Does amino acid sequence determine the properties of A β dimer?

Christopher Lockhart, Seongwon Kim, Rashmi Kumar, and Dmitri K. Klimov^{a)}

School of Systems Biology, George Mason University, Manassas, Virginia 20110, USA

(Received 15 April 2011; accepted 22 June 2011; published online 21 July 2011)

The effect of random reshuffling of amino acids on the properties of dimers formed by A β peptides is studied using replica exchange molecular dynamics and united atom implicit solvent model. We show that thermodynamics of dimer assembly and the dimer globule-like state are not affected by sequence permutation. Furthermore, sequence reshuffling does not change the distributions of non-local interactions and, to a large extent, amino acids in the dimer volume. To rationalize these results, we demonstrate that Gaussian statistics applies surprisingly well to the end-to-end distances of the peptides in the dimer implying that non-bonded interactions between distant along the chain amino acids are effectively screened. This observation suggests that peptides in the dimer behave as ideal chains in polymer melt, in which amino acids lose their “identity” and therefore the memory of sequence position. As a result large-scale properties of the dimer become universal or sequence independent. Comparison of our simulations with the prior theoretical studies and their implications for experiments are discussed. © 2011 American Institute of Physics. [doi:10.1063/1.3610427]

I. INTRODUCTION

A propensity to form ordered amyloid assemblies appears to be a generic property of polypeptide chains and has been observed experimentally for a number of sequences with low homology.¹ Amyloid fibrils are stabilized by a network of backbone hydrogen bonds as well as hydrophobic and electrostatic interactions between side chains.² As a self-assembly process amyloidogenesis has many diverse implications. Amyloid structures possess significant technological potential due to their use in producing conductive nanowires, peptide-metal composites, and a variety of nanoscale objects of different geometry, from nanospheres to nanotubes.^{3,4} From the biomedical perspective, amyloidogenesis is linked to multiple disorders, including Alzheimer's, Parkinson's, and Creutzfeldt-Jakob diseases,⁵ although some amyloid structures have important biological function.⁶

Amyloid assembly involves multiple conformational transitions being initiated with the oligomerization of monomeric peptides and leading to the appearance of fibrils as final products.^{7–9} According to amyloid cascade hypothesis, Alzheimer's disease is related to the aggregation of A β peptides,¹⁰ of which the 40-residue species, A β _{1–40}, are most abundant.¹ At micromolar concentration and normal physiological conditions, these peptides exist as monomers and oligomers up to the fourth (tetrameric) order.^{11–13} Importantly, biomedical studies have indicated that A β oligomers appear to be the primary cytotoxic species in Alzheimer's disease.¹⁴ Furthermore, synaptic structure and function can be impaired even by the smallest A β oligomers, dimers.¹⁵ These observations implicate a crucial role of A β oligomers in amyloidogenesis, which are not only on-pathway intermediates in amyloid fibril assembly, but the cytotoxic agents.

According to multiple experimental studies A β monomers in aqueous solution adopt mostly random coil structure.^{16,17} The final products of amyloidogenic assembly, A β fibrils, are characterized by extensive polymorphic β -sheet structure.^{18,19} However, the structures of A β oligomers remain elusive for experimental investigation due to rapid interconversion of oligomeric species and lack of well-defined structure.¹³ Under these circumstances molecular dynamics (MD) simulations performed at all-atom resolution appear to be a viable alternative.^{20,21} Indeed, numerous computational studies have explored the conformational properties of A β monomers or its fragments.^{22–25} Consistent with the experiments A β _{1–40} conformational ensemble is generally random, although simulations have identified several flickering structured regions, including short β -strands, helices, and β -turns.^{22,25}

Recently, several groups have initiated MD simulations of A β oligomers.^{26–31} In line with the experiments¹³ these studies have shown that aggregation triggers extensive conformational changes in A β peptides. For example, in our previous MD investigations we have reported that the conformational ensembles of A β dimers and tetramers are very similar but sharply distinct from those sampled by the monomers.³⁰ Based on the analysis of atomic density distribution and free energy landscapes, we have argued that A β oligomer resembles a polymer globule in poor solvent.²⁹ In particular, the oligomer was shown to consist of two volume regions – the core with approximately flat density profile and the surface layer with rapidly decreasing density. However, many implications of globular state of A β oligomers have not been explored. For example, statistical physics of polymers postulates that the core of large globule resembles polymer melt and, consequently, long polymer segments have the behavior of ideal (freely jointed) chains. When applied to polypeptide chains, these theoretical results would suggest that particular order of amino acids in A β peptide may be irrelevant for some structural properties of the oligomer.

^{a)} Author to whom correspondence should be addressed. Electronic mail: dklimov@gmu.edu.

To investigate the importance of amino acid sequence and its composition for the properties of A β dimers, we use exhaustive replica exchange molecular dynamics (REMD) and implicit solvent model. We chose to study a heterodimer (HT), in which one chain is a wild-type A β peptide and the other is a mutant A β sequence, in which all amino acids are randomly reshuffled (Fig. 1(a)). The properties of heterodimer are compared with those of A β homodimer (HM) studied earlier, in which both chains are wild-type A β peptides.^{29,30} We show that large-scale energetic and structural properties are largely independent on the specific amino acid sequence, which include the temperature of dimer formation, the number density spatial distributions, and the frequency of interactions formed by amino acid types. To explain these findings, we demonstrate that the peptides in the dimer behave as ideal chains in polymer melt, in which non-bonded interactions between distant along the chain monomers are screened. As a result large-scale properties of the dimer become universal or sequence independent. As expected universality does not hold for local structural properties, such as secondary structure, which are sequence dependent. The paper is concluded with the discussion of the relation of our results to prior theoretical studies and A β amyloidogenesis.

