Atomic resolution structure of full-length human insulin fibrils

Saba Suladze, Riddhiman Sarkar, Natalia Rodina, Krister Bokvist, Manuel Krewinkel, Daniel Scheps, Norbert Nagel, Benjamin Bardiaux, Bernd Reif

PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY OPEN ACCESS

Atomic resolution structure of full-length human insulin fibrils

Saba Suladze[a,b], Riddhiman Sarkar[a,b], Natalia Rodina[a], Krister Bokvist[c,1], Manuel Krewinkel[d], Daniel Scheps[e], Norbert Nagel[f,2], Benjamin Bardiaux[g,h], and Bernd Reif[a,b,3]

Edited by Robert Tycko, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD; received February 7, 2024; accepted May 2, 2024

Patients with type 1 diabetes mellitus who are dependent on an external supply of insulin develop insulin-derived amyloidosis at the sites of insulin injection. A major component of these plaques is identified as full-length insulin consisting of the two chains A and B. While there have been several reports that characterize insulin misfolding and the biophysical properties of the fibrils, atomic-level information on the insulin fibril architecture remains elusive. We present here an atomic resolution structure of a monomorphic insulin amyloid fibril that has been determined using magic angle spinning solid-state NMR spectroscopy. The structure of the insulin monomer yields a U-shaped fold in which the two chains A and B are arranged in parallel to each other and are oriented perpendicular to the fibril axis. Each chain contains two β-strands. We identify two hydrophobic clusters that together with the three preserved disulfide bridges define the amyloid core structure. The surface of the monomeric amyloid unit cell is hydrophobic implicating a potential dimerization and oligomerization interface for the assembly of several protofilaments in the mature fibril. The structure provides a starting point for the development of drugs that bind to the fibril surface and disrupt secondary nucleation as well as for other therapeutic approaches to attenuate insulin aggregation.

amyloid fibril structure | MAS solid-state NMR | molecular modeling | SAXS

Significance

Insulin-derived amyloidosis results in a loss of control of blood sugar levels due to impaired insulin absorption and is accompanied by an insulin dose increase for type I diabetes patients. Although the role of insulin aggregation is appreciated in the treatment of type I diabetes, information reporting on the insulin aggregation mechanism and insulin amyloid structure is lacking. Using magic angle spinning solid-state NMR experiments in combination with molecular modeling, we have obtained a structural model of insulin amyloid fibrils. We have found that insulin fibril polymorphism is induced by lateral association of protofilaments.

Administration of insulin is essential for the survival of individuals diagnosed with type 1 diabetes and is key for patients suffering from type 2 diabetes where production of the hormone is hampered due to β-cell failure (1, 2). Insulin-related amyloidosis is a skin-related complication of insulin therapy. In patients, fibrillar insulin aggregates are found at the site of frequent insulin injections (3–6). In drug manufacturing, insulin fibrillation is complicating the preparation of large quantities of the peptide (7, 8). During production, the pH value is lowered to pH 2 to 4 which facilitates aggregation (9), which results in a reduction of functional activity and unwanted immunological responses in patients (10). In fact, an increase in the autoimmune response to insulin has been observed in Parkinson's disease (PD) patients in comparison to healthy controls, hinting toward a link between insulin aggregation and PD reflecting the impaired insulin homeostasis in PD (11). Moreover, fibrils, which potentially present a reactive surface for secondary nucleation, might induce the formation of toxic intermediates in a fibril-catalyzed reaction in tissue (10). Therefore, elucidating the structural details of insulin fibrils is an important first step toward the understanding of this autocatalytic process. In addition, the insulin fibril structure can open an avenue for the rational design of diagnostic and therapeutic tools for the treatment of insulin amyloidosis.

Amyloid fibrils result from the self-assembly of many copies of a protein into insoluble polymers which are deposited in tissue and organs (12, 13). In recent years, cryoelectron microscopy (cryo-EM) evolved into a powerful tool to investigate the structure of amyloid fibrils (14, 15). Alternatively, solid-state NMR can be employed to yield structural information at atomic level resolution (16, 17). Using magic angle spinning (MAS) solid-state NMR, high-resolution structures of fibrils formed by the Alzheimer's disease β-amyloid peptide (18–20), α-synuclein (21), glucagon (22), and several other disease-relevant proteins have been obtained (23–25). Previously, the physical and chemical parameters that affect insulin amyloid formation including the effect of temperature (9), ionic strength (26), pH (9, 27), the influence of mechanical agitation (28), and contact with hydrophobic surfaces (29) have been investigated. Dobson and colleagues have proposed a hypothetical arrangement of the polypeptide chains in the insulin amyloid protofilament (30). There, each insulin molecule occupies two β-strand layers with the strands parallel to each other. Several studies have implicated that the mature insulin amyloid fibrils are composed of several even-numbered protofilaments wound around one another. The number of subunits in a fibril layer can differ from 2 to 6 (30–32). It is found that all three disulfide bonds are retained in the mature fibrils and each disulfide bond is crucial to the insulin fibril fold and its structural stability

Author contributions: S.S., B.B., and B.R. designed research; S.S., R.S., N.R., K.B., M.K., D.S., and B.B. performed research; K.B., M.K., D.S., and N.N. contributed new reagents/analytic tools; S.S., R.S., N.R., and B.B. analyzed data; and S.S., B.B., and B.R. wrote the paper.

Competing interest statement: K.B., M.K., and D.S. are employees of Sanofi and hold shares and stock options in the company.

This article is a PNAS Direct Submission.

[1]Present address: Biology Research, Neurocrine Biosciences, SanDiego, CA 92130.

[2]Present address: AbbVie Deutschland GmbH, Analytical Research & Development, Ludwigshafen 67061, Germany.

[3]To whom correspondence may be addressed. Email: reif@tum.de.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2401458121/-/DCSupplemental.

Published May 29, 2024.

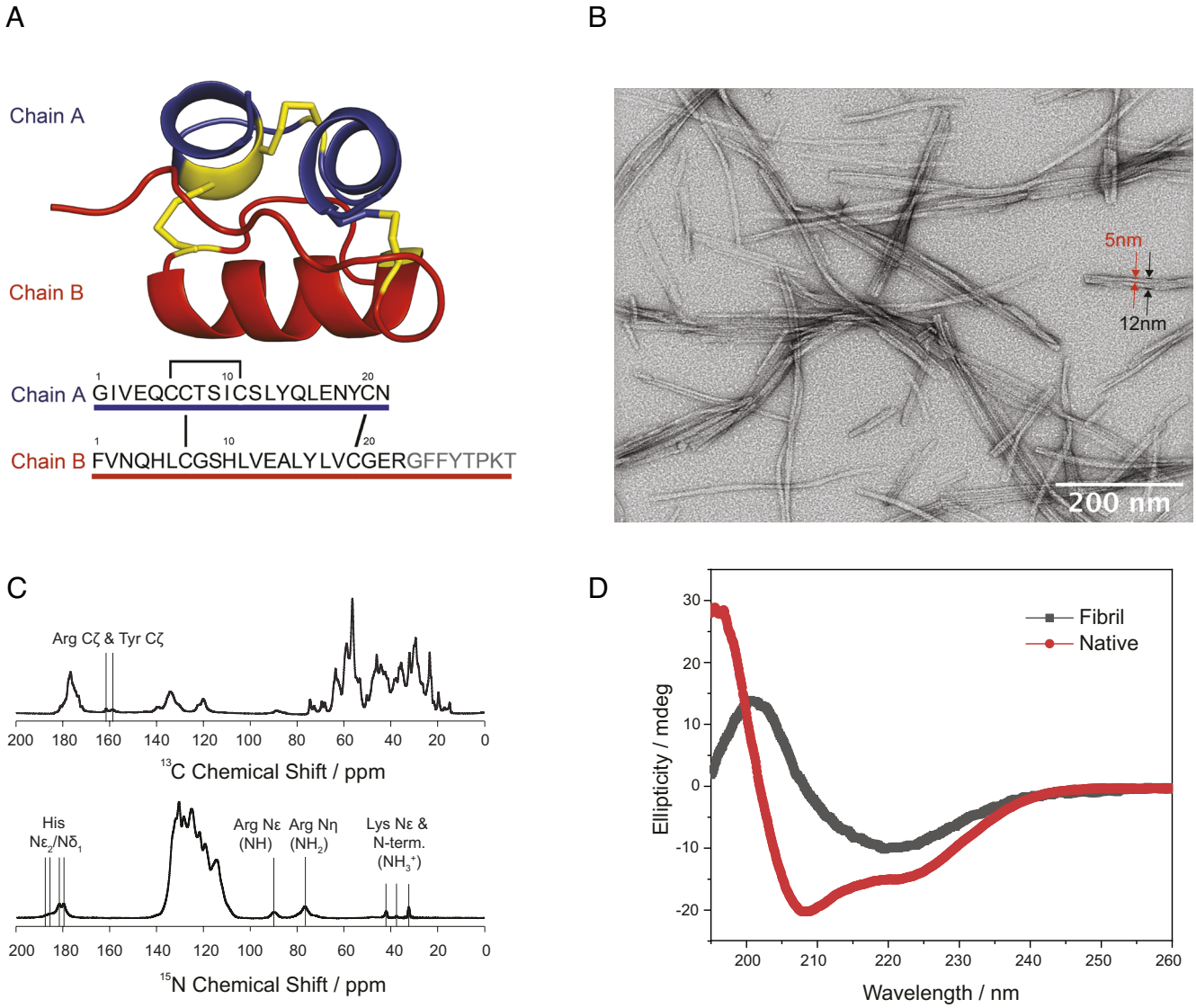

Fig. 1. (*A*) Native human insulin structure (PDB code 2HIU) (41). Chain A, chain B, and disulfide bonds are shown in blue, red, and yellow, respectively. The primary sequence of the individual chains and the position of the three native disulfide bonds are represented at the bottom. (*B*) Negative-stain TEM micrograph of human insulin fibrils used for solid-state NMR. (*C*) 1D ^{13}C and ^{15}N MAS solid-state NMR spectra of human insulin fibrils. Characteristic side chain resonances are indicated. (*D*) CD analysis of insulin fibrils of human insulin before (red) and after (black) 48 h of incubation at 60 °C, in HCl (pH 1.9). The CD spectrum after incubation is indicative of a β-sheet-rich structure.

