THE JOURNAL OF PHYSICAL CHEMISTRY B

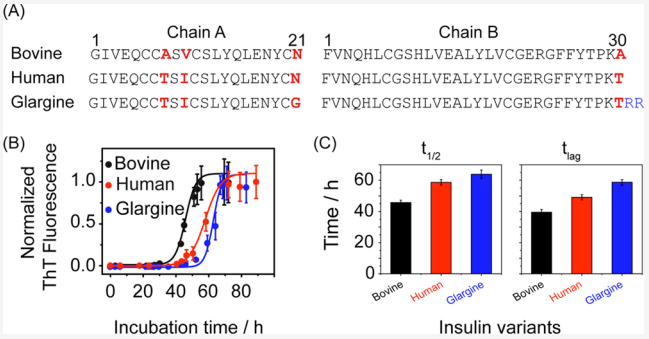
pubs.acs.org/JPCB Article

Molecular Details of a Salt Bridge and Its Role in Insulin Fibrillation by NMR and Raman Spectroscopic Analysis

Bhisma N. Ratha, Rajiv K. Kar, Zuzana Bednarikova, Zuzana Gazova, Samuel A. Kotler, Sreyan Raha, Soumya De, Nakul C. Maiti, and Anirban Bhunia*

Cite This: J. Phys. Chem. B 2020, 124, 1125−1136 Read Online

ABSTRACT: Insulin, a simple polypeptide hormone with huge biological importance, has long been known to self-assemble in vitro and form amyloid-like fibrillar aggregates. Utilizing high-resolution NMR, Raman spectroscopy, and computational analysis, we demonstrate that the fluctuation of the carboxyl terminal (C-ter) residues of the insulin B-chain plays a key role in the growth phase of insulin aggregation. By comparing the insulin sourced from bovine, human, and the modified glargine (GI), we observed reduced aggregation propensity in the GI variant, resulting from two additional Arg residues at its C-ter. NMR analysis showed atomic contacts and residue-specific interactions, particularly the salt bridge and H-bond formed among the C-ter residues Arg31^{B}, Lys29^{B}, and Glu4^{A}. These inter-residue interactions were reflected in strong nuclear Overhauser effects among Arg31$^{B}\delta$H−Glu4$^{A}\delta$H and Lys29$^{B}\delta$Hs−Glu4$^{A}\delta$H in GI, as well as the associated downfield chemical shift of several A-chain amino terminal (N-ter) residues. The two additional Arg residues of GI, Arg31^{B} and Arg32^{B}, enhanced the stability of the GI native structure by strengthening the Arg31^{B}, Lys29^{B}, and Glu4^{A} salt bridge, thus reducing extensive thermal distortion and fluctuation of the terminal residues. The high stability of the salt bridge retards tertiary collapse, a crucial biochemical event for oligomerization and subsequent fibril formation. Circular dichroism and Raman spectroscopic measurement also suggest slow structural distortion in the early phase of the aggregation of GI because of the restricted mobility of the C-ter residues as explained by NMR. In addition, the structural and dynamic parameters derived from molecular dynamics simulations of insulin variants highlight the role of residue-specific contacts in aggregation and amyloid-like fibril formation.



INTRODUCTION

Protein misfolding has become a topic of high-thrust research in biophysics and medicine because of the prevailing number of amyloid diseases.[1] Clinical conditions such as Alzheimer's, Parkinson's, and Huntington's disease, as well as lifestyle disorders such as type II diabetes, are associated with the aggregation of normally soluble proteins into insoluble amyloid fibrils.[2] Importantly, diabetes has become endemic in several regions, affecting about 300 million people globally.[3] Diabetes management strategies include administration of insulin secretagogues and/or subcutaneous injection of insulin. Hence, maintaining insulin stability and identifying reasons behind its misfolding into nonfunctional amyloid fibrils are vital because of its clinical importance. Consequently, the clinical relevance of insulin has garnered the interest of many groups to explore the structural and kinetic aspects of its fibrillation process.[4−8]

Human insulin (HI) is a key hormonal protein and plays an important role in the regulation of glucose homeostasis.[9] Insulin is, however, amyloidogenic in nature and produces fibrillar aggregates under favorable solution conditions such as low pH and high temperature.[10,11] Futhermore, other amylodogenic proteins, like hiapp and ab40 mutants, have shown to alter aggregation behavior.[12−16] The mechanism of insulin fibril formation is a complex process. Previous studies have demonstrated that insulin initially forms a multimeric complex, in dynamic equilibrium with monomers.[17] In the multimers, insulin essentially retains the structure of the monomer;[18] however, long exposure to an environment-favoring aggregation ultimately leads to misfolding of the protein and produces amyloid-like fibrillar aggregates.