II. METHODS

A. Simulation model

To perform simulations of A β dimer we used CHARMM MD program,³² united atom force field CHARMM19, and the SASA implicit solvent model.³³ Full description of the CHARMM19+SASA force field and its testing were reported in our previous studies.^{29,30,34,35} In particular, we have shown that the CHARMM19+SASA force field accurately reproduces the experimental distribution of chemical shifts for C $_{\alpha}$ and C $_{\beta}$ atoms in A β monomers.^{17,35} Furthermore, our simulations of A β monomers, dimers, and tetramers utilizing this force field yielded the estimates of β -structure content in good agreement with the experimental measurements.^{13,30}

For this study, we have selected A β_{10-40} peptide, the N-terminal truncated fragment of the full-length A β_{1-40} (Fig. 1(a)). According to the experiments¹² and simulations³⁶ A β_{1-40} and A β_{10-40} oligomers share significant similarities in terms of their size distributions and structure. In addition, it has been recently shown that as in the amyloid fibril the N-terminal fragment of A β_{1-40} monomer is involved in few long-range interactions and does not form ordered secondary structure.^{18,25} Consequently, we use A β_{10-40} as a model of the full-length A β_{1-40} peptide.

We study a heterodimer, in which one chain is a wild-type A β_{10-40} peptide and the other is a mutant A β_{10-40} sequence, in which all amino acids are randomly reshuffled (Fig. 1). The random distribution of A β amino acids in the reshuffled sequence is illustrated in Fig. 1(b). The peptides in the heterodimer were subject to spherical boundary condition with the radius $R_s = 90$ Å and the force constant $k_s = 10$ kcal/(mol Å²). The peptide concentration in the sphere was, therefore, on the order of approximately mM. To compare the structural ensembles of the heterodimer with those of

FIG. 1. (a) The sequence of wild-type (WT) A β_{10-40} peptide and its random (RND) mutant, in which all amino acids are randomly reshuffled. Note that although the sequence of amino acids in the WT and RND is different, their amino acid compositions are identical. Keeping the composition unchanged allows us to probe the role of sequence in dimer assembly. Amino acids from the N- and C-terminals are shown by large font. (b) Mutation map shows new positions of the WT amino acids in RND sequence. Reshuffling places amino acids at new randomly selected sequence positions. (c) Typical structure of the heterodimer composed of the WT (in purple) and RND (in cyan) peptides sampled at 360 K. The structure is visualized using Chimera (see Ref. 50)

the wild-type A β homodimer, we use the results of our previous simulations.^{29,30}

B. Replica exchange simulations

To sample conformational space we used replica exchange molecular dynamics.³⁷ The REMD implementation is described in our previous studies.^{35,38} Briefly, 24 replicas were distributed linearly in the temperature range from 300 to 530 K with the increment of 10 K. This temperature range

spans the spectrum of peptide conformational states, from aggregated to completely dissociated. Small temperature increment provides significant overlap of energy distributions between neighboring replicas. The exchanges were attempted between all neighboring replicas with the average acceptance rate of 54%. Between replica exchanges the system was simulated for 80 ps using NVT underdamped Langevin dynamics with the damping coefficient $\gamma = 0.15 \text{ ps}^{-1}$ and the integration step of 2 fs. Seven REMD trajectories were produced resulting in the cumulative simulation time of 134 μs or 5.6 μs per replica. Because the initial parts of REMD trajectories are not equilibrated and must be excluded from thermodynamic analysis, the cumulative equilibrium simulation time was reduced to $\tau_{sim} \approx 107 \mu\text{s}$ or 4.5 μs per replica. The REMD trajectories were started with random distributions of peptides in the sphere.

C. Computation of structural probes

To probe the interactions between the peptides, we computed the number of side chain contacts. A side chain contact is formed, if the distance between the centers of mass of side chains is less than 6.5 Å. If the amino acids in contact are apolar, the contact is considered hydrophobic. Two peptides are considered aggregated, if there is at least one hydrophobic side chain contact between them. The radial distribution function $G(r)$ reports the atom number density at the distance r from the dimer center of mass. To compute $G(r)$, we considered non-hydrogen atoms from all peptides. The histogram of atom distances r was constructed using the bin-size of 0.5 Å. The pair correlation function $g(r)$ measures the heavy atom number density at the distance r from the reference atom in the dimer. To compute $g(r)$, we constructed the histogram of distances r using all heavy atoms as reference points. Backbone hydrogen bonds (HBs) between NH and CO groups were assigned according to Kabsch and Sander.³⁹ Secondary structure in the peptides was computed using the distribution of (ϕ, ψ) backbone dihedral angles. Specific definitions of β -strand and helix states can be found in our previous studies.³⁵ The analysis of conformational ensembles sampled by the peptides in the heterodimer was performed using clustering technique, which is described in supplementary material.