(33, 34). Other studies focused mainly on the identification of potential amyloidogenic regions in the insulin sequence and characterized the residues which are crucial for fibril formation (31, 35–38). The fundamental principles of the molecular mechanism of amyloid formation and the details of the molecular structure of insulin amyloid fibrils are not yet understood. In this work, MAS solid-state NMR experiments are employed to identify the residues forming the fibril core. The ^{13}C and ^{15}N chemical shifts were used to yield information on the secondary structure propensity of a particular residue and to determine the exact location of the β-strands in the insulin fibril. We find that the core of the insulin fibril structure is formed by complementary residues from chains A and B. We collected more than 200 unique, nontrivial ^{13}C,^{13}C and ^{15}N,^{15}N distance restraints using long-mixing time PAR experiments (39, 40) to yield a structural model of the insulin fibril. The dimensions of the fibrils were estimated from transmission electron microscope (TEM) and small-angle X-ray scattering (SAXS) experiments. The insulin monomer adopts a U-shaped conformation which is stabilized by three disulfide bonds connecting the two insulin chains. These disulfide bridges largely determine the fold of the monomer. Consequently, each insulin molecule adopts two layers in the fibril topology, where each layer in turn contains two β-sheets. Two hydrophobic clusters further stabilize the U-shaped subunit. Surface-exposed hydrophobic residues facilitate the adherence of filaments and induce assembly of protofilaments into mature fibrils. We observe a single set of resonances suggesting that the minimal fibril subunit is a symmetric oligomer. We cannot differentiate between intra- and intermolecular contacts between atoms of residues in the same region of the fibrils suggesting that the minimal fibril subunit is presumably a dimer with cyclic symmetry.

Results and Discussion

Insulin fibrils formed spontaneously when the protein was incubated at 60 °C and pH 2. Fig. 1*B* shows a TEM image of the fibrils that were obtained this way and that were employed in the solid-state NMR experiments. The fibrils appear very homogeneous in terms of cross-over distance and thickness. The average length of a fibril is on the order of 500 nm. We do not observe any amorphous aggregates. Similar left-handed helical twisted fibrils have

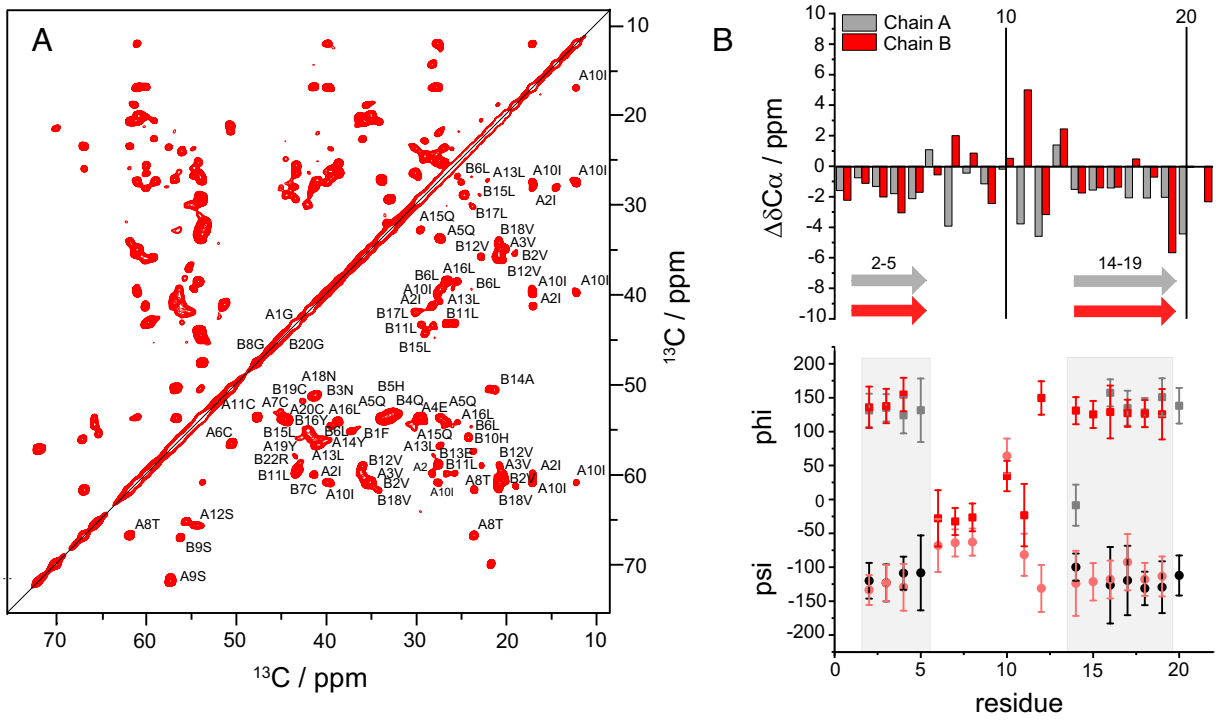

Fig. 2. (*A*) 50 ms mixing DARR spectrum obtained for a u-^{13}C,^{15}N labeled insulin fibril sample. Cross peak assignments are indicated in the figure. Unassigned cross peaks are due to peak doubling for loop residues of chain A and B (see below). (*B*) Residue-specific secondary ^{13}Cα NMR chemical shifts for chains A and B in insulin fibrils. Residues with negative Cα secondary chemical shifts have a high β-sheet propensity (*Top*). Distribution of ϕ & ψ backbone torsional angles according to Talos (45) (*Bottom*).

been reported previously (30). The solid-state NMR experiments yield information on the core of the fibril structure. In SAXS, we observe a fibril thickness of 8 to 10 nm which is consistent with the model that mature fibrils are composed of 2 to 6 intertwined insulin filaments (30).

Following the nomenclature introduced by Tycko, we refer to the "protofilament" as the filamentous unit in an amyloid fibril with minimal width and minimal mass-per-length (12). A protofilament refers to a stacked arrangement of monomers in a cross-β fibril configuration, where a monomer corresponds to an insulin molecule that consists of one A chain and one B chain that are linked by disulfide bridges.

MAS Solid-State NMR 1D Experiments. To confirm the structural homogeneity of the sample, we recorded 1D-^{13}C and ^{15}N spectra (Fig. 1*C*). In cross-polarization (CP) experiments, peaks with high intensity are observed suggesting that the respective residues are part of a well-defined rigid structure within the insulin fibrils. Beyond the amide resonances, the 1D-^{15}N spectra yield lysine amino (B29K, ~33 ppm), arginine guanidinium (B22R, ~72, 74, 84 ppm), histidine imidazole (B5H, B10H ~184, 193 ppm) side chain resonances, as well as the two N-terminal amino NH^{3+} groups (~33, 38, 42 ppm). The ^{15}N spectrum is a sensitive indicator of homogeneity and suggests that the insulin fibril sample is microscopically well ordered.

Chemical Shift Assignments. 2D DARR spectra (Fig. 2*A*) confirm the high structural homogeneity of the fibril sample. In the experiment, the ^{13}C mixing time was set to 50 ms. We observe ^{13}C line widths on the order of ~ 0.7 ppm for resolved resonances. To yield chemical shift assignments, we performed 3D NCOCX, NCACX, and CONCA experiments (42–44). Representative strip plots are shown in (*SI Appendix*, Fig. S3). Essentially, all observable cross peaks in the 3D spectra could be assigned. The most intense resonances in the 3D NCOCX, CONCA, and NCACX spectra are observed for residues 1 to 20 for both chains A and B, indicating that these residues are particularly rigid compared to the C-terminal part of the peptide. All residues of chain A (except A21N) could thus be sequentially assigned (*SI Appendix*, Table S3). The C-terminal residues of chain B (starting from B24F) yield no NMR signal and could not be assigned. Presumably, these residues are dynamic in agreement with previously published data (35). Scalar coupling–based experiments yield no significant intensity, suggesting that the C terminus of the B-chain is structurally heterogeneous (*SI Appendix*, Fig. S1). The assigned backbone and side chain atoms are schematically shown in *SI Appendix*, Fig. S2. Strip plots of the 3D assignment experiments are represented in *SI Appendix*, Fig. S3. Most residues yield a single set of resonances. Exceptions are residues in the loop regions of insulin chains A and B such as A8T, A10I and B6L-B10H, B14A, respectively, where we observe two sets of resonances. The appearance of two sets of resonances is related to the "arrangement of the fibril superstructure and packing of protofilaments" as discussed in the respective paragraph in detail below. This implies that the minimal fibril subunit consists of a symmetric oligomer and that mature insulin amyloid fibrils adopt a single polymorph under the conditions employed here.