While the fibril formation and its structural characterization are well reported, the knowledge of early events during insulin fibrillation is still obscure. Previous reports have shown that early insulin aggregates form a molten globule-like[19,20] state with the partial loss/modification of tertiary structure while

Received: November 4, 2019
Revised: January 15, 2020
Published: January 20, 2020

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

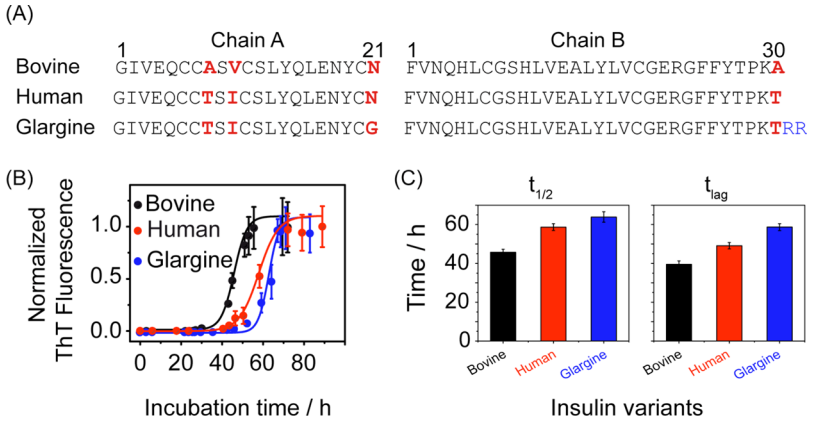


Figure 1. Comparison of kinetics of BI, HI, and GI insulin variants. (A) Sequence alignment showing the residue-specific differences (red) or additional residues (blue). (B) ThT kinetics of (mean ± SD of three independent experiments) 350 μM of BI (black), HI (red), and GI (blue) in 20% acetic acid (pH 1.9) incubated at 335 K. Aliquots were drawn from fibrillating samples at regular intervals, and the normalized fluorescence intensities were plotted against the incubation duration. (C) Bar plot of kinetic parameters of amyloid fibrillization.

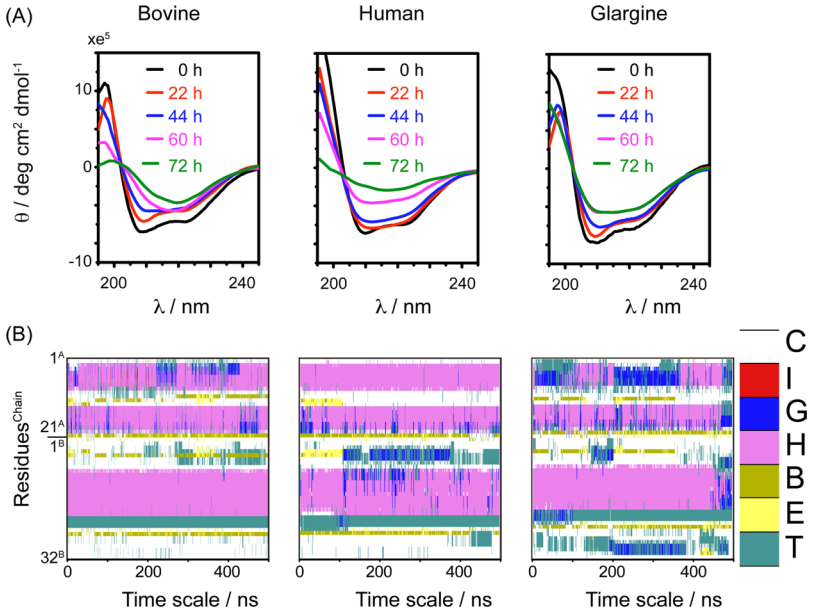


Figure 2. Changes in the secondary structure during fibrillation. (A) Aliquots were drawn at regular intervals during the ThT assay and diluted to 25 μM insulin in PBS (pH 7.4) to record the CD spectra. (B) Residue-specific changes in the secondary structure with respect to the MD simulation time. Color codes used for the assignment of secondary structure denotes T—turn, E—extended β, B—bridge β, H—α-helix, G—3_{10} helix, I—π helix, and C—coil.

retaining most of its secondary structure. This is considered as a key event in the early stage (lag phase) of fibril formation and accelerates hydrophobic collapse of favorable protein segments that lead to the formation of highly stable and well-ordered β-sheet-rich amyloid fibrils.[21,22] However, the formation of the molten globule-like state depends on the native structure of insulin and the severity of the solution condition.

The goal of the present work was to dissect the relationship between the native structures of three insulin variants and explore the factors taking part in their stability: bovine insulin (BI), HI, and modified glargine insulin (GI). We observed how structural instability influences the formation of molten globule-like state and, in turn, fibrillation which proceeded via nucleation through oligomeric intermediates.[23] We addressed this issue by comparing the structural features of BI, HI, and GI in order to identify how structural and sequence similarities determine their thermal stability. Using high-resolution NMR and Raman spectroscopy, we observed the atomic resolution dynamics of the C-ter of insulin B-chain. Importantly, GI has two additional Arg residues at the C-ter end of its B-chain (Figure 1A), which increases its pI to 6.7 (compared to HI which has a pI of 5.4),[24] making it less soluble at physiological pH. While it is known that Arg[25] may increase the thermal stability in proteins of thermophiles,[26−29] there is a lack of knowledge with respect to the effect of C-ter Arg residues on GI structure and the resultant effects that this may have on the prolonged activity in the human body. Moreover, we found that there is a key salt bridge between the residues Glu4^{A} and Lys29^{B} that restricts the flexibility and preserves the stability of insulin. Our results show that GI resists the formation of oligomers via a stabilization of Glu4^{A}−Lys29^{B}, while this is not the case for BI and HI.

