doi:10.1016/j.jmb.2010.02.012 J. Mol. Biol. (2010) 397, 932–946

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Glucagon Fibril Polymorphism Reflects Differences in Protofilament Backbone Structure

Christian Beyschau Andersen[1,2], Matthew R. Hicks[3,4], Valeria Vetri[5], Brian Vandahl[6], Henrik Rahbek-Nielsen[6], Henning Thøgersen[7], Ida Bukh Thøgersen[8], Jan Johannes Enghild[8], Louise C. Serpell[4], Christian Rischel[1] and Daniel Erik Otzen[8]*

[1]*Protein Structure and Biophysics, Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark*

[2]*Institute of Biophysics, Consiglio Nazionale delle Ricerche, Via Ugo La Malfa 153, I-90146 Palermo, Italy*

[3]*Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*

[4]*Department of Biochemistry, John Maynard-Smith Building, School of Life Sciences, University of Sussex, Falmer BN1 9QG, UK*

[5]*Department of Physical and Astronomical Sciences, University of Palermo, Via Archirafi 36, I-90123 Palermo, Italy*

[6]*Protein Science, Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark*

[7]*Structure, Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark*

[8]*Interdisciplinary Nanoscience Centre, Department of Molecular Biology, University of Aarhus, Gustav Wieds Vej 10 C, DK-8000 Aarhus C, Denmark*

Amyloid fibrils formed by the 29-residue peptide hormone glucagon at different concentrations have strikingly different morphologies when observed by transmission electron microscopy. Fibrils formed at low concentration (0.25 mg/mL) consist of two or more protofilaments with a regular twist, while fibrils at high concentration (8 mg/mL) consist of two straight protofilaments. Here, we explore the structural differences underlying glucagon polymorphism using proteolytic degradation, linear and circular dichroism, Fourier transform infrared spectroscopy (FTIR), and X-ray fiber diffraction. Morphological differences are perpetuated at all structural levels, indicating that the two fibril classes differ in terms of protofilament backbone regions, secondary structure, chromophore alignment along the fibril axis, and fibril superstructure. Straight fibrils show a conventional β-sheet-rich far-UV circular dichroism spectrum whereas that of twisted fibrils is dominated by contributions from β-turns. Fourier transform infrared spectroscopy confirms this and also indicates a more dense backbone with weaker hydrogen bonding for the twisted morphology. According to linear dichroism, the secondary structural elements and the aromatic side chains in the straight fibrils are more highly ordered with respect to the alignment axis than the twisted fibrils. A series of highly periodical reflections in the diffractogram of the straight fibrils can be fitted to the diffraction pattern expected from a cylinder. Thus, the highly integrated structural organization in the straight fibril leads to a compact and highly uniform fibril with a well-defined edge. Prolonged proteolytic digestion confirmed that the straight fibrils are very compact and stable, while parts of the twisted fibril backbone are much more readily degraded. Differences in the digest patterns of the two morphologies correlate with predictions from two algorithms, suggesting that the polymorphism is inherent in the glucagon sequence. Glucagon provides a striking illustration of how the same short sequence can be folded into two remarkably different fibrillar structures.

*Corresponding author. E-mail address: dao@inano.dk.

Abbreviations used: FTIR, Fourier transform infrared spectroscopy; TEM, transmission electron microscopy; LD, linear dichroism; MS, mass spectrometry.

0022-2836/$ - see front matter

Received 6 November 2009;
received in revised form
30 January 2010;
accepted 7 February 2010
Available online
12 February 2010

Edited by S. Radford

Keywords: fiber diffraction; linear dichroism; proteolytic patterns; fibril structure

Introduction

Many proteins have a generic ability to form elongated aggregate structures known as fibrils. Fibrils are characterized by a remarkable degree of order often spanning several micrometers, which is several orders of magnitude larger than the dimensions of its molecular constituents.[1] At the same time, fibrils have high mechanical strength and stiffness,[2] a property originating from a protofilament backbone consisting of a single double-pleated β-sheet.[3,4] Based on the crystal structures of short fibrillogenic peptides, fibril backbone interactions have been categorized into eight different classes according to their hydrogen-bonding pattern.[5] Extracellular fibril deposits are tightly linked to neuronal diseases such as Alzheimer's and Parkinson's disease,[6,7] stressing the need to increase our understanding of fibril properties and their formation. Furthermore, as fibrils and other aggregates may increase the immunogenicity of protein therapeutics,[8,9] it is crucial to avoid these structures in drug formulation.[10]

An interesting aspect of fibrillation is the fact that a single protein may give rise to a plethora of different fibril morphologies.[11] Fibrils typically consist of two or more protofilaments associated laterally to form distinct superstructures. These structures may be resolved visually by transmission electron microscopy (TEM) and atomic force microscopy, although the underlying differences in protofilament arrangement may be too subtle to be uniquely determined. The differences in morphology presumably affect fibrillation kinetics and structural properties for reasons that remain unclear. Fibril polymorphism is tightly linked to the strain phenomenon in prion diseases, in which different fibril morphologies of a given protein give rise to distinct heritable states characterized by differences in incubation periods, phenotypes, and the ability to cross transmission barriers.[12,13] This has been shown in a number of studies on the yeast prion Sup35 strains.[14–16] For example, the Weissman group characterized the physical properties of three different Sup35 strains from yeast, showing that phenotype and stability were directly linked to the structural properties of the morphology.[16]

There is currently no consensus as to whether fibril polymorphism is mainly due to the assembly of identical protofilaments or rather reflects polymorphism at a protofilament level.[17] Krishnan and Lindquist observed structural differences in the amyloid core of Sup35.[14] They found that two fibril strains formed *in vivo* at 4 and 25 °C, respectively, differed in the nature of the intermolecular interfaces and the number of residues in the amyloid core, thus demonstrating that these two Sup35 fibril morphologies reflect protofilament polymorphism. Other studies indicate that fibril morphologies are essentially an arbitrary assembly of identical protofilaments that can assemble into various fibril morphologies. For example, it has been reported that certain fibrils may untwine into a number of fibrils with different morphology or even change morphology during growth.[18–21] A structural characterization of different morphologies is often complicated by the intrinsic heterogeneity of fibril samples, and efforts to obtain samples with a distinct morphology often necessitate the addition of salt and excipients, which may complicate the analysis.

In this study, we focus on the 29-residue peptide glucagon, which is involved in the regulation of blood sugar levels and is used therapeutically in cases of hypoglycemia.[22,23] Glucagon's sequence homology to glucagon-like peptide 1, a molecule that forms the basis for recent advances in obesity and diabetes treatment,[24] makes its fibrillation properties highly relevant from a drug formulation perspective. Glucagon morphologies formed under various conditions have been identified and classified by both TEM and atomic force microscopy in previous studies.[25–29] Intriguingly, we previously found that the fibril morphology changes drastically when changing only the glucagon concentration: At low concentration, the fibrils consist of two or several protofilaments with a repetitive twist whereas the fibrils formed at high concentration consist of two protofilaments forming a straight and very homogeneous morphology.[28] We attribute this property to morphology-specific growth inhibition of twisted fibrils through binding of reversible trimers present at high concentrations.[28] Fibrils of either morphology are able to imprint their structural characteristics to daughter fibrils as evidenced by self-seeding experiments in the same study.[28] Here, we exploit this simple model system to better understand the complex nature of the fibril structure. We analyzed the structure of the two types of glucagon fibrils by five different biophysical techniques in order to characterize the secondary structure [far-UV CD and Fourier transform infrared

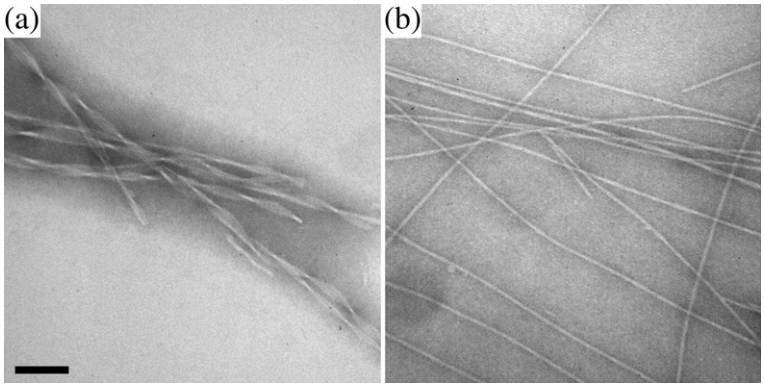

Fig. 1. Morphological differences between glucagon fibrils formed at low and high concentration revealed by TEM. (a) Fibrils formed at 0.25 mg/mL have a distinct twisted appearance with two or more protofilaments twisting repetitively around each other. (b) Fibrils formed at 8 mg/mL are straight and consist of two protofilaments. The scale bar represents 100 nm (adapted from Ref. 28).

spectroscopy (FTIR)], the chromophore alignment relative to the fibril axis [linear dichroism (LD)], fibril architecture and characteristic distances (fiber diffraction), and structural compactness of the backbone (resistance to proteolytic digestion). Our experimental findings are supplemented with insights from aggregation prediction algorithms. We find that structural differences between the two fibril classes are perpetuated at all structural levels, including the backbone region, demonstrating that for this model system, polymorphic subclasses are distinguished by their protofilament structure.