Secondary Structure Analysis. Subsequently, the secondary chemical shifts of the assigned residues were used as input for TALOS+ to estimate the backbone torsion angles (ϕ, ψ) (Fig. 2*B*). In turn, the dihedral angles were employed to predict the location of the secondary structure elements within the insulin fibril. We find four β-strands, two per chain, in each insulin monomer. This is consistent with the cryo-EM insulin fibril model for which two β-strands per layer were reported (30). The exact location of the β-strands has not been described previously in the literature. Our analysis suggests that each insulin chain forms two complementary

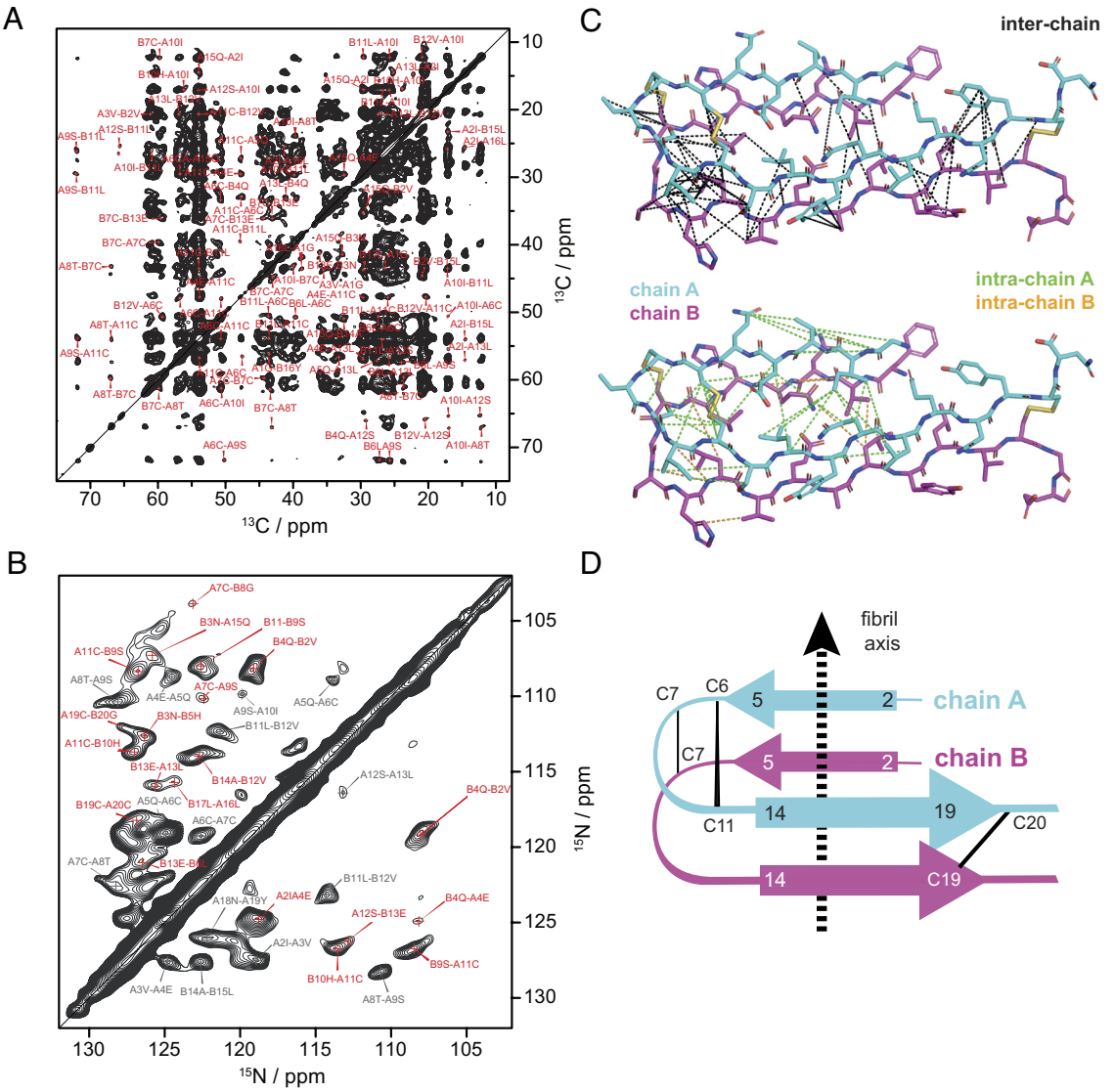

Fig. 3. Solid-state NMR long-range distance restraints and subunit model. (*A*) Aliphatic region of the 2D ^{13}C,^{13}C MAS PAR spectrum of u-^{13}C/^{15}N-insulin fibrils obtained for a mixing time τ_{mix} of 20 ms. Nontrivial long-range correlations are denoted with orange labels, e.g., A2I-A13L, A2I–B15L, B2V-B15L, B2V-A15Q, A6C-A10I, A6C-B12V, and B7C-B13E. (*B*) 2D ^{15}N,^{15}N MAS PAR correlation spectrum obtained for a u-^{13}C,^{15}N labeled insulin fibril sample acquired for a mixing time τ_{mix} of 30 ms. Long-range and sequential connectivities are denoted with red and gray labels, respectively. Long-range contacts such as A4E-B4Q, A11C-B9S, A11C-B10H, A12S-B13E, A13L-B13E, A16L-B17L, and A20C-B19C suggest a parallel arrangement of the two insulin chains in the fibril structure. (*C*) NMR-derived distance restraints. Interchain (*Top*) and intrachain (*Bottom*) connectivities are shown as dashed lines. A and B chain interactions are drawn in green and orange color, respectively. For clarity, only medium- and long-range distance restraints are shown. Chain A and B are colored in cyan and magenta, respectively. (*D*) Schematic representation of the insulin conformation in an amyloid protofilament. Each insulin molecule is arranged in two parallel layers, linked by interchain disulfide bonds. Chain A, chain B, and disulfide bonds are shown in cyan, magenta, and black, respectively.

β-strands involving residues 2 to 5 and 14 to 19. These β-strands are located in the core of the fibril structure and are supposedly rigid. The loop region involving residues 6 to 13 is stabilized by two disulfide bridges (A6C-A11C, A7C-B7C). The rigid configuration of the loop ensures a high quality of the solid-state NMR spectra. The strain induced by the two disulfide bonds also explains the atypical ^{13}C secondary chemical shifts observed for residues 6 to 12 in the chain A (Fig. 2*B*). Observation of unusual chemical shifts for disulfide-stabilized loops is not surprising and was observed previously in solution-state NMR experiments (37, 46). This part of the protein was therefore omitted in the subsequent secondary structure analysis using Talos.

Analysis of the Three-Dimensional Insulin Fibril Structure. Next, we were aiming for the three-dimensional structure of the fibril using solid-state NMR data. Experiments were recorded using a 100% uniformly ^{13}C/^{15}N labeled insulin sample. Proton-assisted recoupling (PAR) experiments were employed to achieve ^{13}C,^{13}C (39) and ^{15}N,^{15}N (40) recoupling and to yield long-range distance restraints. Fig. 3*A* shows a PAR spectrum that was obtained for a mixing time of τ_{mix} = 20 ms. The favorable resolution allowed us to extract numerous long-range distance restraints. Important contacts to determine the fold of the monomer structure include A2I-A13L, A2I–B15L, B2V-B15L, B2V-A15Q, A6C-A10I, A6C-B12V, and B7C-B13E. In addition, we recorded a ^{15}N,^{15}N PAR spectrum shown in Fig. 3*B*, which yields long-range ^{15}N,^{15}N contacts and provides information on backbone-to-backbone interactions. The most important medium- and long-range contacts are illustrated in Fig. 3*C*. We identified contacts between equivalent residues in chains A and B all along the fibril core indicating that the A and B chain β-sheets are arranged in parallel above or below one another. Particularly important cross-peaks include A4E-B4Q, A11C-B9S, A11C-B10H, A12S-B13E, A13L-B13E, A16L-B17L, and A20C-B19C. A complete list of the experimental contacts

is provided in *SI Appendix*, Table S1. The final set of restraints contains 205 unique contacts (178 ^{13}C,^{13}C + 27 ^{15}N,^{15}N), among which 17 restraints are long-range intrachain and 64 restraints are interchain contacts. The monomer structure is defined by 81 long-range and interchain restraints, which corresponds to 3.4 restraints per ordered residues. In the fibril, monomers are stacks on top of each other with the β-strands forming a parallel in-register β-sheet with a slight left-handed twist. We obtain a well-converged ensemble of fibrils with a backbone RMSD of 1.16 Å for three neighboring monomers. The monomer structure is well defined with an ensemble RMSD of 1.28 Å for the core residues when superimposing all nonhydrogen atoms. Burial of hydrophobic side chains appears to be the driving force for fibril formation. In addition, the fibril structure is constrained by the preservation of the disulfide bonds, which assists to induce a beta-roll fold (47). Cross-peaks obtained from a uniformly ^{13}C,^{15}N labeled sample can be due to both intramolecular and intermolecular contacts. Interestingly, in the case of insulin, the contacts observed in the 100 % labeled sample converge well to a single monomer structure during structure calculation, suggesting that insulin adopts a symmetric multimer structure.