Throughout the paper angular brackets $\langle \dots \rangle$ imply thermodynamic averages. The distributions of states produced by REMD were analyzed using multiple histogram method.⁴⁰ The convergence of REMD simulations and error analysis are presented in supplementary material.⁴¹

III. RESULTS

A. Thermodynamics of heterodimer assembly

Using REMD, we have computed the equilibrium conformational properties of the heterodimer and its assembly thermodynamics. The heterodimer thermodynamic quantities are reported at the temperature of 360 K, at which A β peptide locks into fibril-like state during fibril growth.^{34,38} The

impact of random reshuffling of A β sequence is assessed in two ways – (i) by comparing the conformational ensembles of the wild-type (WT) A β and random (RND) peptides in the heterodimer, and (ii) by comparing the properties of the heterodimer with those of the homodimer formed by two identical A β peptides and reported by us earlier.²⁹ Importantly, we observed no qualitative differences in the properties of the heterodimers and homodimers, if the temperature is reduced to 330 K.

The thermodynamics of heterodimer assembly is presented in Fig. 2. The number of interpeptide side chain contacts $\langle C_d(\text{HT}) \rangle$ shown in Fig. 2(a) monotonically increases with the decrease in temperature implying the formation of heterodimer. The same quantity $\langle C_d(\text{HM}) \rangle$ computed for the homodimer shows that the numbers of interpeptide contacts in both dimers are very similar (Fig. 2(a)). For example, at 360 K $\langle C_d(\text{HT}) \rangle \approx 30.7$, which differs from $\langle C_d(\text{HM}) \rangle \approx 30.0$ only by 2%. According to Sec. II, a dimer is operationally assumed formed, if at least one hydrophobic side chain contact is established between the peptides. Then the probability P_d of forming a heterodimer is ≈ 0.98 at 360 K. Same probability P_d has been obtained for the homodimer earlier.²⁹ Fig. 2(a) also displays the numbers of intrapeptide HBs, $\langle N_{ihb}(\text{WT}) \rangle$ and $\langle N_{ihb}(\text{RND}) \rangle$, obtained for the WT and RND peptides in the heterodimer. Over the entire temperature range $\langle N_{ihb}(\text{WT}) \rangle$ slightly exceeds $\langle N_{ihb}(\text{RND}) \rangle$ (at 360 K their values are 8.5 and 7.3, i.e., the difference amounts to 14%). The dimensions of the peptides in the heterodimer are probed by the radius of gyration $\langle R_g \rangle$ (not shown). We found that $\langle R_g \rangle$ for the WT and RND peptides almost coincide being equal to 14.3 and 14.2 Å, respectively ($\approx 1\%$ difference). Therefore, the large-scale structural properties of the WT and RND peptides appear to be similar.

The aggregation free energy landscape is analyzed using the free energy of the heterodimer as a function of the number of interpeptide hydrophobic contacts, $F(C_h)$ (Fig. 2(b)). It follows from this figure that the free energy of heterodimer formation ΔF_{A-D} is $\approx -3.9\text{RT}$ at 360 K (see caption for details). Fig. 2(b) also shows the profile $F(C_h)$ for the homodimer,²⁹ which is almost identical to that for the heterodimer. For example, at 360 K the free energies ΔF_{A-D} differ by merely 0.2RT or 5%. Both free energy profiles $F(C_h)$ in Fig. 2(b) reveal no metastable states or barriers. (Similar results were obtained if the free energy of the heterodimer is considered as a function of the number of interpeptide contacts, C_d .) Taking into account the barrierless free energy profile $F(C_h)$, we have argued in a previous study that dimer formation is described by continuous transition.²⁹ If this is the case, the heterodimer free energy $F(T)$ should follow quadratic dependence on temperature at $T < T_o$, i.e., $F(T) \sim -(T - T_o)^2$, where T_o is the temperature at which dimer formation is completed.⁴² The inset to Fig. 2(b) shows that the heterodimer $F(T)$ is indeed well described by the quadratic function with $T_o = 369 \text{ K}$. For comparison, the temperature of homodimer formation is only 2 K higher ($T_o = 371 \text{ K}$, see Ref. 29). The results presented above suggest that random reshuffling of A β sequence has little impact on the dimer thermodynamics.

FIG. 2. (a) The assembly of heterodimer (HT, thick black line)²⁹ is compared using the number of interpeptide side chain contacts (C_d) as a function of temperature. Intraepitope interactions are probed by the number of intraepitope hydrogen bonds, $\langle N_{ihb} \rangle$, as a function of temperature, in the WT (dashed black line) and RND (dashed grey line) peptides of the heterodimer. (b) Free energy $F(C_h)$ as a function of the number of interpeptide hydrophobic contacts, C_h , computed for the heterodimer (data in black) and homodimer (data in grey) at 360 K. The free energy of dimer formation is $\Delta F_{A-D} = F_A - F_D$, where F_A and $F_D = 0$ are the free energies of the associated (A) and dissociated (D, $C_h = 0$) states. F_A is obtained by integrating over the states, for which $F(C_h) \leq F_{min} + 1.0RT$, where F_{min} is the minimum in $F(C_h)$. Inset: Temperature dependence of the heterodimer free energy $F(T)$ (in black). (Note that $F(T)$ includes the contributions from interpeptide and intraepitope interactions and conformational entropy, but omits the ideal gas term associated with translational entropy.) Quadratic fitting function $F(T) \simeq -\alpha(T - T_o)^2$ with $\alpha = -0.003$ kcal/(mol K²), $T_o = 369$ K is represented by black continuous line. For comparison, $F(T)$ for the homodimer (see Ref. 29) is shown by grey circles. Maximum values of $F(T)$ are set to zero. The figure demonstrates that random reshuffling of amino acids has little impact on the thermodynamics of dimer formation.