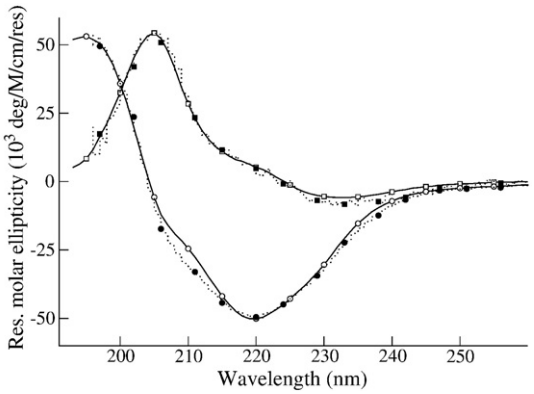

Fig. 2. Glucagon fibril morphologies have vastly different secondary-structure elements as demonstrated by CD spectra of the twisted and straight fibril morphology (squares and circles, respectively). The spectrum of the twisted fibril morphology has a minimum around 233 nm and a maximum at 205 nm and passes through zero at 221 nm. This indicates a structure involving β-turns.[30] The spectrum of the straight fibril morphology has a minimum at 219 nm and a maximum around 195 nm and passes through zero at 202 nm. This spectrum points to a classical β-pleated sheet structure. Neither light scattering nor LD artifacts, due to the presence of fibril floccules and orientational effects, respectively, were significant as evidenced by the close similarity between spectra of sonicated and non-sonicated samples (filled and open symbols, respectively). Spectra for the two sonicated fibril samples are normalized to those of their unsonicated counterpart.

Results

In a previous work, we showed that when viewed by TEM, the fibrils formed at 0.25 mg/mL are characterized by two or more protofilaments twisting repetitively around each other (Fig. 1a), while the fibrils formed at 8 mg/mL consist of two straight protofilaments (Fig. 1b).[28] The samples show a remarkable degree of sample homogeneity. Hence, in the following, we denote these two fibril morphologies twisted and straight, respectively. We have previously shown that the two fibril types are distinguished in at least two ways. Firstly, although both fibril types bind the amyloid-specific dye thioflavin T, the straight fibrils lead to an increase in intensity that is an order of magnitude greater than that induced by the twisted fibrils.[26] Secondly, only the straight fibrils are able to imprint their morphology in cross-seeding experiments.[28] The straight fibrils also form considerably faster, showing shorter lag phases and higher elongation rates.[28] In the following, we examine the structural basis for these intriguing differences.

Far-UV CD spectra indicate fundamentally different peptide conformations

First, secondary-structure elements were measured. Figure 2 shows the far-UV CD spectra of twisted and straight fibrils in the wavelength interval 190–260 nm. The twisted fibril morphology sample has a maximum at 205 nm and a minimum around 233 nm and passes through zero at 221 nm. The spectrum hence resembles protein structures rich in β-turns such as $(Ala_2\text{-}Gly_2)_n$.[30] In contrast, straight fibrils formed at 8 mg/mL have a broad peak with a minimum at 219 nm and a maximum at 195 nm and pass through zero at 202 nm. This spectrum is typical of structures rich in β-sheets. The far-UV CD spectra are hence fundamentally different in accordance with previous CD observations of glucagon fibrils formed under various conditions.[26] This suggests that the secondary-structure elements

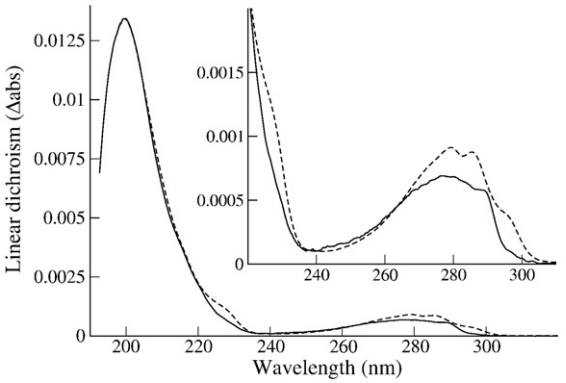

Fig. 3. Orientational differences of the aromatic residues in the two glucagon morphologies as analyzed by LD. Spectra of flow-aligned fibrils of twisted (continuous line) and straight morphology (broken line). The inset shows the spectral differences in the aromatic region. The straight fibrils were diluted 32-fold to take the difference in concentration into account, and the spectra have been normalized for comparison.

in the two morphologies are quite different, although contributions from aromatic chromophores (which can also be in different environments in the two fibril types) may also weigh in. However, FTIR, which is not marred by these potentially confusing overlaps, confirms the existence of β-sheets and β-turns in the straight and twisted fibrils, respectively (see below). CD spectra of an oriented system such as fibrils are sometimes masked by a contribution from LD, which can be several orders of magnitude greater than the CD signal itself.[31] The contribution from LD depends on the axis of orientation and vanishes when the orientation is destroyed.[32] Hence, to ensure that the fibrils were not oriented in the cuvette, we measured the spectra on diluted samples in a cuvette with a relatively long path length. As an additional effort to diminish orientation effects, the fibrils in separate samples were fragmented by sonication (Fig. 2). Upon sonication, the amplitudes of the twisted and straight morphology spectra were reduced by a factor of 26 and 6.5, respectively, but the general spectral features were conserved. A reduction in intensity upon sonication is expected.[32–34] A high level of experimental reproducibility, as well as an insignificant effect of rotating the cuvette in the sample holder, further testifies to a random fibril orientation.

LD spectra show differences in the internal orientation of the aromatic residues

In an aligned sample, LD is the difference between the absorbance of light polarized parallel with and perpendicular to the alignment axis:

$$\mathrm{LD} = A_{//} - A_{\perp} \qquad (1)$$

For fibers under shear flow, as described here, the alignment axis is the long axis of the fiber. In the UV region of the spectrum, electronic transitions (absorbance) occur in the peptide backbone and aromatic side chains of proteins. LD spectra give information on the orientation of the transition moments (the direction of net electron displacement during an electronic transition) relative to the fiber axis as described in Eq. (2):

$$\mathrm{LD}^{\mathrm{r}} = \frac{\mathrm{LD}}{A} = \frac{A_{//} - A_{\perp}}{A} = \frac{3}{2} S\left(3\cos^2\alpha - 1\right) \qquad (2)$$

where the reduced LD (LD^{r}) is equal to the LD divided by the absorbance of an unaligned sample, S is the alignment factor, and α is the angle of the transition moment relative to the alignment (fiber) axis. With fibers, it is often difficult to obtain absorbance spectra of truly isotropic samples; however, the sign of the LD signal alone is a powerful tool for the elucidation of the relative orientations of secondary-structure elements and aromatic side chains. A positive LD signal means that the transition moment is at an angle of <55° from the fiber axis and a negative signal means that the angle is >55°.

Figure 3 shows LD spectra of samples of twisted and straight morphology in the wavelength range from 190 to 320 nm. The straight morphology sample was diluted 32-fold to take the concentration difference into account and then scaled down by a factor of 7.15 to superimpose on the twisted morphology spectrum, based on the peak at 200 nm. In the case of β-structures, there is a strong π-to-π* transition moment for the peptide backbone around 200 nm that is perpendicular to the β-strand. Thus, the positive LD signal, seen here for both fibril types (Fig. 3), indicates that the transition is parallel with the fibril axis and means that the β-strands run across the fiber. This is consistent with the cross-β model of amyloid fibers.[32,35] Although the peptide concentration of the two samples is similar, there is a much larger (~7-fold) LD signal for the fibrils grown at higher concentration. This large difference in LD signal can be due to better alignment from longer and/or more rigid fibers[32] and/or better ordering within the fibrils. In the aromatic region (250–310 nm), there are large differences in spectral shape as well as magnitude. The fact that the signals are positive in this region shows that the transition moments of the aromatic side chains (which are in plane with the aromatic rings of tyrosine and tryptophan) lie along the fiber axis, rather than perpendicular to it. The ~10-fold difference in magnitude is consistent with better aromatic side-chain ordering and alignment in the straight fibrils. The different peak positions are also indicative of a different environment for the aromatic side chains, and the appearance of a shoulder around 295 nm in the spectrum from the straight fibrils is strong evidence for a difference in the packing and/or ordering of the tryptophan side chain in these fibrils compared with the twisted fibrils. Furthermore, the small signal around 230 nm in the straight fibrils may be due to excitation coupling of aromatic side chains, consistent with closer packing in these fibers.[36]

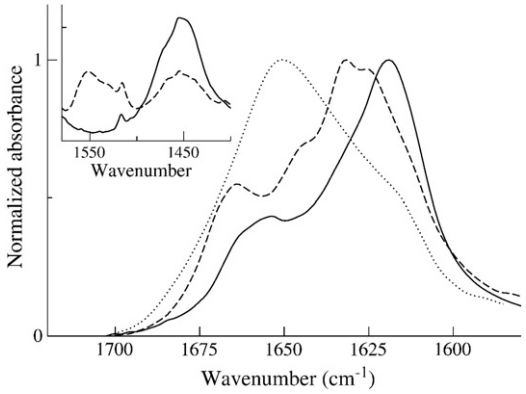

Fig. 4. Glucagon fibril morphologies show differences in the β-structures comprising the fibril backbone as well as backbone compactness as demonstrated by differences in the amide I′ and amide II/II′ region of FTIR spectra. The amide I′ spectrum of the straight fibrils (straight line) has a strong peak at 1619 cm^{-1} indicating intermolecular β-structures typical for amyloid structures. In contrast, the spectrum of twisted fibrils (broken line) has more features including two peaks (1626 and 1632 cm^{-1}) in the β-sheet region and an additional peak (1664 cm^{-1}) indicating β-turn structure elements. The spectrum of non-fibrillated glucagon (dotted line) has a wide peak with maximum at 1651 cm^{-1} associated with the presence of α-helix and random coil in accordance with the glucagon trimer structure. The amide II/II′ region (inset) shows that the spectrum of the straight fibrils has a single intense amide IIv′ peak at 1450 cm^{-1}, while the spectrum of the twisted fibril morphology has an additional amide II peak at 1550 cm^{-1}. This indicates a slower H/D exchange of internal hydrogens due to a more compact twisted structure.