The structural model obtained from the ARIA/CNS calculations is represented in Fig. 4. We observe two distinct hydrophobic cores, one containing residues B2V, B15L, B17L, the other B6L, B11L, A13L. The hydrophobic core regions are interrupted first by the hydrophilic residues B4Q, B13E which are part of the inner core. Next, the hydrophobic core is discontinued in every other layer by a more hydrophilic region in the chain A (residues A6C, A9S, A11C) and (residues A1G, A7E) which prevents an endlessly repeated hydrophobic core along the fibril axis. Hydrophilic residues A7C, B7C, A8T, B9S, B10H, and B12S in the middle of the sequence are solvent exposed and are part of the turn between the two main β-strands in each fibril layer. The arrangement of the disulfide bonds in parallel to the fibril axis is in agreement with Raman experiments (33).

We analyzed the stabilizing energy following the approach suggested by Eisenberg et al. (48, 49). The insulin molecule in the fibril experiences a Gibbs free energy of $\Delta G_0 = -12.9 \pm 1.3$ kcal/mol which corresponds to ~0.31 kcal/mol per residue. The respective energy map is shown in Fig. 4*A*. The overall energy is on the same order of magnitude as what is found for Alzheimer's disease Aβ(1–40) or hIAPP fibrils. Also hIAPP fibrils contain polar residues in the fibril core. However, the average per residue contribution to the stabilization (0.31 kcal/mol/residue) is rather at the lower end of the amyloid spectrum, close to values reported for proteins that undergo reversible fibril formation. Interestingly, the energy map of insulin fibrils has similar properties as glucagon fibrils that bury three negatively charged residues and that are stable only at low pH. In the estimation of the solvation energy, the disulfide-bonding network in insulin fibrils might not be fully considered and, as a consequence, the standard free energy is possibly underestimated.

The most destabilizing residues in the core are three glutamates (A4E, A17E, and B13E) (Fig. 4*C*). In the structural ensemble, hydrogen bonds between side chains of A4E-B4Q and B13E-A13Q, respectively, are occasionally found. In addition, the side chain A17E forms a hydrogen bond with the N-terminal amino group of A1G. Cross-peaks involving A4E-Cβ and A15Q-Cγ confirm the proximity of these residues in the core of the cross-β structure.

Arrangement of the Fibril Superstructure and Packing of Protofilaments. Next, we addressed the question how the fibril protofilaments are arranged to one another to yield a mature fibril. The insulin protofilaments expose seven hydrophobic side chains such as B1F, A3V, A10I, B12V, B14A, A16L, and B18V. These residues likely facilitate the attachment of other protofilaments. Packing of two protofilaments against one another allows to hide those exposed hydrophobic residues. In SAXS experiments, we observe a fibril cross-section diameter on the order of 8 to 10 nm. This diameter allows to accommodate 2 to 8 protofilaments in a twisted fibril. This is confirmed by several other studies (30–32), in which a fibril thickness of 8 to 10 nm was observed. These fibrils can accommodate 2 to 4 protofilaments with a dimension of 2 × 4 nm for each protofilament. The morphology of a mature fibril is obtained by cross-sectional arrangements of the structurally conserved building blocks that constitute the protofibrils (50). Stacking of protofilament pairs perpendicular to the fibril axis implies burial of hydrophobic clusters within the fibril core and allows to further increase fibril stability. We speculate that insulin fibril dimer/oligomers should be formed by parallel arrangement of identical subunits.

Choi et al. have suggested possible quaternary arrangements of protofilaments (32). Potential dimeric arrangements of the basic protofilament substructure obtained with ARIA using ambiguous intra-/interprotofilament restraints are shown in Fig. 4*B* and *SI Appendix*, Fig. S8. We can rule out a stacked arrangement, as this implies an asymmetric arrangement of the molecules in the fibrils. Atoms in the two laterally stacked subunits would experience a different chemical environment which would result in a doubling of the NMR resonances. For the β-strands in chains A and B, however, only one set of peaks is observed. Alternatively, the protofilament subunits can be arranged by implying $C2_Z$ symmetry. In this case, long-range contacts between either residues 1 to 7 or 10 to 18 of each chain should be observed. The NMR experiments do not support this, and rather suggest a parallel arrangement of the β-strands in chain A and B (Fig. 3*D*). Only, $C2_X$ symmetry yields a model that is in agreement with the MAS solid-state NMR data. The $C2_X^{Ct}$ symmetry, in which the C terminus is in close proximity to the symmetry axis, is supported by a higher number of explained NMR connectivities between protofilaments. Additionally, the dimeric arrangement with $C2_X^{Ct}$ symmetry is energetically more favorable as it allows to bury more hydrophobic residues in comparison to the $C2_X^{Nt}$ symmetry. Mixed interfaces involving both chain A and B are energetically not allowed since this arrangement results in unfavorable interactions between polar and nonpolar side chains. A quantitative estimate of the different symmetric ensembles obtained from ARIA calculations with respect to the experimental distance restraints, the stabilizing energies (ΔG^0 in kcal/mol) per molecule for each assembly and the buried surface area are represented in *SI Appendix*, Fig. S10.

To finally yield a mature insulin fibril and to match the results of the SAXS experiments, another symmetry operation is necessary. Application of $C2_Z$ symmetry yields a fibril with four protofilament subunits (Fig. 4*D*). Of note, we observe two sets of resonances in the loop regions of insulin chains A and B (Fig. 4*E*), involving A8T, A10I and B6L-B10H, B14A, respectively. Apparently, both (A·A)∘(A·A)′ and (B·B)∘(A·A)′ interfaces are possible resulting in differences in the chemical environment for the respective loop residues in chain A and B. The resulting fibril has a dimension of 10x5 nm which is in good agreement with the SAXS and TEM data.

Recently, an acid-induced insulin fibril cryo-EM structure (PDB-ID: 8SBD, EMD-40305) has been published (51). Even though fibrils were produced using the same preparation protocol, the cryo-EM structure is in conflict with earlier biophysical data and our solid-state NMR structural model. The cryo-EM structure