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

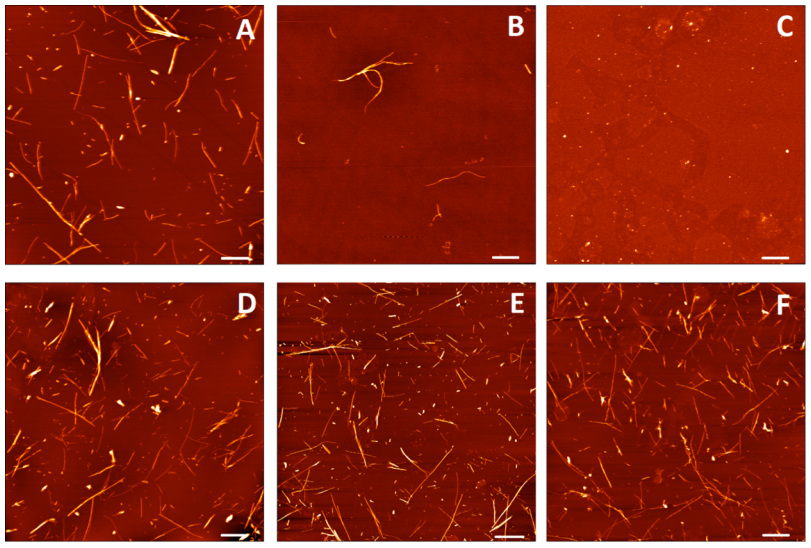


Figure 3. Morphology of amyloid aggregates of different insulin: AFM images of HI (A), BI (B), and GI (C) incubated for 42 h and HI (D), BI (E), and GI (F) incubated for 78 h. Scale bar is 1 μm for all images. The insulin concentration was 50 μM.

RESULTS

In order to realize the role of B-chain C-ter dynamics and salt bridge stability in the initial steps of insulin fibrillation in greater detail, we investigated the structural features and kinetics of fibril formation of three insulin variants: BI, HI, and GI. These three insulin variants have very high sequence similarity (Figure 1A); thus, to compare their fibrillation propensities, we used thioflavin T (ThT) fluorescence to monitor their aggregation kinetics. ThT is a well-known external fluorogenic probe to monitor amyloid fibrillation as it specifically binds to the β-sheet structure of amyloid fibrils, producing an intense fluorescence. The ThT fluorescence intensity is an indirect measure of amyloid fibril mass.[30−32] BI, HI, and GI were incubated at low pH (20% acetic acid, pH 1.9) and high temperature (335 K) at a concentration of 350 μM. Under these conditions, insulin was initially monomeric, yet readily formed amyloid fibrils.[33,34] For all three variants (Figure 1B), the ThT fluorescence response showed sigmoidal growth kinetics typical of nuclei-dependent fibrillization with three stages: (i) the quiescent lag phase, (ii) an elongation phase, and (iii) a final stationary phase. In the lag phase, the monomers of insulin undergo structural changes and form the oligomers, which are the stepping stones for the formation of the protofibrils. Knowles et al. have previously shown that the sigmoidal growth kinetics of insulin involve competing aggregation pathways, which includes orderly structural transition of oligomers and protofibrils, as well as fibril fragmentation at the different stages of the aggregation pathway.[35]

The obtained sigmoidal shapes of growth curves suggest a high probability of mechanistic similarity of insulin fibril formation in all of the examined insulin variants. This is likely due to the fibrillating core of all insulin variants, that is, the sequence and structure of the B-chain central helices are identical. Nevertheless, we observed significant differences in the fibrillation kinetics based on the ThT assay (Figure 1C). BI aggregated faster than the other two variants with a $t_{1/2}$ and t_{lag} of 45.7 ± 0.3 and 39.6 ± 1.4 h, respectively, and for HI, $t_{1/2}$ and t_{lag} were 58.6 ± 0.4 and 49.3 ± 1.2 h, respectively. Interestingly, GI is the slowest aggregating protein as $t_{1/2}$ and t_{lag} are equal to 63.8 ± 0.7 and 58.7 ± 1.6 h, respectively (Figure 1C). The t_{lag} data suggests comparatively higher stability (a relative delay in oligomer formation) of GI than HI and BI. The minor sequence variation probably causes the difference in oligomerization duration for BI, HI, and GI.

Far-UV circular dichroism (CD) spectra were obtained to detect the content of secondary structure of insulin variants. To identify the structural changes over the course of fibrillation, aliquots were drawn at regular intervals from the insulin samples for measurement. As shown in Figure 2A, the helicity of all three insulin isoforms decreases with incubation time. The time-course CD spectroscopy shows similar CD profiles between insulin variants during the lag phase. Although oligomers are known to form during the lag phase,[36,37] the secondary structure of insulin monomers within the oligomers underwent a minimal change, as evidenced by the marginal changes in the minima at 208 and 222 nm with time. These data depict no major secondary structural transition prior entering the growth phase of fibrillation. In contrast, a dramatic difference was observed for the fibril conformation in the CD spectra (Figure 2A), recorded after 72 h of incubation. Retention of helicity in GI could be attributed to the comparatively slower fibrillation observed in the ThT assay than the other two variants. We refrained from deconvoluting the spectra to find structural composition as the deconvolution of the CD spectra using available algorithms often introduce errors, particularly for aggregated samples where scattering varies with the shapes and sizes of aggregates.[38]

The secondary structural changes occurring during fibrillation were further studied in silico. Simulations were run at high temperature and low pH to recapitulate the experimental conditions. The trajectories of 500 ns simulations were analyzed to observe the residue-specific contribution to transitions in secondary structure. The changes in secondary structure with respect to simulation time are shown in Figure 2B. From the simulation, it was found that most changes were in the A-chain between the α-helix (pink) and the 3_{10} helix (blue) for BI, while for HI, the change was observed in the B-chain. However, this transition was found in both A and B chains in the trajectory of GI. Next, to identify the relationship between oligomerization/fibrillation and transitions in the