FTIR spectra show differences in β-sheet backbone structure

FTIR spectroscopy provides information about secondary structure of proteins in H_2O through distinct amide I and amide II absorption peaks centered at approximately 1660 and 1550 cm^{-1}, respectively.[37,38] In D_2O, these peaks are shifted to approximately 1650 and 1450 cm^{-1} (amide I′ and amide II′), respectively[39]. The amide I/I′ absorption contains contributions from the C=O stretching vibration of the amide group with minor contributions from, for example, the out-of-phase C–N stretching vibration, while the amide II/II′ absorption arises from the out-of-phase combination of the N–H (N–D) in-plane bend and the C–N stretching vibrations.[37] Because secondary structures are characterized by specific hydrogen-bonding patterns between amide C=O and N–H groups, it is expected that each secondary-structure element gives rise to characteristic absorption patterns in FTIR spectroscopy. In addition, the amide II/II′ band is particularly sensitive to the H/D exchange of amide groups and may therefore probe structural compactness and flexibility.[38]

FTIR spectra of twisted (broken line) and straight (continuous line) fibril morphology are shown in Fig. 4 together with the spectrum of fresh non-fibrillated glucagon (dotted line). In the amide I′ region (Fig. 4), the straight fibril morphology spectrum has an intense absorption peak at 1619 cm^{-1}, a value typical for the very strong intermolecular hydrogen bonds found in amyloid β-sheet,[37,39,40] and a weaker peak at 1654 cm^{-1}, indicating the presence of disordered structure or α-helix.[37,38] The amide I′ band of the twisted fibril morphology spectrum has several components, most notably two strong peaks at 1626 and 1632 cm^{-1}, respectively, as well as a peak at 1664 cm^{-1}. The two strong peaks are assigned to a β-sheet, whose hydrogen bonding is less strong than that found in the straight morphology at lower wave numbers, and the peak at 1664 cm^{-1} indicates β-turns.[38,41] The shoulder at 1645 cm^{-1} suggests the presence of random coil. For both morphologies, the apparent absence of peaks around 1690 cm^{-1} seems to exclude the possibility of antiparallel β-sheets,[41] but since FTIR relies on mainly empirical correlations between absorption bands and secondary structure, further studies are required to rule out

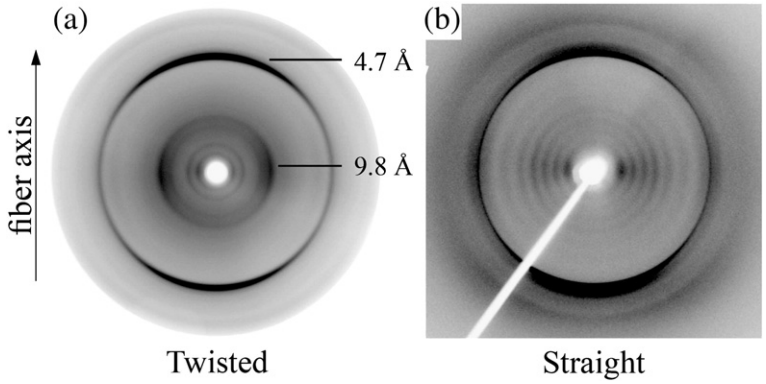

Fig. 5. Fiber diffractograms of partially aligned glucagon fibrils show differences in structural periodicity. (a) Fibrils of twisted morphology. The characteristic cross-β diffraction signals on the meridian and equator at 4.7 and 9.8 Å, respectively, are indicated. (b) Fibrils of straight morphology. The meridional signal at 4.7 Å is clearly visible as is a peculiar periodicity along the equatorial plane. The periodicity is interpreted as scattering from a wave incident on a cylinder as shown in Fig. 6, suggesting that this sample consists of a homogeneous assembly of compact fibrils.

Table 1. Meridional and equatorial peak positions for the twisted and straight glucagon fibril morphology obtained by integration of the diffraction patterns in Fig. 3

	Twisted (Å)	Straight (Å)
Meridional	4.76	4.76
	2.41	2.41
Equatorial	33–37	19.18
	21	13.88
	12.91	10.85
	10	8.95
		7.49
		6.15
		3.81

antiparallel conformations. The spectrum of fresh non-fibrillated 15 mg/mL glucagon has a peak centered at 1651 cm^{-1}, testifying to the presence of α-helix and random coil also found in the solution structure of the glucagon trimer.[42] The shoulder at 1616 cm^{-1} can probably be attributed to the formation of small amounts of aggregates during the measurement.

In the region from 1400 to 1580 cm^{-1} (Fig. 4, inset), the straight morphology spectrum shows a broad peak at 1450 cm^{-1} (amide II) but no peak at 1550 cm^{-1} (amide II′). This indicates that the fibrils of straight morphology, which were formed in H_2O glycine buffer, exchange most amide hydrogens with deuterium during the D_2O spin filtering. In contrast, the twisted morphology spectrum has peaks at both amide II and amide II′ owing to an incomplete exchange of hydrogens, possibly due to the multilayered structure of the fibrils, which decreases their accessibility to exchange. As expected, when grown in D_2O glycine buffer, neither fibril morphology spectrum had an amide II peak (data not shown).

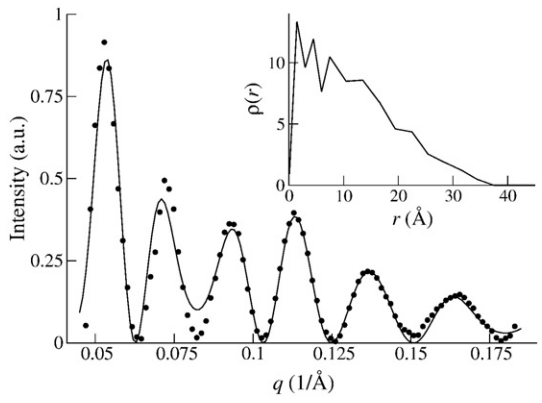

Fig. 6. The diffraction pattern of the straight morphology is similar to scattering from a cylinder with well-defined edges. Profile of the integrated intensity along the equatorial scattering vector for the fiber diffractogram shown in Fig. 5b. The data were fitted to a model based on diffraction from a plane wave incident on a cylinder (continuous line). The good quality of the fit shows that the straight fibrils are very homogeneous. The cylinder density $\rho(r)$ is shown as a function of the cylinder radius r (inset). The cylindrical fibril model predicts a diameter of ~6 nm in accordance with previous TEM observations of glucagon fibrils.[28]

Fiber diffractograms reveal high homogeneity of straight fibrils

X-ray diffraction on partially aligned fibril samples probes repetitive distances within the fibrils and may thus provide a fingerprint of the internal structure.[43] Figure 5a and b show fiber diffractograms of twisted and straight fibrils, respectively. Reflections from both fibrils gave a diffraction pattern consistent with the classic cross-β pattern generally observed for amyloid fibrils. The patterns show a strong, sharp reflection at 4.76 Å (corresponding to the inter-strand distance) on the meridional axis, the axis parallel with the fiber axis. We also observe a much weaker reflection at 2.41 Å, which is probably due to the Bragg condition being fulfilled for $n=2$.[44] Besides these reflections, each sample exhibits a distinct series of equatorial reflections (Table 1). The different patterns of equatorial reflections likely reflect differences in packing of protofilaments in the two samples and may also arise from alternative β-sheet packing arrangements. The twisted fibril diffraction pattern (Fig. 5a) was cross-β in nature and was analyzed using the CLEARER software.[45] A detailed analysis of the diffraction pattern revealed unit cell dimensions of lengths $a=4.69$ Å, $b=29.16$ Å, and $c=37.8$ Å (all angles, 90°). While the 4.69-Å distance reflects the canonical distance between hydrogen-bonded β-strands, the other two dimensions are less easy to assign but are likely to arise from chain length and sheet spacing. The 37.8 and 29.16 Å correspond to 11 and 8–9 residues in a β-sheet conformation, respectively. Interestingly, in a previous alanine-scanning analysis of glucagon fibrillation, we proposed a very tentative strand–loop–strand model of the glucagon fibril,[46] which is consistent with these residue numbers. However, it is possible that this agreement is fortuitous.

It was not possible to obtain satisfactory unit cell dimensions from the diffraction data derived from the straight fibrils using the approach above. However, this sample is observed to have a peculiar periodicity along the equatorial plane (Fig. 5b) reminiscent of the diffraction pattern from a classical multi-slit experiment. We therefore analyzed the diffraction data based on the equivalent diffraction pattern predicted from a plane wave incident on a cylindrical barrier. Figure 6 shows the integrated diffraction profile as a function of scattering vector. The full line shows the theoretical curve from a single fibril with cylindrical symmetry, fitted as explained in Materials and Methods. The satisfactory agreement between the fit and the data testifies to the extraordinary sample homogeneity as well as the high level ordering within individual fibrils.[1] The inset shows the calculated electron density as a function of radius. In this simple model, the fibrils have a diameter of about 6 nm, which is in agreement with previous TEM measurements.[28]

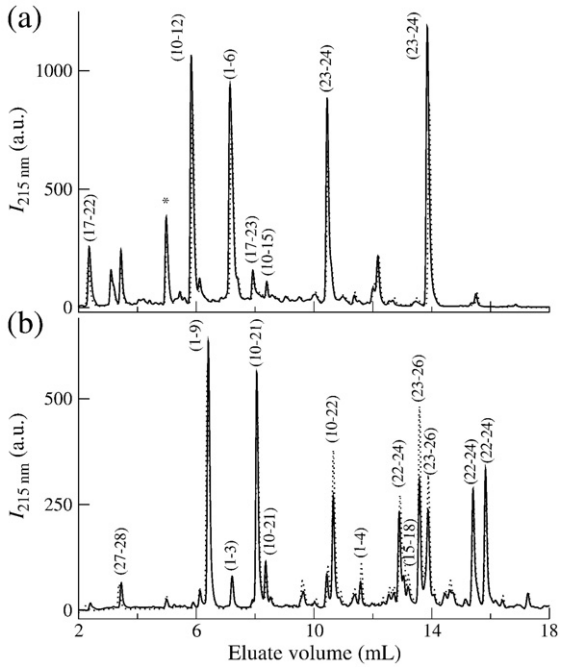

Fig. 7. Identified peptide fragments of glucagon fibrils digested by pepsin. Glucagon fibrils of straight and twisted morphology (continuous and dotted line, respectively) were digested by pepsin leading fragments to be released to the supernatant (a), while peptides from the protease-resistant fibril core were recovered from the pellet (b). Fragments were separated by HPLC and identified by MS and Edman degradation. They are indicated in parentheses on the figure (the asterisk indicates a pepsin fragment). The supernatant fragments tend to be relatively short compared to the pellet fragments. No overlap between identified fragments in supernatant and pellet was observed, and the peak positions appear to be independent of fibril morphology. Data in (a) are after 72 h of digestion and those in (b) are after 1 h of digestion. These times are chosen to give the clearest peaks, but similar behavior is observed at other times.

