interactions by selectively modifying the electrostatic charge state of the residues 22, 23, and 28.

Fig. 4b shows the distributions of the distances between centers of mass of the V24 side chain and the aliphatic part of the K28 side chain, for the WT and the E22Q and D23N mutants. The distribution of the distance between the V24 and A28 is also shown for the K28A mutant. The V24 packs against K28 in 29% of conformations in the WT. In the Dutch mutant, E22Q, the fraction of molecules with V24–K28 hydrophobic contact increases to ≈33%. This observation supports the inference that the major difference between the WT and the Dutch mutant must arise largely from the differences in the charge states. The mutant K28A diminishes the strong electrostatic interaction of the K28 ammonium group with the surrounding water molecules, thus increasing the population of the hydrophobic interactions between the V24 and A28 side chains to 42%. We find that the desolvation of the space between the V24 side chain and the butyl group of K28 (or methyl of A28) is a one-barrier process, in contrast to the two-barrier desolvation observed in the contact between the D23 and K28 side chains (see Fig. 4b).

## K28 makes transient contacts with the peptide backbone

The delicate balance of hydrophilic and hydrophobic forces that act on the K28 side chain can be influenced by the change in the electrostatic charge state of the E22 or D23 side chains. We monitored the disruptions in the hydrophilic/hydrophobic balance by counting the hydrogen-bond contacts between the ammonium group of K28 and the oxygen atoms of the peptide backbone. Figure 5 demonstrates that the number of contacts between the K28 side chain and peptide backbone is slightly larger in the WT and E22Q mutant (≈9%) than in the D23N mutant peptide (≈6%). The contacts between the K28 ammonium group and the peptide backbone has a maximum around the E22, D23, and V24 residues, suggesting a combined influence of the electrostatic (between E22–K28 and D23–K28) and hydrophobic (V24–K28) forces act on the K28 side chain. The increased tendency of K28 to make hydrogen bonds with the peptide backbone in the

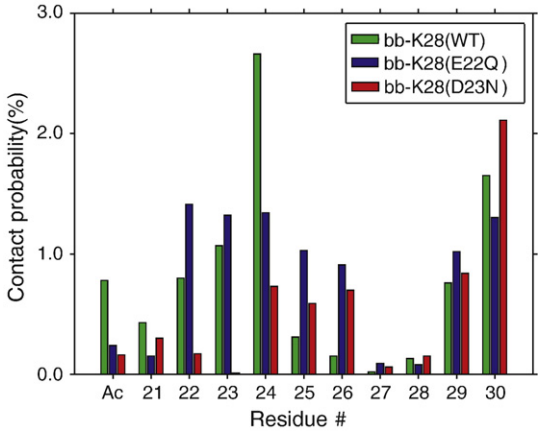


**Fig. 5.** The distance between the $N^{\zeta}$(K28) and the backbone oxygen atoms is used as a measure of interaction between the K28 residue and the peptide backbone. The WT and the E22Q mutant of the $A\beta_{21-30}$ peptide increases the contact probability between the K28 side chain and peptide backbone. The D23N mutation reduces the probability of the K28–backbone contact.

WT and E22Q mutant compared to the D23N mutant provides an explanation for the more favorable p$K_a$ (K28) values for the WT and E22Q peptides (11.2 and 11.1, respectively) and the less favorable p$K_a$ for the D23N mutant peptide (10.4) (see Table 1).

### Sequence determines intrapeptide folding elements

Hydrophobic intramolecular interactions are more probable than intramolecular electrostatic interactions (compare Fig. 4a and b). It is thus reasonable to consider the hydrophobic intrapeptide interactions, i.e. the interaction between the side chain of V24 and the hydrophobic part of the side chain (SC) at position 28, either K in WT, E22Q, and D23N, or A in K28A, as making a dominant contribution to the conformations of the $A\beta_{21–30}$ peptide. The main effect of the intramolecular interactions is a stabilization of the VGSN(24–27) turn. The $C^{\alpha}$(V24)–$C^{\alpha}$(N27) distance was implied to monitor the stability of the VGSN(24–27) turn. Figure 6 shows the free-energy surface projected onto two coordinates, the SC(V24)–SC(K28) and $C^{\alpha}$(V24)–$C^{\alpha}$(N27) distances. A projection of the free energy on the radius of gyration, $R_g$, of the $A\beta_{21–30}$ peptide, and the root-mean-squared displacement (RMSD), from the initial structure used in simulation resulted in a narrow basin with small RMSD values and a wide basin corresponding to large RMSD values (data not shown).

The coordinate $C^{\alpha}$(V24)–$C^{\alpha}$(N27) introduces details to the two superbasins generated by the hydrophobic interaction between the V24 and K28 side chains (see Fig. 4b). The packing of the V24 side chain against the K28 side chain defines two basins, I and II, corresponding to small (5.5–7.0 Å) and larger (8.5–10.0 Å) $C^{\alpha}$(V24)–$C^{\alpha}$(N27) distances, respectively (see Fig. 6). The energy landscape of the $A\beta_{10–35}$ monomer[33] shows a similar bimodal distribution. The mobility around the G25 residue modulates the relative position of the DVG(23–25) and GSN(25–27) fragments. The two fragments of

