

Surface enhanced Raman spectroscopy distinguishes amyloid B-protein isoforms and conformational states

Xinke Yu,^{1†} Eric Y. Hayden ,^{2†} Ming Xia,¹ Owen Liang,¹ Lisa Cheah,¹ David B. Teplow ,^{2,3*} and Ya-Hong Xie^{1,4,5*}

¹Department of Materials Science and Engineering, University of California, Los Angeles, California 90095

²Department of Neurology, David Geffen School of Medicine at UCLA, 635 Charles E. Young Drive South, Room 445, Los Angeles, California 90095

³Molecular Biology Institute and Brain Research Institute, University of California, Los Angeles, California 90095

⁴Jonsson Comprehensive Cancer Center, University of California, Los Angeles, California 90095

⁵California Nano-Systems Institute at UCLA, 570 Westwood Plaza, Building 114, Los Angeles, California 90095

Received 19 February 2018; Accepted 25 April 2018

DOI: 10.1002/pro.3434

Published online 00 Month 2018 proteinscience.org

Abstract: Amyloid β -protein (A β) self-association is one process linked to the development of Alzheimer's disease (AD). A β peptides, including its most abundant forms, A β 40 and A β 42, are associated with the two predominant neuropathologic findings in AD, vascular and parenchymal amyloidosis, respectively. Efforts to develop therapies for AD often have focused on understanding and controlling the assembly of these two peptides. An obligate step in these efforts is the monitoring of assembly state. We show here that surface-enhanced Raman spectroscopy (SERS) coupled with principal component analysis (PCA) readily distinguishes A β 40 and A β 42. We show further, through comparison of assembly dependent changes in secondary structure and morphology, that the SERS/PCA approach unambiguously differentiates closely related assembly stages not readily differentiable by circular dichroism spectroscopy, electron microscopy, or other techniques. The high discriminating power of SERS/PCA is based on the rich structural information present in its spectra, which comprises not only on interatomic resonances between covalently associated atoms and hydrogen bond interactions important in controlling secondary structure, but effects of protein orientation relative to the substrate surface. Coupled with the label-free, single molecule sensitivity of SERS, the approach should prove useful for determining structure activity relationships, suggesting target sites for drug development, and for testing the effects of such drugs on the assembly process. The approach also could be of value in other systems in which assembly dependent changes in protein structure correlate with the formation of toxic peptide assemblies.

Keywords: Amyloid β -protein (A β); Alzheimer disease; Raman spectroscopy; protein aggregation; protein assembly; protein misfolding; biosensor

Abbreviations: AD, Alzheimer's disease; A β , Amyloid β -protein; CD, circular dichroism; EM, electromagnetic field; PCA, principal component analysis; PC, principal component; SERS, surface-enhanced Raman spectroscopy; TEM, transmission electron microscopy

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Alexander von Humboldt-Stiftung.

[†]Xinke Yu and Eric Y. Hayden contributed equally to this work.

*Correspondence to: [Department of Neurology, David Geffen School of Medicine at UCLA, 635 Charles E. Young Drive South, Room 445, Los Angeles, CA 90095]. E-mail: dteplow@mednet.ucla.edu and yahong.xie@gmail.com

Introduction

Amyloid β -protein (A β) assembly into neurotoxic structures appears to be a seminal pathogenetic event in Alzheimer's disease (AD).¹ For this reason, intense efforts have been devoted to understand the physiology of A β production, how the peptide assembles into neurotoxic structures, and the mechanism of amyloid plaque formation. Each of these efforts represents a potential therapeutic approach to prevent or treat AD. Equally important has been the search for biomarkers that would enable accurate diagnosis of disease state and provide metrics for evaluation of clinical trial efficacy. Unfortunately, thus far, no effective therapeutic agents or reliable biomarkers are available for clinical use. One reasonable approach to address these unmet needs is to develop new methods for dissecting the A β assembly process, methods that could reveal structural details of assembly at heretofore unsurpassed resolution and sensitivity and enable identification of novel structural biomarkers, for example, A β structural states that correlate with disease status. Surface enhanced Raman spectroscopy (SERS) is one such method, which is known to have single molecule sensitivity.^{2–4}

Surface enhanced Raman spectroscopy has been applied to the A β system in the past, but no systematic studies of the system have been published, to our knowledge. Beier *et al.* used SERS to detect Congo Red bound to A β , reporting a linear detection regime of 10^{-12} – 10^{-8} M.⁵ Benford *et al.* used a nanofluidic device containing trapped gold colloid particles (60 nm) to physically restrict A β in an illuminated volume.⁶ This allowed them to detect A β 40 at concentrations as low as 11.5 pM. Bhowmik *et al.* used lipid bilayer-coated silver nanoparticles to bind A β 40 and determine its secondary structure.⁷ A number of nanofluidic devices have been fabricated in which A β can be concentrated and its spectra acquired at concentrations as low as 10 fM.^{6,8} Voiciuk *et al.* adsorbed oligomeric forms of A β 42 onto self-assembled monolayers terminated by heptanethiol, octadecanethiol, or N-(6-mercapto)pyridinium groups in an effort to detect unique spectral features.⁹ Changes in carboxylate stretching modes were observed upon binding. Most recently, Nabers *et al.* suggested that spectral shifts in the amide I band of A β in cerebrospinal fluid might discriminate dementia of the Alzheimer's disease type from other types of dementia.¹⁰ This study used FT-IR, not Raman, spectroscopy, but these results illustrate the potential usefulness of SERS because both FT-IR and SERS probe the vibrational modes of molecules.¹¹

Recently, Wang *et al.* developed a graphene-gold hybrid plasmonic SERS platform with intrinsic electromagnetic field enhancement normalization capabilities and extremely high sensitivity.⁴ The platform

is capable of single molecule detection sensitivity and has been shown to detect A β 42 at attomolar (10^{-18}) concentration.¹² We present here results of studies demonstrating that this novel platform readily distinguishes A β 42 from A β 40 and reveals distinct spectral signatures for different conformational and assembly states. These capabilities allow monitoring of time-dependent conformational and assembly changes as well as the potential of defining new disease state biomarkers based on specific spectral signatures.