Differences in fibril morphology affect susceptibility to proteolytic digestion

Protease treatment has been shown to cleave off exposed flexible regions while leaving the fibril core intact.[47–49] In this context, the fibril core is defined as the β-sheet backbone as well as other compact secondary-structure elements. We examined whether the two morphologies showed different resistance to pepsin digestion. Pepsin is a relatively unspecific acidic protease and its promiscuity prevents peptide identifications via cleavage positions alone.[50] While pepsin has weak preferences against, for example, cleavage after glycine (which is the C-terminal residue of one of the fragments we identify), particularly in larger proteins,[50] the long incubation time (up to 72 h) at 21 °C and the short length of glucagon (29 residues) render these small differences insignificant and make it very likely that all accessible peptide bonds will be cleaved. Peptides released during the digest were separated from intact fibrils by ultracentrifugation, and the peptides in the supernatant and pellet were separated by HPLC. Figure 7a compares the chromatograms of peptides released from the fibril core (present in the supernatant) for each morphology, and Fig. 7b compares the chromatograms of peptides in the pepsin-resistant core (recovered from the pellet). Fragments identified by mass spectrometry (MS) and/or Edman degradation are indicated on the figure. Closer inspection reveals clear differences in peak positions between supernatant and pellet; none of the fragments identified in the supernatant were also found in the pellet. The supernatant tends to contain shorter fragments (maximum of 6 residues), probably because released fragments are further digested once in solution. The fragments found in the supernatant revealed cleavage sites (after residues 5, 12, 15, and 17) that were not found in the pellet, suggesting that these bonds are protected in the fibril core. The integrated pellet peak areas are listed in Table 2 sorted in order of descending peak area of the straight fibrils after 1 h of digestion. The five most frequently found peptides in the pellet, ranging from 3 to 13 residues in length, comprised 82% of the total peak area. Also indicated in Table 2 is the relative change in peak area calculated after normalizing the area of the identified peaks. In Fig. 8, the normalized peak areas are plotted for all peaks comprising more than 5% of the peak area. It is readily seen that some fragments found in the straight morphology decrease by about 20% while a few fragments increase by up to 39%. This contrasts with the twisted morphology fibrils, for which the longer fragments decrease by 60–75% and the shorter fragments (1–4 and 22–24) increase 2- and 10-fold, respectively. Interestingly, these short fragments are located in opposite ends of the glucagon peptide. Figure 9 compares the relative occurrence of individual residues in the backbone for each morphology after 72 h of digest. It is clearly seen that in the straight fibrils, residues spanning the entire sequence except residue 29 are found to approximately the same extent, hence demonstrating the compact nature of the conformation. In contrast, the twisted fibrils have two regions comprising residues 1–4 and 22–26, which are overrepresented. Residues 27–29 are almost completely absent from the backbone of the twisted morphology.

Insights from fibril propensity algorithms

The digest patterns were compared with the TANGO and PASTA algorithms, which can predict the aggregation propensity of a given protein sequence. The TANGO prediction has been applied to glucagon previously[46] and is shown in the top part of Fig. 10. Residues 22–27 (FVQWLM) were all assigned a β-aggregation score higher than 5%, while the remainder of the glucagon structure was assigned scores equal to or close to zero. The PASTA algorithm is shown in the bottom part of Fig. 10. In comparison with the TANGO algorithm, which only predicts the region containing residues 22–27 to be

Table 2. Complete list of backbone fragments identified in each of the two morphologies after 1 and 72 h of digest

	Straight			Twisted		
	Peak area percentage		Relative change (%)	Peak area percentage		Relative change (%)
Fragment	1 h	72 h		1 h	72 h	
22–24	22.7	14.9	−18.1	17.6	38.8	97.3
1–9	16.5	11.1	−15.7	14.2	6.6	−58.4
10–21	15.9	10.2	−19.7	11.9	3.7	−72.3
23–26	14.4	13.9	−20.1	17.4	12.3	−36.6
10–22	6.3	7.0	39.0	8.0	2.2	−75.6
1–3	2.0*	1.0	—	1.7*	0.3	—
15–18	1.5*	3.5	—	2.1*	0.0	—
27–28	1.4*	1.0	—	1.2*	0.7	—
1–4	1.3*	3.1	—	1.8	20.1	876.6
Sum	82.0	65.7		76.0	84.7	

The data are sorted in order of descending peak area in the straight fibrils, and the bottom row shows the summed peak areas. For each peak, the relative change in peak area is indicated after normalizing the identified peak areas. Most of the straight morphology fragments change very little over time (about 20%), while long fragments in the twisted morphology (e.g., fragment 1–9) are drastically reduced in intensity leading to the increase in shorter fragments (e.g., fragment 1–4). Fragment 1–4 is found in significant amounts only in the twisted morphology after 72 h. The data presented in this table form the basis of Fig. 8. Peak areas indicated by an asterisk are too small to be accurately determined and have been omitted from Fig. 8.

fibril-prone, PASTA indicates that regions involving residues 2–10 are also prone to fibrillation.

Discussion

In contrast to folded proteins, misfolded protein structures have not been under evolutionary pressure to adopt a specific three-dimensional structure. This is probably the reason for the high number of polymorphic fibril ensembles present under specific conditions.[10,11] Detailed structural information obtained by microcrystallography and solid-state NMR has emerged during the last decade,[5,51] revealing that a number of different backbone packing arrangements are compatible with the canonical cross-β amyloid structure. Polymorphism has attracted considerable attention in part because the ability of prion strains to cross species barriers is linked to the physical properties of individual morphologies. Furthermore, morphologies differ in structural properties with obvious implications for future applications of fibrils as nanomaterials[52] and silks.[53]

A number of studies have shown that backbone structure correlates with morphology. For example, Petkova *et al.* showed that amyloid β-peptide had different morphologies when visualized by TEM depending on the growth conditions (agitated *versus* quiescent).[54] The visual differences in superstructure were also present in intermolecular contacts as probed by solid-state NMR, and the morphology was imprinted on daughter fibrils as shown by self-seeding experiments. Other studies suggest that morphologies are essentially an arbitrary assembly of filaments, which can split apart and rearrange during growth,[18–21] hence suggesting that the polymorphism of protofilaments does not necessarily impact on fibril superstructure.

In this article, we demonstrate how different biophysical techniques can be used to elucidate the structural basis of fibril polymorphism. By applying techniques that address structural properties of large molecules, we have investigated whether glucagon polymorphism manifests itself all the way to the protofilament level or if the twisted morphology fibrils are essentially structurally identical with the straight fibrils. Our approach was to probe secondary-structure elements (far-UV CD and FTIR), alignment of the secondary-structure elements and the chromophores present in the glucagon sequence along the fiber axis (flow-aligned LD), fibril architecture and characteristic distances (fiber diffraction), and the backbone structure (resistance to proteolytic digestion). Our spectroscopic investigations revealed a clear difference in the structural properties, and we found that the compact and well-ordered appearance of the straight fibrils was confirmed by LD, fiber diffraction, and proteolytic digestion. After 72 h of proteolytic digestion, differences in proteolytical susceptibility of the backbone regions emerged, where the compactness of the straight morphology is reflected by a highly stable core comprising almost all of glucagon's 29 residues. In contrast, the twisted morphology has two highly susceptible areas at opposite ends of the peptide as well as an additional cleavage site after residue 4. This implies that the origin of glucagon's fibril polymorphism extends all the way to the regions forming the protofilament backbone. The

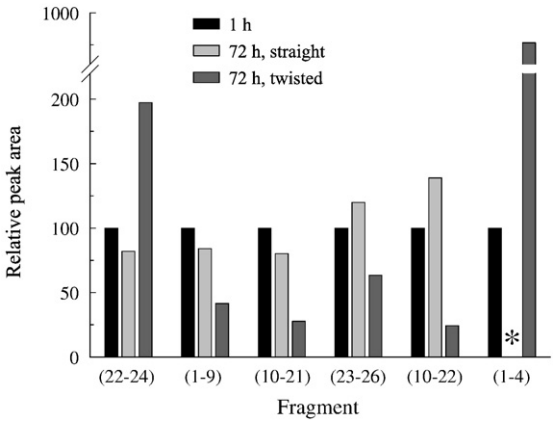

Fig. 8. Relative changes in the peak height of individual glucagon fibril fragments over time. Straight fibrils are resistant to extended proteolytic digestion, while twisted fibrils are continuously degraded leading to smaller backbone fragments. Peptide fragments found in the fibril pellet are sorted in order of descending area in the straight morphology after 1 h of proteolytic digest (cf., Table 2). The peak areas after 1 h are assigned a value of 100 and the peak areas after 72 h are shown relative to this value. For the straight morphology, most fragments changed by ca 20% after 72 h. For the twisted morphology, the longer fragments decrease by ~60–75%, and the short fragments 22–24 and 1–4 increase ~2- and 10-fold, respectively. Fragment 1–4 was not significantly present in the straight morphology as indicated by an asterisk.

