Unraveling VEALYL Amyloid Formation Using Advanced Vibrational Spectroscopy and Microscopy


DOI
10.1016/j.bpj.2020.05.026

Publication date
2020

Document Version
Final published version

Published in
Biophysical Journal

License
Article 25fa Dutch Copyright Act

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Unraveling VEALYL Amyloid Formation Using Advanced Vibrational Spectroscopy and Microscopy

Steven J. Roeters, 1,* Mathias Sawall, 2 Carl E. Eskildsen, 3 Matthijs R. Panman, 1 Gergely Tordai, 1 Mike Koeman, 4 Klaus Neymeyr, 2,5 Jeroen Jansen, 4 Age K. Smilde, 6 and Sander Woutersen 1
1Van’t Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, the Netherlands; 2Institut für Mathematik, Universität Rostock, Rostock, Germany; 3Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, the Netherlands; 4Institute for Molecules and Materials, Radboud University, Nijmegen, the Netherlands; 5Leibniz-Institut für Katalyse, Rostock, Germany; and 6Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands

ABSTRACT Intermediate species are hypothesized to play an important role in the toxicity of amyloid formation, a process associated with many diseases. This process can be monitored with conventional and two-dimensional infrared spectroscopy, vibrational circular dichroism, and optical and electron microscopy. Here, we present how combining these techniques provides insight into the aggregation of the hexapeptide VEALYL (Val-Glu-Ala-Leu-Tyr-Leu), the B-chain residue 12–17 segment of insulin that forms amyloid fibrils (intermolecularly hydrogen-bonded β-sheets) when the pH is lowered below 4. Under such circumstances, the aggregation commences after approximately an hour and continues to develop over a period of weeks. Singular value decompositions of one-dimensional and two-dimensional infrared spectroscopy spectra indicate that intermediate species are formed during the aggregation process. Multivariate curve resolution analyses of the one and two-dimensional infrared spectroscopy data show that the intermediates are more fibrillar and deprotonated than the monomers, whereas they are less ordered than the final fibrillar structure that is slowly formed from the intermediates. A comparison between the vibrational circular dichroism spectra and the scanning transmission electron microscopy and optical microscope images shows that the formation of mature fibrils of VEALYL correlates with the appearance of spherulites that are on the order of several micrometers, which give rise to a “giant” vibrational circular dichroism effect.

INTRODUCTION

The formation of amyloid aggregated by incorrectly folded proteins is associated with over 50 diseases in which (transient) oligomeric species are thought to play a particularly toxic role. Obtaining a detailed biophysical and structural understanding of such intermediate species is challenging with conventional techniques. Here, we present a combination of three vibrational spectroscopy techniques and two microscopy techniques (each sensitive to a different length scale), analyzed with novel multivariate curve resolution techniques that allow an unprecedented molecular view over length scales ranging from nano- to millimeters and timescales ranging from picoseconds to weeks. We apply the approach to reveal and characterize intermediates in the amyloid aggregation by the hexapeptide VEALYL, the most amyloidogenic segment of insulin.

SIGNIFICANCE Amyloid aggregation by incorrectly folded proteins is associated with over 50 diseases in which (transient) oligomeric species are thought to play a particularly toxic role. Obtaining a detailed biophysical and structural understanding of such intermediate species is challenging with conventional techniques. Here, we present a combination of three vibrational spectroscopy techniques and two microscopy techniques (each sensitive to a different length scale), analyzed with novel multivariate curve resolution techniques that allow an unprecedented molecular view over length scales ranging from nano- to millimeters and timescales ranging from picoseconds to weeks. We apply the approach to reveal and characterize intermediates in the amyloid aggregation by the hexapeptide VEALYL, the most amyloidogenic segment of insulin.

Submitted July 25, 2019, and accepted for publication May 11, 2020.
*Correspondence: s.j.roeters@chem.au.dk

Steven J. Roeters’s present address is Kemisk Institut, Aarhus Universitet, 8000 C Aarhus, Denmark.
Matthijs R. Panman’s present address is Department of Chemistry and Molecular Biology, University of Gothenburg, 40530 Gothenburg, Sweden.

