J|A|C|S COMMUNICATIONS

Published on Web 01/29/2009

Observation of β-Sheet Aggregation in a Gas-Phase Tau-Peptide Dimer

Timothy D. Vaden,* Sally A. N. Gowers, and Lavina C. Snoek*

Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, U.K.

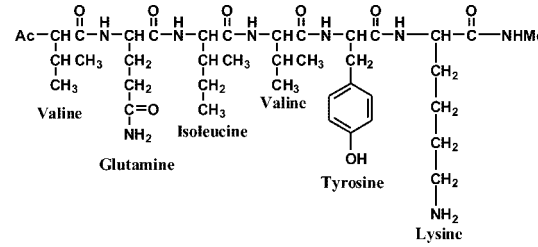
Received October 1, 2008; E-mail: timothy.vaden@chem.ox.ac.uk; lavina.snoek@chem.ox.ac.uk

The formation of insoluble amyloid fibrils from native-state proteins is linked to diseases such as Alzheimer's disease (AD) and type-II diabetes. Amyloid deposits develop when proteins misfold out of their native conformations and aggregate into insoluble fibrils,[1] characterized by a cross-β motif in which β-strands run orthogonal to the fibril axis and repetitive hydrogen bonding extends parallel to the axis.[2] This cross-β structure may be the global minimum energy conformation for a wide variety of proteins,[3] and identifying this structural motif in small model peptide systems and characterizing it under different conditions can yield valuable clues about its stability, both in general and in specific systems, and about the molecular-level details of amyloid formation. In this communication, we characterize an amyloidogenic peptide dimer in the isolated environment of a cold molecular beam using IR spectroscopy and density functional theory (DFT) calculations. These first results suggest the formation of a β-sheet conformation and highly motivate further analysis of this interesting system.

In AD, the tau protein forms intracellular amyloid fibrils in neurons.[4,5] In these fibrils, the ^{306}VQIVYK^{311} sequence adopts a β-strand conformation and forms parallel β-sheets with other ^{306}VQIVYK^{311} sequences that combine in a face-to-face interdigitating dry "steric zipper" structure.[6] The hexapeptide VQIVYK models this key amyloidogenic peptide sequence and forms amyloid-like fibrils with the same cross-β structure found in full tau amyloid fibrils.[6] This peptide has a clear tendency to form amyloid-like aggregates and is therefore a good model system for studying amyloid formation in different environments. In this context, understanding the inherent structure and stability of the ^{306}VQIVYK^{311} segment in an isolated (i.e., gas-phase) aggregate is of fundamental importance. We recently characterized the capped, single gas-phase Ac−VQIVYK−NHMe peptide (Scheme 1) isolated in a cold molecular beam using IR/UV hole-burning spectroscopy combined with DFT calculations.[7] The isolated peptide adopts a β-hairpin (Figure 1B) with two extended halves stabilized by backbone hydrogen bonds, demonstrating its inherent preference for extended backbone regions. We also characterized the peptide in its protonated state (on the K side chain) and showed that $NH_3^+ \rightarrow$ O=C interactions destroy the β-hairpin conformation.[7] In the "real" fibril, the K NH_3^+ is probably bound to other molecules and cannot interact with the backbone;[8] thus, the neutral peptide is likely a more relevant amyloid model. To form an amyloid fibril, the β-hairpin would have to unfold, adopt a β-strand conformation, and then aggregate into a β-sheet unit cell, where noncovalent interactions between β-strands compensate for the energetic cost of unfolding. In this context, Fricke et al.[9] have reported gas-phase β-sheet dimer formation in tripeptides.

To explore the transition from the hairpin-like monomer to the amyloid-like dimer for the Ac−VQIVYK−NHMe peptide, we studied the gas-phase (Ac−VQIVYK−NHMe)$_2$ complex by recording the IR spectrum in the structurally diagnostic OH/NH/CH stretch region and comparing it with quantum-chemical predictions

Scheme 1. Structure of the Capped Model Peptide Representing the Key Amyloidogenic Region ^{306}VQIVYK^{311} of the Tau Protein



and with the structures found in the disease-related tau protein. The Ac−VQIVYK−NHMe peptides (Celtek) were vaporized by laser desorption[7] and seeded into a pulsed supersonic argon expansion, where they subsequently formed the noncovalent dimeric complex upon cooling to ~10−15 K[10] (the vibrational temperature may be slightly higher because of inefficient cooling for this large molecule). Experimental details are provided in the Supporting Information. The OH/NH/CH stretch IR spectrum was recorded using the IR/UV hole-burning technique[11] as described previously.[7]