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

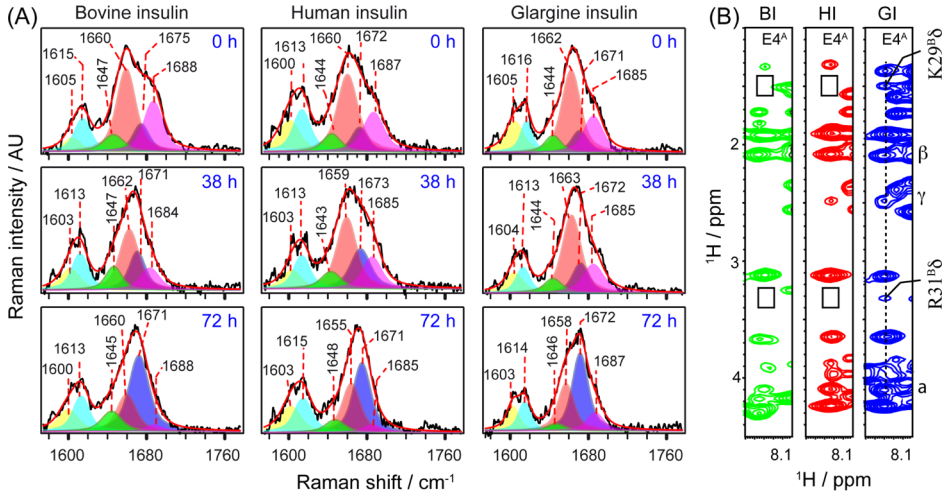


Figure 4. Secondary structure analysis from the amide I region of the Raman spectra obtained from incubated samples at different time points of incubation and NOEs showing interaction of B-chain C-ter with A-chain N-ter: (A) expanded amide I region (1590−1700 cm^{-1}) of the Raman spectra was fitted with six different curves/bands. Details of fits are given in Material and Methods section. The component bands were obtained after the fitting analysis, and the color codes are as follows: α-helix: orange, ∼1660 cm^{-1}; organized β-sheet: blue, ∼1673 cm^{-1}; loose β-strand, PPII, ∼1685 cm^{-1}: pink, undefined structure: green, ∼1640 cm^{-1}, and two for the aromatic residues (yellow, ∼1603 cm^{-1} and cyan peaks, ∼1615 cm^{-1}). The area under each curve represents the fraction of the secondary structure assigned to the bands. Among the six bands, four represent different secondary structure components. The incubation duration and variants are denoted in Figure 1. (B) Strip plot showing the interaction between $Arg31^{B}$ and $Glu4^{A}$ or $Lys29^{B}$ and $Glu4^{A}$ from the 2D $^{1}H-^{1}H$ NOESY experiments. The experiment was performed with a Bruker Avance III 600 MHz and at 298 K.

secondary structure and, as seen in the ThT assay and time-course CD spectra, respectively, we utilized atomic force microscopy (AFM) to observe changes in the morphology of aggregates.

MORPHOLOGY OF AGGREGATES COMPLEMENTS THE KINETIC STUDY RESULTS

The mutations in the protein sequence demonstrated the formation of different aggregation end products.[15,39−42] AFM was used for imaging all three insulin intermediates at different incubation times (Figure 3) to observe the higher-order fibril morphology. The prominent differences in the biophysical properties of BI, HI, and GI aggregates can lead to the microscopic changes in structure.[43−46] The AFM results confirmed the formation of fibrils in all the three cases and hence reinforced the "on-pathway" fibrillation.[47] As expected, the morphology and physical dimensions of aggregates are not identical for all the three insulin variants. After 42 h of incubation, HI formed fibrillar aggregates (Figure 3A), while BI formed a mixture of prefibrillar and fibrillar aggregates (Figure 3B) and GI formed small globular aggregates (Figure 3C), justifying its longer lag time. After 78 h of incubation, all insulin variants showed amyloid fibrils with different morphologies. We further characterized the fibrils with Gwyddion and FiberApp (Figure S1 and Table S1). The analysis revealed that HI fibrils are 9.6 ± 1.6 nm high with 45.5 ± 6.9 nm wide. The large width derives from the bigger clusters of the fibrils that laterally associate with one another. Likewise, BI fibrils also displayed a lateral association with individual fibrils of 11.0 ± 0.3 height and 40.4 ± 5.6 nm width. However, the BI fibrillar aggregates were longer (656.9 ± 53.0 nm) than the HI aggregates (528.7 ± 32.8 nm). However, GI formed extensive twisted fibrils (9.2 ± 0.4 nm high and 828.7 ± 45.4 nm long) after 78 h. GI also exhibited mixed fibril population with two distinct twist distances of 136.5 ± 6.3 and 163.4 ± 5.3 nm (Figure S1 and Table S1). The morphology of the aggregates agrees with the ThT fluorescence and CD spectroscopic findings;[48] however, it does not explain the structural basis of the difference in fibrillation propensities of BI, HI, and GI. To further understand the structural changes during fibrillation, we turned to Raman spectroscopy.