B. Density distributions and interactions in heterodimer

We next investigate the density distribution in the heterodimer. Fig. 3(a) shows the radial distribution functions for the number density $G(r)$ computed at 360 K for the heterodimer and homodimers.²⁹ Both density profiles are nearly identical and as for the homodimer there are two volume regions in the heterodimer – the core with relatively constant density and the surface layer, in which the density rapidly decays. If the core region is defined from the condition $G(R_c) = 0.7G(0)$, then its radius for the heterodimer is $R_c \approx 11.0$ Å. The same value of R_c was determined previously for the homodimer.²⁹

The inset to Fig. 3(a) shows the number density pair correlation function $g(r)$ computed for heavy atoms in the heterodimer. The pronounced peaks at $r = 1$ and 2 Å are associated with the atoms adjacent along the covalently bonded structure. Indeed, if the computation of $g(r)$ is restricted to the pairs of atoms from non-bonded amino acids (i.e., $|i - j| > 1$, where i and j are amino acid indices), then the two peaks in $g(r)$ disappear (Fig. 3(a)). Therefore, starting with the third maximum at $r \approx 5$ Å the function $g(r)$ describes non-bonded interactions. Applying exponential fit to the tail in $g(r)$ at $r > 5$ Å, one can estimate the correlation radius λ for density fluctuations (caption to Fig. 3(a)). We found that for the heterodimer $\lambda = 7.3$ Å, which is very close to the homodimer value ($\lambda = 7.2$ Å, plot of $g(r)$ is not shown). In the previous study we argued that A β homodimer resembles polymer globule in poor solvent.²⁹ Given almost identical atom density distributions in the homodimer and heterodimers and the fact that the correlation radius λ is significantly smaller than the dimer size ($2R_c$), heterodimer has globular characteristics (see Sec. IV).

It is instructive to consider the statistics of end-to-end distances for the heterodimer WT and RND peptides. Fig. 3(b) shows the WT probability distribution $P(r_{1N})$ of the distances r_{1N} between the N- and C-terminals (namely, C $_{\alpha}$ atoms of Tyr10 and Val40). According to statistical mechanics of polymers, the end-to-end distance r of a long ideal chain acquires a Gaussian distribution, i.e.,

$$P(r) \simeq 4\pi \left[3/2\pi a^2 \right]^{\frac{3}{2}} r^2 \exp \left(-\frac{3r^2}{2a^2} \right). \quad (1)$$

Fig. 3(b) shows that the probability distribution Eq. (1) with $a = 17.4$ Å provides a strikingly good fit to $P(r_{1N})$. If short peptide segments between amino acids i and j are considered, then the distribution of end-to-end distances r_{ij} clearly becomes non-Gaussian. For example, Fig. S3 in supplementary material presents the distribution $P(r_{ij})$ for the WT amino acids ten monomers apart ($|i - j| = 10$). The fit based on Eq. (1) fails to describe this distribution. However, according to the inset to Fig. 3(b) the end-to-end distances $\langle r_{ij}^2 \rangle$ for the longer segments ($14 < |i - j| \leq N = 31$, N is the number of amino acids in the peptide) scale linearly with $|i - j|$. The linear scaling is an indication of applicability of Gaussian statistics in this range of $|i - j|$ (see Sec. IV). Similar results were obtained for the RND peptide. For example, its end-to-end distance is well described by Eq. (1) with $a = 16.5$ Å and

the linear fit to $\langle r_{ij}^2 \rangle$ holds at $23 < |i - j| < N = 31$ (Fig. 3(b)). From the theoretical perspective, Gaussian statistics of end-to-end distance suggests cancellation of excluded volume interactions between monomers, a somewhat surprising conclusion further examined in Sec. IV.

Distribution of amino acids in the dimer volume has been probed by computing the probabilities P_c for amino acid to occur in the core. Because the peptide sequences in the heterodimer are different, we redefine P_c as the probability $P_c(k)$ for an amino acid type k to occur in the core. Because mutation only reshuffles residues without introducing new or deleting existing amino acids, $P_c(k)$ allows us to compare spatial amino acid distributions for the WT and RND. In all, 15 amino acid types are present in each of the peptides and their $P_c(k)$ are displayed in Fig. 3(c). It is seen that the probabilities of occurrence of amino acids in the core are similar for both peptides (the correlation coefficient between $P_c(k)$ is 0.6). Furthermore, the average absolute difference in $P_c(k)$ for all amino acid types is 0.06 or 10%.