Results and Discussion

SERS analysis of unassembled A β 40 and A β 42

Our initial experiments sought to establish the spectral characteristics of low molecular weight (LMW; see Materials and Methods) A β 40 and A β 42. To do so, these peptides were prepared at concentrations of 20 μ M in 10 mM sodium phosphate, pH, 7.4, and applied to a unique hybrid SERS platform. This platform consists of a hexagonal array of gold pyramids overlaid with a single molecular layer of graphene.⁴ SERS spectra were acquired in the wavenumber range of 550–1800 cm^{-1} . Graphene produces two characteristic peaks, the D- and G-peaks, at ≈ 1350 and ≈ 1580 cm^{-1} , respectively.¹³ The graphene G-peak height depends directly on the local electromagnetic field (EM) intensity (within the area of illumination of a tightly focused laser beam ~ 1 μ m in diameter), which can vary substantially among Raman active locations (hot spots) on the platform. Normalization of peak heights at a particular hot spot to the graphene G-peak height thus provides the means to accurately determine analyte signal intensities. The graphene D-peak arises from the disordered structure of graphene, including broken carbon-carbon bonds and folds formed from the nearly planar graphene overlaid on the pyramidal platform surface, both of which can be byproducts of the fabrication process. The presence of disorder in the sp^2 -hybridized carbon system leads to the appearance of the D-peak peak.^{13,14} The D-peak intensity depends on the polarization direction of the laser beam relative to that of the graphene fold direction,¹⁴ thus its provenance differs from that of the G-peak. Neither the D-peak nor the G-peak arise from the protein analyte. However, these peaks do occur in the higher wavenumber portion of the amide II band (1510–1580 cm^{-1}) and the lower wavenumber portion of the amide I band (1600–1700 cm^{-1}) and thus can obscure some protein vibrations related to secondary structure. As will be shown below, our method of data analysis does not depend upon these obscured signals.

Spectra for A β 40 and A β 42, which have been baseline subtracted and normalized to the G-peak (see Materials and Methods), are shown in Figure 1.

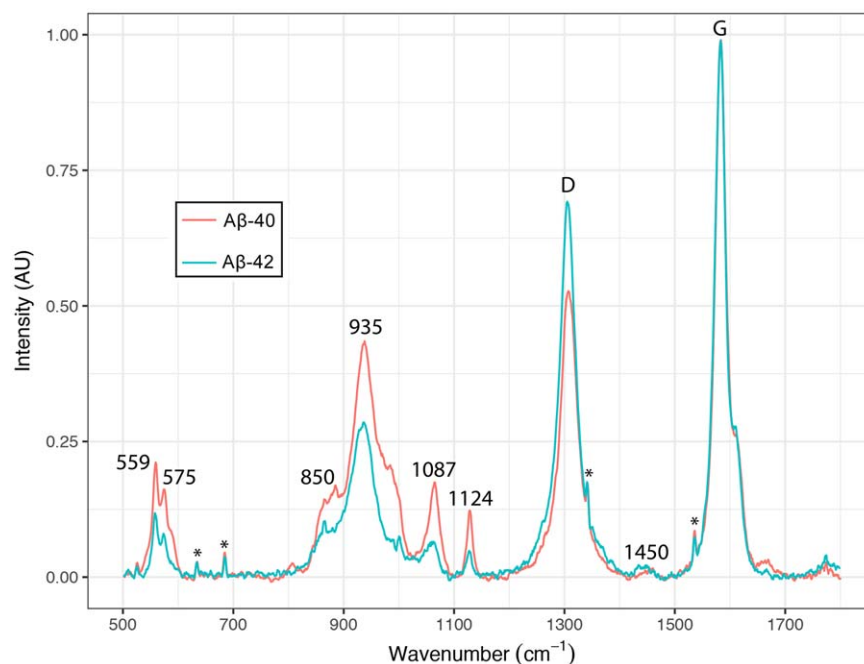


Figure 1. SERS analysis of A β 40. SERS spectra of A β 40 (red) and A β 42 (turquoise) are shown. Wavenumbers are listed above each peak. Graphene D- and G-peaks, at 1350 and 1580 cm^{-1} , respectively, are signified by letters. Peaks signified by asterisks are likely due to cosmic rays, as the peak height-to-width ratios are extremely large. The data presented are representative of two independent experiments.

Predominant peaks in the A β 40 spectrum occurred at 935, 1087, and 1124 cm^{-1} . Less intense peaks at 559, 575, 823, 850, 982, and 1450 cm^{-1} were observed reproducibly in the A β 40 spectra. The A β 42 spectrum had clearly observable, but smaller, peaks at 935, 1087, and 1124 cm^{-1} . The 823 cm^{-1} peak observed in the A β 40 spectrum was not seen in the A β 42 spectrum and the 935 cm^{-1} peak shoulders at 850 and 982 cm^{-1} were substantially smaller. However, the peak at 1450 cm^{-1} was more pronounced.

The observation in the A β 40 and A β 42 spectra of peaks at identical wavenumbers is expected because the primary structures of the two peptides also are identical, with the exception of the Ile-Ala dipeptide at the C-terminus of A β 42. However, the conformational states of the peptides during oligomerization and fibril formation have been shown to differ.¹⁵ In addition, different conformers may be oriented differently with respect to the graphene surface.¹⁶ These factors likely explain the fact that the peak intensity profiles of A β 40 and A β 42 are distinct. To more fully understand the significance of the distinct patterns of spectral intensities, we performed unbiased multivariate analysis using principal component analysis (PCA), reasoning that PCA might enable the differentiation of A β isoforms and assembly states.^{17,18} We parameterized the analysis using nine normalized peak intensities (Table I). We then performed PCA with each vector having the

same variance and found that principal components (PC—Here and in subsequent analyses, PC1 and PC2 are those components representing the first and second highest percentages of the total variance, respectively. PC1 and PC2 will vary among different experiments.) 1 and 2 accounted for 57.8% and 15.0%, respectively, of the variance in the data. The cumulative percentage of 72.8% means that the first two principal components account for the majority of the variance in the system. We note that other components account for <5% each of the total variance, so their inclusion in our analyses would not alter our conclusions. PCA analysis can produce statistically significant results when $n > 5p$, where n is number of nodes and p is number of vectors used.¹⁹ In our case, using 80 spectra and nine vectors, $n > 11p$, so we are confident that the first two components accurately account for $\sim 75\%$ of the total system variance.