Editor: Elsa Yan.
https://doi.org/10.1016/j.bpj.2020.05.026
© 2020 Biophysical Society.
Alzheimer’s), α-synuclein (4,5) (related to Parkinson’s), and islet amyloid polypeptide (6,7) (related to type-II diabetes). Monitoring the formation of amyloids from monomers through intermediates to mature fibrils is challenging with conventional techniques because they are not able to probe the many time- and length scales over which the process occurs. Here, we demonstrate how a combination of microscopy and vibrational spectroscopy techniques allows for a holistic comprehension of the amyloid formation process. We apply a combination of Fourier transform infrared spectroscopy (FTIR, also known as one-dimensional infrared spectroscopy (1D-IR)), nonlinear two-dimensional infrared spectroscopy (2D-IR), and vibrational circular dichroism (VCD) and perform optical microscopy and scanning transmission electron microscopy (STEM) to study the amyloid aggregation of the hexapeptide Val-Glu-Ala-Leu-Tyr-Leu (VEALYL). The three vibrational spectroscopy techniques complement each other because they are each sensitive to a different structural length scale. We focus on spectra of the amide-I band (1600–1700 cm\(^{-1}\)) because it is sensitive to the conformation of proteins. The combination of the omnipresence of the amide group in proteins (as peptide linkage between amino acids), and its large amide-I transition dipole moment, which leads to strong coupling between the local modes, results in unique spectral features across the bands that are sensitive to the tertiary structure and environment of proteins. In this frequency range, the IR frequencies are strongly influenced by the strengths of the hydrogen bonds the peptide groups engage in but also by the coupling between the amide-I modes (8). With 2D-IR, these couplings can specifically be determined, and one is more sensitive to structures that lead to large, delocalized normal modes (9), making the technique particularly suited to study amyloid fibrils (10–23). The technique has been applied previously to reveal the structure of intermediates formed in amyloid aggregation (24,25). VCD is, yet again, sensitive to even bigger structures because of the strong enhancement that can take place when the amyloid fibrils start to form bundles, twisting around each other (26–29). By combining these vibrational spectroscopy techniques with microscopy, we can correlate the vibrational spectra with morphology.

VEALYL, the hexapeptide segment of insulin that has the highest propensity for amyloid formation (30), is believed to play an important role in the aggregation of insulin (31,32). This 51-residue protein is an important glucose-uptake-mediating hormone, which is widely used to treat diabetes. When insulin aggregates during the manufacturing, purification, storage, or infusion, its activity is lost, which renders it therapeutically ineffective, and injected protein aggregates can even trigger an immune response (31).

A recent ion mobility mass spectrometry and gas-phase infrared spectroscopy study has provided insight into the structure of the early oligomers (up to nine monomers), revealing the presence of β-sheet structure in oligomers composed of more than three monomers, whereas smaller oligomers are composed of random coil structure (33). Amyloid VEALYL aggregates formed after 3 weeks of incubation have been studied in their crystalline form with x-ray crystallography (32) and in situ (in the incubation solution) form with NMR spectroscopy (34), revealing a nearly complete β-sheet conformation with little variation in the unit cell structure for various macromolecular polymorphs. The vibrational spectroscopy and microscopy methods employed here allow for an in situ and nondestructive investigation of VEALYL aggregation all the way from separate monomers to mature fibrils. By applying two multivariate curve resolution techniques (FACPACK and multivariate curve resolution (MCR) with alternating least squares (MCR-ALS)) to the recorded spectroscopic data, to our knowledge, new structural characteristics of the intermediate species are revealed.

**MATERIALS AND METHODS**

**Sample preparation**

VEALYL was bought in lyophilized form from GL Biochem (Shanghai, China). To exchange the protons present in the amide groups of the backbone of the peptide and to remove the remaining amount of trifluoroacetic acid that was used in solid-phase peptide synthesis, VEALYL was dissolved at a concentration of 18.5 mM in D\(_2\)O, after which 0.7 μL of 35 wt% DCl was added for each milliliter of solution (equivalent to a 0.084 M solution of DCl). The samples were then vortexed at 1400 revolutions per minute for 1 min to properly dissolve the peptides, after which the peptide solution was left to exchange \(^1\)H for \(^1\)H for 20 min. Subsequently, the samples were rapidly cooled to 77 K by immersion of the Eppendorf tubes in liquid nitrogen. After this, the samples were lyophilized by placing them in vacuum and freeze-drying for 8 h in a ScanVac CoolSafe (LaboGene, Allerød, Denmark), in which 1 mL of peptide solution was fully lyophilized.