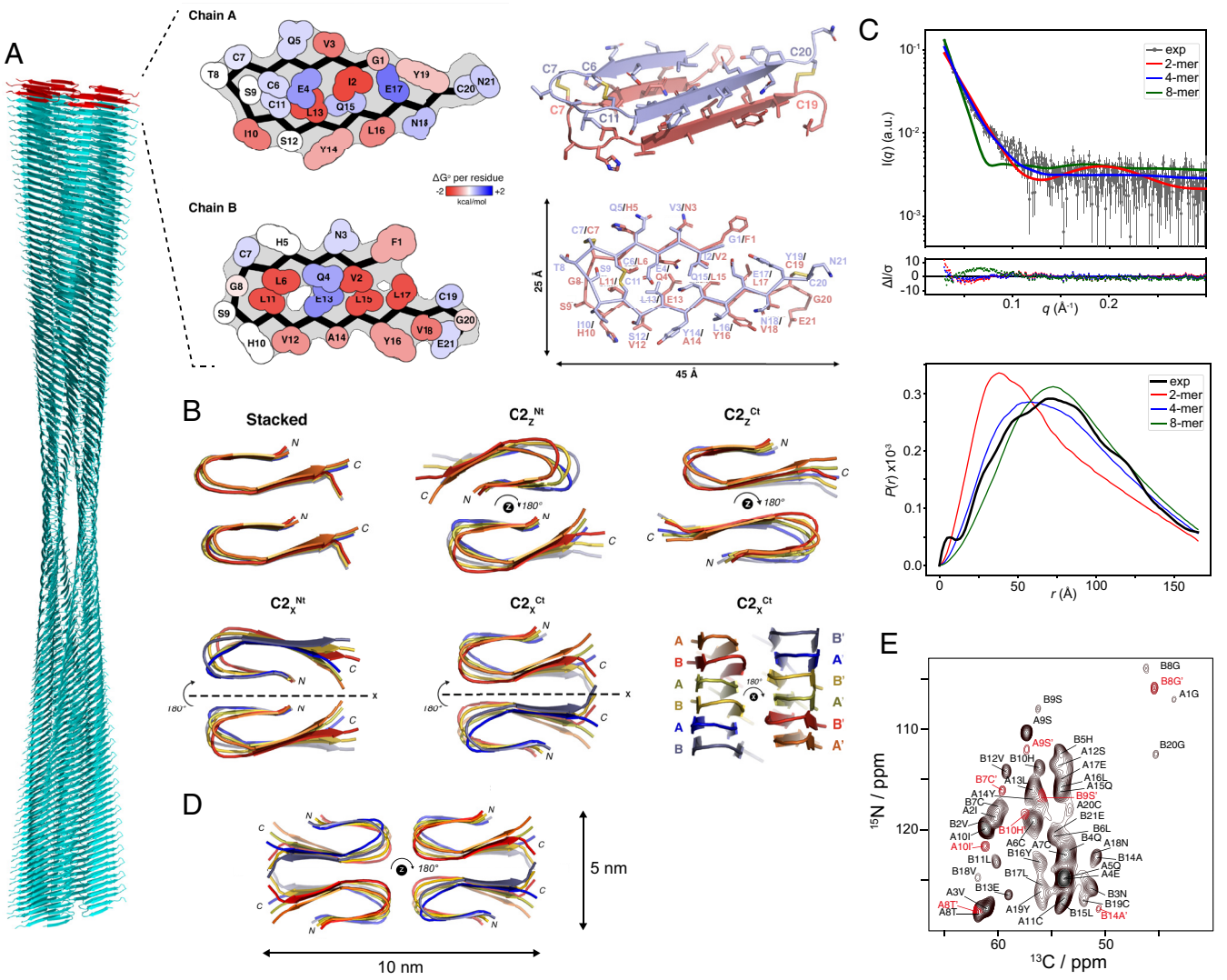

Fig. 4. Insulin fibril structure. (*A*) Cartoon representation of an extended insulin fibril. A tetrameric filament arrangement is shown that is in agreement with the experimental NMR and SAXS data. The first peptide layer is color in red. In addition, the side chain packing arrangement of chain A (*Top*) and chain B (*Bottom*) in insulin fibrils is shown. The fibril axis is oriented perpendicular to the peptide plane. Each residue is color coded using its standard free energy according to Sawaya et al. (48) Red and blue indicate stabilizing and destabilizing residues, respectively. The backbone trace is visualized as a thick black line. The Connolly surface is drawn in gray. In addition, the insulin monomer inside the fibril is represented in a cartoon and stick representation (*Right*, *Top*). To visualize hydrophobic matching, a superposition of chain A and B in the fibril structural model is shown on the bottom right. The first buried hydrophobic side chain cluster involves residues B2V, B15L, B17L, the second cluster includes B6L, B11L, A13L. Disulfide bridges are shown in yellow. Dimensions of the protofilament are indicated. All molecular figures were generated using the PyMOL software package. (*B*) Potential dimeric protofilament arrangements in the insulin fibril structure. The basic building block is represented in (*A*). Protofilaments can be either stacked upon each other laterally yielding an asymmetric arrangement (stacked). Alternatively, C2 symmetry can be applied with the symmetry axis aligned either in parallel ($C2_z$) or perpendicular ($C2_x$) to the fibril axis, with the N- or C-termini in proximity to the symmetry axis. The symmetry $C2_x^{Ct}$ matches the NMR-derived distance restraints. (*C*) Experimental and theoretical SAXS scattering curves and pairwise distribution function of the insulin fibril sample used in this study. The theoretical scattering curves were calculated assuming a dimeric, tetrameric, or octameric arrangement of insulin protofilaments. (*D*) Tetrameric bundle of an insulin fibril. The conformer is obtained after application of an additional $C2_z$ symmetry operation and is in agreement with the experimental EM, SAXS, and solid-state NMR data. (*E*) 2D-^{15}N,^{13}C correlation spectrum for a uniformly ^{13}C,^{15}N labeled insulin fibril with residue specific assignments. We observe two sets of resonances for residues in the loop regions of insulin chains A and B, involving residues A8T, A10I and B6L-B10H, B14A, respectively. The second set of resonances is color coded in red and labeled using an apostrophe.

contains a very small amount of cross-β topology involving only 19% of all residues. The interface of the two protomers in the fibril dimer adopts a non-β-sheet conformation. As such the fibril structure should suffer from a decreased stability. The cryo-EM dimer interface involves furthermore the C-terminus of the B-chain which has been shown to be flexible and dynamic in the fibril state (35). A possible fit of the solid-state NMR structural model into the cryo-EM map is shown in *SI Appendix*, Fig. S4. While isolated segments of the solid-state NMR monomeric model fit into the electron density, the overall topology is very different. The disulfide bonds in the cryo-EM structure are oriented perpendicular to the fibril axis. This is in contrast to Raman experiments which showed that the disulfide bonds are arranged in parallel to the fibril axis (33). Further, insulin fibrils appear to have diameters greater than 10 nm in negative stain TEM experiments. It is difficult to imagine how larger assemblies can be formed from an asymmetric dimer structure as the basic building block. The fibrils in the cryo-EM study were prepared at relatively low concentrations (1 mg/mL or 172 μM, respectively). By contrast, fibril formation for the samples described here and in previous biophysical studies was induced at a 10-fold higher protein concentration. The differences in concentration might result in a different fibril morphology which could explain the discrepancies between the two studies.

Structural Polymorphism of Insulin Fibrils. In vitro, insulin readily forms amyloid-like fibrils under a variety of conditions which promote dissociation of the native oligomer and which induces unfolding of the insulin monomer. These conditions include low protein concentration, acidic pH, high ionic strength, denaturant such as urea, guanidine hydrochloride, or organic solvents, high temperature, and extensive agitation. The overall shape of the amyloid aggregates depends on the conditions of their formation. However, sometimes different morphologies are observed in similar insulin fibril preparations (13, 30, 50, 52). Variability of amyloid fibril morphology is observed as well in ex vivo fibrils extracted from amyloidogenic tissue (53) This morphological variability is attributed to the number and arrangement of protofilaments. For fibrils formed by antibody light chains (50), transthyretin (54), and β2-microglobulin (13), differential arrangement of protofilaments has been shown to contribute to polymorphism, which was suggested to be a consequence of the lateral arrangement of protofilaments. Thereby, the molecular structure at the protofilament level is not considerably changed (13, 50, 52).

To address polymorphism, we prepared insulin fibrils using 20% acetic acid and compared the fibrils obtained this way with fibrils that were grown in the presence of 0.1 M HCl. A TEM image of the corresponding fibril preparation is displayed in *SI Appendix*, Fig. S5. In contrast to the samples prepared in HCl, we observe laterally associated insulin fibrils that are forming bundles. These results are consistent with prior findings showing that in the presence of 0.1 M HCl, insulin generally forms helical fibrils with a diameter of up to 10 nm (55, 56). In 20 % acetic acid, similar fibrils are formed but with less twist and a larger diameter on the order of 60 to 80 nm. Despite the apparent morphological differences, the insulin fibril spectra yield almost identical 2D $^{13}C,^{13}C$ correlation spectra (*SI Appendix*, Fig. S6*A*) indicating a preserved fibril core structure. Interestingly, the N-terminal residues of chain B are not observable in the acetic acid preparation (residues B1F-B5H) (*SI Appendix*, Table S4). This is puzzling since the N-terminus of the B chain is an integral part of the insulin fibril structural model. In the A chain, the loop connecting the two β-strands is stabilized by a disulfide bond involving residues Cys-6 and Cys-11. In chain B, this disulfide bond is missing and we speculate that this results in a decreased stability of the fibril. In fact, we find that the chemical shifts of the observable atoms in both preparations are highly correlated (*SI Appendix*, Fig. S6*B*), suggesting that the conformational differences are very small in the two preparations. In the presence of 20 % acetic acid, thicker fibrils are observed in EM images suggesting that larger fibril assemblies are obtained by packing of tetrameric filaments. Again, assemblies in which chains A and B form mixed-chain interactions are not populated, since we only observe one set of resonances for the first β-strands of chain A and B, respectively. Chain B contains a histidine residue that is protonated under the conditions that are employed in this study. The lack of a disulfide bond, and repulsion between positively charged histidine side chains is presumably responsible for the disappearance of the N-terminal residues of chain B.

In patients, injection-localized amyloidosis forms under physiological conditions. In addition, nearly all-commercial injectable insulin formulations (except insulin glargine) are currently formulated at neutral pH. Here, we characterized the insulin fibril structure at low pH. It will be interesting to see whether the fibril structure presented here will be preserved, and whether the glutamate side chains (A4E, B13E) will be protonated at neutral pH. Interestingly, in a recent cryo-EM structural model of Alzheimer's disease Aβ42 fibrils that was solved at acidic pH, stabilizing interactions between lysine and aspartate side chains were observed (57). It has been shown that the fold of the Aβ(1 – 42) fibrils does not change over the complete pH range from pH 2 to 7 (58). Similarly, the increase in pH from 2 to 8 induces a partial disruption of the cross-β network in insulin fibrils leading to a reduction of stability without disrupting the amyloid fibril structure (59). Further structural studies under conditions that resemble more closely the physiological environment are needed to extend our understanding of the nature of injection-localized amyloidosis.