While we were unable to resolve UV features between 281−283 nm, the IR spectrum of (Ac−VQIVYK−NHMe)$_2$ does not change with different UV probe wavelengths. The IR spectrum measured at 282 nm is shown in Figure 1A (upper trace). On the basis of successful structural assignments in previous IR investigations of neutral amino acids and peptides,[12−15] bands above 3500 cm^{-1} correspond to free OH stretch modes [here, the only candidate is from Y, observable at 3652 cm^{-1} in the monomer spectrum (Figure 1A, lower trace)], bands between 3100 and 3500 cm^{-1} to free NH stretch and hydrogen-bonded OH/NH modes, and bands below 3100 cm^{-1} typically to CH stretch modes.

A few structural clues can immediately be inferred by comparing the dimer spectrum to the monomer spectrum. The absence of the free OH band in the (Ac−VQIVYK−NHMe)$_2$ spectrum indicates

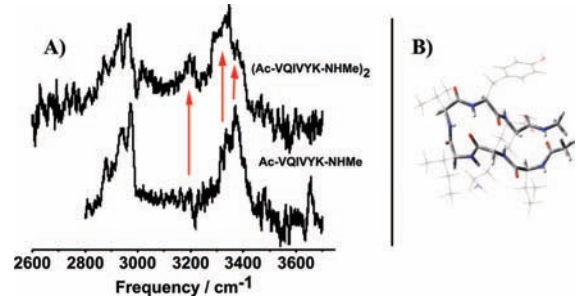


Figure 1. (A) IR spectrum in the OH/NH/CH stretch region for the dimer (Ac−VQIVYK−NHMe)$_2$ (upper trace) compared with the spectrum of the monomer (lower trace, from ref 7). (B) Conformation of the gas-phase monomer Ac−VQIVYK−NHMe peptide, from ref 7. The peptide backbone is shown as a thick tube to highlight its conformation, while the side chains are shown as thin wires.

10.1021/ja807760d CCC: $40.75

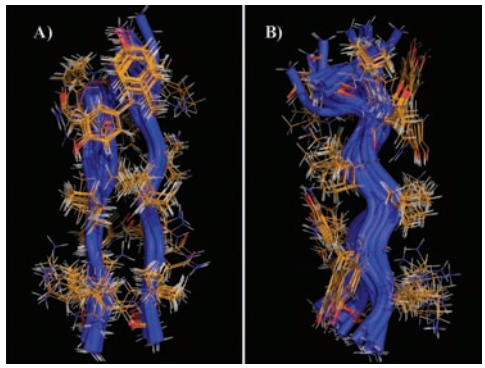


Figure 2. Two views of the 20 lowest-energy MM conformations within the first XCluster family (20 kJ/mol), superimposed to illustrate their similarity: (A) front view; (B) side view. The backbones are displayed as ribbons.

that Y is involved in hydrogen bonding upon dimerization. The NH band at 3370 cm^{-1} in the monomer spectrum changes in the dimer spectrum: the low-frequency side of the band is clearly more intense, and the peak is at lower frequency (3348 cm^{-1}). Also, a new IR band appears at 3192 cm^{-1} for $(Ac-VQIVYK-NHMe)_2$. Intermolecular interaction strengths in the dimer may be qualitatively inferred from spectral band shapes and positions through the correlation between hydrogen-bond strength and OH/NH stretch frequency: stronger hydrogen bonds have longer covalent OH/NH bond lengths and hence lower vibrational frequencies. Figure 1A thus strongly suggests a major conformational change upon dimerization, accompanied by the formation of NH and OH hydrogen bonds stronger than those in the isolated peptide.