A graph of the results using PC1 and PC2 as axes revealed that the data from A β 40 and A β 42 clustered in two distinct regions (Fig. 2). The A β 40 cluster was substantially smaller than the A β 42 cluster, which suggests that its component conformers were more homogenous structurally than the conformers in the A β 42 cluster. This is apparent from examination of overlays of the 80 spectra submitted to cluster analysis for each peptide isoform (see Fig. S1). The A β 40 and A β 42 clusters display similar variance in the PC1 dimension. The variance

Table I. A β 40 and A β 42 Raman peak positions, assignments, and intensities. Peaks have been assigned to specific bond resonances based on published data.^{22,31,32} We note that formal confirmation of the assignments requires additional experimentation involving site-specific labeling of amino acids or bond pairs. Wavenumber assignments (cm^{-1}) are: 559, aliphatic; 575, C-C bond stretching mode; 823, out-of-plane ring breathing vibration or Tyr; 850, single bond stretching for Tyr and Val; 935, number of carbon-carbon bonds of protein backbone or Gly; 982, C-C stretching in β -sheets or part of Phe; 1000, Lys or Asn; 1124, Val or Ile; 1450, CH_2 bending or Phe. An illustration of vibrational modes is shown in Figure 7.

Wavenumber (cm^{-1})	559	575	823	850	935	982	1087	1124	1450
Peak assignment	Aliphatic	C-C stretching	Out-of-plane ring breathing or Tyr	Amino acid single bond stretching, Tyr or Val	n(C-C) of protein backbone or Gly	C-C stretching β -sheet or Phe	Lys or Asn	Val or Ile	CH_2 bending or Phe
A β 40 average intensity (SD)	(0.06)	(0.07)	(0.05)	(0.06)	(0.07)	(0.03)	(0.02)	(0.04)	(0.03)
A β 42 average intensity (SD)	(0.03)	(0.04)	(0.00)	(0.03)	(0.06)	(0.02)	(0.03)	(0.05)	(0.04)
A β 42(SD)	0.0319	0.0392	0.0000	0.0291	0.0613	0.0233	0.0329	0.0519	0.0411

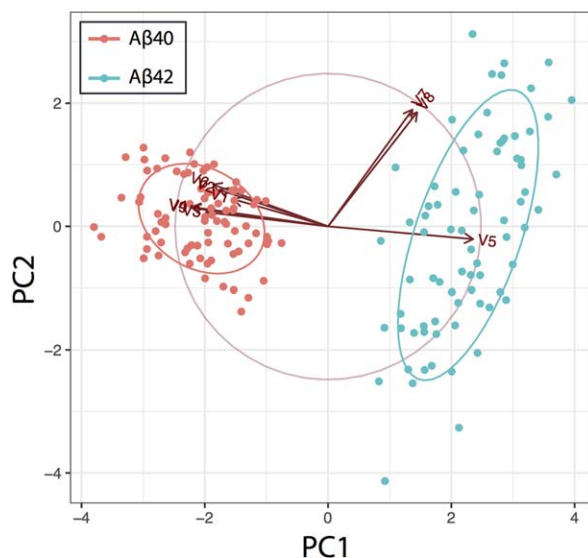


Figure 2. PCA analysis. Plot of principal components 1 and 2 from analysis of unassembled A β 40 (salmon) and A β 42 (turquoise). Ellipses surrounding clusters of the data, indicating the majority of the data points are in the ellipse. The brown arrows are the projections of the vectors in PC space.

of A β 42 in the PC2 dimension was approximately twice that of A β 40. The equations specifying PC1 ($\vec{C}_1$) and PC2 ($\vec{C}_2$) provide an explanation for the cluster locations and shapes.

$$\vec{PC}_1 = -0.27\vec{V}_1 - 0.31\vec{V}_2 - 0.35\vec{V}_3 - 0.38\vec{V}_4 + 0.42\vec{V}_5 \\ - 0.33\vec{V}_6 + 0.24\vec{V}_7 + 0.26\vec{V}_8 - 0.39\vec{V}_9$$

$$\vec{PC}_2 = 0.15\vec{V}_1 + 0.21\vec{V}_2 + 0.08\vec{V}_3 + 0.11\vec{V}_4 - 0.08\vec{V}_5 \\ + 0.23\vec{V}_6 + 0.65\vec{V}_7 + 0.64\vec{V}_8 + 0.10\vec{V}_9$$

Vectors ($\vec{V}$) 1–9 represent the peak intensities of SERS peaks at 559, 575, 823, 850, 935, 982, 1000, 1124, and 1450 cm^{-1} , respectively. A key difference between the two principal component vector equations is the absolute value of the coefficients of vectors $\vec{V}_7$ and $\vec{V}_8$ (the peak heights at 1000 and 1124 cm^{-1} , respectively), which are substantially larger in the case of A β 42 compared to A β 40. This observation, which is not immediately apparent from analysis of peak intensities alone (Table I), explains why the variance in PC2 space is twice as large for A β 42 than it is for A β 40. Resonances at 1000 and 1124 cm^{-1} are produced by Lys and Asn, and by Val and Ile, respectively (Fig. 7). The presence of the additional Ile at the C-terminus of A β 42 likely is an explanation for at least a portion of the increased magnitude of $\vec{V}_8$. Conformational effects due to the distinct conformational dynamics of A β 42 may also contribute to the differences in vector magnitudes.

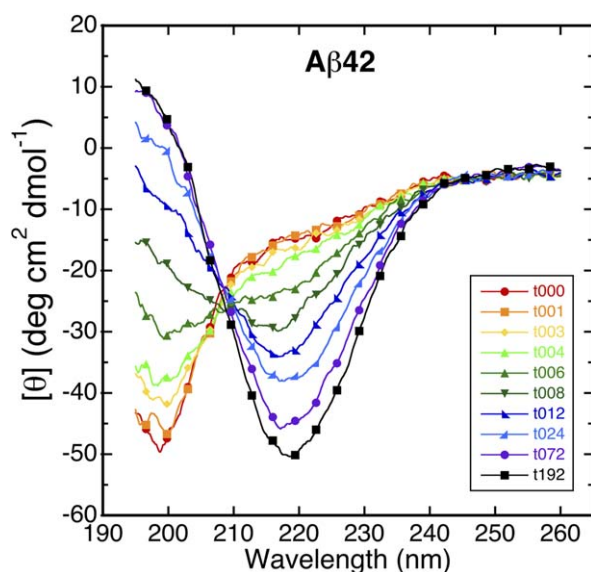


Figure 3. Circular dichroism spectroscopy. A β 42 was incubated at a concentration of 20 μ M in 10 mM sodium phosphate, pH 7.4, at 37°C. CD spectra then were acquired periodically. Overlapping spectra have not been presented so as to make visualization of the time-dependence of spectral changes easier. The data shown are representative of four independent experiments. Spectral colors represent different time points, which are specified in hours in the box to the right.