Conclusions

So far, the structural features of insulin fibrils were yet unknown. We performed MAS solid-state NMR experiments using uniformly $^{13}C,^{15}N$ labeled protein to yield three-dimensional models of insulin fibril structures at low pH. These fibrils consist of a single polymorph. Insulin molecules are arranged in two layers, adopt a U-shaped conformation for both the chain A and B, respectively, and fold into four β-strands within the fibril unit cell. Interchain $^{15}N,^{15}N$ distance restraints suggest that the amyloid fibrils are arranged in parallel and in register. Residues 1 to 20 of each chain are part of the amyloid fibril core, while the C-terminal residues of chain B (21 to 31) appear to be unstructured. The monomer structure has a heavy atom backbone RMSD of ~1.4 Å. The structure of the basic subunit has a dimension of ~2 × 4 nm. We identified two hydrophobic clusters buried in the interior of the fibril, while two hydrophobic patches exposed on the surface potentially serve as a dimerization interface and a site for seeding and secondary nucleation. In the mature fibril, protofilaments are packed into a dimer-of-dimers arrangement which allows to efficiently bury solvent-exposed hydrophobic amino acids consistent with SAXS and TEM analysis. Our structure represents a valuable resource for the understanding of insulin aggregation and opens an avenue for the rational design of small molecules that have high selectivity for the fibril surface and that can potentially disrupt secondary nucleation.

Materials and Methods

Sample Preparation. Isotopically labeled insulin was obtained from Sanofi-Aventis and was produced as described previously (60). Briefly, recombinant human insulin was produced by *Escherichia coli* into inclusion bodies. During cultivation ^{13}C labeled glucose was used as carbon source and ^{15}N labeled ammonia as nitrogen source. The peptide was purified from inclusion bodies by chromatographic steps including preparative CEX and RP resins as summarized before (61).

Preparation of the Solid-State NMR Sample. The fibril sample was prepared by incubating the insulin monomer solution at pH 1.9 in the presence of 5 % seeds at 60 °C at a concentration of 10 mg/mL for 24 h. Seeds were prepared from preformed fibrils by sonication for 30 s in a water bath. To prepare one solid-state NMR sample, 10 mg of protein was used. A Beckman Coulter ultracentrifuge equipped with a SW32Ti swinging bucket rotor (run at 28,000 rpm), in combination with a packing device from Giotto Biotech (https://www.giottobiotech.com) was employed to transfer amyloid fibrils into a Bruker 1.9 mm MAS rotor. This way, fibrils were directly sedimented into the MAS solid-state NMR rotor.

TEM. Fibrils were characterized by TEM, using Formvar/Carbon 300 mesh copper-coated grids from Electron Microscopy Sciences (Hatfield, PA). To charge a grid, 5 μL of sample was pipetted on the surface for 2 min. The excess material was blotted off using filter paper. Subsequently, the grid was washed two times with distilled water. The excess liquid was wiped off using filter paper. Afterward, 5 μL of 2 % uranyl acetate was applied to the grid. After 30 s of incubation time, the staining solution was removed with a filter paper. Finally, the grid was dried for 10 to 15 min. A JEOL 1400 plus microscope (JEOL) operating at 120 kV was employed to take micrographs, using a Ruby camera. The nominal magnification was set to 60 k, which yields a pixel size of 0.275 nm/px. IMAGEJ2 software was used to apply the scale bar.

CD Spectroscopy. CD experiments were carried out at room temperature employing a JASCO spectropolarimeter. Spectra were collected at a final insulin concentration of 20 μM. Quartz cells with a path length of 0.1 cm were used. Spectra were recorded between 190 nm and 260 nm at 0.1 nm intervals with a response time of 1 s. Each spectrum represents the average of five scans. To avoid instrumental baseline drift between measurements, the background buffer signal was subtracted.

Solid-State NMR. Solid-state fibril samples were measured at a magnetic field strength of 17.6 T which corresponds to a ^{1}H Larmor frequency of 750 MHz. The measurements were acquired setting the MAS frequency to 16.65 kHz. The effective sample temperature was adjusted to 0 °C. 2D ^{13}C,^{13}C correlation experiments were acquired using DARR (62) for mixing, employing a mixing time of 50 ms. To assign the NMR chemical shifts, conventional 3D NCACX, 3D NCOCX, and 3D CONCA were recorded (42, 43). For ^{13}C,^{15}N transfers, specific CP-based experiments were employed (63). In addition, 3D CANCO experiments were performed to confirm and to assign ambiguous residues. In these experiments, optimal control CP (OC-CP) was used to improve sensitivity (64, 65). Long-range interactions were obtained from long mixing time as well as PAR experiments (39) (^{13}C,^{13}C: 20 ms, ^{15}N,^{15}N: 30 ms). CCPN (66) was employed to assign, compare, and characterize all solid-state NMR spectra.

SAXS. SAXS experiments were performed on a Rigaku BioSAXS1000 instrument attached to a Rigaku HF007 microfocus rotating anode with a copper target (40 kV, 30 mA). Transmissions were measured with a photodiode beam stop. The scattering vector calibration was done with a silver behenate sample. Fibril samples with concentrations of 2.5 mg/mL and 5 mg/mL as well as respective buffers were measured in single capillaries with approx. 50 μL sample capacity. A typical measurement consisted of eight 900 s frames, which were compared to check for radiation damage, for a total measurement duration of 7,200 s. Sample temperature was controlled using a Julabo F25-MA thermostat with a specified temperature stability of ±0.02 K. Circular averaging and solvent subtraction were done using Rigaku SAXSLab software. Respective pair distance distribution functions were obtained with the GNOM program from the ATSAS 3.0.4 software suite (67). Backcalculation of scattering profiles was performed using CRYSOL from the ATSAS suite (68).

Structure Calculation of Insulin Fibrils. The structure of insulin fibers was calculated using ARIA2.3 (69), in conjunction with the software CNS (70) in an iterative way with several cycles of restraints analysis. Distance restraints derived from ^{13}C,^{13}C and ^{15}N,^{15}N cross-peaks were employed with an upper bound of 7.5 Å and allowed to be automatically discarded by ARIA across iterations. In the calculation, five insulin molecules (chain A and chain B$\Delta_{22\text{-}30}$) were modeled, while stacking of protomers was imposed using symmetry-distance restraints between neighboring molecules (71, 72). An additional CNS energy term was added to minimize the RMSD between insulin molecules (71). Canonical backbone dihedral angles for β-strand topology ($\phi = -135° \pm 20°$; $\psi = 135° \pm 20°$) were imposed for residues 1 to 6 and 14 to 19 of chain A and B. In each ARIA iteration, 100 conformers were calculated. The number of simulated annealing steps was increased to 20,000/120,000 for the high-temperature and cooling stages, respectively, to ensure convergence. Three disulfide bridges (A6C–A11C, A7C–B7C, and A20C–B19C) were imposed during structure calculation since preliminary calculations without any disulfide-bond restraints converged to the same topology and spatial proximity of the cysteine side chains.

An initial round of ARIA calculation was performed using NMR-derived distance restraints and ambiguous Cα-Cα restraints between residue *i* of a particular chain and residues [*i*-1, *i*, *i*+1] of a neighboring chain (intra- or intermolecular), imposing an upper-bound of 4.8 Å, expected for parallel β-strands. Additionally, an empirical hydrogen-bonding potential (73) was incorporated into the energy scoring function to favor hydrogen bond formation of backbone donors and acceptors that are in close proximity and which adopt favorable geometry. This strategy was designed to objectively determine intra- and intermolecular β-strand arrangements without imposing explicit hydrogen-bonding restraints between specific residues, nor register, nor orientation of buried/exposed side chains of residues in β-strands.

In parallel, a series of ARIA calculations was carried out with NMR restraints and explicit hydrogen bonding restraints (intra- and intermolecular) imposing different orientations (inside/outside) and register (*i*, *i*-1, *i*+1) of β-strands of chain A and B along the fiber axis. The comparison of the structure and restraints statistics among these 12 calculations (restraints, energy, convergence, etc.) shows that the calculation with in-register β-strand pairing and buried Q15(A)/L15(B) side chains is most favorable (*SI Appendix*, Figs. S5–S7 and S9). The resulting topology of the insulin monomer structure in the fibril is highly similar to the one obtained without explicit hydrogen-bonding restraints. The monomer structure is well defined with an ensemble RMSD of 1.28 Å for well-ordered residues when superimposing all nonhydrogen atoms (*SI Appendix*, Fig. S5 and Table S2). The fibril ensemble converged well with a backbone RMSD of 1.16 Å for three neighboring monomers. Finally, after manual inspections of unsatisfied restraints, a final round of ARIA calculation was performed with NMR restraints and the most favorable hydrogen-bond restraints, followed by refinement in a shell of water molecules of the 10 lowest-energy conformers (74). The quality of the final structural ensemble of insulin fibers was evaluated with PROCHECK (75) and PSVS (76). The restraints and structure statistics is given in *SI Appendix*, Tables S1 and S2.