To further analyze the impact of amino acid sequence on dimer properties, we considered inter- and intrapeptide side chain contacts. As in Fig. 3(c), we examined the interactions formed by amino acid types. Fig. 4(a) shows the numbers of interpeptide side chain contacts $\langle C_d(k; \text{HT}) \rangle$ formed by the heterodimer amino acid types k . It is clear that $\langle C_d(k; \text{HT}) \rangle$ is very similar to the same quantity $\langle C_d(k; \text{HM}) \rangle$ obtained for the homodimer (the correlation coefficient computed for $\langle C_d(k; \text{HT}) \rangle$ and $\langle C_d(k; \text{HM}) \rangle$ is 0.82). To analyze intrapeptide interactions, we divided them into local (amino acids i and j are less than six positions apart, $|i - j| < 6$) and non-local ($|i - j| \geq 6$). Fig. 4(b) demonstrates that there is little difference in the distributions of non-local contacts for the WT and RND peptides, $\langle C_{nl}(k; \text{WT}) \rangle$ and $\langle C_{nl}(k; \text{RND}) \rangle$ (the respective correlation coefficient is 0.87). In contrast, the WT and RND distributions of local contacts, $\langle C_l(k; \text{WT}) \rangle$ and $\langle C_l(k; \text{RND}) \rangle$, shown in Fig. 4(c) are strikingly different. For example, for $k = \text{ASP}$ or TYR the numbers of contacts in the WT and RND peptides differ by about two-fold and the correlation coefficient for $\langle C_l(k; \text{WT}) \rangle$ and $\langle C_l(k; \text{RND}) \rangle$ is approximately zero. Therefore, specific connectivity of amino acids plays minor role in the distributions of interpeptide and non-local intrapeptide contacts. This observation is consistent with the Gaussian statistics of end-to-end distances (see Sec. IV).

C. Secondary structure in heterodimer

It follows from Fig. 4(c) that sequence does determine the formation of local interactions. Then, one may expect to observe differences in the secondary structure formed in the WT and RND peptides. Lower panel of Fig. 5(a) displays the fractions of β -strand structure in both peptides in the heterodimer, $\langle S(i; \text{WT}) \rangle$ and $\langle S(i; \text{RND}) \rangle$. Although the average fractions of β structure are very similar ($\langle S(\text{WT}) \rangle = 0.40$ and $\langle S(\text{RND}) \rangle = 0.42$), their distributions along the peptide sequences are very different. Same conclusion follows from the upper panel in Fig. 5(a) showing the distribution of helix structure, $\langle H(i; \text{WT}) \rangle$ and $\langle H(i; \text{RND}) \rangle$. In contrast to the WT, RND β -content is elevated in the C-terminal (0.45 vs

FIG. 3. (a) Normalized radial distribution function for the number density $G(r)$ in the heterodimer (in black) and homodimer (in grey). The distance r is measured from the dimer center of mass. $G(0)$ represents the number density in the center of the heterodimer or homodimer. Inset: the number density pair correlation function $g(r)$ (circles) computed for the heterodimer. Correlation radius λ is estimated by applying the exponential fit $g(r) = g_0 \exp(-(r - r_0)/\lambda)$ at $r > 5$ Å (continuous line). The fitting parameters are $g_0 = 0.03 \text{ Å}^{-3}$, $r_0 = 5.3$ Å, and $\lambda = 7.3$ Å. The pair correlation function $g(r)$ computed excluding adjacent amino acids is shown by squares. For comparison, the parameters of the exponential fit to the homodimer $g(r)$ are very similar ($g_0 = 0.03 \text{ Å}^{-3}$, $r_0 = 5.2$ Å, and $\lambda = 7.2$ Å). (b) Probability distribution $P(r_{1N})$ of the distances r_{1N} between the N- and C-terminals of the heterodimer WT peptide. Gaussian function (Eq. (1)) shown by the black curve provides an excellent fit to $P(r_{1N})$. Inset: the peptide segment end-to-end distance $\langle r_{ij}^2 \rangle$ as a function of the number of amino acids in the segment $|i - j|$. Filled and open circles are for the WT and RND peptides. Linear fits $\langle r_{ij}^2 \rangle = l^2 |i - j|$ are shown by thick lines ($l = 1.45$ Å and 2.19 Å for the WT and RND peptides, respectively). Panels (a, b) suggest that the heterodimer has globule-like properties. (c) Probability $P_c(k)$ for amino acid type k to occur in the dimer core: WT peptide (filled circles), RND peptide (open circles). Similarities in $P_c(k)$ for the two peptides suggest that spatial distribution of amino acids in the dimer is not affected by sequence reshuffling. All plots are obtained at 360 K.

FIG. 4. (a) The number of interpeptide side chain contacts $\langle C_d(k) \rangle$ formed by amino acids of the type k : heterodimer (HT, in black), homodimer (HM, in grey). (b,c) The numbers of non-local $\langle C_{nl}(k) \rangle$ (b) and local $\langle C_l(k) \rangle$ (c) intrapeptide contacts formed by amino acids of the type k in the WT (filled circles) and RND (open circles) peptides. The figure demonstrates that in contrast to local interactions amino acid sequence does not affect interpeptide and non-local intrapeptide interactions. All plots are obtained at 360 K.

0.38) and the RND helix fraction is suppressed in the N-terminal (0.13 vs 0.26). Fig. 5(b) shows the propensities to form β -strand and helix structure by the amino acid types k . Some similarities are observed between the propensities to form β -structure by the same amino acid types in the WT and RND, $\langle S(k; \text{WT}) \rangle$ and $\langle S(k; \text{RND}) \rangle$ (the correlation coefficient is 0.70). However, little correlation is seen in the propensities to form helix structure by the same amino acid types, $\langle H(k; \text{WT}) \rangle$ and $\langle H(k; \text{RND}) \rangle$ (the correlation coefficient decreases to 0.33). This result is not unexpected because assembly of helical structure is intrinsically cooperative and involves at least four consecutive residues.

The analysis of conformational ensembles formed by the WT and RND peptides in the heterodimer is reported in supplementary material.⁴¹ We found that RND has fewer conformational clusters and its distribution of secondary structure is sharply distinct from that seen in the WT. On the other hand, there is little variation in the distribution of interpeptide interactions between the WT and RND clusters. These findings are consistent with the data presented above.