SERS analysis of A β assembly

We next sought to establish whether SERS could distinguish different stages of A β assembly. We characterized assembly stages by performing SERS in parallel with circular dichroism (CD) spectroscopy and transmission electron microscopy (TEM). We thus obtained CD, TEM, and SERS data from the same sample aliquots. The data shown are representative of four independent experiments. CD complements SERS by providing information in spectral regions obscured by the graphene D- and G-peaks. It

also allows real time monitoring of secondary structure changes in assembly reactions *in hydro*. CD spectra were acquired immediately after initiation of assembly reactions of 20 μ M A β 42 in 10 mM sodium phosphate, pH 7.4, at 37°C. The spectra were consistent with statistical coil (SC) structure, as indicated by a minimum in molar ellipticity $[\Theta]$ at 198 nm and a gradual increase in $[\Theta]$ as wavelength increased toward 260 nm (Fig. 3). A concerted time-dependent increase in $[\Theta]_{198}$ and decrease in $[\Theta]_{218}$ were consistent with β -sheet formation (Fig. S2). Maximal β -sheet content was observed at 120 h (this spectrum was essentially superimposable with the 192 h spectrum shown in Fig. 3). These data were consistent with the SC $\rightarrow\beta$ -sheet secondary structure transitions that occur during A β fibril formation.²⁰ Negative stain EM done in parallel with the CD studies confirmed that fibril assembly was occurring (Fig. 4). The starting peptide solution (0 h), which displayed statistical coil secondary structure, contained only small globular structures of $\approx$ 8 nm diameter. Short fibrils of width $\approx$ 10 nm were observed at 6 h, during the coil $\rightarrow\beta$ -sheet transition period. When maximal β -sheet was observed, long fibrils were present. The morphologies of these structures did not change substantially after 24 h.

SERS spectra were acquired periodically from 0 to 168 h. Eighty spectra were taken at each time point and the intensities of the peaks were normalized to the graphene G peak (1584 cm^{-1}) and then averaged. Figure 5 shows a plot of the averaged spectra from nine different time points. Spectra acquired after 24 h were identical to those at 24 h. The graphene G peaks for each spectrum have the same intensities because of normalization. Time-dependent spectral changes were observed for peaks produced by A β . At 0 h, the spectrum (red) was dominated by the graphene D and G peaks, but some intensity at 935 cm^{-1} was observed. At 1 h

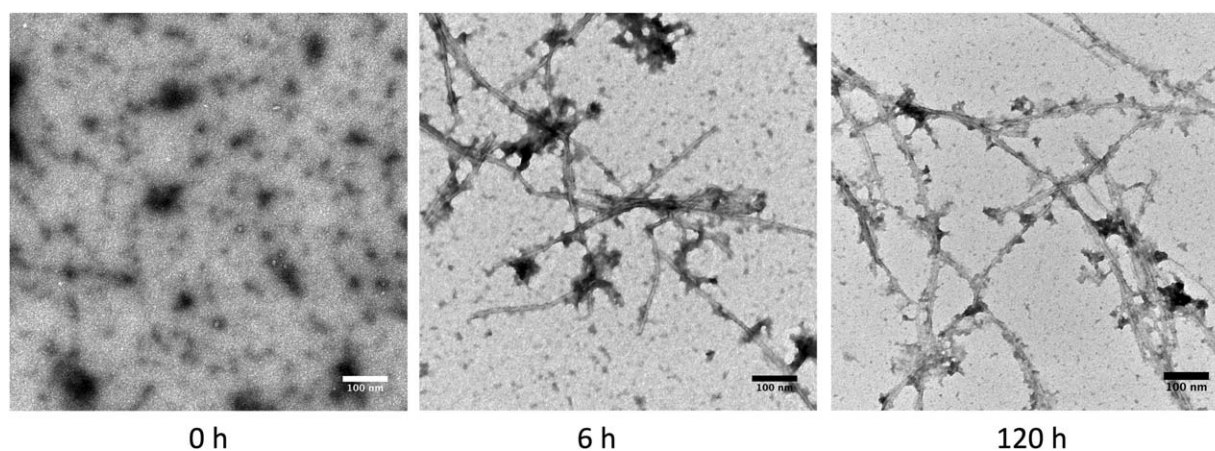


Figure 4. Transmission electron microscopy. A β 42 (20 μ M in 10 mM sodium phosphate, pH 7.4) was incubated at 37°C for a total of 192 h. Panels shown are representative of the sample morphologies observed at the indicated time point. Scale bars are 100 nm.

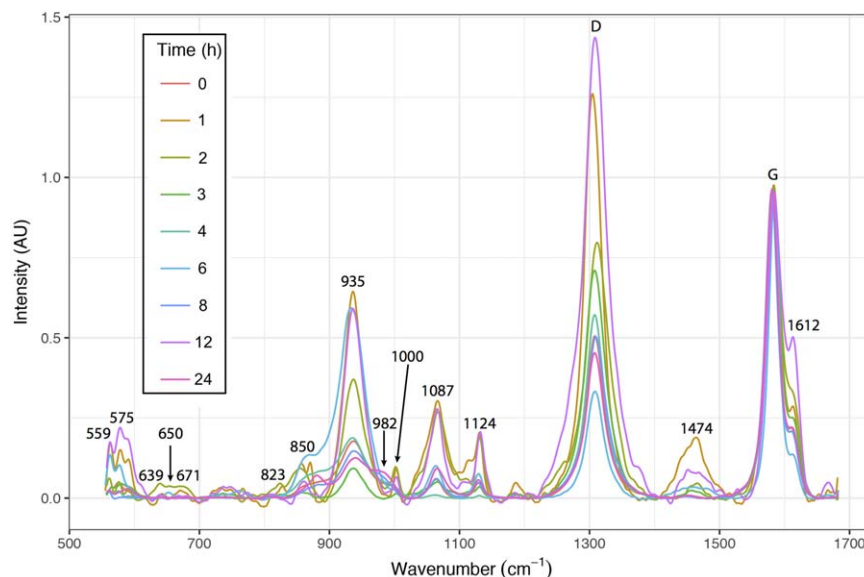


Figure 5. SERS analysis of A β 42. Spectra were acquired periodically from the same samples used for CD and EM. Axes are wavenumber (cm^{-1}) and intensity (AU). We show spectra up to and including 24 h, after which all spectra were identical, within experimental error. Numbers above peaks specify their wavenumbers. Graphene peaks are denoted by letters. The data shown are representative of four independent experiments.