Modeling of 2-Protofilament Assemblies. Models of 2-protofilament assemblies were generated using a similar procedure as described above for the single filament calculation. Now, 10 copies of the insulin molecule were calculated in ARIA (2 × 5 molecules, each with two chains, corresponding to two protofilaments). To include possible contacts between the two protofilaments, each assigned cross-peak was considered to be originating from either intra-protofilament or interprotofilament contacts, thus resulting in ambiguous distance restraints. Hydrogen-bond restraints and dihedral angle restraints were added similarly as in the single filament calculation. On top of the restraints maintaining the symmetry of each protofilament (a group of five molecules), a noncrystallographic symmetry (NCS) term was added to minimize the RMSD between the atomic coordinates of the two groups of five molecules. In addition, a packing restraint between the centers of mass of the two groups was included. Two types of topologies were tested implying either C2 symmetry or no symmetry between the two protofilaments. Point symmetry between the protofilaments was imposed using distance restraints in the calculations, thus avoiding to specify any explicit symmetry axis (69, 71). Upon visual inspection of the C2 symmetric calculation, the pool of generated conformers was clustered based on Cα atom coordinates (77). This resulted in four clusters which are referred to as C_{2Z}^{Nt}, C_{2Z}^{Ct}, C_{2X}^{Nt}, C_{2X}^{Ct}. For the stacked assembly, no C2 symmetry was imposed. However, ambiguous interprotofilament restraints were allowed between the N-terminal part of a protofilament (residues A1–A9+B1–B9) and the C-terminal region of the adjacent one (residues A9–A21+B9–B21). The initial calculations without symmetry resulted in two virtually identical set of conformers in which the two protofilaments were simply exchanged. Statistical analyses (cross-peaks distances, free energy, and buried surface area) were computed using the eight lowest energy conformers from each cluster.

Energetic and Interface Analysis. Standard free energy of insulin fibrils was calculated using the program developed by Sawaya et al. which was downloaded at https://doi.org/10.5281/zenodo.6321286 (48, 49). The calculations were performed using the coordinates of the pentameric NMR structure of insulin fibrils. The indicated energy values refer to the energy determined for the center molecule in the pentamer. Energies were averaged over all conformers of the NMR ensemble. For models with two protofilaments, the reported energy corresponds to the mean energy of the central monomer from each protofilament. The buried surface area of the 2-protofilament assembly models was computed using NACCESS software (78). The buried surface area, i.e., the interface area between two protofilaments, was calculated as the difference between the sum of the accessible surface area of the central monomer in each isolated protofilament and the surface area of the two central monomers in the assembly, divided by two.

Data, Materials, and Software Availability. The ^{13}C/^{15}N chemical shifts for the insulin fibrils are deposited in the BioMagResBank (BMRB) under the ID is 51867 (79). The fibril structure is uploaded in the PDB under the accession code 8RVT (80). All other data are included in the manuscript and/or *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) with the grant Re1435/25-1. We acknowledge financial support from the Helmholtz-Gemeinschaft. We acknowledge Matthias Brandl for support with in-house SAXS data acquisition.

Author affiliations: [a]Bavarian Nuclear Magnetic Resonance Center at the Department of Biosciences, School of Natural Sciences, Technische Universität München, Garching 85747, Germany; [b]Helmholtz-Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Institute of Structural Biology, Neuherberg 85764, Germany; [c]Sanofi-Aventis Deutschland GmbH, Diabetes Research, Industriepark Höchst, Frankfurt 65926, Germany; [d]Sanofi-Aventis Deutschland GmbH, Manufacturing Science and Technology, Industriepark Höchst, Frankfurt 65926, Germany; [e]Chemistry Manufacturing & Controls Microbial Platform, Sanofi-Aventis Deutschland GmbH, Microbial Platform, Industriepark Höchst, Frankfurt 65926, Germany; [f]Sanofi-Aventis Deutschland GmbH, Tides Platform, Industriepark Höchst, Frankfurt 65926, Germany; [g]Institut Pasteur, Department of Structural Biology and Chemistry, Structural Bioinformatics Unit, CNRS UMR 3528, Université Paris Cité, Paris 75015, France; and [h]Institut Pasteur, Department of Structural Biology and Chemistry, Bacterial Transmembrane Systems Unit, CNRS UMR 3528, Université Paris Cité, Paris 75015, France