To minimize hydrogen exchange from the air with the sample, all subsequent steps were performed inside a glove box that is continuously flushed with N\(_2\). The lyophilized VEALYL powder was resolvated in a solution containing 150 mM NaCl and 50 mM phosphate, leading to a pH of ~6. The aggregation was started according to the procedure described in (32): by lowering the pH to 2.5 with the addition of a few microliters of 1 M DCl. The solution was then vortexed at 1400 rpm for exactly 1 min. Subsequently, the pH was measured with a Thermo Scientific Orion 2-Star Plus pH Meter (Thermo Fisher Scientific, Waltham, MA). Unless noted differently, a VEALYL concentration of 11 mM was used.

**1D- and 2D-IR and VCD spectroscopy**

A Bruker Vertex 70 FTIR spectrometer (Bruker, Billerica, MA) is used to measure the 1D-IR (FTIR) spectra of 5-μL drops of sample solution in between two CaF\(_2\) windows. A grease 50-μm spacer prevents the two CaF\(_2\) windows from touching each other and seals the sample solution to prevent atmospheric water vapor from changing the H/D ratio. The FTIR data presented in Fig. 1 are normalized on the integrated area of the amide-I region (1600–1700 cm\(^{-1}\)); without this normalization step, the effect of the slow acidic hydrolysis of the peptide bonds is even more pronounced.

The setup used to measure the 2D-IR spectra has been described in detail before (35). In short, we generate 1 kHz of 3 mJ pulses with a central wavelength of 794 nm and convert this in an optical parametric amplifier into mid-IR (~20 μL, ~6100 nm) that has a spectral full width at half max (FWHM) of 150 cm\(^{-1}\). The beam is then split into a probe and reference beam (each 5%), and a pump beam (90%) that is aligned through a Fabry-Pérot interferometer.
Subsequently, the pump beam is narrowed to a FWHM of 10 cm$^{-1}$. The pump and probe beams are overlapped in the sample with a 1.5-ps delay in order to monitor the amyloid aggregation of an 11-mM VEALYL solution after lowering the pH. The first sign of amyloid aggregation are the $\beta$-sheet peaks in the 1D-IR and 2DIR spectra that are observed after a lag phase of 30 min to 3 h (this variation was observed within five repeated experiments) of incubation at a VEALYL concentration of 11 mM, whereas at neutral pH, the amide-I spectrum remains unchanged with respect to the monomeric random coil spectrum, even after 3 months. We find that increasing VEALYL concentration (with a minimum concentration of 11 mM, whereas at neutral pH, the amide-I spectrum remains unchanged with respect to the monomeric random coil spectrum, even after 3 months. We find that increasing VEALYL concentration (with a minimum concentration of 11 mM) greatly enhances the amyloid fibril formation by VEALYL (30,32). In Fig. 7, an overview is presented of the results of the five experimental techniques employed to monitor the amyloid aggregation of an 11-mM VEALYL solution after lowering the pH. The time-dependent 1D-IR spectra of the amide-I and adjacent sidechain-absorption region (Fig. 1) show a conformational transition after the pH is lowered from 7 to 2.5, which protonates the sidechain of the glutamic acid residue ($pK_a = 4.15$), thereby removing its negative charge. Lowering the pH greatly enhances the amyloid fibril formation by VEALYL (30,32).

**RESULTS AND DISCUSSION**

**General observations**

As in previous studies of amyloid formation by VEALYL (30,32,34), we trigger the aggregation by lowering the pH to 2.5, which protonates the sidechain of the glutamic acid residue ($pK_a = 4.15$), thereby removing its negative charge. Lowering the pH greatly enhances the amyloid fibril formation by VEALYL (30,32). In Fig. 7, an overview is presented of the results of the five experimental techniques employed to monitor the amyloid aggregation of an 11-mM VEALYL solution after lowering the pH. The first sign of amyloid aggregation are the $\beta$-sheet peaks in the 1D-IR and 2DIR spectra that are observed after a lag phase of 30 min to 3 h (this variation was observed within five repeated experiments) of incubation at a VEALYL concentration of 11 mM, whereas at neutral pH, the amide-I spectrum remains unchanged with respect to the monomeric random coil spectrum, even after 3 months. We find that increasing VEALYL concentration (with a minimum concentration of 2 mM), temperature, and agitation decreases the aggregation tendency.

**Early aggregation: intermediate structures observed with time-dependent 1D- and 2D-IR spectroscopy**

**1D-IR (FTIR) spectroscopy**

Time-dependent 1D-IR spectra of the amide-I and adjacent sidechain-absorption region (Fig. 1) show a conformational transition after the pH is lowered from 7 to 2.5 at $t = -5$ min (these 5 min are required to prepare the IR sample and to completely purge the sample chamber of the IR spectrometer from atmospheric water vapor). After a lag phase of 30 min, the peptides aggregate during a growth phase of ~1000 min into mature fibrils that are mainly present during the final plateau regime that is subsequently reached.