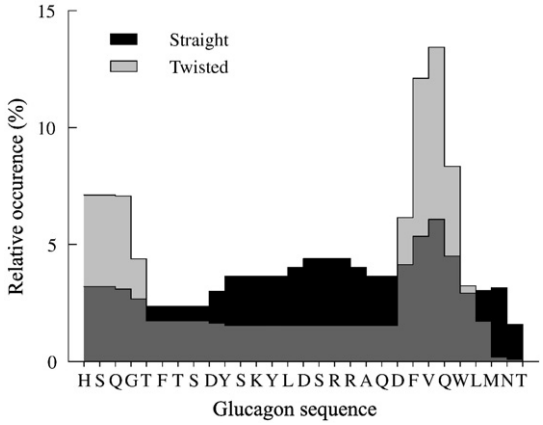

Fig. 9. Glucagon morphologies show substantial differences in residues resistant to proteolytic attack. The figure shows the relative occurrence of individual peptides along the 29-residue glucagon primary sequence obtained from the peptide peak areas after 72 h of digest (Table 2). For the straight fibrils (black bars), all residues except residue 29 are found in the fibril backbone and all are approximately equally prevalent. The backbone of the twisted fibrils (gray bars), on the other hand, is dominated by a short stretch of protease-resistant residues at the N- and C-termini.

predictions of two different aggregation propensity algorithms offer some explanation to this behavior as detailed below.

Packing of structural elements differs between the two fibril morphologies

CD is affected by the local organization of the peptide backbone and the asymmetry of the aromatic chromophores (in glucagon's case, 2 Tyr, 2 Phe, and 1 Trp), while LD is sensitive to the orientation of the β-sheet and the aromatic chromophores relative to the fibril axis.[32,35] Both CD and LD showed significant differences between the two morphologies. The twisted morphology had an intriguing and unusual CD spectrum (Fig. 2), which could indicate the presence of β-turns.[30] The straight morphology, on the other hand, was that of classical β-sheet. These findings agree with previous findings on glucagon as well as insulin fibrils.[26] LD confirms the cross-β model of amyloid fibrils and shows that the aromatic rings of Trp and Tyr lie along the fiber axis but with different environments in the two morphologies (Fig. 3). Furthermore, LD suggests that the straight fibril morphology is much more rigid and/or ordered. FTIR conveys information about conformational differences in secondary structure, and due to the sensitivity towards β-sheet conformations in particular, FTIR has been applied in numerous studies to probe amyloid formation and amyloid polymorphism.[17,26,39,41,55–57] The FTIR amide I′ spectra (Fig. 4) of fibrils of twisted (continuous line) and straight (broken line) morphology show large differences in number of peaks and peak positions. Most notably, the differences in the peak positions in the β-sheet region (approximately 1611–1645 cm^{-1}) demonstrate that the very backbone structure depends on the morphology, with the hydrogen bonds in the twisted morphology being weaker than those of the straight morphology. A similar correspondence between the shift in amide I peak position and the β-sheet twist angle has been found in a previous study.[58] Another important observation in the spectrum of the twisted morphology is the peak at 1664 cm^{-1}, which, in accordance with the CD spectrum of these fibrils, indicate the presence of β-turns. The amide II/II′ region (Fig. 4, inset) shows that despite hours of H/D exchange in a spin filter, a peak due to N–H bending mode of the amide group is still present in the twisted morphology. The presence of this peak, which is not present in the straight morphology sample or in samples prepared in D_2O glycine buffer (data not shown), hence suggests that H/D exchange is substantially slower in the twisted morphology or that water is trapped inside the fibril. In any case, the observation that a complete H/D exchange did not occur clearly indicates that the twisted morphology protects a part of the fibril structure, possibly due to a more convoluted fibril structure than the straight morphology.

Fiber diffraction

Fiber diffraction is a useful tool to elucidate repetitive distances within fibrils[59] and, in this study, also the overall fibril thickness of the straight

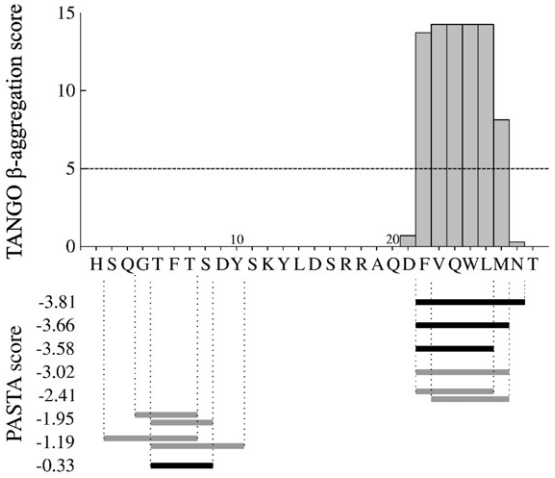

Fig. 10. Prediction of glucagon fibrillation propensity based on TANGO and PASTA aggregation scores identifies fibrillation-prone regions at both the N-terminus and the C-terminus. TANGO (top) assigns a β-aggregation score above 5% (broken line) to six consecutive residues (residues 22–27), thus identifying this part of the peptide as prone to aggregation. PASTA (bottom) predicts the same part of glucagon to form parallel or antiparallel β-sheets (black and gray bars, respectively) and suggests that the N-terminus may be involved in fibril formation. The differences in aggregation behavior predicted by TANGO and PASTA algorithms suggest that polymorphism is inherent in the glucagon sequence.

fibrils. The diffraction patterns from fibers made from straight and twisted fibrils were clearly different, illustrating that fiber diffraction can be an efficient tool to distinguish morphologies. Twisted fibrils gave evidence of a unit cell with dimensions 4.69 Å×29.16 Å×37.9 Å. These unit cell dimensions could accommodate the glucagon molecule in a turn configuration as proposed previously based on an Ala scan of the fibrillation propensity.[46] Although probably fortuitous, we find it intriguing that the CD and FTIR experiments give some support to this turn model as well. The multiple rings in the diffraction pattern from the straight fibrils formed at 8 mg/mL glucagon can, in principle, be due to a multitude of internal periodicities in the fibers (Fig. 5). The high uniformity of these fibers observed previously by TEM[28] prompted us to consider an alternative explanation of higher-order scattering from identical fibers with cylindrical symmetry. The good agreement between the experimental data and the theoretical fit (Fig. 6) indeed shows that this interpretation is also consistent with data. To our knowledge, this is the first time it has been possible to model a fibril sample as identical cylinders. However, note that the fitting procedure is almost certainly an underdetermined problem; hence, the density profile we obtain is not unique. We arrived at the profile by a heuristic procedure, and a different approach probably would have given an at least somewhat different result. However, if the fibers in the sample had a broad distribution of widths, it would not be possible to observe higher-order diffraction rings at all, since the pattern would be smeared. Hence, the fibers in the sample must have quite uniform width if the interpretation in terms of higher-order diffraction is correct. This implies that the fibrils are homogeneous and have a well-defined edge, rather than an amorphous halo whose average density decreases smoothly away from the fibril. The uniform structure and the electron density radius of about 3 nm are consistent with previous TEM observations showing the homogeneous appearance of these fibrils with a width of about 6 nm. Furthermore, the main assumption behind the diffraction analysis, that the fibrils are homogeneous and compact, is consistent with the LD observations described above.

Polymorphism originates from differences in the protofilament core

The fibril core is generally believed to consist of an extended double-pleated β-sheet, which is compact and stable compared to the residues residing outside the core. This view of the fibril structure has emerged through, for example, solid-state NMR,[51] electron microscopy image analysis,[60] hydrogen exchange,[61] and proteolytic digestion.[47] For example, when exposed to proteolytic digestion, the peptide fragments outside the core will be released quickly and the core will remain intact. This property has been exploited to characterize the fibril structure of various proteins, for example, lysozyme,[48] apolipoprotein,[62] and β_2-microglobulin,[49,63] as well as fibrils of the $A\beta_{1-40}$ peptide.[47,64] The proteolytic digest technique has been used to distinguish different folding patterns present in various fibril morphologies of β_2-microglobulin[49] and is hence of special interest to our work on glucagon fibrils. In the work on protein fibrils, the fibril backbone is typically found to contain only a fraction of the residues, while large regions are at least partially unstructured and thus exposed to proteolytic digestion. In contrast, in the work on the 40-residue Aβ peptide, almost the entire sequence was recovered after digest, although cleavage sites were also found inside the fibril core.[47,61,64] In fibrils made from peptides, where there is a need to have contiguous stretches of peptides incorporated into the fibril for stability reasons, it is likely that a very large part of the peptide will be involved in the fibril backbone structure. This could potentially make it difficult to distinguish between morphologies by proteolytical digestion.

While the backbone fragments of the two glucagon morphologies were for the most part similar (Fig. 7), there are differences especially when comparing their distribution after 72 h of proteolytic digestion (Fig. 8). The straight morphology backbone fragments changed relatively little (Fig. 8) and were equally distributed over the entire primary sequence with Thr29 as the only exception (Fig. 9). This is consistent with a very compact and stable fibril morphology with almost the entire peptide forming a well-protected core as evidenced by LD and fiber diffraction. Numerous cleavage sites at the C-terminus (after residues 21, 22, 24, and 26) but only one at the N-terminus (after residue 9, cf., Table 2) indicate that an especially resistant part extends from residues 1–21. This observation is the exact opposite of the TANGO prediction (Fig. 10), which assigns a high β-aggregation score to residues 22–27, where most cleavages are observed, and an aggregation score close to zero for the rest of the sequence including the most pepsin-resistant part. The reason for this discrepancy may be that at this high concentration, residues 19–27 in soluble glucagon may form not only a β-sheet but also an α-helical structure.[65,66] This prediction is supported by NMR studies on non-aggregated glucagon in which residues 17–29 are α-helical and hence not readily available to aggregation.[42] A trimer involving an α-helical structure at the C-terminus would give credence to PASTA, which predicts β-sheet also at the N-terminal.