(orange spectrum), the intensity of the 935 cm^{-1} peak had increased substantially and peaks now also could be seen at 850, 1000, 1087, 1124, ≈ 1175 , and 1460 cm^{-1} . The intensity of the 935 cm^{-1} peak relative to the other peaks was lower at 2 h (dark green spectrum). Relative peak intensities differed at a number of time points, which suggested that unique populations of assemblies were being detected.

To determine whether the spectral differences observed could distinguish different assembly states, we performed PCA analyses to visualize the data. We employed decision trees to identify peaks that would be most useful in differentiating among time points.²¹ Decision trees are particularly suitable for this purpose as they provide simple measures of attribute importance that can be used to rank the importance of the peaks. The C4.5 algorithm was used to prioritize key peaks after examination of all non-graphene peaks with signal-to-noise ratios >10 , which comprised peaks at wavenumbers 559, 575, 639, 650, 671, 823, 850, 935, 982, 1000, 1087, 1124, 1190, 1474, and 1612. Nodes in the decision trees (Fig. S3) were generated by choosing the attribute of the data that most effectively split the tree with the highest normalized information gain. These nodes comprised peak intensities at 823, 850, 935, 1000, and 1124 cm^{-1} .

We performed PCA analysis using intensities observed at the five different wavenumbers. PC1 and PC2 contributed 77.7% and 8.5%, respectively, to the total variance of the data. Cluster analysis in the PC1:PC2 plane (Fig. 6) revealed a striking time-dependence, and thus assembly stage dependence, of

the locations of the data clusters. Each cluster comprises 67% of the data from a particular time point. Centroids for each cluster in the PC1 dimension were determined by averaging *all* data points. For ease of examination, the clusters have been delimited by color-coded elliptical boundaries, each color representing a specific time. To determine whether the differences in centroid positions were significant, paired Student's *t*-tests were performed among all pairs of centroids at all times (Table II). With the exception of differences between 4 and 6 h, 6 and 8 h, and 24 and 48 h, all differences were highly significant ($10^{-39} < P < 10^{-3}$). However, if we consider differences in the PC2 dimension, the *P*-values of 4 h versus 6 h and 6 versus 8 h were significant ($P = 0.002$ and $P = 0.037$, respectively). The centroids at 24 and 48 h were almost identical ($P = 0.99$), as they were in the PC1 dimension ($P = 0.94$). This suggests that the assembly process was complete by 24 h. Taken together, these statistical analyses show that each of the 9 clusters represents a different assembly stage.

The fact that nine clusters were located at different positions in the PC1:PC2 plane shows that at least nine different assembly states were differentiated by SERS/PCA. It is possible that more unique states exist. These could be determined by sampling the assembly process at additional times prior to 24 h. We interpret the clustering as indicative of populations of assemblies that differ between themselves in both the types and relative amounts of different conformers present. Each centroid thus represents a population-average conformer. Overlaps among clusters indicate some population-average

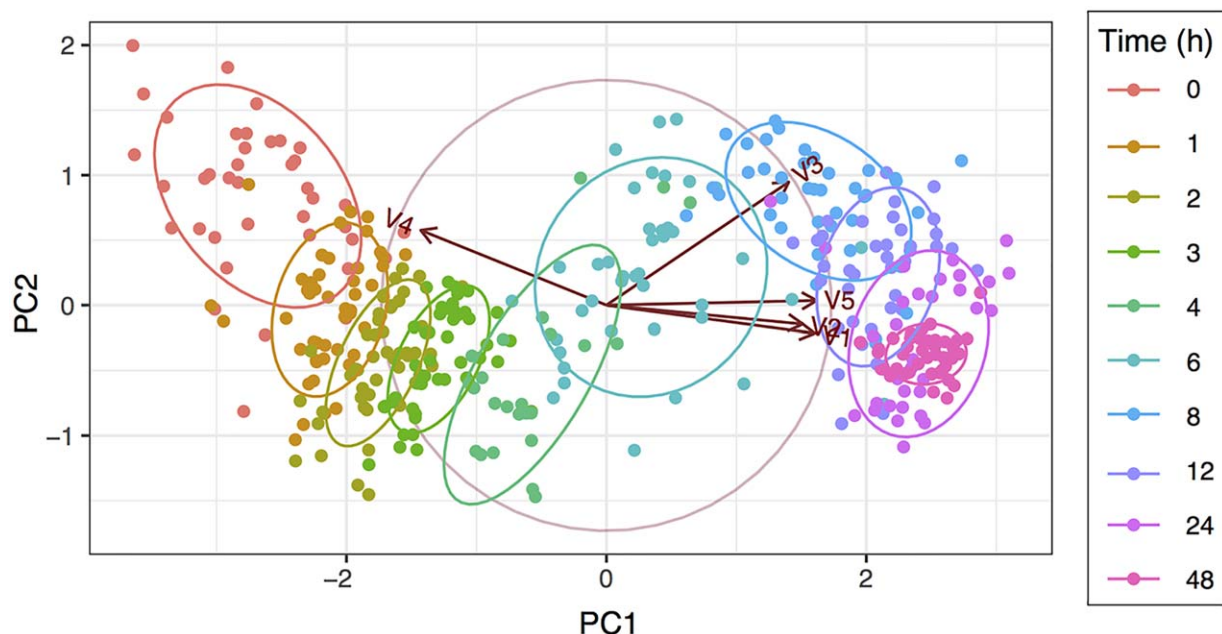


Figure 6. PCA analysis. Plot of principal components 1 and 2 from analysis of spectra acquired during A β 42 assembly. Ellipses surrounding clusters enclose 2/3 (1 z-unit from the average) of the data. Vectors $\vec{v}_1$ – $\vec{v}_5$ (brown arrows) signify the peak intensities at 800, 850, 935, 1000, and 1124 cm^{-1} , respectively. Colors signify the incubation time at which the points and boundaries of each ellipse were obtained (see box at right).

conformational similarity. The time-dependence of ellipse centroid position displayed an amplitude in the PC1 dimension that was ~ 3 -fold larger than that in the PC2 dimension.