The largest change is observed in the amide-I band, where a single broad peak centered at ~1645 cm$^{-1}$ transforms into...
an intense peak at \( \approx 1622 \text{ cm}^{-1} \) and a small peak at \( \approx 1682 \text{ cm}^{-1} \). This change indicates a molecular transition from random coil to amyloid structure (42). The sidechain absorptions of Tyr-OD (\( \nu(\text{CC}) \), \( \delta(\text{CH}) \)) and deprotonated and protonated Glu (\( \nu_{\text{as}}(\text{COO}^-) \) and \( \nu(\text{C}=\text{O}) \)) are visible at \( \approx 1515, \approx 1575, \) and \( \approx 1720 \text{ cm}^{-1} \), respectively (8). The fact that the frequency of the Glu (\( \nu(\text{C}=\text{O}) \)) sidechain vibration is higher than \( 1706 \text{ cm}^{-1} \) indicates that the C=O is not fully hydrogen bonded (43), which can be expected in the dry steric-zipper (32) environment inside the fibril core (Fig. S2). The intensities (normalized to the total amide-I area) of the \( \beta \)-sheet peaks reach a maximum at \( \approx 70 \text{ h} \), after which they decrease again (see also fifth row of Fig. 7). We ascribe this decrease to slow acidic hydrolysis of the peptide bonds. As a result of this process, the pH increases, which leads to a partial back conversion of amyloid to random coil structure (Fig. S1).

Singular value decomposition analysis. To estimate how many spectroscopic species are present in the data, we first perform a singular value decomposition (44) (SVD; see Fig. 2). When such an analysis is applied to a time-dependent spectral data set, the data are decomposed in orthogonal spectral and temporal profiles, and the singular value of each component indicates how much the component contributes to the total data set.

An SVD is a factorization of the two-dimensional data set \( D \) of the following form:

\[
D = U \Sigma V^T = \sum_{i=1}^{\min(m,n)} u_i \sigma_i v_i^T,
\]

where the columns of \( U(m \times m) \) are the left singular vectors of \( D \), \( \Sigma (m \times n) \) is a diagonal matrix containing the singular values \( \sigma_i \), on its diagonal in a decreasing order, and the columns of \( V(n \times n) \) are the right singular vectors. Both \( U \) and \( V \) are orthonormal. Moreover, \( u_i \) is the \( i \)-th column of \( U \) and hence the \( i \)-th left singular vector of \( D \). Analogously, \( v_i \) is the \( i \)-th column of \( V \) and hence the \( i \)-th right singular vector of \( U \).

The singular values can be viewed as weights and are multiplied with the left and right singular vectors when estimating \( D \). A large singular value indicates that the corresponding singular vectors carry systematic information, whereas a small singular value indicates that the corresponding singular vectors carry noise.

In the SVD of the time-dependent 1D-IR data of aggregating VEALYL (with \( m = 417 \) time points and \( n = 300 \) frequencies) presented in Fig. 2, the left singular vectors describe the temporal profiles, and the right singular vectors describe the spectral profiles of the species. Before the SVD, \( D \) was column-wise mean-centered. The first three singular values are clearly larger than the remaining singular values. Therefore, it appears that with 1D-IR spectroscopy, only three species can be resolved in the amyloid aggregation process of VEALYL. The right singular vector of the fourth component especially shows that this component mostly contains experimental noise that cannot be attributed to absorptive features of the protein.

Because of the fact that both \( U \) and \( V \) are orthonormal, an SVD provides a unique solution. However, this orthogonality constraint causes the left and right singular vectors to not, respectively, resample the pure temporal and spectral profiles of the physical species.

Determination of the rotational ambiguity with FACPACK. To determine the concentration and spectral profiles of physical species that give rise to a time-dependent spectral data set, MCR techniques can be applied. However, such techniques suffer from the so-called rotational ambiguity of the solution. This means that, in general, a factorization is not unique, and a continuous range of possible factorizations exists. In other words, there are many different linear combinations of temporal and spectral profiles that can reproduce the experimental data. In this respect, non-SVD decompositions differ from an SVD decomposition; because of orthogonality between the singular vectors, an SVD decomposition gives only one unique...
decomposition. Due to the orthogonality constraint, SVD decompositions do not provide a “proper” pure component decomposition, and return temporal and spectral profiles with negative entries. To reconstruct all possible concentration and spectral (absorption) profiles, $C$ and $A$, respectively, of the pure components, the first $s$ components of a truncated SVD are transformed such that both $C$ and $A$ only contain nonnegative entries (45). Using the software package FACPACK (46,47) the area of feasible solutions (AFS; see e.g., (48–53)) is determined according to