The twisted fibrils' backbone fragments were subject to much larger pepsin-induced changes over time than the straight fibrils. In general, the longer fragments (1–9, 10–21, and 10–22) decreased by 58–76%, while some shorter fragments increased in intensity (Fig. 8 and Table 2). Most notably, residues 1–4 increased approximately 10-fold, perhaps at the expense of fragment 1–9. The cleavage after residue 4 was not significant in the straight morphology most likely owing to different structural elements at the N-terminus. After extended digestion, two regions at each end of the peptide

(1–4 and 22–26) stand out as particularly resistant, while residues 5–21 and especially 27–29 are readily digested (Fig. 9). Fibrils formed at low peptide concentration, where glucagon is predominantly monomeric, hence follow the predictions of TANGO and PASTA concerning the C-terminal residues. Intriguingly, these digestion data also lend some support to a hairpin model of glucagon fibrils with a large fraction of β-turns, which is further supported by our CD, FTIR, and fiber diffraction data and also suggested by an alanine scan.[46]

In conclusion, we find that glucagon's polymorphism is perpetuated at all structural levels down to the structural elements of the protofilament backbone. Glucagon's intrinsic polymorphism is nicely highlighted by the differences in predicted aggregation behavior from the two prediction programs TANGO and PASTA, which suggest that this polymorphism is inherent in the glucagon sequence. However, a full elucidation of the molecular details underlying this polymorphism must await the advent of atomic-resolution structures of the different glucagon fibrils. Efforts to achieve this are currently underway.

Materials and Methods

Glucagon fibrillation

In all experiments, pharmaceutical-grade glucagon (Novo Nordisk, Bagsværd, Denmark) was dissolved in 0.22 μm sterile-filtered glycine/HCl buffer adjusted to pH 2.5. To remove aggregates, we centrifuged the samples at 32,000*g* for 15 min. The peptide concentration was adjusted to 0.25 or 8 mg/mL and measured by absorbance at 280 nm using a theoretical molar extinction coefficient of 8250 M^{-1} cm^{-1}. Samples were fibrillated at room temperature in 96-well microtiter plates (Corning Inc., Corning, NY) sealed with tape pads (3M, St. Paul, MN), and the fibrillation was measured by fluorescence. The fluorescence plate reader (SpectraMax Gemini EM; Molecular Devices, Sunnyvale, CA) was programmed to read every 10 min. The samples were not shaken except for the motion of the plate during reading. Fibrillation was followed using the intrinsic fluorescence from Trp25 and from the fluorescent dye ThT (Sigma-Aldrich, St. Louis, MO) added to 40 μM from a 2-mM stock solution to a separate glucagon sample in the microtiter plate. Tryptophan fluorescence was measured by excitation at 300 nm, and the blue-shifted emission spectrum was quantified by the 330/355-nm ratio. ThT fluorescence was measured by excitation at 444 nm and emission at 488 nm. Fibril formation was typically completed within 48 h. The results presented in this study are from freshly prepared fibril solutions, but judging from a large number of experiments using the techniques applied in this study, the fibrillated solutions remain stable for several weeks.

CD spectra

CD spectra of freshly prepared fibrils were collected on a spectropolarimeter (Chirascan, Applied Photophysics, Surrey, UK) in the range from 195 to 260 nm with a step size of 0.5 nm, a bandwidth of 1 nm, and a scan rate of 7.5 nm/min. Six repeats made up the low-concentration spectra, while three repeats made up the high-concentration spectra, in order to improve the signal-to-noise ratio. For both CD and LD, spectra were truncated at the point where the high-tension voltage on the photomultiplier tube exceeded 650 V. This is where the response for the instrument used becomes nonlinear. The temperature was fixed at 20 °C using the built-in temperature control. Fibrillated samples of glucagon are generally viscous; thus, transferring the samples into narrow cuvettes may lead to some orientation of the fibrils. Furthermore, stray light from large fibril floccules may distort the CD spectra. To minimize these effects, we used quartz cuvettes with a 1-mm light path, diluted the fibrillated samples to 0.1 mg/mL, and compared the spectra obtained to those of ultrasonicated samples. Buffer background was subtracted from each spectrum, and no additional filtering or noise reduction was applied. The signal is reported in mean residue ellipticity.

LD spectra

LD measurements were performed on a Jasco J715 spectropolarimeter adapted for LD measurement (Jasco UK, Great Dunmow, UK). The sample alignment was carried out using a micro-volume Couette cell that was built in-house (equivalent models were available from Kromatek, Great Dunmow, UK), using a rotation speed of 3000 rpm. All spectra were recorded using 0.25 mg/mL glucagon; that is, the sample containing 8 mg/mL glucagon was diluted 32-fold just prior to recording. Spectra were measured in the range 350 to 180 nm with a data pitch of 0.2 nm, a 1-nm bandwidth, a scanning speed of 50 nm/min, and a 2-s response time. Sixteen spectra were averaged for each sample or buffer baseline. Buffer spectra were measured under the same conditions and subtracted from the sample spectra.

FTIR

For FTIR measurements, glucagon fibrils formed at 0.25 and 8 mg/mL were concentrated and washed in D_2O (Cambridge Isotope Laboratory, Andover, MA) using spin filters (Sartorius Vivascience, Hannover, Germany) with a 30-kDa cutoff. The filtrate did not contain measurable amounts of protein. A 20-μL drop of the gel-like substance remaining after several hours of filtering was transferred to the cell consisting of two CaF_2 windows separated by a 0.025-mm Teflon spacer. Spectra were acquired at room temperature on a Bruker (Ettlingen, Germany) Vertex 70 FTIR single-beam spectrometer equipped with a MIR light source and a DLATGS pyroelectric detector (SALEX Galileo, Florence, Italy) after purging the sample compartment with nitrogen to reduce water vapor. Sample and solvent absorption spectra were calculated with respect to the spectrum of the empty cell. Each measurement reported is an average of 200 spectra in the 400–4000 cm^{-1} range with a spectral resolution of 2 cm^{-1}. Spectra were normalized to the D_2O peak at 1200 cm^{-1}, and the buffer spectrum was subtracted. Absorption of the amide I′ band was larger than 1, avoiding the need for smoothing the spectra. Spectra were scaled to the amide I′ peak. Although the data reported here are from fibril samples made in H_2O glycine buffer and subsequently washed in D_2O, spectra measured on samples made in D_2O glycine buffer and extracted from the pellet after centrifugation were similar except for small differences in the amide II/II′ region as detailed in Results. For comparison, a spectrum was acquired on a freshly prepared

15 mg/mL glucagon sample dissolved in a D_2O glycine/HCl buffer using a 0.100-mm Teflon spacer. For this spectrum, the buffer spectrum was subtracted and the spectrum was smoothed using a 9-point Savitzky and Golay function. Peak positions were read from the spectrum.

Fiber diffraction

A drop (10 μl) of fibril sample solution was placed between the wax-sealed ends of two glass capillary tubes (1 mm diameter) and allowed to dry. Within a day, the drop had dried and a narrow fiber bundle had formed. X-ray fiber diffractograms were obtained on an R-AXIS IV++ imaging plate area detector (Rigaku Corp., Tokyo, Japan) mounted on a rotating anode with a Cu-Kα wavelength of 1.5419 Å. Data were collected with a sample-to-detector distance of 160 or 300 mm, and the sample was rotated 0.5° during the 20-min data acquisition. In order to avoid diffraction from crystallized buffer salts, we collected fiber diffractograms on freshly prepared moist fibers. Buffer salts were also removed by washing the fibrils repeatedly in milli-Q water (Millipore). The position of the reflections obtained in this way was similar to the position of the reflections observed from newly formed fibers. However, as diffraction was generally lower from desalted fibers, possibly due to inferior alignment of the fibrils, data analysis was performed using the freshly prepared fibers. Diffraction patterns were initially observed using the X-ray crystallography program MOSFLM (CCP4). Centering and accurate measurement of diffraction signal peaks were performed using CLEARER.[45] Exploration of possible unit cell parameters was performed using the unit cell determination function within CLEARER.

The diffraction pattern from straight fibers as a function of scattering vector q was fitted to the following expression based on the diffraction pattern expected from a plane wave incident on a cylinder:[67,68]

$$I(q) = F(q)^2 \tag{3}$$

$$F(q) = \int_0^\infty \rho(r) J_0(2\pi r q) 2\pi r dr \tag{4}$$

In this expression, J_0 is the Bessel function of the first kind of order zero and $\rho(r)$ is the electron density as a function of the distance from the fiber center. Variations in electron density along the fiber axis and as a function of the angle around the fiber perimeter are ignored, as is the interference of scattering between different fibers. Equation (4) is obtained by casting the Fraunhofer approximation in cylindrical coordinates and evaluating the angular integral.[69] In practice, we found that a reasonable fit was obtained when $\rho(r)$ was taken to be a piecewise linear function in the interval 0–37.5 Å and 0 outside, and the integral was evaluated by summation over 750 equidistant points in the range 0–37.5 Å. The values at the following values of r were allowed to vary: 0, 1.5, 3, 4.5, 6, 7.5, 10.5, 13.5, 16.5, 19.5, 22.5, 25.5, 28.5, 31.5, and 34.5 Å. The value at 37.5 Å was set to zero; all other values between $r=0$ Å and $r=37.5$ Å were obtained by linear interpolation. Nonlinear least-squares fitting was performed by the Optim routine in the freely available R statistics program (The R Project for Statistical Computing†).