Analysis of the component vectors comprising PC1 and PC2 shows that, from 0 to 4 h, the positions of ellipse centroids within PC2 space are determined primarily by decreases in $\vec{V}_4$, corresponding to peak intensity at 1000 cm^{-1} . This suggests that Lys or Asn residues, which resonate at 1000 cm^{-1} , are oriented with the graphene surface in a manner that is sub-optimal with respect to Raman signal production. This orientation difference likely reflects conformational changes during peptide assembly. Figure 7 highlights regions of the A β structure wherein vibrational mode intensity differences are noted. Between 4 and 8 h, ellipse position is

determined primarily by $\vec{V}_3$, indicating an increase in the intensity of the 935 cm^{-1} Raman peak, which is produced by carbon-carbon bond resonances.²² As discussed above, this peak intensity change likely reflects changes in the interaction of the peptide with the graphene due to peptide assembly. After 8 h, the predominant contributor to ellipse position again is $\vec{V}_4$, which shows that the peak intensity of the 1000 cm^{-1} Raman peak decreases. After 24 h, ellipse position and shape do not change substantially, suggesting that the structures of the assemblies producing the Raman spectra are end-state products. This supposition was consistent with results of the EM analysis, which showed no substantial morphological changes after 24 h. A general feature of the time-dependence of ellipse position is that it increases monotonically in the PC1

Table II. Significance of differences in centroid positions. Paired Student's *t*-tests were performed to determine if positions of cluster centroids in the PC1 dimension were significantly different ($P < 0.05$). The table lists the *P*-values obtained. Insignificant differences are bolded. "X" indicates positions of self-comparison

Time (h)	0	1	2	3	4	6	8	12	24	48
0	X									
1	5.7×10^{-10}	X								
2	5.6×10^{-17}	5.1×10^{-14}	X							
3	6.0×10^{-22}	7.1×10^{-27}	4.9×10^{-12}	X						
4	1.5×10^{-28}	6.7×10^{-34}	1.1×10^{-25}	1.5×10^{-16}	X					
6	9.0×10^{-29}	5.2×10^{-32}	1.9×10^{-24}	1.4×10^{-16}	0.17	X				
8	4.6×10^{-30}	3.5×10^{-35}	6.7×10^{-28}	6.1×10^{-20}	0.02	0.49	X			
12	7.3×10^{-35}	1.3×10^{-47}	2.4×10^{-42}	3.6×10^{-35}	1.2×10^{-9}	2.2×10^{-5}	1.6×10^{-4}	X		
24	1.2×10^{-39}	3.3×10^{-63}	2.9×10^{-62}	9.0×10^{-57}	5.6×10^{-23}	2.0×10^{-15}	2.7×10^{-15}	2.2×10^{-10}	X	
48	1.2×10^{-39}	2.9×10^{-63}	2.4×10^{-62}	7.3×10^{-57}	4.8×10^{-23}	1.8×10^{-15}	2.4×10^{-15}	1.8×10^{-10}	0.94	X

3 h, the sizes of the ellipses decrease as well, which is consistent with a folding process that decreases the conformational space of the peptide. Taken together, the data suggest that the decrease in PC2 during this time period is indicative of initial A β self-association leading to small oligomers and fibril nuclei. As fibril growth occurs, the heterogeneity of assemblies increases, which explains why the 4 h cluster is larger in area than those at 1–3 h. This growth period is reflected in a modest increase in the rate of change in $[\Theta]_{198}$ (~ 4 deg cm² dmol⁻¹ h⁻¹ compared to an initial rate of ~ 3 deg cm² dmol⁻¹ h⁻¹; Fig. S1). In addition, between 6 and 8 h, a negative inflection at $[\Theta]_{216}$ appears, which monotonically decreases over time, consistent with the increased β -sheet secondary structure produced by fibril formation. Ellipse position in the PC2 dimension rises concurrently (Fig. 6). From 8 to 24 h, progressive increases in β -sheet (CD) and fibril content (TEM) correlated with decreases in cluster position in the PC2 dimension. Increasing V_4 magnitudes were the prime contributor to the monotonic decrease in ellipse position in the PC2 dimension. We note that the centers of the ellipses of 24 and 48 h were in essentially the same position in PC1:PC2 space, but ellipse area decreased markedly during this time period, which suggests increased structural homogeneity of the peptide assemblies. This effect may be related to fibril aggregation, which commonly is observed following fibril growth phases.²⁴

Conclusions

We show that Raman spectra obtained using a graphene-gold hybrid plasmonic platform, in combination with PCA analysis, enables facile distinction between A β 40 and A β 42, the peptide isoforms associated with classical vascular AD (A β 40) and parenchymal (A β 42) plaques, respectively, in AD. We further show that the approach is capable of revealing assembly dependent changes in peptide conformation and self-association. Correlation of these spectral changes with CD and TEM data allow regions in PCA space to be linked to specific populations of A β assemblies. What may be particularly important is the observation of a minimum of nine differentiable clusters within PCA space, which reflect at least nine differentiable assembly states in the fibril formation pathway. Because spectral changes can be linked to changes in resonances of specific amino acids within the A β peptide, future sited-directed amino acid substitution studies of these individual states may provide new insights into the roles of different amino acids in stabilizing or destabilizing these states. Thus, coupled with the label-free, single molecule sensitivity of SERS, the SERS/PCA approach should prove useful for determining structure activity relationships, suggesting

target sites for drug development, and for testing the effects of such drugs on the assembly process. The approach also could be of value in other systems in which assembly dependent changes in protein structure correlate with the formation of toxic peptide assemblies.

Materials and Methods

Preparation of A β

A β was synthesized and characterized, as described previously.²⁵ Briefly, peptide synthesis was performed on an automated peptide synthesizer (model 433A, Applied Biosystems, Foster City, CA, USA) using 9-fluorenylmethoxycarbonyl-based methods on preloaded Wang resins. Peptides were purified, using reverse-phase high-performance liquid chromatography (HPLC). Quantitative amino acid analysis and mass spectrometry yielded the expected compositions and molecular weights, respectively, for each peptide. Purified peptides were stored as lyophilizates at -20°C . Peptides were prepared by dissolution in a 1:4.5:4:5 ratio of 60 mM NaOH:Milli-Q water:22.2 mM sodium phosphate, pH 7.4, to yield a nominal A β concentration of 1 mg/mL in 10 mM sodium phosphate, pH 7.4. The peptide solution then was sonicated for 1 min in a bath sonicator (Branson Model 1510, Danbury, CT, USA) and filtered through a 30,000 molecular weight cut-off Microcon centrifugal filter device (Millipore, Billerica, MA, USA) for 15 min at 16,000g. The A β concentration in the eluate was determined by UV absorbance ($\epsilon_{280} = 1280$ cm⁻¹ M⁻¹) using a Beckman DU-640 spectrophotometer (Beckman Instruments, Fullerton, CA, USA). This protocol reproducibly yields aggregate-free A β monomer in rapid equilibrium with low order oligomers, which is termed “low molecular weight” (LMW) A β .²⁶

Circular dichroism spectroscopy (CD)

LMW A β was prepared at a concentration of 20 μM in sodium phosphate, pH 7.4, and incubated without agitation at 37°C in a 1 mm path-length quartz cuvette (Hellma, Forest Hills, NY, USA). CD spectra then were acquired periodically with a J-810 spectropolarimeter (JASCO, Tokyo, Japan). Spectra were recorded from 195 to 260 nm at 0.2 nm resolution with a scan rate of 100 nm/min. Ten scans were acquired and averaged for each sample.