$$D = U \cdot \Sigma \cdot T^{-1} \cdot T \cdot V^T,$$

with $U \in \mathbb{R}^{n \times s}$, $\Sigma \in \mathbb{R}^{s \times s}$, and $V \in \mathbb{R}^{n \times s}$ as the truncated factors of the SVD and $T \in \mathbb{R}^{n \times n}$ as the transformation matrix that completely represents all possible nonnegative concentration and spectral profiles ($C \in \mathbb{R}^{n \times s}$ and $A \in \mathbb{R}^{s \times n}$, respectively), where $D = CA$ follows from the Lambert-Beer law (the AFS computation is explained in detail in the Supporting Materials and Methods). The AFS sets cover all possible (concentration and spectral) profiles that can be part of an MCR solution.

Based on the SVD analysis, we assume that there are three components with different absorption spectra in the calculation of the AFS. The data set might include contributions from more species, but because of the noise and linear dependencies between the components, it is only possible to identify three linearly independent components. Hence, we compute decompositions of $D$ with respect to three components ($C$ is an $n \times 3$ matrix and $A$ is a $3 \times n$ matrix). Furthermore, we compute the AFS also with respect to three components.

The resulting feasible profiles are shown in Fig. 3, and the AFS sets are shown in Fig. S3. With the applied constraints (nonnegativity for both the spectral and temporal profiles) and a tolerance for negative entries of 0.0086 OD, there is a range of feasible spectra and associated concentration profiles (time dependencies) that are all equally compatible with the experimental data. The spectral shapes of all of the possible intermediate species indicate that with time, the amount of $\beta$-sheet increases, and the amount of random coil structure decreases. The sidechain $\nu_{\text{as}}$(COO$^-$) and $\nu$(C=O) peaks show that the Glu residues of the intermediates (green curves) are more deprotonated than the Glu residues of the monomers. The presence of deprotonated Glu residues in the amyloid aggregates formed by VEALYL under these experimental conditions has been observed before in NMR experiments (34). This deprotonation might be due to a lower $pK_a$ of the Glu sidechain in the fibrillar structure: the formation of a hydrogen bond between deprotonated Glu side chains and theNH$_3^+$ (Val) terminus of the peptide stabilizes the carboxylate ion form of Glu (54,55). Alternatively, the aforementioned pH increase due to the slow acidic hydrolysis of peptide bonds could explain the deprotonation of the Glu side chains in the more mature intermediates.

**MCR-ALS analysis.** Similar solutions are found when an MCR-ALS analysis is performed (56,57). With this method, the physical components that give rise to the variations measured in the experiment are determined in a different manner: the data are fitted while various constraints and assumptions are applied (58) (see Supporting Materials and Methods for the computational methods).

By performing an MCR-ALS analysis on the 1530–1800 cm$^{-1}$ data assuming nonnegative temporal and spectral profiles, we find evidence for the appearance and disappearance of intermediate species during the aggregation process, similar to the FACPACK analysis (see Fig. S4–S6). The intermediate spectra obtained with the MCR-ALS analysis reflect structures that are more fibrillar and deprotonated than the monomeric species.

**2D-IR spectroscopy**

2D-IR spectroscopy of the amide-I region is a very sensitive tool to study the secondary structure of (aggregating) proteins (9). The spectra are recorded by exciting the protein solution at a specific frequency $\nu_{\text{pump}}$ and probing the resulting changes in the absorption over the entire IR frequency range. Scanning $\nu_{\text{pump}}$ and plotting the absorption change as a function of $\nu_{\text{pump}}$ and $\nu_{\text{probe}}$, one can obtain 2D-IR spectra that can be regarded as the vibrational analog of 2D-NMR spectra. When two normal modes $A$ and $B$ are coupled, exciting mode $A$ ($\nu_{\text{pump}} = \nu_A$) causes an absorption change at the frequency of mode $B$ ($\nu_{\text{probe}} = \nu_B$) and vice versa, and this gives rise to crosstalk patterns in the 2D-IR spectrum. 2D-IR spectroscopy is particularly suited to study the formation of amyloid fibrils (9–25) Because delocalized vibrational eigenmodes—which are abundant in amyloid fibrils—give a stronger 2D-IR signal than localized eigenmodes, the 2D-IR spectrum can provide additional structural insight that cannot be deduced from the 1D-IR (FTIR) spectrum, crosstalks between eigenmodes give additional information on the fibril structure, and the shape of the peaks also gives additional information about the heterogeneity of the fibril structure (21), also through the anisotropy $R = (\Delta A_{\text{par}} - \Delta A_{\text{perp}})/(\Delta A_{\text{par}} + 2\Delta A_{\text{perp}})$, with $\Delta A_{\text{par, perp}}$ as the crosstalk intensity for parallel and perpendicular pump versus probe polarization, from which the relative orientation of the associated modes can be derived with $R = (3\cos(\theta)^2 - 1)/5$ (9).