RAMAN SPECTROSCOPY AGREES WELL WITH CD SPECTROSCOPY

Raman spectroscopy is a unique nondestructive technique for the structural analysis of proteins and is a powerful tool for studying protein aggregation and related phenomena. The flexibility of Raman spectroscopy allows for the measurement of conformational changes in proteins whether in solution or in solid/semisolid states, that is, soluble monomers/oligomers or fibrils, respectively. This method measures molecular vibrations and, consequently, directly reports on the structure of the polypeptide backbone based on the amide mode, for example, amide I and III bands. The amide I Raman band of a protein provides explicit secondary structural information as it is largely associated with the C═O stretching mode and a lesser extent of C−N stretching and Cα−C−N deformation nodes.[48−54] Consequently, the band position is highly dependent on the nature/type of protein secondary structure. Further, decomposition of the amide I Raman band allows one to determine the overall composition of protein secondary structure.

The amide I region of Raman spectra recorded for different samples were fitted using a previously established protocol (please see details in Material and Methods section) by Maiti et al. and Dong et al.[50,52,55] Figure 4A shows the amide I region of the Raman spectra (1590−1700 cm^{-1}) for BI, HI, and GI of freshly prepared and incubated samples. For the 0 h sample, in all the three cases, a broad amide I band at ∼1659 cm^{-1} was observed with a shoulder at ∼1685 cm^{-1} corresponding to α-helix and PPII+ β-strand (extended) conformations, respectively.[38,52,55,56] The bandwidth at half

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

Table 1. Raman Spectroscopic Analysis of Aggregation Products of Insulin Variants

			α-helix			organized β-sheet			loose β-strand and PPII			undefined		
	amide I band width (cm^{-1})	time (h)	peak (cm^{-1})	width (cm^{-1})	[a] area %	peak (cm^{-1})	width (cm^{-1})	[a] area %	peak (cm^{-1})	width (cm^{-1})	[a] area %	peak (cm^{-1})	width (cm^{-1})	[a] area %
BI	46.7	0	1660	23.0	49	1675	19.7	13	1688	22.3	27	1647	25.0	11
	39.6	38	1662	23.3	44	1671	19.6	25	1684	22.5	16	1647	21.2	15
	34.6	72	1660	23.4	26	1671	23.0	55	1688	19.9	4	1645	25.0	15
HI	49.0	0	1660	24.5	51	1672	20.1	12	1687	23.0	26	1644	22.1	11
	45.3	38	1659	23.8	44	1673	24.5	27	1685	22.7	18	1643	24.4	11
	41.2	72	1659	20.7	35	1670	20.3	50	1681	19.8	5	1643	22.1	10
GI	45.6	0	1662	22.4	54	1671	19.7	13	1685	22.4	23	1644	19.7	10
	41.1	38	1663	22.7	51	1672	19.7	18	1685	24.3	21	1644	21.2	10
	39.6	72	1658	19.6	32	1672	19.1	52	1687	20.5	11	1647	21.3	5

[a]For area percentage calculation, only the given peaks were considered, i.e. ranging from 1633 to 1690 cm^{-1}.

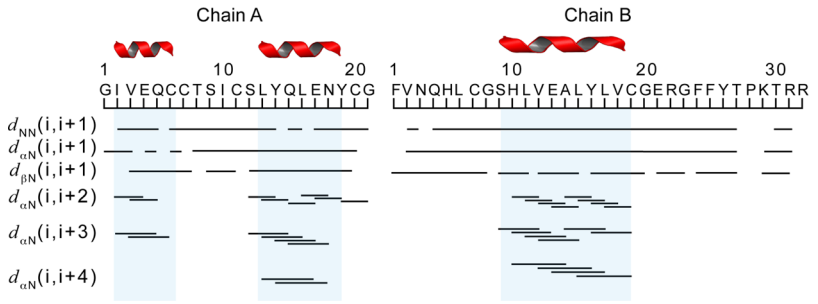


Figure 5. Summary of NOEs of GI: bar diagram summarizing NOEs observed for 350 μM GI in 20% acetic acid-d_4 (pH 1.9). The experiment was performed with a Bruker Avance III 600 MHz and at 298 K.

maxima of amide I band was ∼49 cm^{-1}, suggesting the presence of multiple conformations.[38] Figure 4A also displays fitted bands (details are given in the Material and Methods section) associated with distinct protein conformations and the integrated areas under each fitted curve quantify relative abundance of the particular type of secondary structure. The component band at ∼1660 cm^{-1} indicates α-helical conformation and the contribution was ∼50% for all the three insulin variants at 0 h, whereas cross-β-sheet component at ∼1673 cm^{-1} accounted for ∼13%. The combined intensity of bands appeared at ∼1685 and ∼1640 cm^{-1} represents polyproline II (PPII) conformations, β-strands, and other disordered structures. The amide I band in the experimentally obtained Raman spectra of insulin samples in the fibrillar state was relatively narrow compared to the spectra obtained for monomeric samples (0 h incubation) and appeared at ∼1673 cm^{-1} (Figure 4A). This indicated the transformation of the peptide bonds (some if not all) into a compact cross-β-sheet structure. The band fitting analysis revealed that ∼52% peptide bonds of the protein attained this conformation in the fibrillar state (Table 1).