Transmission electron microscopy (TEM)

Five microliters of A β 42 (20 μM) were removed at the time of each CD measurement and then spotted onto carbon-coated Formvar grids (Electron Microscopy Sciences, Hatfield, PA, USA). After 2 min, the droplet was displaced with an equal volume of 1% (w/v) filtered (0.2 μM) uranyl acetate in water (Electron Microscopy Sciences). This solution was wicked

off and then the grid was air-dried. All grids were coded at the end of the time course so that the operator of the electron microscope did not know what samples were being imaged. Electron microscopy was done using a JEOL 1200 EX transmission electron microscope with an accelerating voltage of 80 kV, which is typical for protein examination.²⁷ Digital images were analyzed with ImageJ 1.50d, using the “measure tool” to calculate dimensions, and unblinded after the analysis was complete.

Fabrication of graphene-gold hybrid plasmonic platform

Fabrication of the platform is based on sphere lithography.²⁸ Briefly, periodic gold nanopillar structures with tunable size and sharpness can be fabricated by a wafer-scale, bottom-up templating technology. Spin-coated on (001) silicon wafers, close-packed monolayer polystyrene balls with a diameter of 200 nm serve as templates. Monolayer graphene is synthesized by chemical vapor deposition (CVD) and transferred to coat the gold tip surface by Poly(methyl methacrylate) (PMMA).

Raman measurement

Immediately following solubilization, 20 μ L aliquots of A β 40 or A β 42 were applied to a graphene-coated, pyramidal gold hybrid platform and dried *in vacuo*. Spectra were acquired using a Renishaw inVia microscope under ambient conditions. The excitation wavelength was 785 nm and the He-Ne laser power was 0.5 mW. The 785 nm laser was chosen due to the relatively lower photon energy of excitation, which avoids thermal degradation of biomaterials. The grating used was 1800 lines/mm, and the objective lens used was 50 $\times$. We scanned the entire region on the platform occupied by the samples ($\approx 24 \mu\text{m} \times \approx 30 \mu\text{m}$) using Raman mapping with a step size of 3 μm (i.e., independent areas of 9 μm^2 each). Eighty spectra were acquired for each sample. This process controls for acquisition of spectra unrepresentative of the average spatial orientation or assembly state of the peptide, two factors that can affect peak location and intensity, and which become problematic if spectra are acquired from only one or a few areas of the platform.²⁹ For Raman measurements done in parallel with CD, 10 μ L aliquots were taken from the CD cuvette at the time of each CD measurement, applied to the platform, and then spectra were acquired, essentially as described above. Raman data were analyzed using Renishaw WiRE 4.2 software, which automatically subtracts the baseline signal and removes noise. Peak intensities in each spectrum were normalized to the graphene G peak to enable spectral comparisons among samples.

Principal component analysis (PCA)

R³⁰ was used to analyze and visualize the multi-dimensional SERS data sets¹⁷ by using the built in PCA function “prcomp” and open source library “ggbiplot” (with extended functionality for labeling groups, drawing a correlation circle, and adding normal probability ellipsoids). Dimensions in the analyses were the intensities at five different wavenumbers (823, 850, 935, 1000, and 1124 cm^{-1}) determined by decision tree analysis to be necessary and sufficient for differentiating among time points.²¹ PCA transforms the original variables into a set of linear combinations (principal components: PC), which allows the retention of the data variability, while examining them independently in a weighted fashion of decreasing order of variance. Variability in the vectors gathered by the PCs was calculated and the largest two PCs were plotted. All time points for the first 48 h were included in the analyses. After that time, CD spectroscopy revealed no spectral changes, so no additional SERS spectra were analyzed. Data from each time point was considered a separate group and PCA was done to maximize the between-group variation.

Decision trees

Decision trees were produced using R with the “rpart” package (generating classification and regression trees). The classification method in the “rpart” function was used to produce the tree information and the “prp” function in “rpart.plot” library was used to visualize the decision trees.

t-test

Paired *t*-tests, as implemented in R as “t.test” and using the key parameter “P.value,” were done to assess the statistical significance of differences between the centroids of clusters in the PC1 dimension, in which the largest variances were observed. The analyses were performed on centroids in the PC2 dimension if no significant differences were observed in PC1. Significance was defined as $P < 0.05$.

Supplementary Material

1. A figure and figure legend showing overlays of all 80 spectra acquired for A β 40 and A β 42.
2. A figure and figure legend showing CD intensity changes over time at 198 and 218 nm.
3. A figure and figure legend showing the decision tree obtained through analysis of PCA data for the assembly states of A β 40 and A β 42.

Acknowledgments

We thank Margaret M. Condrón for peptide synthesis. We acknowledge the support of the Alexander von Humboldt Foundation (Y.H.X.), Zhejiang University Cao Guangbiao High-Tech Talent Fund (Y.H.X.),

the National Institutes of Health (NS038328 and AG041295; DBT), and the State of California, Department of Public Health, Alzheimer's Disease Program grant #16-10322 (DBT). This work was funded in part (Y.H.X., X.Y., O.L.) by a grant from the United States Government and the generous support of the American people through the United States Department of State and the United States Agency for International Development (USAID) under the Pakistan - U.S. Science & Technology Cooperation Program. The contents do not necessarily reflect the views of the United States Government.

Conflict of Interest

The authors declare no conflicts of interest.

Author Contributions

Y.-H.X. and D.B.T. conceived and coordinated the study, analyzed the data, and wrote the paper. X.Y. and E.H. performed the experiments, analyzed the data, prepared figures and tables, and wrote the paper. M.X., O.L., and L.C. performed experiments.