As expected, the transition from a broad distribution of random coil modes centered around 1645 cm$^{-1}$ into two sharp amyloid $\beta$-sheet modes at $\sim$1622 and $\sim$1682 in (mature) fibrils (Fig. 1) is also visible in the 2D-IR spectra (Fig. 4). The anisotropy of the crosstalk at ($\nu_{\text{probe}}, \nu_{\text{pump}} = 1682, 1622$) cm$^{-1}$ is $R = -0.21 \pm 0.03$, indicating that there is a 90 $\pm$ 11$^\circ$ angle between the modes that the 1622 and the 1682 cm$^{-1}$ peaks are composed of, which is consistent with the assignment of the modes to the
and $\nu_1$ of an antiparallel $\beta$-sheet (59,60). This angle is reproduced with one-exciton spectral calculations (9,23) based on the crystal structure of the amyloid fibrils formed by VEALYL (Protein Data Bank, PDB: 2OMQ), by which the 2D-IR spectrum can also be reproduced well (see Fig. S7), indicating that the structures formed at the end of the aggregation experiment are similar to those obtained by Ivanova et al. (30).

Also, the ratio of the 1D-IR/2D-IR spectrum can be used to determine the magnitude of the transition dipole moment of a given isolated mode (61). However, for amyloids, the vibrational modes extend over at least a dozen strands, which approaches the asymptotic number of strands after which the delocalization no longer influences the $\beta$-sheet absorption (20), thereby limiting the discriminating ability of the transition dipole strength determination. Subtle variations in interlocal mode differences and dihedral angles that are present in the fibril structure lead to a distribution of couplings whose average value determines the exact frequency of the $\nu_\perp$ $\beta$-sheet mode. Lomont and co-workers (20) ascribe the fact that they measured a variation $>2$ in the transition dipole moment magnitude to variations in the couplings. Because of the difficulty of reproducing the transition dipole magnitudes in heterogeneous amyloid samples, which we also
observed in previous measurements in our laboratory of the transition dipole moment of the low-frequency β-sheet mode of α-synuclein (unpublished data), we did not determine them here for VEALYL. The procedure we follow here to prepare the amyloid VEALYL samples (i.e., lowering the pH right before pipetting the sample volume into an IR cell) results in a sample heterogeneity that affects time-dependent measurements of VEALYL aggregation with 2D-IR spectroscopy. Because the overlap of the focused pump and probe beam is typically of the order of 250 μm, there is a certain probability that the beams are not focused at a nucleation site (see Fig. 7 for an estimate of the typical sizes of the aggregates in the different aggregation phases), as can be seen by comparing the lower two spectra in Fig. 4 that have been recorded 500 μm from each other, which contain different amounts of random coil and amyloid β-sheet structure.

The line shape of the 2D-IR peaks shows that there is a large spectral inhomogeneity, as indicated by the inverted slope of the line through the zero crossings. A small (close to 0), inverted slope through these points indicates a mode with a small spectral inhomogeneity, whereas an inverted slope close to unity indicates a large spectral inhomogeneity (62–64). When the VEALYL molecules are in their monomeric state, the inverted slope is close to unity; with increasing aggregation, the inverted slope becomes smaller, indicating a more homogeneous structure, as can be expected for the amyloid crystals that are formed.

An SVD analysis of the combined 2D-IR data set presented in Fig. 4 in its unrolled, or flattened, 122 (time points) × 1024 (frequency points) form (see Fig. S8) also shows that there are three components that describe absorptive features of the protein, whereas other components describe noise contributions (see Fig. S9), corroborating the conclusions derived from the SVD analysis on the 1D-IR data set.