† http://www.r-project.org

Enzymatic digest and peptide sequencing

Fibrillated 0.25 mg/mL glucagon samples were centrifuged at 32,000***g*** in Amicon Ultra-4 spin filters with a 30-kDa cutoff (Millipore) and resuspended to ~1 mg/mL in sample buffer. Fibrillated 8-mg/mL samples were diluted to 1 mg/mL in sample buffer. The fibril samples were then digested proteolytically by mixing the 1-mg/mL fibril solution 1:1 (w/w) with a 1-mg/mL porcine pepsin solution (Sigma-Aldrich). The pepsin concentration was determined by absorption measurements at 280 nm using a theoretical molar extinction coefficient of 51,715 M^{-1} cm^{-1}. Fibril samples of each morphology were digested in 96-well microtiter plates in a fluorescence plate reader at 21 °C. Due to the low thermal stability of glucagon fibrils formed at high (>3 mg/mL) concentrations (melting temperature, ca 32 °C),[26] heating of the fibrils during digestion must be avoided. After 1 and 72 h of pepsin digest, samples were extracted from the microtiter plate, and the digested fibrils were washed four times with 4 mL sample buffer by centrifugation at 32,000***g*** in Amicon Ultra-4 spin filters with a 50-kDa cutoff (Millipore). This step separates the 34.6-kDa pepsin molecules and soluble peptide fragments from the digested fibrils. Subsequently, the fibril pellet was resuspended in sample buffer to a concentration of ~1 mg/mL, and the fibrils were dissolved by heating at 75 °C for 15 min in order to quench the pepsin digest process. Three percent TFA was added to a final concentration of 0.2%, and the fractions were stored immediately at −80 °C. Peptides from both the digestion supernatant and dissolved fibrils were separated by RP-HPLC (Agilent Technologies, Palo Alto, CA, model 1200) on a 15-cm reversed-phase Symmetry300 C18 column (Waters Corp., Milford, MA) with a gradient of 5–80% aqueous acetonitrile (0.1% TFA) over 40 min at a flow of 1 mL/min. A 100-μL sample volume was injected on the column. The eluent was monitored at 215 nm, and fractions were collected based on peak detection and dried by vacuum centrifugation. A fresh sample of 1 mg/mL monomeric undigested glucagon eluted after 18.5 mL, and all digested samples eluted before this point. Only minute amounts of uncleaved glucagon were detected in the digested samples. Chromatograms were integrated using the Agilent ChemStation software.

The peptides were identified by a combination of MS and Edman degradation. For the MS analysis, the fractionated peptides were redissolved in 70% acetonitrile and analyzed by matrix-assisted laser desorption/ionization MS using α-cyano-4-hydroxy cinnamic acid as matrix and an Ultraflex mass spectrometer (Bruker Daltonics, Billerica, MA) for acquisition of MS and LIFT MS/MS spectra. Peptides were identified by comparison to a theoretical digest of glucagon and porcine pepsin using the BioTools software (Bruker Daltonics). For N-terminal amino acid sequence analysis, peptides were applied to Biobrene precycled glass-fiber filters (Applied Biosystems, Foster City, CA) and subjected to automated Edman degradation in an Applied Biosystems PROCISE 494 HT sequencer with on-line HPLC (Applied Biosystems Model 120A) analysis of phenylthiohydantoins.

TANGO and PASTA prediction algorithms

TANGO incorporates five different conformational states of the protein—β-turn, α-helix, β-sheet, the folded state, and β-aggregates—and different energy terms, taking into account hydrophobicity and solvation energies, electrostatic interactions, and hydrogen bonding. For

each residue in a peptide, TANGO computes the percent occupancy of the β-aggregation conformation. As a guideline, five consecutive residues each with a score higher than 5 have been found to be a good indicator of aggregation propensity independently of the size of the peptide or protein.[70] TANGO scores for glucagon were calculated using the TANGO web interface.[71] A more recent prediction algorithm, PASTA, was also applied to the glucagon sequence. PASTA is a computational approach based on the propensities of two residues to be facing each other on neighboring strands in a β-sheet.[72] Hence, the basic assumption is that the same interactions found in structures of known globular proteins are also found in the fibril backbone. The method assigns energy scores to fragments of the same length from the peptide in both parallel and antiparallel orientations. Compared to other algorithms, PASTA is able to predict the registry of the intermolecular hydrogen bonds between fibril-prone sequences and also to discriminate between parallel and antiparallel β-strand configurations. PASTA was implemented as an SPL (SYBYL programming language) script in the commercial software SYBYL 8.0 (Tripos, St. Louis, MO) according to the method described by Trovato *et al.*[72] The energy terms $\varepsilon^{p}_{i,j}(L)$ and $\varepsilon^{a}_{i,j}(L)$ in Eqs. (2) and (3) of Ref. 72 were the PASTA energy scores for a particular peptide sequence in a parallel or antiparallel configuration, respectively.

Acknowledgements

We thank Pawel Sikorski for extensive help with the fiber diffraction analysis. For helpful discussions that improved this work as well as experimental expertise, we thank Anders Svensson, Mathias Norrman, Gerd Schluckebier, Valeria Militello, and Maurizio Leone. Maria Mårtensson, Lasse Gudmand Lading, Mikkel Melchior Rasmussen, and Klaus Drachmann Ottosen are gratefully acknowledged for technical assistance. C.B.A. acknowledges a Novo Nordisk International Postdoc Fellowship. D.E.O. is supported by the Danish Research Foundation through inSPIN.

References

1. Knowles, T. P. J., Smith, J. F., Craig, A., Dobson, C. M. & Welland, M. E. (2006). Spatial persistence of angular correlations in amyloid fibrils. *Phys. Rev. Lett.* **96**, 238301.
2. Smith, J. F., Knowles, T. P. J., Dobson, C. M., MacPhee, C. E. & Welland, M. E. (2006). Characterization of the nanoscale properties of individual amyloid fibrils. *Proc. Natl Acad. Sci. USA*, **103**, 15806–15811.
3. Eanes, E. D. & Glenner, G. G. (1968). X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem.* **16**, 673–677.
4. Makin, O. S. & Serpell, L. C. (2005). Structures for amyloid fibrils. *FEBS J.* **272**, 5950–5961.
5. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I. *et al.* (2007). Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature*, **447**, 453–457.
6. Chiti, F. & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366.
7. Lansbury, P. T. & Lashuel, H. A. (2006). A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature*, **443**, 774–779.
8. Rosenberg, A. S. (2006). Effects of protein aggregates: an immunologic perspective. *AAPS J.* **8**, E501–E507.
9. De Groot, A. S. & Scott, D. W. (2007). Immunogenicity of protein therapeutics. *Trends Immunol.* **28**, 482–490.
10. Frokjaer, S. & Otzen, D. E. (2005). Protein drug stability: a formulation challenge. *Nat. Rev., Drug Discov.* **4**, 298–306.
11. Pedersen, J. S. & Otzen, D. E. (2008). Amyloid—a state in many guises: survival of the fittest fibril fold. *Protein Sci.* **17**, 1–9.
12. Collinge, J. (2003). Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519–550.
13. Chien, P., Weissman, J. S. & DePace, A. H. (2004). Emerging principles of conformation-based prion inheritance. *Annu. Rev. Biochem.* **73**, 617–656.
14. Krishnan, R. & Lindquist, S. L. (2005). Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature*, **435**, 765–772.
15. Cobb, N. J. & Surewicz, W. K. (2007). Prion strains under the magnifying glass. *Nat. Struct. Mol. Biol.* **14**, 882–884.
16. Tanaka, M., Collins, S. R., Toyama, B. H. & Weissman, J. S. (2006). The physical basis of how prion conformations determine strain phenotypes. *Nature*, **442**, 585–589.
17. Kodali, R. & Wetzel, R. (2007). Polymorphism in the intermediates and products of amyloid assembly. *Curr. Opin. Struct. Biol.* **17**, 48–57.
18. Anderson, M., Bocharova, O. V., Makarava, N., Breydo, L., Salnikov, V. V. & Baskakov, I. V. (2006). Polymorphism and ultrastructural organization of prion protein amyloid fibrils: an insight from high resolution atomic force microscopy. *J. Mol. Biol.* **358**, 580–596.
19. Kad, N. M., Myers, S. L., Smith, D. P., Smith, D. A., Radford, S. E. & Thomson, N. H. (2003). Hierarchical assembly of β2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J. Mol. Biol.* **330**, 785–797.
20. Goldsbury, C., Kistler, J., Aebi, U., Arvinte, T. & Cooper, G. J. S. (1999). Watching amyloid fibrils grow by time-lapse atomic force microscopy. *J. Mol. Biol.* **285**, 33–39.
21. Goldsbury, C. S., Wirtz, S., Muller, S. A., Sunderji, S., Wicki, P., Aebi, U. & Frey, P. (2000). Studies on the in vitro assembly of Aβ 1–40: implications for the search for Aβ fibril formation inhibitors. *J. Struct. Biol.* **130**, 217–231.
22. Drucker, D. J. (1998). Glucagon-like peptides. *Diabetes*, **47**, 159–169.
23. Drucker, D. J. (2005). Biologic actions and therapeutic potential of the proglucagon-derived peptides. *Nat. Clin. Pract. Endocrinol. Metab.* **1**, 22–31.
24. Meier, J. J. & Nauck, M. A. (2005). Glucagon-like peptide 1 (GLP-1) in biology and pathology. *Diabetes Metab. Res. Rev.* **21**, 91–117.
25. Dong, M., Hovgaard, M. B., Xu, S., Otzen, D. & Besenbacher, F. (2006). AFM study of glucagon fibrillation via oligomeric structures resulting in interwoven fibrils. *Nanotechnology*, **17**, 4003–4009.
26. Pedersen, J. S., Dikov, D., Flink, J. L., Hjuler, H. A., Christiansen, G. & Otzen, D. E. (2006). The changing