To help interpret these results, we predicted structures, energies, and IR spectra for plausible conformers of $(Ac-VQIVYK-NHMe)_2$. We generated ~10 000 conformers using an iterative molecular mechanics (MM)-based conformational search with the Amber* force field.[16] Within discrete structural families, the backbone conformations are the same, and side-chain conformational differences are either negligible (e.g., slight differences in the Q position) or not important (e.g., differences in V CH_3 group rotations). As an illustration, the 20 lowest-energy MM structures (<20 kJ/mol) within the first family are overlaid in Figure 2. The high degree of structural similarity predicted for these conformers suggests that the β-sheet motif is a very stable arrangement for this isolated system. For our computational examination of this system, we applied XCluster analysis[17] to the 200 lowest-energy MM structures to rearrange them into initial, conformationally distinct families. Five such families were identified and selected for subsequent investigation by DFT. Since this method favored only antiparallel β-sheet conformers, we also selected a parallel β-sheet structure for investigation by DFT. The conformational search procedure was preliminary rather than complete; a rigorous, exhaustive search would have required repeated long-time molecular dynamics simulations with a suitable potential. We computed structures and IR spectra at the B3LYP/3-21G* level (and at HF/3-21G* for one conformer) and relative energies at the B3LYP/6-311++G**//B3LYP/4-31G* level, corrected for basis-set superposition error (BSSE) and for zero-point energy (ZPE) using harmonic B3LYP/3-21G* frequencies.[18] Structures did not qualitatively change between the 3-21G* and 4-31G* basis sets. While these calculations were of at best modest accuracy, computational resources for larger-basis-set frequency calculations were not available to us.

The most stable members from the five clusters are presented in Figure 3A. While conformers **1**, **3**, and **4** look similar, there are

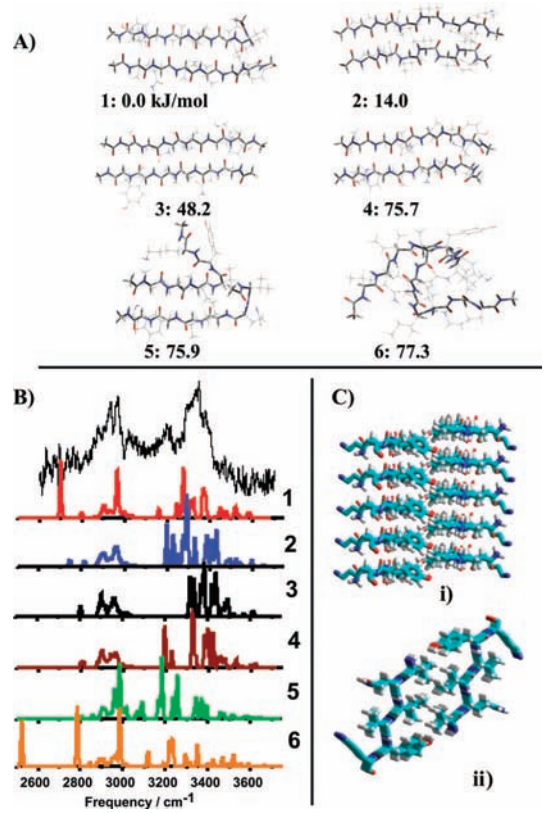


Figure 3. (A) DFT structures of five conformer families from the 200 lowest-energy MM structures and one parallel structure. Relative DFT B3LYP/6-311++G**//B3LYP/4-31G* energies corrected for BSSE and B3LYP/3-21G* ZPE are shown below each structure. (B) Predicted IR spectra (convolved with a 10 cm^{-1} fwhm Gaussian line shape) for conformers **1**−**6** compared with the experimental spectrum. (C) Two views of the crystal structure exhibit (i) parallel β-sheet interactions along the parallel coordinate and (ii) antiparallel zipper interactions along the perpendicular coordinate.

important differences that assigned them to different clusters. **1**, **3**, and **4** represent antiparallel β-sheets: **3** presents a straight, fully closed β-sheet structure, while in **1** one end is opened at the V residue to allow a hydrogen bond between a backbone NH and the Q side chain. **4** is similar to **3** but appears slightly bent at one end. **2** is a slightly bent parallel β-sheet (the only parallel conformer considered). **5** and **6** represent two different unfolded states: **5** is a halfway-opened antiparallel β-sheet in which one strand forms intramolecular hydrogen bonds, and **6** shows two somewhat intertwined backbone conformations, reminiscent of the monomer hairpin, where most of the intermolecular interactions involve side chains as opposed to backbone groups.