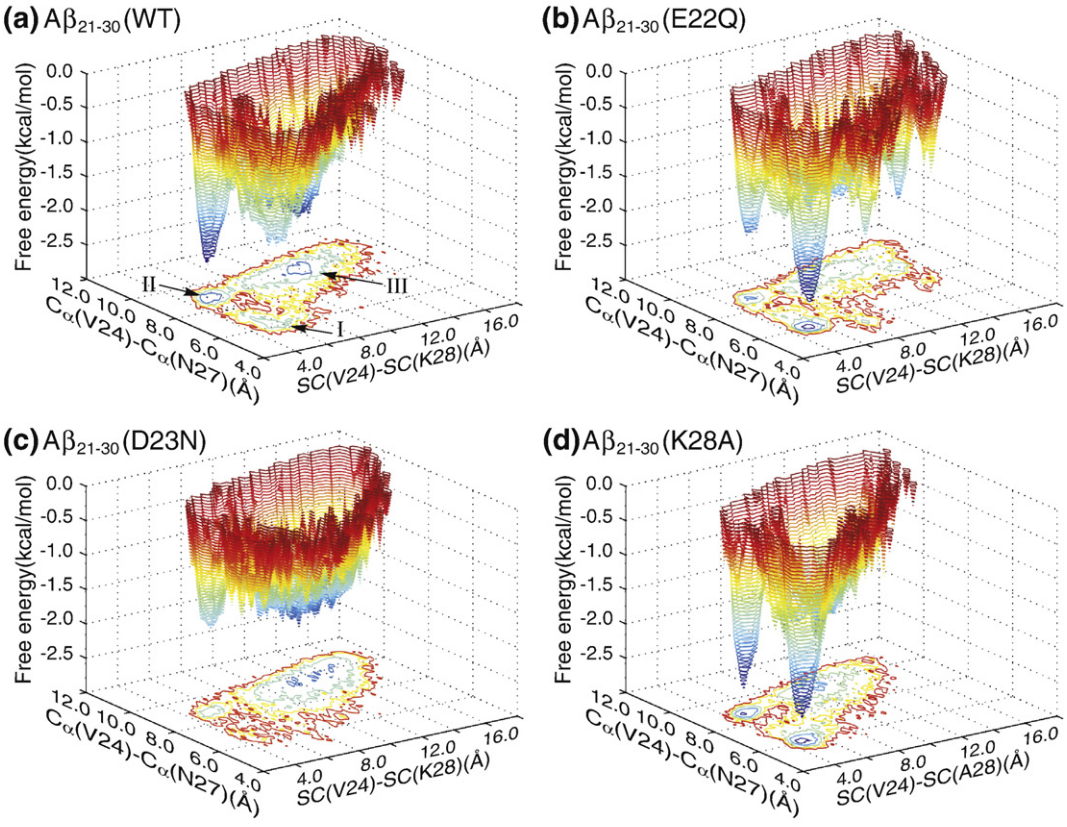


**Fig. 6.** The free-energy hypersurface is projected on the space defined by the distance between the center of mass of the V24 side chain (SC) and the hydrophobic part of the K28 side chain (SC) (*x*-axis) and the distance between the $C^{\alpha}$ atoms of the residues V24 and N27 (*y*-axis). Three hyperbasins are observed in the $C^{\alpha}$(V24)–$C^{\alpha}$(N27) *versus* SC(V24)–SC(K28) projection of the free energy. Basin I is characterized by a compact structure of the VGSN(24–27) segment of the $A\beta_{21–30}$ peptide. The peptide structures in basin II are described by strong hydrophobic interaction between the V24 and K28 side chains, while the ends of the VGSN(24–27) backbone are well separated. The flexible structures of the $A\beta_{21–30}$ peptide have large values of the *x* and *y* coordinates and define Basin III. Basin II is deeper than Basin I in the WT (a), in contrast to the E22Q mutant (b). The D23N mutation makes the decapeptide more flexible, with Basin I almost canceled and Basin III deeper and wider (c). The K28A mutation enhances the hydrophobic interaction between the V24 side chain and the side chain of A28, particularly through cancellation of the electrostatic interaction of K28 side chain with the neighboring water molecules. Basins I and II are deep and equally populated (d).

structures in Basin I define a loop centered on the G25 residue. The maximum stability (minimum free energy) in Basin I is associated with a helix-like structure. The two fragments DVG(23–25) and GSN (25–27), shifted out of plane, are stabilized by hydrogen bonds between the O(D23)–N(K28) and N(G25)–O(K28) atoms, hydrophobic interaction between V24 and K28, and salt bridge between D23 and K28. The DVG(23–25) and GSN(25–27) fragments of the structures in Basin II converge to an "S-like" peptide structure with a bend about the G25 residue. Basin III is broader than Basin I or Basin II, with the two fragments DVG(23–25) and GSN(25–27) oscillating about residue G25.

### *Electrostatic interactions subtly influence the hydrophobic effect*

The electrostatic charge of the 23–27 fragment influences the populations and the positions of the three basins of attraction in the free-energy landscape. Basin II is deeper in the WT (Fig. 6a), with a minimum free energy $E_{min}^{WT,II} = -2.5$ kcal/mol and a population of 13%; Basin I has $E_{min}^{WT,I} = -1.7$ kcal/mol and a population of 9%. The population in Basin III in the WT is 39% with $E_{min}^{WT,III} = -2.2$ kcal/mol. The electrostatic neutralization of the residue in position 22 in the E22Q mutant peptide (Fig. 6b) shifts the balance between Basin I, with $E_{min}^{E22Q,I} = -2.7$ kcal/mol