**Advanced aggregation: STEM and optical microscopy and VCD spectroscopy reveal that the appearance of higher-order structures correlates with the presence of mature fibril structures**

**VCD spectroscopy**

VCD measures the difference in absorbance of left- and right-handed circularly polarized light (65,66). When applied to the amide-I region, the chiral properties of proteins can be probed. Time-dependent VCD spectra of aggregating VEALYL show maxima at 1567 and 1621 cm\(^{-1}\) and minima at 1462, 1510, 1635 (with a shoulder at 1649), 1678, and 1710 cm\(^{-1}\), which all grow with time (see Fig. 5). We assign these peaks to \(\nu_{as}(\text{COO}^-)\) modes of deprotonated Glu residues, amide-I modes of amyloid β-sheets, \(\nu_{as}(\text{COO}^-)\) modes of deprotonated Glu residues, \(\nu(\text{CC}), \delta(\text{CH})\) modes of Tyr-OH residues, other amide-I modes of amyloid β-sheets, and to \(\nu_{as}(\text{C}=\text{O})\) of protonated Glu residues, respectively. Interestingly, the amide-I VCD spectra of VEALYL fibrils are “reverse fibril” spectra, which are nearly mirror images with respect to the canonical “normal fibril” spectra published by Kurouski et al. for insulin at the same pH (~2.5) (28,29,67). The reversed VCD
The spectrum of VEALYL compared to the VCD spectrum of full-length insulin fibrils indicates a difference in \( \beta \)-sheet conformation and/or stacking geometry. The evidence presented in (29) indicates that differences in the supramolecular organization of filaments are the principal underlying cause of reversed VCD spectra of amyloid fibrils. The lack of an observable twist in the STEM images (see Figs. 7 and S10) is consistent with this explanation, although only large-diameter twists would be observable with this technique. But besides a difference in the supramolecular organization, the origin of the phase reversal may also lie in the different \( \beta \)-sheet conformation, which is thought to differ considerably between VEALYL (32) and insulin (68).

The VCD signal still grows by a factor of \( \sim 7 \) after \( t > 50 \) h, whereas the shape of the 1D-IR spectra hardly changes from this point onwards. This indicates that the VCD intensity is not determined by structural features at the nanometer scale (as these would also affect the 1D-IR spectra).

**Optical and electron microscopy**

On the other hand, the light and electron microscopy images (see Figs. 7 and S10) do reveal a large conformational change that occurs from this time point onwards: the appearance of higher-order aggregates known as “spherulites” (69,70). This makes it likely that the observed “giant” VCD effect (29,71,72) is caused by the appearance of chiral aggregates that are of the order of the wavelength of the probe light, possibly due to a resonance effect. The twist of in silico VEALYL fibrils observed in molecular dynamics simulations in which the crystal structure that was resolved in ref. (32) was relaxed (34) may also contribute to the strong VCD signal.

Additionally, the STEM images reveal the appearance of curved structures after 2 h of incubation, which are probably protofibrils: intermediate structures that already have some fibrillar features but are not yet crystalline. Given the coexistence of these structures with the intermediate species observed in the 1D-IR data (see Fig. 7), these are probably composed partly of random coil structure and partly of intermolecular \( \beta \)-sheet.

An SVD and MCR-ALS analysis of the time-dependent VCD spectra of aggregating VEALYL (data not shown) shows that only a single component is present in the data set. The data can be fitted well with the Avrami equation, which describes crystallization in three dimensions (73–75):

\[
f(t) = \frac{I_\infty - I(t)}{I_\infty - I_0} = 1 - e^{-k^n},
\]

where \( f(t) \) is the fraction of crystallized material, \( I_\infty \) is the asymptotic VCD intensity at infinite time, \( I(t) \) is the time-dependent VCD intensity, \( I_0 \) is the intensity at \( t = 0 \), \( k \) is a constant containing only time-independent factors, and the integer exponent \( n \) yields information about the mode of nucleation and the shape of the developing structures (76), with \( n = 1 \) indicating a needle-like growth, \( n = 2 \) indicating a disk-like growth, \( n = 3 \) indicating a sphere-like growth, and \( n = 4 \) indicating three dimensions of growth and a constant nucleation rate (77). The fact that the data are fitted well with the noninteger value of \( n = 2.5 \) (see Fig. 6) indicates a dimensionality that is in between disk-like and sphere-like growth, which is consistent with the STEM images (Fig. 7 and Fig. S10). However, interpretation of the Avrami constant is difficult; there have been almost as many cases in which \( n \) did match the observed crystal growth as when it did not (78).