IV. DISCUSSION

This study explores the effect of amino acid sequence on the properties of dimers formed by A β peptides. To this end, we have designed a heterodimer in which one peptide retains wild-type sequence, whereas the sequence of the other is randomly reshuffled (Fig. 1). By comparing the properties of the heterodimer and the homodimer composed of WT A β peptides and of the WT and RND peptides in the heterodimer, the following conclusions can be drawn.

Mutation changes the free energy of dimer formation ΔF_{A-D} by only 0.2RT (or 5%) and does not affect qualitatively the dimer free energy $F(C_h)$ computed as a function of the number of interpeptide hydrophobic contacts C_h , which remains barrierless (Fig. 2(b)). This figure also suggests that the mutation does not change the temperature of dimer formation T_o . The free energy $F(C_h)$ and the temperature dependence of the dimer free energy $F(T)$ are consistent with the dimer assembly as continuous transition (Fig. 2(b)).²⁹ As known from polymer physics continuous transition also describes the formation of globule for flexible chains.⁴³

We have further shown that the atom number density distributions $G(r)$ for the heterodimer and homodimers are almost identical (Fig. 3(a)) implying that reshuffling does not alter the dimer radius R_c ($= 11$ Å). Furthermore, there is no difference in the correlation radius λ for the heterodimers and homodimers (Fig. 3(a)). It is important that, because $\lambda/2R_c \approx 0.3$, the scale of density fluctuations is limited and they do not extend over the entire dimer. Local scale of fluctuations in the dimers also follows from the computations of the radius of gyration $\langle R_g \rangle$. The ratio of the standard deviation δR_g to $\langle R_g \rangle$ is about 0.1 for both heterodimers and homodimers suggesting that the dimer dimensions experience minor fluctuations. This fluctuation regime is consistent with the globular state, which, according to statistical mechanics of polymers is characterized by local fluctuations.⁴³ The profile of radial atom number density $G(r)$ in Fig. 3(a) reveals

FIG. 5. (a) The fractions of β -strand ($\langle S(i) \rangle$) and helix ($\langle H(i) \rangle$) structure formed by amino acids i in the WT (filled circles) and RND (open circles) peptides. (b) The fractions of β -strand ($\langle S(k) \rangle$) and helix ($\langle H(k) \rangle$) structure formed by amino acid types k in the WT (filled circles) and RND (open circles) peptides. The figure shows that amino acid sequence not only determines the distribution of secondary structure but also the propensities of amino acid types k to adopt helix conformations. All plots are obtained at 360 K.

the dense core and the surface layer with rapidly decreasing density. This density profile is typical of polymer globules.⁴³ Taking into account the free energy and density calculations, it appears that reshuffling of amino acids in A β peptide has little impact on the thermodynamics of dimer assembly and does not change its globular state.

To understand this result and its physical implications we consider the distributions of dimer interactions. Random reshuffling has little effect on the number of non-local interactions, including interpeptide side chain contacts $\langle C_d(k) \rangle$ and intrapeptide contacts $\langle C_{nl}(k) \rangle$ occurring between amino acids distant along the chain. Specifically, Figs. 4(a) and 4(b) demonstrate that an amino acid of the type k forms almost equal number of non-local side chain contacts independent on specific sequence, i.e., when k appears in the WT or RND peptides (the respective correlation coefficients > 0.8). To rationalize this observation we computed the probability distribution functions of the end-to-end distance $P(r_{1N})$. We showed that $P(r_{1N})$ obtained for the WT and RND peptides can fit remarkably well with the Gaussian distribution given by Eq. (1) (Fig. 3(b)). Furthermore, as shown in supplementary material, Eq. (1) describes well the distributions of end-to-end distances for the peptides in the homodimer and tetramer (Figs. S5 and S6). Importantly, Gaussian function fails to describe the distributions of distances r_{ij} between the amino acids i and j with small sequence separation $|i - j|$ (e.g., when $|i - j| = 10$, see supplementary material). According to statistical mechanics of polymers the applicability of Gaussian statistics implies cancellation of interactions between monomers.⁴³ As a result, a polypeptide appears as “ideal” chain of non-interacting beads. This effect is known from Flory theorem, which states that in a dense polymer melt (or solution) a given chain adopts “ideal” conformation due to effective screening of excluded volume interactions by neighboring chains.⁴⁴ However, the end-to-end distance r_{1N} of an ideal polymer chain scales with the number of monomers N as $\langle r_{1N}^2 \rangle = l^2 N$, where l is the effective length of the monomer.⁴³ Then, if Gaussian statistics is applicable, the end-to-end distance $\langle r_{ij}^2 \rangle$ for the long segments of $|i - j|$ amino acids should scale linearly with $|i - j|$. The inset to Fig. 3(b) shows that this is indeed the case for sufficiently long segments ($|i - j| > 14$ for the WT peptide and $|i - j| > 23$ for the MT). The linear scaling of $\langle r_{ij}^2 \rangle$ with $|i - j|$ supports our conjecture that long polypeptide chains in the dimer acquire “ideal” properties.