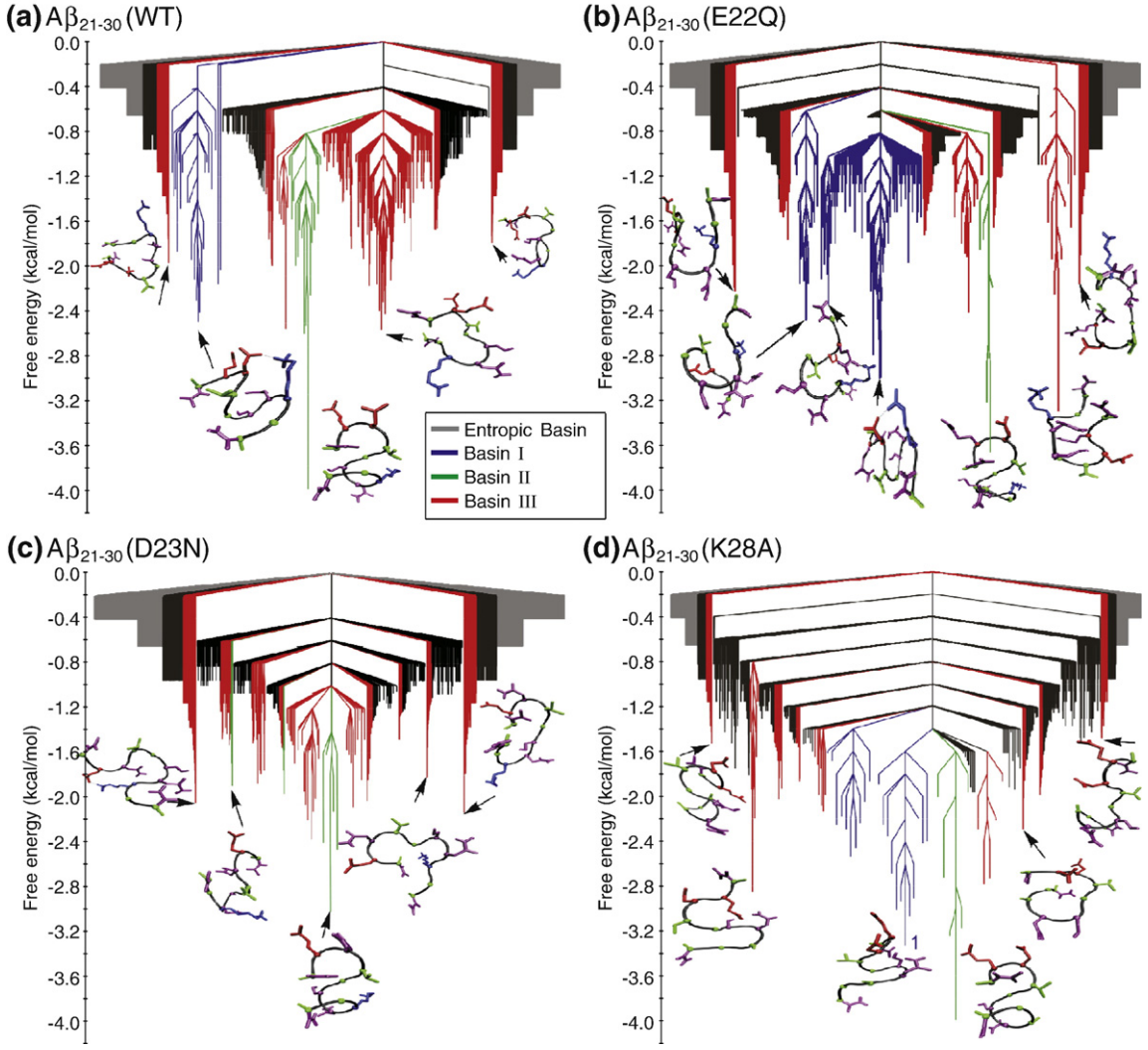


**Fig. 7.** The change in the electrostatic properties of the E22, D23, and K28 side chains in the $A\beta_{21-30}$ peptide induces pronounced changes in the TRDS representations of the peptide free energy. The gray color defines the entropic basin of unstructured configurations. The colors blue, green, and red correspond to structures belonging to Basins I, II, and III, respectively, defined in Fig. 6. The hydrophobic interaction between the V24 and K28 side chains and the electrostatic interactions within the E22–K28 and D23–K28 side-chain pairs generate a rough free-energy surface (multiple on-path basins) in the $A\beta_{21-30}$(WT) peptide (a). The roughness aspect of the free-energy surface is accentuated by the E22Q mutation (b), especially due to the increased electrostatic interaction between D23 and K28 side chains (see Fig. 4a). The D23N mutation (c) does not increase the intrapeptide electrostatic interactions (see Fig. 4a), generating a smoother free-energy surface, with a weak convergence towards a configuration stabilized by the V24–K28 hydrophobic interaction (Basin II in Fig. 6c). The K28A mutation (d) generates a deep and relatively smooth free-energy funnel.

and a population of 19%, and Basin II, with $E_{min}^{E22Q,II} = -2.1$ kcal/mol and a population of only 7%. Basin III is depopulated compared to the WT, with a population of 30% and $E_{min}^{E22Q,III} = -2.0$ kcal/mol. The D23N mutation (Fig. 6c) dramatically reduces the populations in Basin I to only 3% and $E_{min}^{D23N,I} = -1.0$ kcal/mol, and Basin II, with $E_{min}^{D23N,II} = -1.7$ kcal/mol and 5% population. Basins I and II have the same minimum free energy, $E_{min}^{K28A,I} = E_{min}^{K28A,II} = -2.7$ kcal/mol in the K28A mutant (Fig. 6d). However, Basin I is broader, with a population of 23%, than Basin II, with a population of 12%. The population of Basin III in the K28A mutant is 33% and $E_{min}^{K28A,III} = -2.2$ kcal/mol. It should be stressed, as in the case with all MD simulations, that the equilibrium populations of the various basins are not quantitative. However, a number of measures suggest that the relative shifts occurs in the probability of D23–K28 in the Dutch mutant.

### Electrostatic interactions roughen the free-energy surface

In addition to computing the reduced dimensional representation of the free-energy landscape (Fig. 6), we have also obtained a free-energy representation that uses a clustering procedure. Following Krivov and Karplus,[36,37] we used a transition disconnectivity graph (TRDG) representation of the free-energy hypersurface to analyze the combined influence of both electrostatic and hydrophobic intrapeptide interactions on the Aβ$_{21–30}$ peptide behavior. The TRDG representations of the free-energy hypersurface are shown in Fig. 7. The nodes with a free energy higher than −0.6 kcal/mol and separated by barriers lower than 0.6 kcal/mol form an entropic basin (gray color). Table 2 presents the total number of nodes $N$, direct transitions $N_{dir}$, and nodes in the entropic basin associated with the WT, E22Q, D23N, and K28A peptides. Peptide structures corresponding to local minima energies in the TRDG representation form the association with the basins I (blue), II (green), and III (red) defined in Fig. 6. The global minimum free energy of the WT peptide (Fig. 7a), −4.0 kcal/mol, corresponds to an S-like peptide structure of the 23–27 fragment, associated with Basin II. Basin I has a minimum free energy of −2.5 kcal/mol. The WT peptide structure, corresponding to the local minimum free energy in Basin I, has a helix-like shape and is separated by a barrier of 2.5 kcal/mol from the remaining basins in the TRDG representation. Basin III, with a minimum free energy of −2.6 kcal/mol, is separated from Basin II by a barrier of 2.1 kcal/mol.

The energy landscape of the E22Q mutant peptide is characterized by a similar number of nodes, direct transitions, and weight of the entropic basin (see Table 2). Note that Basin I in the E22Q peptide (in blue in Fig. 7b), with a minimum of −3.0 kcal/mol, is broader compared to the WT and is divided into three subbasins, unidentified in the projected representation of the free energy in Fig. 6b. As in the WT free-energy transition disconnectivity graph (FE-TRDG), the local minimum (−3.6 kcal/mol) corresponding to Basin II defines the global minimum in the E22Q FE-TRDG. A barrier of 2.4 kcal/mol separates the local minimum in Basin I from the local minimum in Basin II. The minimum free energy in Basin III of the E22Q FE-TRDG is lower (−3.1 kcal/mol) compared to the corresponding minimum in the WT FE-TRDG (−2.6 kcal/mol). Both the FE-TRDG representations for the WT and E22Q peptides consist of multiple subbasins (Fig. 7a and b).

The D23N mutation (Fig. 7c) produces a flat, funnel-like, free-energy surface. The global minimum free energy, −3.0 kcal/mol, is associated to Basin II and is connected through multiple transitions to Basin III. The increased number of total nodes, direct transitions, and nodes in the entropic basin associated with the D23N peptide (Table 2) indicates diminished intrapeptide interaction energies compared to either WT or E22Q peptides.