The accuracy of the energies in Figure 3A is suspect, as the predicted energy difference between **1** and **3** (representing a minor conformational change) is much larger than that predicted between **3** and **5** (a major conformational change). Therefore, we focus more on the agreement between the IR spectra than on the associated relative structure stabilities. The B3LYP energies do allow us to estimate (e.g., for **1**) the energetic cost of unfolding the monomer (60 kJ/mol) and the binding energy of the dimer (166 kJ/mol), which verifies the hypothesis that dimerization compensates for peptide unfolding (with a net energy gain of 106 kJ/mol, based on our calculations).[19]

The most accurate frequency calculations available to us for this system are those at the B3LYP/3-21G* level. While this is a low level of theory, Gerhards et al.[20] have demonstrated satisfactory experiment−theory agreement for HF/3-21G* NH stretch frequen-

cies of a β-sheet dimer due to error cancelation. We computed HF/3-21G* frequencies for **1** (Figure S1 in the Supporting Information), but scaling behaviors cause HF/3-21G* to be much more time-consuming than B3LYP/3-21G*, and resources for further HF/3-21G* frequencies were not available to us. In Figure 3B, the theoretical IR spectra are compared with the experimental spectrum. On the basis of an inspection of the experimental and theoretical results, we scaled CH modes by 0.95 and left the OH and NH modes unscaled. The comparison ultimately relies on the predicted overall band pattern, focusing in particular on relative NH stretch frequency shifts rather than absolute frequency positions. In the first four theoretical spectra, at least one Y OH stretch band appears at high frequency, in the range 3500−3620 cm^{-1} (in **2**, one OH band is at 3225 cm^{-1}), while in the spectra of **5** and **6**, this band is shifted to below 3200 cm^{-1}. Free NH bands are computed to lie in the 3400−3600 cm^{-1} range (bands between 3500 and 3600 cm^{-1} thus may correspond to either NH or OH modes). The intense bands in the 3200−3400 cm^{-1} range correspond to hydrogen-bonded NH stretch transitions, except for the 3225 cm^{-1} band in **2**. The CH bands are computed to lie in the 2800−3100 cm^{-1} range (some NH stretch bands also appear in this region).

From the comparison between the experimental spectrum and the six predictions, one can immediately conclude that the globular conformer **6** and partly opened **5** are not good fits. The β-sheet conformer **1** matches fairly well with the experiment, especially regarding the NH bands between 3150 and 3500 cm^{-1}; however, it does not show the weak band at 3040 cm^{-1}, and it predicts an intense band at 2700 cm^{-1}, corresponding to the NH−Q interaction, that is not observed. This interaction may be poorly computed with B3LYP/3-21G* or could be undetectable because of dilution through band broadening, or **1** could simply be absent. The similar conformer **3** does predict a feature at 3040 cm^{-1} but fails to predict the stronger band at 3192 cm^{-1}. The overall spectrum of **4** agrees best with the experimental one, reproducing the NH bands quite well, but does not predict the small 3040 cm^{-1} band. Interestingly, the spectrum of the parallel conformer **2** could also agree with the experimental spectrum, although the agreement is not as good as that for **4**. On the basis of only the intensity and relative frequency patterns of the predicted NH bands, we cannot distinguish between parallel and antiparallel β-sheet conformations, but our results *do* suggest an assignment to β-sheet rather than unfolded or globular structures. In this context, spectroscopic data in the amide I/II region would be very helpful for making more unambiguous assignments.

The tau-protein ^{306}VQIVYK^{311} sequence is the crucial region for amyloid fibril formation. In the fibrils, peptide backbones at this sequence adopt parallel β-sheet conformations, enabling fibril formation via cross-β steric zippers.[6] The gas-phase Ac−VQIVYK−NHMe monomeric peptide adopts a β-hairpin conformation.[7] On the basis of the comparison between our experimental and preliminary computational results, it appears that $(Ac-VQIVYK-NHMe)_2$ aggregates adopt a β-sheet structure represented by structures **2**−**4**, indicating that this secondary structure is thermodynamically most attractive and readily forms without any "guiding" influence of a solvent or protein environment. Our key result is that the β-sheet conformers agree with the experimental IR spectrum while the unfolded β-sheets do not (Figure 3B). From our preliminary analysis, we cannot conclude whether the aggregated gas-phase VQIVYK segment forms parallel or antiparallel β-sheets; they may adopt a mixture of both, as has been observed for segments in amylin fibrils.[21] It is interesting to compare the cross-β structures from the crystal[22] (shown in Figure 3C) with structures **1**−**4**: the gas-phase data compare favorably with the characteristic β-sheet interstrand distances and NH → O═C hydrogen-bonding interactions observed in the extended β-strand crystal structures. In conclusion, it appears that in the absence of any environment, $(Ac-VQIVYK-NHMe)_2$ adopts a β-sheet structure stabilized by strong interstrand hydrogen bonding, showing that β-sheet aggregation alone can compensate for the energetic cost of protein unfolding.