face of glucagon fibrillation: structural polymorphism and conformational imprinting. *J. Mol. Biol.* **355**, 501–523.
27. De Jong, K. L., Incledon, B., Yip, C. M. & DeFelippis, M. R. (2006). Amyloid fibrils of glucagon characterized by high-resolution atomic force microscopy. *Biophys. J.* **91**, 1905–1914.
28. Andersen, C. B., Otzen, D., Christiansen, G. & Rischel, C. (2007). Glucagon amyloid-like fibril morphology is selected via morphology-dependent growth inhibition. *Biochemistry*, **46**, 7314–7324.
29. Andersen, C. B., Yagi, H., Mauro, M., Martorana, V., Ban, T., Christiansen, G. *et al.* (2009). Branching in amyloid fibril growth. *Biophys. J.* **96**, 1529–1536.
30. Brahms, S., Brahms, J., Spach, G. & Brack, A. (1977). Identification of β, β-turns and unordered conformations in polypeptide chains by vacuum UV circular dichroism. *Proc. Natl Acad. Sci. USA*, **74**, 3208–3212.
31. Davidsson, Å., Nordén, B. & Seth, S. (1980). Measurement of oriented circular dichroism. *Chem. Phys. Lett.* **70**, 313–316.
32. Adachi, R., Yamaguchi, K., Yagi, H., Sakurai, K., Naiki, H. & Goto, Y. (2007). Flow-induced alignment of amyloid protofilaments revealed by linear dichroism. *J. Biol. Chem.* **282**, 8978–8983.
33. Dzwolak, W., Loksztejn, A., Galinska-Rakoczy, A., Adachi, R., Goto, Y. & Rupnicki, L. (2007). Conformational indeterminism in protein misfolding: chiral amplification on amyloidogenic pathway of insulin. *J. Am. Chem. Soc.* **129**, 7517–7522.
34. Chatani, E., Lee, Y. H., Yagi, H., Yoshimura, Y., Naiki, H. & Goto, Y. (2009). Ultrasonication-dependent production and breakdown lead to minimum-sized amyloid fibrils. *Proc. Natl Acad. Sci. USA*, **106**, 11119–11124.
35. Dafforn, T. R., Rajendra, J., Halsall, D. J., Serpell, L. C. & Rodger, A. (2004). Protein fiber linear dichroism for structure determination and kinetics in a low-volume, low-wavelength couette flow cell. *Biophys. J.* **86**, 404–410.
36. Marshall, K. E., Hicks, M. R., WIlliams, T. L., Hoffmann, S. V., Rodger, A. & Serpell, L. C. (2010). Characterizing the assembly of the Sup35 yeast prion fragment, GNNQQNY: structural changes accompany a fiber-to-crystal switch. *Biophys. J.* **98**, 330–338.
37. Jackson, M. & Mantsch, H. H. (1995). The use and misuse of FTIR spectroscopy in the determination of protein structure. *Crit. Rev. Biochem. Mol. Biol.* **30**, 95–120.
38. Barth, A. (2007). Infrared spectroscopy of proteins. *Biochim. Biophys. Acta*, **1767**, 1073–1101.
39. Militello, V., Casarino, C., Emanuele, A., Giostra, A., Pullara, F. & Leone, M. (2004). Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering. *Biophys. Chem.* **107**, 175–187.
40. Yan, Y. B., Zhang, J., He, H. W. & Zhou, H. M. (2006). Oligomerization and aggregation of bovine pancreatic ribonuclease A: characteristic events observed by FTIR spectroscopy. *Biophys. J.* **90**, 2525–2533.
41. Bouchard, M. A. R. I., Zurdo, J., Nettleton, E. J., Dobson, C. M. & Robinson, C. V. (2000). Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy. *Protein Sci.* **9**, 1960–1967.
42. Wagman, M. E., Dobson, C. M. & Karplus, M. (1980). Proton NMR studies of the association and folding of glucagon in solution. *FEBS Lett.* **119**, 265–270.
43. Makin, O. S., Sikorski, P. & Serpell, L. C. (2006). Diffraction to study protein and peptide assemblies. *Curr. Opin. Struct. Biol.* **10**, 417–422.
44. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. F. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273**, 729–739.
45. Makin, O. S., Sikorski, P. & Serpell, L. C. (2007). CLEARER: a new tool for the analysis of X-ray fibre diffraction patterns and diffraction simulation from atomic structural models. *J. Appl. Crystallogr.* **40**, 966–972.
46. Pedersen, J. S., Dikov, D. & Otzen, D. E. (2006). N- and C-terminal hydrophobic patches are involved in fibrillation of glucagon. *Biochemistry*, **45**, 14503–14512.
47. Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B. & Wetzel, R. (2001). Structural features of the Aβ amyloid fibril elucidated by limited proteolysis. *Biochemistry*, **40**, 11757–11767.
48. Frare, E., Mossuto, M. F., de Laureto, P. P., Dumoulin, M., Dobson, C. M. & Fontana, A. (2006). Identification of the core structure of lysozyme amyloid fibrils by proteolysis. *J. Mol. Biol.* **361**, 551–561.
49. Myers, S. L., Thomson, N. H., Radford, S. E. & Ashcroft, A. E. (2006). Investigating the structural properties of amyloid-like fibrils formed in vitro from β2-microglobulin using limited proteolysis and electrospray ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* **20**, 1628–1636.
50. Hamuro, Y., Coales, S. J., Molnar, K. S., Tuske, S. J. & Morrow, J. A. (2008). Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. *Rapid Commun. Mass Spectrom.* **22**, 1041–1046.
51. Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F. & Tycko, R. (2002). A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl Acad. Sci. USA*, **99**, 16742–16747.
52. Knowles, T. P., Fitzpatrick, A. W., Meehan, S., Mott, H. R., Vendruscolo, M., Dobson, C. M. & Welland, M. E. (2007). Role of intermolecular forces in defining material properties of protein nanofibrils. *Science*, **318**, 1900–1903.
53. Rammensee, S., Slotta, U., Scheibel, T. & Bausch, A. R. (2008). Assembly mechanism of recombinant spider silk proteins. *Proc. Natl Acad. Sci. USA*, **105**, 6590–6595.
54. Petkova, A. T., Leapman, R. D., Guo, Z. H., Yau, W. M., Mattson, M. P. & Tycko, R. (2005). Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science*, **307**, 262–265.
55. Radovan, D., Smirnovas, V. & Winter, R. (2008). Effect of pressure on islet amyloid polypeptide aggregation: revealing the polymorphic nature of the fibrillation process. *Biochemistry*, **47**, 6352–6360.
56. Caughey, B., Raymond, G. J. & Bessen, R. A. (1998). Strain-dependent differences in β-sheet conformations of abnormal prion protein. *J. Biol. Chem.* **273**, 32230–32235.
57. Dzwolak, W., Smirnovas, V., Jansen, R. & Winter, R. (2004). Insulin forms amyloid in a strain-dependent manner: an FT-IR spectroscopic study. *Protein Sci.* **13**, 1927–1932.
58. Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G. & Fändrich, M. (2004). FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils. *Protein Sci.* **13**, 3314–3321.
59. Fandrich, M. & Dobson, C. M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* **21**, 5682–5690.

60. Sachse, C., Fändrich, M. & Grigorieff, N. (2008). Paired β-sheet structure of an Aβ(1–40) amyloid fibril revealed by electron microscopy. *Proc. Natl Acad. Sci. USA*, **105**, 7462–7466.
61. Kheterpal, I., Zhou, S., Cook, K. D. & Wetzel, R. (2000). Aβ amyloid fibrils possess a core structure highly resistant to hydrogen exchange. *Proc. Natl Acad. Sci. USA*, **97**, 13597–13601.
62. Wilson, L. M., Mok, Y. F., Binger, K. J., Griffin, M. D. W., Mertens, H. D. T., Lin, F. *et al.* (2007). A structural core within apolipoprotein C-II amyloid fibrils identified using hydrogen exchange and proteolysis. *J. Mol. Biol.* **366**, 1639–1651.
63. Monti, M., Amoresano, A., Giorgetti, S., Bellotti, V. & Pucci, P. (2005). Limited proteolysis in the investigation of β2-microglobulin amyloidogenic and fibrillar states. *Biochim. Biophys. Acta*, **1753**, 44–50.
64. Kheterpal, I., Chen, M., Cook, K. D. & Wetzel, R. (2006). Structural differences in Aβ amyloid protofibrils and fibrils mapped by hydrogen exchange–mass spectrometry with on-line proteolytic fragmentation. *J. Mol. Biol.* **361**, 785–795.
65. Chou, P. Y. & Fasman, G. D. (1974). Prediction of protein conformation. *Biochemistry*, **13**, 222–245.
66. Chou, P. Y. & Fasman, G. D. (1975). Conformation of glucagon: predictions and consequences. *Biochemistry*, **14**, 2536–2541.
67. Burge, R. E. (1959). X-ray scattering by bundles of cylinders. *Acta Crystallogr.* **12**, 285–290.
68. Harget, P. J. & Krimm, S. (1971). Direct analysis of small-angle equatorial X-ray scattering from fibrous systems. I. Expressions for the intensity and Patterson function. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen Crystallogr.* **A27**, 586–596.
69. Jackson, J. D. (1998). *Classical Electrodynamics.* John Wiley & Sons, Hoboken, NJ.
70. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **22**, 1302–1306.
71. Serrano, L., Schymkowitz, J. & Rousseau, F. (2007). Tango—a computer algorithm for prediction of aggregating regions in unfolded polypeptide chains. http://tango.crg.es/.
72. Trovato, A., Chiti, F., Maritan, A. & Seno, F. (2006). Insight into the structure of amyloid fibrils from the analysis of globular proteins. *PLoS Comput. Biol.* **2**, 1608–1618.