The K28A mutation generates the lowest number of nodes, direct transitions, and nodes in the entropic basin of the four peptide variants studied (Table 2). The free-energy surface (Fig. 7d) converges to two associated subbasins, I and II, with free-energy minima of −3.3 and −4.0 kcal/mol, respectively. Figure 7d shows that Basin I is divided into two subbasins, separated by a transition state of −1.2 kcal/mol. Overall, the K28A mutant produces a deep and smooth funnel-like free-energy surface, with reduced transitions into Basin III.

**Table 2.** The number of nodes, $N$, and direct transitions among nodes, $N_{dir}$, associated with the FE-TRDGs of the WT, E22Q, D23N, and K28A peptides. $N_{dir}$ is reduced to $N-1$ using the minimum-cut algorithm[38]

| | $N$ | $N_{dir}$ | $N_S$ |
|---|---|---|---|
| Ac-Aβ$_{21–30}$-NH$_2$(WT) | 1942 | 2799 | 751 |
| Ac-Aβ$_{21–30}$-NH$_2$(E22Q) | 1932 | 2979 | 750 |
| Ac-Aβ$_{21–30}$-NH$_2$(D23N) | 2739 | 3599 | 1195 |
| Ac-Aβ$_{21–30}$-NH$_2$(K28A) | 1515 | 2561 | 527 |

$N_S$ nodes with free energies higher than −0.6 kcal/mol define the entropic basin.

## Discussion and Conclusions

### Balance between hydrophobic and electrostatic forces determines the structures in the dominant basins

In order to elucidate the balance between hydrophobic and electrostatic interactions in the most structured region of Aβ peptide, we determine the energy landscape of Aβ$_{21–30}$(WT) peptide and three mutants, E22Q, D23N, and K28A. Low p$K_a$ shift values of the E22, D23, and K28 residues indicate that the titratable residues are mostly exposed to solvent, suggesting a flexible structure for the Aβ$_{21–30}$ peptide, in accord with recent experimental and computational studies.[26–29] A key finding of this work is that the monomer conformations of the WT

and the mutants are highly mobile and do not adopt a unique structure; that is, there is no folding transition as implied in recent studies.[28,29] Rather, the conformations can be clustered into distinct basins depending on the extent of order in some regions and the probability of being in these clusters. The nature of the energy landscape varies depending on the balance between electrostatic and hydrophobic interactions. The shift in the population among the various basins can be manipulated by point mutations as well as changes in the external conditions.

Rank-ordered intrapeptide hydrophobicity increases with D23N, WT, E22Q, and K28A variants. The differences in the intrapeptide hydrophobic interactions were explained using estimates of the peptide free energy. The free energy was projected on two coordinates associated with the hydrophobic intrapeptide interactions, SC(V24)–SC(K28) and $C^{\alpha}$(V24)–$C^{\alpha}$(N27) distances. A third "coordinate," which measures the effect of successive point mutations of the titratable residues, revealed the role of the intrapeptide electrostatic interactions on the peptide behavior. Hydrophobic interactions cluster the peptide structures into two basins, I and II, with close contacts between the V24 and K28 side chains. Fluctuating structures with solvent-separated SC (V24)–SC(K28) and $C^{\alpha}$(V24)–$C^{\alpha}$(N27) pairs are clustered in Basin III.

The G25 residue divides the DVGSN(23–27) fragment into two halves, DVG(23–25) and GSN(25–27). Mobility of the backbone about the residue G25 causes the difference between Basins I and II. In Basin I, the DVG(23–25) and GSN(25–27) fragments form a loop centered on the G25 residue. The two fragments form a plane at the top edges of Basin I. An out-of-plane shift in the two fragments creates a helix-like motif stabilized by the D23–K28 salt bridge, V24–K28 hydrophobic interaction, and hydrogen bonds between O(D23)–N(K28) and N(G25)–O(K28) atoms. In Basin II, the fragments DVG(23–25) and GSN(25–27) adopt an S-like structure centered on the G25 residue, stabilized by hydrophobic interaction between the V24 and K28 residues. The hydrophobic interaction between the V24 and K28 side chains is modulated by the hydrogen bonds between the ammonium group of K28 with water. The titratable residues E22, D23, and K28 are involved in the intrapeptide interactions. The role of these three residues in the peptide stability was dissected by modifying their electrostatic charge through successive mutations. The hydrogen bonds between the E22 side chain and water control the desolvation around the ED(22–23) fragment. The hydrogen bonds between the K28 side chain and water oppose the V24–K28 hydrophobic interaction. The electrostatic charge of the D23 side chain controls the strength of the D23–K28 salt bridge.

## Point mutations shift the populations between basins

The relative populations of Basins I and II are controlled by the intrapeptide electrostatic and hydrophobic interactions. The E22Q mutation decreases the strength of the electrostatic interactions between the E22 side chain and water, reducing the desolvation barrier for the 22–23 fragment. Consequently, the hydrogen bond between the D23 and K28 side chains is more populated in Basin I. Electrostatic neutralization of D23 in the D23N mutant peptide completely destabilizes Basin I and drastically reduces the depth of Basin II. We conclude that the hydrophobic interaction between V24 and K28 and the hydrogen bonds between the N23 and K28 side chains are destabilized by the solvation of the E22 and K28 side chains in the D23N mutant peptide. Moreover, the E22–K28 contact does not occur frequently enough in the D23N mutant peptide to populate either Basin I or Basin II. The K28A mutation enhances the hydrophobic interactions between the side chains of the V24 and residue A28. Consequently, Basins I and II are observed to be more populated in the K28A mutant.

Rough FE-TRDG surfaces obtained for the $A\beta_{21-30}$ (WT) and E22Q peptides are due to the intrapeptide electrostatic interactions E22–K28 and D23–K28. The more frequent D23–K28 electrostatic contacts in the E22Q peptide increase the population in the associated basin compared to that in the WT. The lack of intrapeptide interaction in the D23N mutant generates a smooth and flat FE-TRDG profile. The lack of favorable intrapeptide electrostatic interactions and enhancement of the hydrophobic interactions in the K28A peptide generate a smoother and deeper funnel-like profile of the FE-TRDG.