Time-dependent Raman spectroscopic analysis (Table 1) further indicated no major changes in the secondary structural fold prior to the start of the fibril formation, involving compact cross-β-sheet structure. GI reached an equilibrium state of fibril formation over a longer time frame, and the content of β-sheet structure at 38 h of incubation was lower than that observed for BI and HI, which agrees with observations from the ThT assay (Figure 1B). Collectively, time-dependent Raman analysis, CD experiments, and the ThT fluorescence assay demonstrated that the lag phase time was longer for GI. These data suggest some alteration/realignment in the tertiary structure. This comparatively slow transition from α-helix to β-sheet in the case of GI could be due to greater structural stabilization, resulting from the two additional residues ($Arg31^B$ and $Arg32^B$) at the C-ter of GI. The Arg residues may protect the monomer from extensive thermal distortion and retard the aggregation and further fibril formation. Further, the unstructured B-chain C-ter may play a crucial role in the oligomerization/nucleation process. To find robust structural evidence in support of this rationale, we carried out high-resolution two-dimensional (2D) NMR spectroscopy for verification at the atomic level.

NMR STUDIES DEMONSTRATE THE INTERACTION OF B-CHAIN C-TER WITH A-CHAIN N-TER

NMR Spectroscopy. To probe what makes GI more structurally stable than BI and HI, 2D $^1H-^1H$ nuclear Overhauser effect spectroscopy (NOESY) was used for all the three isoforms. 2D $^1H-^1H$ NOESY spectra of all the three Zn-free insulin in 20% d_4-acetic acid were well resolved with a sufficient number of nuclear Overhauser effect (NOE) cross-peaks (Figure S2). We were able to assign all residues for BI and HI, while for GI, ∼52 residues were assigned out of 53 residues by sequential inter-residue NOE cross-peaks; Table S2A−C shows the assigned chemical shifts of BI, HI, and GI, respectively. The signature region of the spectrum is displayed in Figure S2.

For BI, the chemical shifts of residues $Cys11^A$, $Leu13^A$, $His5^B$, $Leu6^B$, $Gly8^B$, $Ser9^B$, and $Leu11^B$ shifted upfield from those of HI, while $Ser12^A$ exhibited a downfield shift. Mutations at $Thr8^A \rightarrow Ala8^A$ and $Ile10^A \rightarrow Val10^A$ from HI to BI likely give rise to the observed chemical differences between these two insulin variants.[57] Though GI has higher sequence identity with the HI, the C-ter modification of both A-chain and B-chain has introduced differences.

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

Upon comparison of the NOESY spectra of GI with HI, we observed the downfield shift of Ile2^{A}, Val3^{A}, Glu4^{A}, Tyr19^{A}, Ala14^{B}, Glu21^{B}, Gly23^{B}, Tyr26^{B}, and Thr27^{B} (Table S2). A closer inspection of the NOESY spectrum of GI suggests a strong NOE between Arg31$^{B}\delta$Hs−Glu4^{A}H, which would explain the downfield chemical shifts of the A-chain N-ter residues. Critically, we observed NOEs between Lys29$^{B}\beta$Hs−Glu4^{A}H and Arg31$^{B}\gamma$Hs−Glu4^{A}H NOEs in GI; however, these interactions were absent in the other two isoforms (Figure 4B). Because the crystal structures for HI and BI are available, we determined the structure of GI with the above NOEs (Figures 5 and S2) using simulated annealing. Amide chemical shifts are very sensitive to the chemical environment, and mutation at a certain position or addition of certain residues can result in chemical shift perturbation of the surrounding residues.[58] To examine this, we considered the amide proton chemical shifts of HI, GI, and BI. Results show that the natural mutations in A-chain, that is, the Ala to Thr and Val to Ile at positions 8 and 10 of A-chain (Figure S3A,B), respectively, had a minimal effect on the adjacent residues. This suggests that there are minimal structural differences due to these point mutations. However, residues 25−29 of GI showed chemical shift perturbations of 0.05−0.14 ppm (Figure S3A,B). Comparative analysis of the NOESY spectra of all three insulin variants points to a unique interaction present in GI not present in BI or HI, namely, between residues Lys29^{B}−Glu4^{A} and Arg31^{B}−Glu4^{A}. Furthermore, those NOEs between Glu4^{A} and both Lys29^{B} and Arg31^{B} point to a transient structure that provides added structural stability of GI over BI or HI.

Structure of GI by Simulated Annealing. The assigned NOEs of GI (Figure 5) were used to compute the complete structure using a simulated annealing method. Of note, the available crystal structures of GI (accession code: 5VIZ) lack the C-ter residues "Thr30ArgArg32" in the B-chain. Our NMR data provide the complete structure and explain the additional reluctance of GI to undergo fibrillation when compared to BI and HI. Moreover, our insights can also help to develop better aggregation inhibitors by restricting the dynamics of B-chain C-ter of GI. The structure calculation statistics are given in Table S3. The ensemble structure showed a root-mean-square deviation (rmsd) of 0.58 Å when aligned on the B-chain N-ter residues from 4 to 18. The obtained structure shows the important contact between the Arg31^{B} and Lys29^{B} with the Glu4^{A}, shown by the dotted red line (Figure S3D) with the corresponding NOEs from the NOESY spectra shown in Figure S3E. The obtained PDB was analyzed with "protein contact map" tools[59] on nanoHUB.org, which assists to observe the overall distances and contacts within the protein. The contact map (Figure S3F) reflects the similarities shown in Figure 6D. From the contact map, it can be observed that there are ample number of contacts in the corner region (shaded red) and the corresponding distance map (Figure 6D), shown in blue in the corner region; both suggest the proximity of the C-ter of B-chain and the N-ter of A-chain in GI.