**CONCLUSIONS**

By combining three vibrational spectroscopy techniques with two microscopy techniques, the aggregation of
FIGURE 7 An overview of the data, analysis, and interpretation of the vibrational spectroscopy and imaging techniques used to study the amyloid formation of VEALYL. Row 1 shows the qualitative description and artist impressions of the structural interpretations of the various aggregation phases. Row 2 shows the 1D-IR (FTIR) spectra. Row 3 shows the 2D-IR spectra. Row 4 shows the VCD spectra. Row 5 shows the normalized profiles of the intensity at VEALYL’s amyloid β-sheet frequency (1622 cm⁻¹) for the various techniques and concentration profiles of the three species from the FACPACK analysis of the 1D-IR data. Row 6 shows the light microscope images. Row 7 shows the STEM images. The 2D-IR spectra were recorded only up to 44 h of the same sample due to experimental procedural reasons (see main text). The third and fourth STEM images are recorded of the same sample with a VCD intensity that is 50% of the maximum, and an amyloid β-sheet IR signal that is already at its maximum because the different species (as shown in the fifth row) coexist when the intermediates and the mature fibrils are formed. To see this figure in color, go online.
VEALYL can be monitored from solvated monomers to intermediates and to insoluble mature fibrils. In Fig. 7, we give an overview of the different stages of this process, as observed in our experiments. The figure shows that each technique is sensitive to a specific range of length scales. By applying this combination of methods to other, potentially more complex, amyloid formation processes, detailed descriptions of the mechanism of the formation could be obtained. In the case of the amyloid aggregation of VEALYL, there appears to be limited structural variation among the observed molecular species that play a role in the amyloid aggregation process; they are composed of random coil structure and antiparallel β-sheets with an increasing amount of β-sheet structure with increasing aggregation. This is in line with previous ion mobility-spectrometry-mass spectrometry and gas-phase infrared spectroscopy observations for VEALYL (33). The molecular sensitivity of the vibrational techniques has now revealed that there are differences in the protonation state of the different molecular species that play a role in the aggregation. Also, the eventual observation of homogeneous and bundled fibrils, as indicated by the 2D-IR line shape and the VCD intensity, respectively, is in line with previous studies (32,34). By using the here-employed vibrational spectroscopy techniques to observe the transformation of monomers to oligomers and finally to mature fibrils in an in situ manner, we can continuously monitor the molecular evolution.

We show how MCR analyses of 1D- and 2D-IR amide-I spectra of amyloid proteins can provide insight into the mechanism of molecular processes. In the case of the amyloid aggregation of VEALYL, the analysis revealed intermediate species with an increased deprotonation state and an increased intermolecularly hydrogenbonded β-sheet structure, even though the exact spectrum is subject to rotational ambiguity. It should be noted that the spectrum of the intermediate species (depicted in Fig. 3) could reflect a group of ambiguities. It should be noted that the spectrum of the intermediate species (depicted in Fig. 3) could reflect a group of relatively stable intermediates that share certain distinct spectral features, or a group of relatively stable intermediates in a continuum of incrementally growing aggregates.

When vibrational spectra are complemented by spectral calculations (9,71), e.g., on molecular structures obtained with molecular dynamics simulations, more detailed structural information can be derived. Also, the application of advanced analysis methods on the time-dependent 2D-IR data (like parallel factor analysis (PARAFAC) or Tucker modeling techniques that have been developed especially for three-dimensional data sets) could allow an even more detailed analysis of amyloid systems (79).

For amyloid systems whose aggregation cannot be triggered like that of VEALYL, a large variation in the so-called “lag phase” might hinder connecting the results from the different techniques straightforwardly. However, in that case, an internal calibration of the different spectroscopic measurements can be obtained by deriving the 1D-IR spectrum from the 2D-IR and VCD spectra, as has been done in Fig. S11 for the data presented in this article.

Overall, we present a, to our knowledge, novel combination of techniques to study amyloid formation and show that many structural details can thus be revealed on length scales ranging from nano- to millimeters and timescales ranging from picoseconds to weeks.

SUPPORTING MATERIAL
Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.05.026.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS
We thank Gertien Smits for suggesting the slow hydrolysis reaction as an explanation for the loss of peptide bonds at long timescales.

S.J.R. acknowledges funding from the Villum 2515 Foundation (Experiment Grant 22956) and the 2516 Lundbeck Foundation for a postdoctoral fellowship.

REFERENCES


52. Golshan, A., H. Abdollahi, ... R. Tauler. 2016. A review of recent methods for the determination of ranges of feasible solutions resulting in...