## Sequence context determines the energetics of salt bridges

The simulations show that the E22 and D23 side chains make infrequent electrostatic interactions with the K28 side chain in the WT. The energies of the independent E22–K28 and D23–K28 electrostatic interactions revealed differences in the energy range preferences. Low energy levels, between −3.0 to −4.0 and −5.0 to −8.0 kcal/mol, are more populated with the E22–K28 salt bridge in the WT (see the inset in Fig. 3). Energy levels −1.5 to −2.5 and −4.0 to −5.0 kcal/mol are more populated with the D23–K28 salt bridge in the WT (see the inset in Fig. 3). It appears that the D23–K28 salt bridge is more probable in a "wet" environment, while the E22–K28 salt bridge is more probable in a "dry" environment. The lack of clear-cut preference for either of the salt bridges is in accord with recent experiments and implicit-solvent simulations.[26,29]

The E22–K28 and D23–K28 residue pairs were found to equally populate hydrogen-bonded salt-bridge states in the WT, while the water-mediated E22–K28 salt bridges are energetically inactivated and more frequent compared to the water-mediated D23–K28 salt bridges. A one-barrier desolvation process was observed for the E22–K28 pair in the WT peptide, in contrast to the D23N peptide, where a two-barrier desolvation was observed. The D23–K28 pair desolvates in two activated steps in both the WT and E22Q peptides.

### Charge states and shifts in population in aggregation-prone structures explain enhanced fibrillization rate of E22Q over the WT

Recently,[26] it was suggested that the E22–K28 salt bridge and the V24–K28 hydrophobic interaction stabilize the VGSN(24–27) loop in a polar environment, while a hydrophobic environment enhances the D23–K28 salt bridge and destabilizes the V24–K28 hydrophobic contact. Based on this result, it was speculated[26] that the enhanced fibrillization of the E22Q mutant is induced by a reorganization of the VGSN(24–27) loop followed by a destabilization of the V24–K28 hydrophobic contact and an enhancement of the D23–K28 salt bridge. Such an interpretation fails to take into account the desolvation barrier of the D23 side chain associated with the protofibril formation.

Our results indicate that there is a tendency to form a D23–K28 salt bridge in the isolated monomer. The E22 side chain makes infrequent salt bridges with the K28 side chain, especially modulated by an increase in the local hydrophobicity. This explains the lack of the first desolvation barrier of the E22 residue. The tendency of E22 is to be solvated, thus increasing the desolvation barrier around the 22–23 region. The E22Q mutation decreases the desolvation barrier around the 22–23 fragment, enhancing the D23–K28 contacts, while there is no significant change in the V24–K28 hydrophobic contact. In our previous study, we rationalized the enhanced rate of fibril formation of E22Q based on this physical picture.[24] The results in Figs. 6 and 7, which show a greater population of the aggregation-prone structures in Basin I in E22Q compared to the WT, provide a structural explanation for increased fibrillization rate in the Dutch mutant. The structures within Basin I are prone to aggregation based on both the increased hydrophobicity[39] and on the increased probability of the D23–K28 electrostatic interaction. Such a structure is topologically similar with that observed in the amyloid fibril model.[30–32] Thus, mutations (E22Q) that enhance the probability of sampling the aggregation-prone structures will have a higher fibril formation rate.

Given that the intrapeptide hydrophobic interactions in E22Q and the WT are similar, it follows that the differences in the charge states[24] between the two sequences account for the rate of fibril growth in the two sequences. The larger tendency of the D23 side chain to make a salt bridge with the K28 side chain compared to E22 is supported by the depopulation of Basin I in the D23N mutant. The K28A mutation enhances the intrapeptide hydrophobic interactions between the V24 and A28 side chain, increasing the population in Basin I characterized by a helix-like structure. The observed intermediary helix structure in our simulations, stabilized by the V24–K28 hydrophobic interaction, is in accord with the experimental α-helix structures observed during the fibrillization process of the Aβ$_{1–40}$ protein.[40] The V24–K28 hydrophobic interaction is destabilized by removal of the solvation water in the putative protofibril structures of Aβ$_{1–40}$[30,31] and Aβ$_{1–42}$[32] proteins.

The succession of helix structure formation, driven by the V24–K28 hydrophobic interaction and stabilized by intrapeptide hydrogen bonds and electrostatic interaction, followed by destabilization in low-polar environment, is an exemplification of the principle of amyloid self-assembly.[33] According to our results, the V24–K28 hydrophobic interaction decreases the distance between the D23 and K28 residues, enhancing the salt bridge between the side chains. Our work shows that the VGSN(24–27) loop, centered on S26 and stabilized by the V24–K28 hydrophobic interaction, is a transitory state between Basin I and Basin II, with small SC(V24)–SC(K28) and medium-range C$^{\alpha}$(V24)–C$^{\alpha}$(N27) distances. The flexibility about the G25 residue causes the peptide to fall into either Basin I or Basin II. The strong V24–K28 hydrophobic interaction forces the 23–25 and 25–27 fragments to shift out of plane, converging to a helix-like structure. In a reduced polar environment, such as the protofibril structure, the hydrophobic interaction is diminished, and the two fragments 23–25 and 25–27 are stabilized in a planar loop structure by the D23–K28 salt bridge. The increased amyloidogenic propensity of the E22Q mutant compared to the WT peptide appears to be correlated with the increased D23–K28 contact probabilities in the isolated monomer. The V24–K28 hydrophobic interaction, which drives the D23–K28 salt bridges, is "dissolved" in the reduced-solvent environment of the putative protofibril structures.[30–32]

The multitude of the amyloid-like conformers, which contain mainly β-sheets, suggested that formation of amyloid fibrils is a general property of any polypeptide chain,[39] while the sequence of the protein and the environmental conditions control the rate of aggregation.[39] Increased hydrophobicity, through single point mutation, of the residues not involved in the folding core of the acylphosphatase protein was shown to increase the rate of the aggregation process.[41] The hydrophobicity of the side chains of the quadrupole mutated S6 protein was proven to be responsible for the protein aggregation in a tetramer, with the Aβ homologous fragment forming interpeptide antiparallel β-sheets.[42] In contrast, a decrease in the hydrophobicity in the Aβ$_{1–42}$ protein using random screening mutations proved to make the protein less prone to aggregate.[43] The net charge of proteins also influences the aggregation of proteins. The E22Q mutation increases the amyloidogenic propensity of the Aβ$_{10–35}$ protein through a decrease of the desolvation barrier in the proximity of the hydrophobic core LVFFA(17–21).[44] Inverse correlation of the aggregation propensity with the net charge of the muscle acylphosphatase protein was found in denaturing conditions.[45] In contrast, increase in the net charge of the RNase Sa at pH close to the p*I* of the protein was shown to destabilize the protein and further convert it into amyloid.[46] A larger net charge compared to the hydrophobicity of a protein at physiological conditions was suggested to stabilize natively unstructured proteins.[47] We expect that the

increased propensity for α-helix formation compared to propensity for β-sheet should diminish the amyloidogenicity of the protein.[41]