The well-defined secondary structural motifs (Figures 5 and 6) are comparable to the previously reported insulin structures; however, the three polar residues Thr30^{B}, Arg31^{B}, and Arg32^{B} of GI have affected the random coil segment at C-ter of B-chain. The medium range αN (i, i + 3/4) NOEs from Ile2^{A} to Thr8^{A} and Leu13^{A} to Tyr19^{A} imply a helical structure, and the two helices were joined by a loop between Ser9^{A} to Ser12^{A}. In contrast, only a weak αN (i, i + 3) NOE Gln15^{A}−Asn18^{A} was

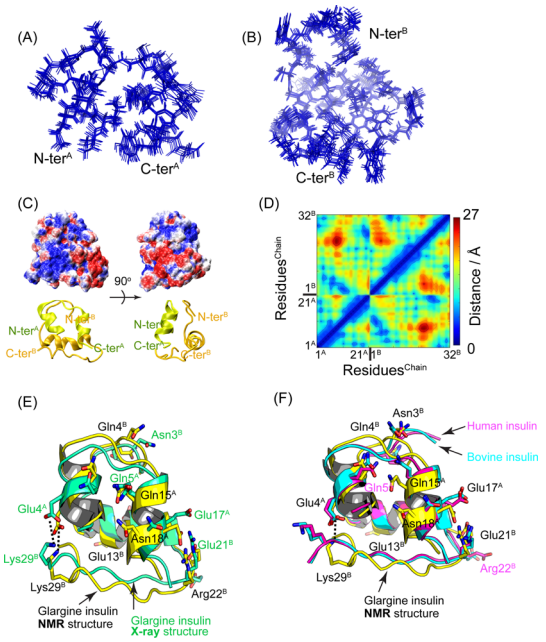


Figure 6. 3D solution structure of GI (PDB: 6K59). Superposition of backbone atoms (N, Cα, and C′) of the 10 lowest energy structure of GI for (A) A-chain and (B) B-chain. (C) APBS-calculated electrostatic surface potential of GI, where red and blue represent the electronegative and electropositive potential, respectively. (D) Amino acid contact map of GI. Comparison of NMR GI structure with (E) X-ray structure of GI (PDB: 5VIZ), (F) HI (PDB: 3W7Y), and BI (PDB: 2ZP6) structures. The black dotted line indicates the hydrogen bond or salt bridge between Lys29^{B} and Glu4^{A}.

observed in the C-ter of A-chain, suggesting an extended conformation in this segment. In B-chain, Ser9^{B} to Cys19^{B} displays a dominant helical content, followed by a turn from Gly20^{B} to Gly23^{B}. Random coil structure is found in the N-ter (Phe1^{B}−Gly8^{B}) as well as in the C-ter (Phe24^{B}−Arg32^{B}) segment. Overall, the lowest energy ensemble structures show conserved backbone and side-chain atoms with a low rmsd value of 0.8 ± 0.3 Å for backbone atoms (N, Cα, and C′). The structural stability of GI was also found to be mediated either by salt bridges or by hydrogen bonds between Glu4^{A}−K29^{B}, Glu4^{A}−R31^{B}, and Glu17^{A}−Arg22^{B} (Figure 4B). In addition, the salt bridge or hydrogen bond between Glu13^{B}/Arg32^{B} also contributes to the structural stability.

When compared with the X-ray structure of GI (accession code: 5VIZ), the backbone and side-chain rmsd values between the two structures were found to be 1.84 ± 0.13 and 1.91 ± 0.14 Å, respectively. Comparing the backbone rmsd values for individual chains, the values are 1.55 Å for A-chain and 2.55 Å for B-chain, suggesting that the inclusion of "T30RR32^{B}" segment at the C-ter decreases the dynamicity in the NMR structure at the C-ter of B-chain. More specifically, positively charged T30RR32^{B} residues restricted the flexibility either by salt bridge or hydrogen bond interaction between Glu4^{A} and Arg31^{B}/Lys29^{B}. However, we cannot rule out that the additional positive charges may also prevent intermolecular interactions that would lead to fibril formation.

The electrostatic surface potential of GI shows that the electropositive residues Gln5^{A}, Gln15^{A}, and Asn18^{A} are solvent-exposed. Additionally, the electropositive characteristic of this surface is also supported from the B-chain with exposure of residues Lys29^{B} and Arg31^{B}. On the other hand, the surface

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

rotated by 90° shows electronegative characteristics with polar residues Ser12^{A}, Tyr14^{A}, Glu17^{A}, and Glu21^{B}. Structural statistics of the refined models are shown in Table S3. The spatial distribution of amino acids in GI is shown as a residue contact map (Å) in Figure 6D.