1. S. G. Swinnen, J. B. Hoekstra, J. H. DeVries, Insulin therapy for type 2 diabetes. *Diabetes Care* **32**, S253–S259 (2009).
2. M. G. Minze, L. M. Chastain, Combination therapies in the management of type 2 diabetes: The use of insulin degludec/liraglutide. *Ther. Clin. Risk Manage.* **12**, 471–478 (2016).
3. B. Swift, P. N. Hawkins, C. Richards, R. Gregory, Examination of insulin injection sites: An unexpected finding of localized amyloidosis. *Diabetes Med.* **19**, 881–882 (2002).
4. M. R. Nilsson, Insulin amyloid at injection sites of patients with diabetes. *Amyloid* **23**, 139–147 (2016).
5. T. Nagase *et al.*, Insulin-derived amyloidosis and poor glycemic control: A case series. *Am. J. Med.* **127**, 450–454 (2014).
6. F. E. Dische *et al.*, Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. *Diabetologia* **31**, 158–161 (1988).
7. S. Frokjaer, D. E. Otzen, Protein drug stability: A formulation challenge. *Nat. Rev. Drug Discov.* **4**, 298–306 (2005).
8. M. Akbarian, Y. Ghasemi, V. N. Uversky, R. Yousefi, Chemical modifications of insulin: Finding a compromise between stability and pharmaceutical performance. *Int. J. Pharm.* **547**, 450–468 (2018).
9. L. Nielsen *et al.*, Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochemistry* **40**, 6036–6046 (2001).
10. K. Iwaya *et al.*, Toxicity of insulin-derived amyloidosis: A case report. *BMC Endocr. Disord.* **19**, e61 (2019).
11. K. R. Wilhelm *et al.*, Immune reactivity towards insulin, its amyloid and protein S100B in blood sera of Parkinson's disease patients. *Eur. J. Neurol.* **14**, 327–334 (2007).
12. R. Tycko, Molecular structure of amyloid fibrils: Insights from solid-state NMR. *Q. Rev. Biophys.* **39**, 1–55 (2006).
13. M. G. Iadanza *et al.*, The structure of a beta(2)-microglobulin fibril suggests a molecular basis for its amyloid polymorphism. *Nat. Commun.* **9**, e4517 (2018).
14. X. C. Bai, G. McMullan, S. H. W. Scheres, How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57 (2015).
15. D. S. Eisenberg, M. R. Sawaya, "Structural studies of amyloid proteins at the molecular level" in *Annual Review of Biochemistry*, Kornberg, Ed. (Stanford University, California, 2017), vol. 86, pp. 69–95.
16. C. Wasmer *et al.*, Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526 (2008).
17. B. H. Meier, R. Riek, A. Böckmann, Emerging structural understanding of amyloid fibrils by solid-state NMR. *Trends Biochem. Sci.* **42**, 777–787 (2017).
18. J. X. Lu *et al.*, Molecular structure of beta-amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **154**, 1257–1268 (2013).
19. Y. L. Xiao *et al.*, A beta(1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat. Struct. Mol. Biol.* **22**, 499–505 (2015).
20. M. T. Colvin *et al.*, Atomic resolution structure of monomorphic A beta(42) amyloid fibrils. *J. Am. Chem. Soc.* **138**, 9663–9674 (2016).
21. M. D. Tuttle *et al.*, Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
22. M. D. Gelenter *et al.*, The peptide hormone glucagon forms amyloid fibrils with two coexisting beta-strand conformations. *Nat. Struct. Mol. Biol.* **26**, 592–598 (2019).
23. D. T. Murray *et al.*, Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell* **171**, 615–627 (2017).
24. C. Seuring *et al.*, The three-dimensional structure of human beta-endorphin amyloid fibrils. *Nat. Struct. Mol. Biol.* **27**, 1178–1184 (2020).
25. X. Wu *et al.*, The structure of a minimum amyloid fibril core formed by necroptosis-mediating RHIM of human RIPK3. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2022933118 (2021).
26. A. K. Buell *et al.*, Electrostatic effects in filamentous protein aggregation. *Biophys. J.* **104**, 1116–1126 (2013).
27. J. Haas *et al.*, Primary steps of pH-dependent insulin aggregation kinetics are governed by conformational flexibility. *ChemBioChem* **10**, 1816–1822 (2009).
28. R. Malik, I. Roy, Probing the mechanism of insulin aggregation during agitation. *Int. J. Pharm.* **413**, 73–80 (2011).
29. L. Nault, C. Vendrely, Y. Brechet, F. Bruckert, M. Weidenhaupt, Peptides that form beta-sheets on hydrophobic surfaces accelerate surface-induced insulin amyloidal aggregation. *FEBS Lett.* **587**, 1281–1286 (2013).
30. J. L. Jimenez *et al.*, The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9196–9201 (2002).
31. M. I. Ivanova, S. A. Sievers, M. R. Sawaya, J. S. Wall, D. Eisenberg, Molecular basis for insulin fibril assembly. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18990–18995 (2009).
32. J. H. Choi, B. C. H. May, H. Wille, F. E. Cohen, Molecular modeling of the misfolded insulin subunit and amyloid fibril. *Biophys. J.* **97**, 3187–3195 (2009).
33. D. Kurouski *et al.*, Disulfide bridges remain intact while native insulin converts into amyloid fibrils. *PLoS ONE* **7**, e36989 (2012).
34. V. Sereda, M. R. Sawaya, I. K. Lednev, Structural organization of insulin fibrils based on polarized raman spectroscopy: Evaluation of existing models. *J. Am. Chem. Soc.* **137**, 11312–11320 (2015).
35. L. Nielsen, S. Frokjaer, J. Brange, V. N. Uversky, A. L. Fink, Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* **40**, 8397–8409 (2001).
36. G. L. Devlin *et al.*, The component polypeptide chains of bovine insulin nucleate or inhibit aggregation of the parent protein in a conformation-dependent manner. *J. Mol. Biol.* **360**, 497–509 (2006).
37. Y. W. Yang *et al.*, An Achilles' heel in an amyloidogenic protein and its repair. Insulin fibrillation and therapeutic design. *J. Biol. Chem.* **285**, 10806–10821 (2010).
38. A. K. Surin, S. Y. Grishin, O. V. Galzitskaya, Determination of amyloid core regions of insulin analogues fibrils. *Prion* **14**, 149–162 (2020).
39. G. De Paepe, J. R. Lewandowski, A. Loquet, A. Böckmann, R. G. Griffin, Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* **129**, e245101 (2008).
40. J. R. Lewandowski, G. De Paepe, M. T. Eddy, R. G. Griffin, 15)N-(15)N proton assisted recoupling in magic angle spinning NMR. *J. Am. Chem. Soc.* **131**, 5769–5776 (2009).
41. Q. X. Hua *et al.*, Structure of a protein in a kinetic trap. *Nat. Struct. Biol.* **2**, 129–138 (1995).
42. A. McDermott *et al.*, Partial assignments for uniformly (^{13}C,^{15}N)-enriched BPTI in the solid state. *J. Biomol. NMR* **16**, 209–219 (2000).
43. J. Pauli, M. Baldus, B.-J. Van Rossum, H. De Groot, H. Oschkinat, Backbone and side-chain ^{13}C and ^{15}N signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. *ChemBioChem.* **2**, 272–281 (2001).
44. A. Böckmann *et al.*, Solid state NMR sequential resonance assignments and conformational analysis of the 2 x 10.4 kDa dimeric form of the *Bacillus subtilis* protein Crh. *J. Biomol. NMR* **27**, 323–339 (2003).
45. Y. Shen, A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
46. K. Krizkova *et al.*, Structural and functional study of the GlnB22-insulin mutant responsible for maturity-onset diabetes of the young. *PLoS ONE* **9**, e112883 (2014).
47. A. V. Kajava, A. C. Steven, beta-rolls, beta-helices, and other beta-solenoid proteins. *Adv. Prot. Chem.* **73**, 55–96 (2006).
48. M. R. Sawaya, M. P. Hughes, J. A. Rodriguez, R. Riek, D. S. Eisenberg, The expanding amyloid family: Structure, stability, function, and pathogenesis. *Cell* **184**, 4857–4873 (2021).
49. D. Eisenberg, A. D. McLachlan, Solvation energy in protein folding and binding. *Nature* **319**, 199–203 (1986).
50. W. Close *et al.*, Physical basis of amyloid fibril polymorphism. *Nat. Commun.* **9**, 699 (2018).
51. L. W. Wang *et al.*, Structural basis of insulin fibrillation. *Sci. Adv.* **9**, eadi1057 (2023).
52. O. M. Selivanova, O. V. Galzitskaya, Structural polymorphism and possible pathways of amyloid fibril formation on the example of insulin protein. *Biochemistry (Mosc.)* **77**, 1237–1247 (2012).
53. J. L. Jimenez, G. Tennent, M. Pepys, H. R. Saibil, Structural diversity of ex vivo amyloid fibrils studied by cryo-electron microscopy. *J. Mol. Biol.* **311**, 241–247 (2001).
54. A. W. Fitzpatrick *et al.*, Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5468–5473 (2013).
55. J. L. Whittingham *et al.*, Insulin at pH 2: Structural analysis of the conditions promoting insulin fibre formation. *J. Mol. Biol.* **318**, 479–490 (2002).
56. L. Nielsen, S. Frokjaer, J. F. Carpenter, J. Brange, Studies of the structure of insulin fibrils by Fourier transform infrared (FTIR) spectroscopy and electron microscopy. *J. Pharm. Sci.* **90**, 29–37 (2001).
57. L. Gremer *et al.*, Fibril structure of amyloid-beta(1–42) by cryo-electron microscopy. *Science* **358**, 116–119 (2017).
58. N. Becker *et al.*, Atomic resolution insights into pH shift induced deprotonation events in LS-shaped Abeta(1–42) amyloid fibrils. *J. Am. Chem. Soc.* **145**, 2161–2169 (2023).
59. S. L. Shammas *et al.*, Perturbation of the stability of amyloid fibrils through alteration of electrostatic interactions. *Biophys. J.* **100**, 2783–2791 (2011).
60. J. Sandow, W. Landgraf, R. Becker, G. Seipke, Equivalent recombinant human insulin preparations and their place in therapy. *Eur. Endocrinol.* **11**, 10–16 (2015).
61. Y. Y. Siew, W. Zhang, Downstream processing of recombinant human insulin and its analogues production from *E. coli* inclusion bodies. *Bioresour. Bioprocess.* **8**, e65 (2021).
62. K. Takegoshi, S. Nakamura, T. Terao, C-13–H-1 dipolar-driven C-13–C-13 recoupling without C-13 rf irradiation in nuclear magnetic resonance of rotating solids. *J. Chem. Phys.* **118**, 2325–2341 (2003).
63. M. Baldus, A. T. Petkova, J. Herzfeld, R. G. Griffin, Cross polarization in the tilted frame: Assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **95**, 1197–1207 (1998).
64. Z. Tošner *et al.*, Overcoming volume selectivity of dipolar recoupling in biological solid-state NMR. *Angew. Chem. Int. Ed. Engl.* **57**, 14514–14518 (2018).
65. Z. Tošner, M. J. Brandl, J. Blahut, S. J. Glaser, B. Reif, Maximizing efficiency of dipolar recoupling in solid-state NMR using optimal control sequences. *Sci. Adv.* **7**, eabj5913 (2021).
66. T. J. Stevens *et al.*, A software framework for analysing solid-state MAS NMR data. *J. Biomol. NMR* **51**, 437–447 (2011).
67. K. Manalastas-Cantos *et al.*, ATSAS 3.0: Expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
68. D. Franke *et al.*, ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).
69. B. Bardiaux, T. Malliavin, M. Nilges, "ARIA for solution and solid-state NMR" in *Protein NMR Techniques*, Shekhtman, Burz, Eds. (ed. 3, Humana Press, New York, 2012), vol. **831**, pp. 453–483.
70. A. T. Brünger *et al.*, Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D* **54**, 905–921 (1998).
71. M. Nilges, A calculation strategy for the structure determination of symmetric dimers by ^{1}H NMR. *Proteins* **17**, 297–309 (1993).
72. A. Daskalov *et al.*, Structural and molecular basis of cross-seeding barriers in amyloids. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014085118 (2021).
73. A. Grishaev, A. Bax, An empirical backbone–backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J. Am. Chem. Soc.* **126**, 7281–7292 (2004).
74. J. P. Linge, M. A. Williams, C. A. Spronk, A. M. Bonvin, M. Nilges, Refinement of protein structures in explicit solvent. *Proteins* **50**, 496–506 (2003).
75. R. A. Laskowski, J. A. C. Rullmann, M. W. MacArthur, R. Kaptein, J. M. Thornton, AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
76. A. Bhattacharya, R. Tejero, G. T. Montelione, Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778–795 (2007).
77. X. Daura *et al.*, Peptide folding: When simulation meets experiment. *Angew. Chem. Int. Ed. Engl.* **38**, 236–240 (1999).
78. S. J. Hubbard, J. M. Thornton, *NACCESS, Computer Program* (Department of Biochemistry and Molecular Biology, University College London, 1993).
79. S Suladze *et al.*, Assignment of insulin fibril chemical shifts. BMRB. https://doi.org/10.13018/BMR51867. Deposited 3 March 2023.
80. B. Bardiaux, S. Suladze, B. Reif, Structure of full-length human insulin fibrils. RCSB PDB. http://doi.org/10.2210/pdb8rvt/pdb. Deposited 2 February 2024.