## Computational Methods and Models

### Molecular dynamics

The MD simulations were carried out with the program CHARMM (Chemistry at Harvard Molecular Mechanics)[48] version c29b1 with the PARAM22[49] all-atom force field. The three-point (TIP3P) model for water molecules was used to treat the solvent explicitly.[50] The structures of amyloid fibrils of the $A\beta_{10–35}$ peptide form fibrils with structures that are similar to the ones formed by the $A\beta_{1–40}$ and $A\beta_{1–42}$ proteins. A loop centered on the 21–30 segment brings the β-strands of the central hydrophobic core LVFFA(17–21) and the C-terminal region 29–35[30,51] into contact. This observation and the fact that $A\beta_{10–35}$ peptide forms parallel fibrils justify the use of the $A\beta_{10–35}$ protein as a model for the $A\beta_{1–40}$ and $A\beta_{1–42}$ proteins. The coordinates of the fragment 21–30 of the NMR structure of $A\beta_{10–35}$ protein[20] were used as initial coordinates of the $A\beta_{21–30}$ peptide. The peptide N- and C-termini were acetylated and amidated, respectively. The protonation state of the titratable amino acids was fixed to the expected values at neutral pH in all simulations. The monomer was centered in a truncated octahedron cell that was carved from a larger pre-equilibrated cell of pure water. The size of the primary cell was set according to the minimum-image convention and periodic boundary conditions. To remove steric clashes between atoms, the steepest descent energy minimization algorithm was used for an initial 200 minimization steps, while the peptide atoms were fixed in their positions, achieving an RMS gradient of the potential energy of 0.9 kcal $mol^{-1}$ $Å^{-1}$. The adopted basis Newton–Raphson algorithm was used to further minimize the potential energy of the system until the RMS gradient of the potential energy was less than 0.1 kcal $mol^{-1}$ $Å^{-1}$.

The system was heated linearly to 300 K over 120 ps followed by an equilibration phase involving two steps: the system was equilibrated for 80 ps using NVE molecular dynamics with a leapfrog integrator, followed by an additional 70 ps of NPT molecular dynamics. The pressure was restrained to 1 atm by using a variant of the extended system method, the Langevin piston algorithm.[52] The temperature was restrained to 300 K with the Nosé–Hoover thermostat.[53] In order to prevent any conformational change of the peptide during the heating and equilibration phases, the monomer atoms were restrained about the starting structure using a harmonic restraint with force constant of 20 kcal $mol^{-1}$ $Å^{-2}$. To ensure a gentle equilibration of the water surrounding the monomer, the restraints were gradually reduced to zero. The center of mass of the monomer was constrained to the center of mass of the box of water using the miscellaneous mean field potential commands implemented in CHARMM. The electrostatic interactions were calculated with no truncation using the particle mesh Ewald summation algorithm[54] with a fast Fourier transform grid point spacing of 0.95 Å, and a fifth-degree B-spline interpolation. The width of the Gaussian distribution in real space was 0.32 $Å^{-1}$. In order to reach the convergence of the electrostatic energy, the system was electrostatically neutralized by randomly adding one or two sodium ions in the WT or K28A systems, respectively. The real-space electrostatic and van der Waals interactions were smoothly shifted to zero at 10 Å, using an atom-based cutoff. The list of the nonbonded interactions was truncated at 12 Å. The lengths of the bonds containing hydrogen atoms were fixed with the SHAKE algorithm[55] and the equation of motion was iterated using a time step of 2 fs in the leapfrog integrator.

### Computational estimation of p$K_a$

The effect of electrostatic interactions on a titratable residue at atomistic resolution and short time scales can be obtained from computationally derived p$K_a$ values. The electrostatic influence of the environment and sequence context result in a shift of the p$K_a$ value of the $i$th titratable residue, $\Delta pK_{a,i}$, relative to the $pK_{a,i}^{model}$ value appropriate for the residue isolated in solution.[56] The shift, $\Delta pK_{a,i}$, was estimated by computing the $pK_{a,i}^{protein}$ for a given protein conformation as

$$pK_{a,i}^{protein} = pK_{a,i}^{model} + \Delta pK_{a,i} \quad (1)$$

A negative value of $\Delta pK_a$ corresponds to a relatively favorable deprotonation of the ionizable residue, while a positive value implies a relatively favorable protonation state. The p$K_a$ shift is expressed as[56]

$$\Delta pK_a = -\frac{1}{2.303 k_B T}(\Delta G_{Born} + \Delta G_{back} + \Delta G_{inter}), \quad (2)$$

where $\Delta G_{Born}$ is the difference between the free-energy change associated with growing the distribution of the atomic partial charges in the protein and model structures, respectively. $\Delta G_{back}$ is the background electrostatic free energy of interaction between the charged residue $i$ and the atomic partial charges of the neutralized protein structure relative to the free energy in the model structure. $\Delta G_{inter}$ is the electrostatic free energy of interaction between residue $i$ and all the other $j \neq i$ titratable residues in the protein structure. In order to compute the terms in Eq. (2) one must first characterize the protonation state of each titratable residue in the protein structure. The probability of finding ionizable residue $i$ in a protonated state is[56]

$$\theta_i = \frac{\sum_{\{x\}} x_i A_x}{\sum_{\{x\}} A_x}, \quad (3)$$

where $\sum_{\{x\}}$ denotes a summation over all possible protonation states. For $N$ ionizable residues in a protein structure, there are $2^N$ protonation states. $A_x$ depends on the interaction of titratable residue $i$ with the solvated protein at a given pH. The p$K_a$ of residue $i$ is defined to be the pH at which $\theta_i = 0.5$.

The Macroscopic Electrostatics with Atomic Detail (MEAD)[57,58] package was used to compute the p$K_a$ values of the titratable groups (E22, D23, and K28) in the Aβ$_{21–30}$ peptide structure. The dimension of the grid box was obtained by multiplying the maximum Cartesian dimension of the peptide structure by a factor of 2. A grid size of 1.0 Å was used in the first iteration step. A second "focusing" iteration step used a grid size of 0.25 Å, while the grid box was centered on the ionizable residue. A value of 80 was used as the dielectric constant for aqueous solution, while for the peptide structure, values of both 4 and 20 were used, and the results were compared. The p$K_a$ values were computed for Aβ$_{21–30}$ structures separated by 10 ps in the simulated trajectories.