We argued above that A β dimer forms a dense globular state characterized by local density fluctuations. Because the globule core has approximately constant density, it resembles a polymer melt in which interactions between amino acids placed far apart along the chain become effectively screened.⁴⁴ Consequently, from the perspective of non-local properties amino acids lose their “identity” and their specific position in a sequence becomes irrelevant. It is important to note that Gaussian statistics cannot be applied to the distribution of monomer end-to-end distances (Fig. S4 in supplementary material). This suggests a fundamental difference in the properties of A β oligomers and monomers, i.e., only oligomeric peptides, but not monomers, display “ideal” chain behavior.

As the screening of interactions vanishes for the pairs of amino acids proximal along the sequence, local interactions and properties become sequence dependent. In line with this argument, Fig. 4(c) shows that there are dramatic differences in the numbers of local intrapeptide contacts $\langle C_l(k) \rangle$ formed by the same amino acid types in the WT and RND peptides. Fig. 5(b) further reveals that amino acids of the type k demonstrate different helix $\langle H(k) \rangle$ propensities. The cluster analysis in supplementary material, which takes into account intrapeptide interactions, also shows that the local conformational ensemble is determined by a polypeptide sequence. For example, two differences can be discerned in the WT and RND conformational ensembles. First, the RND peptide has fewer clusters than the WT, of which one is dominant (75% of all structures). Second, new secondary structure distribution is seen in the RND peptide, in which β - or helix structures tend to localize in the C-terminal. Such distribution of local structure is not observed in the WT peptide. Therefore, as expected from standard bioinformatics analysis secondary structure, particularly helix, as well as local interactions depend on a sequence context. It is likely that local structural properties of the WT and RND peptides are different due to different persistence lengths l (caption to Fig. 3(b)). At the same time, melt-like behavior leads to the similarity in the large scale properties of the heterodimers and homodimers and WT and RND peptides (Figs. 2, 3, 4(a), and 4(b)). Small differences in those properties are expected due to finite size effects.

Our finding that the peptide end-to-end distance in the dimer follows Gaussian probability distribution is consistent with the previous theoretical studies of proteins. For example, Banavar *et al.* have analyzed the statistics of end-to-end distances for the fragments of native proteins.⁴⁵ They found that the distribution given by Eq. (1) is applicable to the fragments with the number of amino acids from 48 to 80. Qualitatively, similar results were obtained by Grosberg and co-workers for the unbiased ensemble of random walks representing proteins on a cubic lattice.⁴⁶ However, our study suggests that the Gaussian statistics is applicable not only to chain fragments but to the entire chains within the oligomeric structures. It is likely that this outcome is related to the similarity between oligomeric structures and dense polymer melts.

Previous computer simulations have suggested that generic amyloid assembly is a two-step process, in which peptides initially coalesce into disordered oligomers resembling molten globules, which are later reorganized into ordered β -sheet structures.⁴⁷ Because A β dimers are the smallest oligomers and according to our cluster analysis lack stable secondary (or tertiary) structure, they may represent the molten globule species populated during the first step of amyloidogenesis. In our study, by employing the results of polymer physics we show that A β dimers indeed resemble polymer globules. Furthermore, using the analogy with polymer melts we propose that their global structural properties are independent on the order of amino acids in a sequence.

Our simulations predict that the involvement of amino acids in the aggregation interface, i.e., in the interpeptide interactions (Fig. 4(a)), is not significantly affected by the sequence. This observation has two implications. First, if a set of amino acids is important for aggregation of one peptide

(such as central hydrophobic cluster in A β 1-40 Ref. 48), then the same amino acids are expected to determine the aggregation of another peptide with similar amino acid composition, even if these residues are located at new sequence positions. However, if mutation alters sequence composition, then the distribution of interpeptide interactions formed by amino acid types is likely to change as it occurs, for example, in the case of Arctic E22G mutation.⁴⁹ Second, if a ligand blocks oligomer assembly for a given peptide, then the same ligand should also be effective in preventing the aggregation of other peptides, which retain the amino acid composition of the given one. If validated, this observation will accelerate the identification of anti-aggregation agents for amyloidogenic peptides.

ACKNOWLEDGMENTS

This work was supported by the Grant R01 AG028191 from the National Institute on Aging (NIH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or NIH.