Acknowledgment. We thank Mr. Tjalling de Boer for experimental assistance, Prof. John P. Simons for insightful discussions and the use of equipment, and the Oxford Supercomputing Centre for use of their facilities for the calculations. We are grateful for the support provided by the Royal Society USA/Canada Research Fellowship (T.D.V.) and University Research Fellowship (L.C.S.), Corpus Christi College, Oxford (S.G., L.C.S.), and Linacre College, Oxford (T.D.V.).

Supporting Information Available: Complete experimental details, Figure S1, complete ref 16, and Cartesian coordinates for conformers **1**−**6**. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hamley, I. W. *Angew. Chem., Int. Ed.* **2007**, *46*, 8128–8147.
(2) Makin, O. S.; Serpell, L. C. *FEBS J.* **2005**, *272*, 5950–5961.
(3) Perczel, A.; Hudaky, P.; Palfi, V. K. *J. Am. Chem. Soc.* **2007**, *129*, 14959–14965.
(4) Teplow, D. B.; Lazo, N. L.; Bitan, G.; Bernstein, S.; Wyttenbach, T.; Bowers, M. T.; Baumketner, A.; Shea, J.-E.; Urbanc, B.; Cruz, L.; Borreguero, J.; Stanley, H. E. *Acc. Chem. Res.* **2006**, *39*, 635–645.
(5) Berriman, J.; Serpell, L. C.; Oberg, K. A.; Fink, A. L.; Goedert, M.; Crowther, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9034–9038.
(6) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. Ø.; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453–457.
(7) Vaden, T. D.; Gowers, S. A. N.; de Boer, T. S. J. A.; Steill, J. D.; Oomens, J.; Snoek, L. C. *J. Am. Chem. Soc.* **2008**, *130*, 14640–14650.
(8) Inouye, H.; Sharma, D.; Goux, W. J.; Kirschner, D. A. *Biophys. J.* **2006**, *90*, 1774–1789.
(9) Fricke, H.; Funk, A.; Schrader, T.; Gerhards, M. *J. Am. Chem. Soc.* **2008**, *130*, 4692–4698.
(10) Meijer, G.; de Vries, M. S.; Hunziker, H. E.; Wendt, H. R. *Appl. Phys. B: Lasers Opt.* **1990**, *51*, 395–403.
(11) De Vries, M. S.; Hobza, P. *Annu. Rev. Phys. Chem.* **2007**, *58*, 585–612.
(12) Brenner, V.; Piuzzi, F.; Dimicoli, I.; Tardivel, B.; Mons, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 2463–2466.
(13) Dian, B. C.; Longarte, A.; Mercier, S.; Evans, D. A.; Wales, D. J.; Zwier, T. S. *J. Chem. Phys.* **2002**, *117*, 10688–10702.
(14) Inokuchi, Y.; Kobayashi, Y.; Ito, T.; Ebata, T. *J. Phys. Chem. A* **2007**, *111*, 3209–3215.
(15) Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321*, 49–56.
(16) McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7743–7746.
(17) Shenkin, P. S.; McDonald, D. Q. *J. Comput. Chem.* **1994**, *8*, 899–916.
(18) Frisch, M. J.; et al. *Gaussian 03*; Gaussian, Inc: Pittsburgh, PA, 2003.
(19) The single-point energy of the unfolded monomer is an inherent ingredient of the BSSE calculation.
(20) Gerhards, M.; Unterberg, C.; Gerlach, A.; Jansen, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2682–2690.
(21) Madine, J.; Jack, E.; Stockley, P. G.; Radford, S. E.; Serpell, L. C.; Middleton, D. A. *J. Am. Chem. Soc.* **2008**, *130*, 14990–15001.
(22) http://www.doe-mbi.ucla.edu/~sawaya/chime/xtalpept/.

JA807760D