### Free-energy analysis

The disconnectivity graph representation of the free-energy hypersurface (FEDG)[36] is a complement to the potential energy disconnectivity graph.[59,60] A 2 Å all-atom RMSD of structures separated by at least 10 ps in the simulated trajectories was used as a definition of distinct configurations in a cluster-based algorithm.[37] The free energy is expressed as[36]

$$F_{ij} = -k_B T \ln(N_{ij}), \quad (4)$$

where $k_B$ is the Boltzmann's constant and $T$ is the absolute temperature of the system. $N_{ii}$ represents the number of microstates in basin $i$, while $N_{ij}$ is the number of transitions between basins $i$ and $j$. The free energy of basin $i$ is obtained for $i=j$. The free energy of transition between basins $i$ and $j$ is acquired for $i \neq j$. The number of transitions between basins $i$ and $j$ was maximized using the Gomory–Hu "minimum cut" algorithm.[38]

The free-energy surface was estimated by a similar approach and projected on two coordinates: (1) the distance between the centers of mass of the V24 side chain and the hydrophobic portion of the K28 side chain and (2) the distance between the C$^\alpha$ atoms of the V24 and N27 residues. Here, $i$ and $j$ define the number of structures within 0.2 Å bins along the two coordinates. Two consecutive analyzed structures are separated by 10 ps.

## Acknowledgements

We thank Dr. David Wales and Dr. Mark Miller for stimulating discussions about TRDG. This work was supported by a grant from the National Institutes of Health (R01 GM076688-05).

## References

1. Glenner, G. G. & Wong, C. W. (1984). Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Comm.* **120**, 885–890.
2. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L. & Beyreuther, K. (1985). Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc. Natl. Acad. Sci. USA*, **82**, 4245–4249.
3. Selkoe, D. J. (1994). Alzheimer's disease: a central role for amyloid. *J. Neuropathol. Exp. Neurol.* **53**, 438–447.
4. Lambert, M. P., Barlow, A. K., Chromy, B. A., Edwardsdagger, C., Freed, R., Liosatos, M. *et al.* (1998). Diffusible, nonfibrillar ligands derived from Aβ$_{1–42}$ are potent central nervous system neurotoxins. *Proc. Natl. Acad. Sci. USA*, **95**, 6448–6453.
5. Walsh, D. M., Hartley, D. M., Kusumoto, Y., Fezoui, Y., Condron, M. M., Lomakin, A. *et al.* (1999). Amyloid β-protein fibrillogenesis: structure and biological activity of protofibrillar intermediates. *J. Biol. Chem.* **274**, 25945–25952.
6. Hardy, J. & Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*, **297**, 353–356.
7. Yong, W., Lomakin, A., Kirkitadze, M. D., Teplow, D. B., Chen, S. H. & Benedek, G. B. (2002). Structure determination of micelle-like intermediates in amyloid β-protein fibril assembly by using small angle neutron scattering. *Proc. Natl. Acad. Sci. USA*, **99**, 150–154.
8. Harper, J. D. & Lansbury, P. T., Jr. (1997). Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* **66**, 385–407.
9. Bitan, G., Kirkitadze, M. D., Lomakin, A., Vollers, S. S., Benedek, G. B. & Teplow, D. B. (2003). Amyloid β-protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci. USA*, **100**, 330–335.
10. Massi, F. & Straub, J. E. (2001). Energy landscape theory for Alzheimer's amyloid β-peptide fibril elongation. *Proteins*, **42**, 217–229.
11. Thirumalai, D., Klimov, D. K. & Dima, R. I. (2003). Emerging ideas on the molecular basis of protein and peptide aggregation. *Curr. Opin. Struct. Biol.* **13**, 146–159.
12. Kirschner, D. A., Inouye, H., Duffy, L. K., Sinclair, A., Lind, M. & Selkoe, D. J. (1987). Synthetic peptide homologous to beta protein from Alzheimer disease forms amyloid-like fibrils *in vitro*. *Proc. Natl. Acad. Sci. USA*, **84**, 6953–6957.
13. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**, 125–142.
14. Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C. L. & Beyreuther, K. (1991). Aggregation and secondary structure of synthetic amyloid beta A4 peptides of Alzheimer's disease. *J. Mol. Biol.* **218**, 149–163.
15. Barrow, C. J. & Zagorski, M. G. (1991). Solution structures of β peptide and its constituent fragments: relation to amyloid deposition. *Science*, **253**, 179–182.
16. Sticht, H., Bayer, P., Willbold, D., Dames, S., Hilbich, C., Beyreuther, K. *et al.* (1995). Structure of amyloid A4-(1–40)-peptide of Alzheimer's disease. *Eur. J. Biochem.* **233**, 293–298.
17. Coles, M., Bicknell, W., Watson, A. A., Fairlie, D. P. & Craik, D. J. (1998). Solution structure of amyloid beta-