- ¹C. M. Dobson, *Nature (London)* **426**, 884 (2003).
- ²F. Chiti, M. Stefani, N. Taddei, G. Ramponi, and C. M. Dobson, *Nature (London)* **424**, 805 (2003).
- ³Y. Song, S. R. Challa, C. J. Medforth, Y. Qiu, R. K. Watt, D. Pea, J. E. Miller, F. van Swolab, and J. A. Shelnutt, *Chem. Commun.*, 1044 (2004).
- ⁴M. Rechtes and E. Gazit, *Phys. Biol.* **3**, S10 (2006).
- ⁵D. J. Selkoe, *Nature (London)* **426**, 900 (2003).
- ⁶D. M. Fowler, A. V. Koulov, W. E. Balch, and J. W. Kelly, *Trends Biochem. Sci.* **32**, 217 (2007).
- ⁷M. D. Kirkitadze, G. Bitan, and D. B. Teplow, *J. Neurosci. Res.* **69**, 567 (2002).
- ⁸R. M. Murphy, *Annu. Rev. Biomed. Eng.* **4**, 155 (2002).
- ⁹O. S. Makin, E. Atkins, P. Sikorski, J. Johansson, and L. C. Serpell, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 315 (2005).
- ¹⁰J. Hardy and G. A. Higgins, *Science* **256**, 184 (1992).
- ¹¹S. Narayanan and B. Reif, *Biochemistry* **44**, 1444 (2005).
- ¹²G. Bitan, S. S. Vollers, and D. B. Teplow, *J. Biol. Chem.* **278**, 34882 (2003).
- ¹³K. Ono, M. M. Condron, and D. B. Teplow, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 14745 (2009).
- ¹⁴C. Haass and D. J. Selkoe, *Nat. Rev. Mol. Cell. Biol.* **8**, 101 (2007).
- ¹⁵G. M. Shankar, S. Li, T. H. Mehta, A. Garcia-Munoz, N. E. Shepardson, I. Smith, F. M. Brett, M. A. Farrell, M. J. Rowan, C. A. Lemere, C. M. Regan, D. M. Walsh, B. L. Sabatini, and D. J. Selkoe, *Nat. Med.* **14**, 837 (2008).
- ¹⁶R. Riek, P. Guntert, H. Dobeli, B. Wipf, and K. Wuthrich, *Eur. J. Biochem.* **268**, 5930 (2001).
- ¹⁷L. Hou, H. Shao, Y. Zhang, H. Li, N. K. Menon, E. B. Neuhaus, J. M. Brewer, I.-J. L. Byeon, D. G. Ray, M. P. Vitek, T. Iwashita, R. A. Makula, A. B. Przybyla, and M. G. Zagorski, *J. Am. Chem. Soc.* **126**, 1992 (2004).
- ¹⁸A. T. Petkova, W.-M. Yau, and R. Tycko, *Biochemistry* **45**, 498 (2006).
- ¹⁹A. K. Paravastua, R. D. Leapman, W.-M. Yau, and R. Tycko, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18349 (2008).
- ²⁰B. Ma and R. Nussinov, *Curr. Opin. Chem. Biol.* **10**, 445 (2006).
- ²¹A. Cafisch, *Curr. Opin. Chem. Biol.* **10**, 437 (2006).
- ²²N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang, and A. E. Garcia, *J. Mol. Biol.* **368**, 1448 (2007).
- ²³A. Baumketner and J.-E. Shea, *J. Mol. Biol.* **362**, 567 (2006).
- ²⁴B. Tarus, J. E. Straub, and D. Thirumalai, *J. Am. Chem. Soc.* **128**, 16159 (2006).
- ²⁵A. Vitalis and A. Cafisch, *J. Mol. Biol.* **403**, 148 (2010).
- ²⁶Y. Lu, P. Derreumaux, Z. Guo, N. Mousseau, and G. Wei, *Proteins: Struct., Funct., Bioinf.* **75**, 954 (2009).
- ²⁷G. Bellesia and J.-E. Shea, *Biophys. J.* **96**, 875 (2009).
- ²⁸F. S. Nandel, P. Anand, and U. H. E. Hansmann, *J. Chem. Phys.* **129**, 195102 (2008).
- ²⁹S. Kim, T. Takeda, and D. K. Klimov, *J. Chem. Phys.* **132**, 225101 (2010).
- ³⁰S. Kim, T. Takeda, and D. K. Klimov, *Biophys. J.* **99**, 1949 (2010).
- ³¹Y. Chebaro, N. Mousseau, and P. Derreumaux, *J. Phys. Chem. B* **113**, 7668 (2009).
- ³²B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1982).
- ³³P. Ferrara, J. Apostolakis, and A. Cafisch, *Proteins: Struct., Funct., Bioinf.* **46**, 24 (2002).
- ³⁴T. Takeda and D. K. Klimov, *Biophys. J.* **96**, 4428 (2009).
- ³⁵T. Takeda and D. K. Klimov, *Proteins: Struct., Funct., Bioinf.* **77**, 1 (2009).
- ³⁶T. Takeda and D. K. Klimov, *J. Phys. Chem. B* **113**, 6692 (2009).
- ³⁷Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **114**, 141 (1999).
- ³⁸T. Takeda and D. K. Klimov, *Biophys. J.* **96**, 442 (2009).
- ³⁹W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
- ⁴⁰A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **63**, 1195 (1989).
- ⁴¹See supplementary material at <http://dx.doi.org/10.1063/1.3610427> for applicability of implicit solvent model and REMD convergence.
- ⁴²L. D. Landau and E. M. Lifshitz, *Statistical Physics, Course of Theoretical Physics Vol. 5* (Butterworth-Heinemann, Oxford, 1984).
- ⁴³A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (American Institute of Physics, Woodbury, 1994).
- ⁴⁴P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, 1953).
- ⁴⁵J. R. Banavar, T. X. Hoang, and A. Maritan, *J. Chem. Phys.* **122**, 234910 (2005).
- ⁴⁶R. Lua, A. L. Borovinskiy, and A. Y. Grosberg, *Polymer* **45**, 717 (2004).
- ⁴⁷M. Cheon, I. Chang, S. Mohanty, L. M. Luheshi, C. M. Dobson, M. Vendruscolo, and G. Favrin, *PLOS Comput. Biol.* **3**, 1727 (2007).
- ⁴⁸B. Urbanc, M. Betnel, L. Cruz, G. Bitan, and D. B. Teplow, *J. Am. Chem. Soc.* **132**, 4266 (2010).
- ⁴⁹A. R. Lam, D. B. Teplow, H. E. Stanley, and B. Urbanc, *J. Am. Chem. Soc.* **130**, 17413 (2008).
- ⁵⁰E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, *J. Comput. Chem.* **25**, 1605 (2004).