References

- Roychaudhuri R, Yang M, Hoshi MM, Teplow DB (2009) Amyloid β -protein assembly and Alzheimer disease. *J Biol Chem* 284:4749–4753.
- Kneipp K, Wang Y, Kneipp H, Perelman LT, Itzkan I, Dasari R, Feld MS (1997) Single molecule detection using surface-enhanced Raman scattering (SERS). *Phys Rev Lett* 78:1667–1670.
- Liu H, Zhang L, Lang X, Yamaguchi Y, Iwasaki H, Inouye Y, Xue Q, Chen M (2011) Single molecule detection from a large-scale SERS-active $\text{Au}_{79}\text{Ag}_{21}$ substrate. *Sci Rep* 1:112.
- Wang P, Liang O, Zhang W, Schroeder T, Xie YH (2013) Ultra-sensitive graphene-plasmonic hybrid platform for label-free detection. *Adv Mater* 25:4918–4924.
- Beier HT, Cowan CB, Chou I-H, Pallikal J, Henry JE, Benford ME, Jackson JB, Good TA, Coté GL (2007) Application of surface-enhanced Raman spectroscopy for detection of beta amyloid using nanoshells. *Plasmonics* 2:55–64.
- Benford ME, Chou I-H, Beier HT, Wang M, Kameoka J, Good TA, Cot GL (2008) In vitro detection of β amyloid exploiting surface enhanced Raman scattering (SERS) using a nanofluidic biosensor. *Proc SPIE* 6869, Plasmonics in Biology and Medicine V, 68690W. doi: 10.1117/12.763913.
- Bhowmik D, Mote KR, MacLaughlin CM, Biswas N, Chandra B, Basu JK, Walker GC, Madhu PK, Maiti S (2015) Cell-membrane-mimicking lipid-coated nanoparticles confer Raman enhancement to membrane proteins and reveal membrane-attached amyloid- β conformation. *ACS Nano* 9:9070–9077.
- Choi I, Huh YS, Erickson D (2012) Ultra-sensitive, label-free probing of the conformational characteristics of amyloid β aggregates with a SERS active nanofluidic device. *Microfluid Nanofluid* 12:663–669.
- Voiciuk V, Valincius G, Budvytytė R, Matijoška A, Matulaitienė I, Niaura G (2012) Surface-enhanced Raman spectroscopy for detection of toxic amyloid β oligomers adsorbed on self-assembled monolayers. *Spectrochim Acta A Mol Biomol Spectrosc* 95:526–532.
- Nabers A, Ollesch J, Schartner J, Kotting C, Genius J, Hafermann H, Klafki H, Gerwert K, Wiltfang J (2016) Amyloid- β -secondary structure distribution in cerebrospinal fluid and blood measured by an immuno-infrared-sensor: a biomarker candidate for Alzheimer's disease. *Anal Chem* 88:2755–2762.
- Larkin P (2011) *Infrared and Raman spectroscopy: principles and spectral interpretation*. San Diego, Oxford, Amsterdam: Elsevier, Waltham.
- Yu X, Hayden EY, Wang P, Xia M, Liang O, Bai Y, Teplow DB, Xie Y-H. Quantification characteristics of a graphene-gold hybrid plasmonic SERS platform and its use in studies of Alzheimer amyloid β -protein. Manuscript in preparation.
- Ferrari AC, Meyer JC, Scardaci V, Casiraghi C, Lazzeri M, Mauri F, Piscanec S, Jiang D, Novoselov KS, Roth S, Geim AK (2006) Raman spectrum of graphene and graphene layers. *Phys Rev Lett* 97:187401.
- Wang P, Zhang W, Liang O, Pantoja M, Katzer J, Schroeder T, Xie YH (2012) Giant optical response from graphene-plasmonic system. *ACS Nano* 6:6244–6249.
- Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB, Teplow DB (2003) Amyloid β -protein ($\text{A}\beta$) assembly: $\text{A}\beta_{40}$ and $\text{A}\beta_{42}$ oligomerize through distinct pathways. *Proc Natl Acad Sci USA* 100:330–335.
- Medek A, Hajduk PJ, Mack J, Fesik SW (2000) The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands. *J Am Chem Soc* 122:1241–1242.
- Abdi H, Williams LJ (2010) Principal component analysis. *Wiley Interdisc Rev Comput Statist* 2:433–459.
- You ZH, Lei YK, Zhu L, Xia J, Wang B (2013) Prediction of protein-protein interactions from amino acid sequences with ensemble extreme learning machines and principal component analysis. *BMC Bioinform* 14:S10.
- Reynaldo J, Santos A (1999) Cronbach's alpha: A tool for assessing the reliability of scales. *J Extension* 37:1–5.
- Hayden EY, Yamin G, Beroukhim S, Chen B, Kibalchenko M, Jiang L, Ho L, Wang J, Pasinetti GM, Teplow DB (2015) Inhibiting amyloid β -protein assembly: Size-activity relationships among grape seed-derived polyphenols. *J Neurochem* 135:416–430.
- Quinlan JR (1993) *C4.5: programs for machine learning*. San Mateo, California: Morgan Kaufmann Publishers.
- Movasaghi Z, Rehman S, Rehman IU (2007) Raman spectroscopy of biological tissues. *Appl Spectr Rev* 42: 493–541.
- Movasaghi Z, Rehman S, Rehman IU (2007) Raman spectroscopy of biological tissues. *Appl Spectrosc Rev* 42:493–541.
- Qiang W, Yau WM, Tycko R (2011) Structural evolution of Iowa mutant β -amyloid fibrils from polymorphic to homogeneous states under repeated seeded growth. *J Am Chem Soc* 133:4018–4029.
- Walsh DM, Lomakin A, Benedek GB, Condron MM, Teplow DB (1997) Amyloid β -protein fibrillogenesis. Detection of a protofibrillar intermediate. *J Biol Chem* 272:22364–22372.
- Walsh DM, Hartley DM, Kusumoto Y, Fezoui Y, Condron MM, Lomakin A, Benedek GB, Selkoe DJ, Teplow DB (1999) Amyloid β -protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *J Biol Chem* 274:25945–25952.

27. Petkova AT, Leapman RD, Guo Z, Yau WM, Mattson MP, Tycko R (2005) Self-propagating, molecular-level polymorphism in Alzheimer's β -amyloid fibrils. *Science* 307:262–265.
28. Wang P, Xia M, Liang O, Sun K, Cipriano AF, Schroeder T, Liu H, Xie YH (2015) Label-free SERS selective detection of dopamine and serotonin using graphene-Au nanopyramid heterostructure. *Anal Chem* 87:10255–10261.
29. Siddhanta S, Narayana C (2012) Surface enhanced Raman spectroscopy of proteins: Implications for drug designing invited review article. *Nanomater Nanotech* 2:1–13.
30. R Core Team (2013) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
31. Gremlich H-U, Yan B (2001) Infrared and Raman spectroscopy of biological materials. New York and Basel: Marcel Dekker.
32. Tu AT, Tu A (1982) Raman spectroscopy in biology: principles and applications. New York: John Wiley and Sons Inc.