peptide(1–40) in a water–micelle environment. Is the membrane-spanning domain where we think it is? *Biochemistry*, **37**, 11064–11077.
18. Shao, H., Jao, S., Ma, K. & Zagorski, M. G. (1999). Solution structures of micelle-bound amyloid β-(1–40) and β-(1–42) peptides of Alzheimer's disease. *J. Mol. Biol.* **285**, 755–773.
19. Crescenzi, O., Tomaselli, S., Guerrini, R., Salvadori, S., D'Ursi, A. M., Temussi, P. A. & Picone, D. (2002). Solution structure of the Alzheimer amyloid β-peptide (1–42) in an apolar microenvironment. Similarity with a virus fusion domain. *Eur. J. Biochem.* **269**, 5642–5648.
20. Zhang, S., Iwata, K., Lachenmann, M. J., Peng, J. W., Li, S., Stimson, E. R. *et al.* (2000). The Alzheimer's peptide Aβ adopts a collapsed coil structure in water. *J. Struct. Biol.* **130**, 130–141.
21. Massi, F., Peng, J. W., Lee, J. P. & Straub, J. E. (2001). Simulation study of the structure and dynamics of the Alzheimer's amyloid peptide congener in solution. *Biophys. J.* **80**, 31–44.
22. Zhang S. (1999). *Studies of Beta Amyloid Congeners Directed Toward Understanding the Molecular Mechanism Underlying the Formation Of Amyloid Deposits in Alzheimer's Disease.* PhD thesis, Boston University, Boston, MA.
23. Massi, F. & Straub, J. E. (2003). Structural and dynamical analysis of the hydration of the Alzheimer's β-amyloid peptide. *J. Comput. Chem.* **24**, 143–153.
24. Massi, F., Klimov, D., Thirumalai, D. & Straub, J. E. (2002). Charge states rather than propensity for β-structure determine enhanced fibrillogenesis in wild-type Alzheimer's β-amyloid peptide compared to E22Q Dutch mutant. *Protein Sci.* **11**, 1639–1647.
25. Chaney, M. O., Webster, S. D., Kuo, Y. M. & Roher, A. E. (1998). Molecular modeling of the Aβ1–42 peptide from Alzheimer's disease. *Protein Eng.* **11**, 761–767.
26. Borreguero, J. M., Urbanc, B., Lazo, N. D., Buldyrev, S. V., Teplow, D. B. & Stanley, H. E. (2005). Folding events in the 21–30 region of amyloid β-protein (Aβ) studied in silico. *Proc. Natl. Acad. Sci. USA*, **102**, 6015–6020.
27. Cruz, L., Urbanc, B., Borreguero, J. M., Lazo, N. D., Teplow, D. B. & Stanley, H. E. (2005). Solvent and mutation effects on the nucleation of amyloid β-protein folding. *Proc. Natl. Acad. Sci. USA*, **102**, 18258–18263.
28. Baumketner, A., Bernstein, S. L., Wyttenbach, T., Lazo, N. D., Teplow, D. B., Bowers, M. T. & Shea, J. E. (2006). Structure of the 21–30 fragment of amyloid β-protein. *Protein Sci.* **15**, 1239–1247.
29. Lazo, N. D., Grant, M. A., Condron, M. C., Rigby, A. C. & Teplow, D. B. (2005). On the nucleation of amyloid β-protein monomer folding. *Protein Sci.* **14**, 1581–1596.
30. Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F. & Tycko, R. (2002). A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. USA*, **99**, 16742–16747.
31. Petkova, A. T., Yau, W. M. & Tycko, R. (2006). Experimental constraints on quaternary structure in Alzheimer's β-amyloid fibrils. *Biochemistry*, **45**, 498–512.
32. Lührs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Döbeli, H. *et al.* (2005). 3D structure of Alzheimer's amyloid-A$\beta_{1–42}$ fibrils. *Proc. Natl. Acad. Sci. USA*, **102**, 17342–17347.
33. Tarus, B., Straub, J. E. & Thirumalai, D. (2006). Dynamics of Asp23-Lys28 salt bridge formation in A$\beta_{10–35}$ monomers. *J. Am. Chem. Soc.* **128**, 16159–16168.
34. Hernandez, G. (2005). *Unusual Chemical and Conformational Properties in the Alzheimer's Disease Aβ Peptide.* PhD thesis, Boston University, Boston, MA.
35. Forsyth, W. R., Antosiewicz, J. M. & Robertson, A. D. (2002). Empirical relationships between protein structure and carboxyl p$K_a$ values in proteins. *Proteins Struct. Funct. Genet.* **48**, 388–403.
36. Krivov, S. V. & Karplus, M. (2002). Free energy disconnectivity graphs: application to peptide models. *J. Chem. Phys.* **117**, 10894–10903.
37. Krivov, S. V. & Karplus, M. (2004). Hidden complexity of free energy surfaces for peptide (protein) folding. *Proc. Natl. Acad. Sci. USA*, **101**, 14766–14770.
38. Gomory, R. E. & Hu, T. C. (1961). Multi-terminal network flows. *J. Soc. Ind. Appl. Math.* **9**, 551–570.
39. Chiti, F. & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366.
40. Kirkitadze, M. D., Condron, M. M. & Teplow, D. B. (2001). Identification and characterization of key kinetic intermediates in amyloid β-protein fibrillogenesis. *J. Mol. Biol.* **312**, 1103–1119.
41. Chiti, F., Taddei, N., Baroni, F., Capanni, C., Stefani, M., Ramponi, G. & Dobson, C. M. (2002). Kinetic partitioning of protein folding and aggregation. *Nat. Struct. Biol.* **9**, 137–143.
42. Otzen, D. E., Kristensen, O. & Oliveberg, M. (2000). Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. *Proc. Natl. Acad. Sci. USA*, **97**, 9907–9912.
43. Wurth, C., Guimard, N. K. & Hecht, M. H. (2002). Mutations that reduce aggregation of the Alzheimer's Aβ42 peptide: an unbiased search for the sequence determinants of Aβ amyloidogenesis. *J. Mol. Biol.* **319**, 1279–1290.
44. Massi, F. & Straub, J. E. (2001). Probing the origins of increased activity of the E22Q "Dutch" mutant Alzheimer's β-amyloid peptide. *Biophys. J.* **81**, 697–709.
45. Chiti, F., Calamai, M., Taddei, N., Stefani, M., Ramponi, G. & Dobson, C. M. (2002). Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases. *Proc. Natl. Acad. Sci. USA*, **99**, 16419–16426.
46. Schmittschmitt, J. P. & Scholtz, M. J. (2003). The role of protein stability, solubility, and net charge in amyloid fibril formation. *Protein Sci.* **12**, 2374–2378.
47. Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **11**, 739–756.
48. Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaninathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Phys.* **4**, 187–217.
49. MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J. *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
50. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Physiol.* **79**, 926–935.
51. Antzutkin, O. N., Leapman, R. D., Balbach, J. J. & Tycko, R. (2002). Supramolecular structural constraints on Alzheimer's beta-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance. *Biochemistry*, **41**, 15436–15450.

52. Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. R. (1995). Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621.
53. Hoover, W. G. (1985). Canonical dynamics: equilibrium phase–space distributions. *Phys. Rev. A*, **31**, 1695–1697.
54. Darden, T., York, D. & Pedersen, L. (1993). Particle mesh Ewald: an *N-log*(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
55. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical-integration of Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341.
56. Bashford, D. & Karplus, M. (1990). p$K_a$'s of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, **29**, 10219–10225.
57. Bashford, D. & Gerwert, K. (1992). Electrostatic calculations of the p$K_a$ values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* **224**, 473–486.
58. Bashford, D. (1997). An object-oriented programming suite for electrostatic effects in biological molecules. In (Ishikawa, Y., Oldehoeft, R. R., Reynders, J. V. W. & Tholburn, M., eds), *Scientific Computing in Object-Oriented Parallel Environments. Lecture Notes in Computer Science ISCOPE97*, vol. 1343, pp. 233, Springer, Berlin.
59. Wales, D. J. & Scheraga, H. A. (1999). Global optimization of clusters, crystals, and biomolecules. *Science*, **285**, 1368–1372.
60. Becker, O. M. & Karplus, M. (1997). The topology of multidimensional potential energy surfaces: theory and application to peptide structure and kinetics. *J. Chem. Phys.* **106**, 1495–1517.