DISCUSSION

Although insulin aggregation has been extensively studied, there are still gaps in our understanding of this process. Environmental factors such as variations in pH or salt concentrations alter the aggregation kinetics, resulting in polymorphic structures of insulin fibrils. For example, in a Fourier transform infrared spectroscopy study, researchers demonstrated the insulin fibril polymorphism formed at variable pH (pH 1.6 and 2).[60] Further, salt concentration plays an important role in the formation of supramolecular assembly of insulin fibrils which are termed spherulites.[11]

Here, using zinc-free forms of BI, HI, and GI with identical buffer conditions provided a fair comparison of their fibrillation, ensuring that the differences in fibrillation kinetics were due to the internal factors of protein. Gibson and Murphy reported that the equimolar concentration of a peptide with sequence RRRRRRLVEALYL (where LVEALYL is part of the B-chain central helix) can attenuate the insulin fibrillation.[61] Additionally, it was demonstrated that point mutations (H10D and L17Q) in B-chain can slow down the insulin fibrillation.[62] These reports demonstrate the importance of the B-chain central helix in the context of insulin fibrillation. As shown in our results, GI fibrillates slower than BI and HI, likely owing to site-specific differences in their amino acid sequences. Although BI, HI, and GI have identical sequences in their B-chain central helix, the marked difference in the fibrillation kinetics of these three proteins can be attributed to their flexible regions. Interestingly, truncation at the B-chain C-ter can accelerate the insulin fibrillation by exposing the fibrillation core or the hydrophobic pocket of globular insulin.[63,64] Hence, we can deduce that the B-chain C-ter plays a vital role in progression of the insulin fibrillation.

The structural study of GI with NMR spectroscopy herein sheds light on the forces stabilizing the B-chain C-ter with the A-chain N-ter. Our findings are supported by crystallographic evidence, which suggested that the presence of Gly21^{A}, Arg31^{B}, and Arg32^{B} helped in the formation of additional polar contacts with nearby residues in the crystal packing, an important comparison made with respect to HI. A key distinction, however, is that the crystal structure was studied in the presence of phenol and divalent zinc ions to facilitate crystallization.[65,66] Hence, to fully understand the mechanistic details that govern insulin dynamics, an atomic level NMR study provides a unique advantage over other techniques.

Residues Arg31^{B} and Arg32^{B} in GI increase its pI and render the solubility of the protein at physiological pH, increasing the duration of its activity in blood. The overall positive charge at the B-chain C-ter increases the probability of salt bridge formation with the A-chain N-ter residues, namely, Glu4^{A} and Ile2^{A}. The proximity of the B-chain C-ter and A-chain N-ter also increases the possibility of interchain hydrogen bond formation, as shown by the solution structure of GI in this study. Together, the salt bridge and hydrogen bond increase the stability of the monomeric protein. Additionally, the computational analysis of the insulin variants also supports the crucial differences in the pI of side-chain residues (Table S4). The residues His10^{B} and Glu13^{B} are part of the B-chain helix and are known to participate in fibril formation. The nearby Glu21^{B} in GI also shows a comparatively lower pK_a (shown in Table S4), which may also play a vital role in hindering the formation of initial oligomeric species by electrostatic repulsion as shown in amyloid-β[67] and TDP-43.[68]

The present work thus demonstrates that the increased flexibility in the C-ter of insulin generates conformational states conducive to structural rearrangement and hydrophobic association, resulting in the formation of molten-globule structures prone to aggregate into high-order fibrillar assemblies.[17] Of particular importance is the formation of a salt bridge in GI that makes its tertiary unfolding more difficult, staving off nucleation events occurring more readily in the wild-type proteins.

CONCLUSIONS

Based on our results, we hypothesized that despite the identical sequence of the fibrillating core, the difference in the fibrillation kinetics arises from the difference in structural features of the insulin variants. Our results point to higher stability of GI because of the presence of the T30RR32^{B} residues, which increase the strong electrostatic interaction at the B-chain C-ter. AFM images show the formation of protofibrillar and fibrillar species of HI and BI; however, only globular oligomers of GI are observed at the same time point. Additionally, the computational analysis of the insulin variants also supports the crucial differences in the pK_a of side-chain residues. Finally, NMR-derived structures demonstrate that the structural stability of insulin is regulated by C-ter residues Arg31^{B} and Lys29^{B}, forming a salt bridge and/or a H-bond with Glu4 of the A-chain. Collectively, our results provide structural and dynamic evidence for the initial misfolding events that trigger and stave off insulin aggregation.

MATERIALS AND METHODS

Bovine pancreatic insulin and HI were procured from Sigma (St. Louis, Missouri, USA), while GI was purchased from marketed drug Basalog of Biocon (Bangalore, India). Tetradeutero acetic acid and D_2O were purchased from Cambridge Isotope Laboratories. Reagents used in the experiments were of high grade and purchased from either Sigma (St. Louis, Missouri, USA), SRL (Mumbai, India), or Merck (Darmstadt, Germany).

ThT Fluorescence Assay. The insulin samples were made zinc-free by extensive dialysis in a volatile buffer, followed by lyophilization.[69] For GI, after extensive dialysis for removing the excipients from the pharmaceutical formulation, the sample was passed through the Superdex peptide 10/300 GL (GE Healthcare Life Sciences, Chicago, Illinois, USA), followed by lyophilization to ensure high purity. The zinc-free samples, viz., BI, HI, and GI, were solubilized in 20% acetic acid at a final concentration of 350 μM and pH adjusted to 1.9 with HCl. The UV absorbance (276 nm) was used to determine the concentration of insulin. ThT was dissolved in Milli-Q water and centrifuged at high speed to get rid of the undissolved solvent. The concentration of ThT solution was determined by using a molar extinction coefficient[70] of 36,000 M^{-1} cm^{-1}. Insulin samples (350 μM) were incubated at 335 K in a water bath and aliquots were drawn at regular time intervals. The aliquots were diluted in 10 mM phosphate buffer (pH 7.4) with 100 mM NaCl.[70,71] The ThT fluorescence was measured in a quartz cuvette by setting the spectrofluorometer (PTI) at

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

an excitation wavelength of 450 nm. The emission was observed at 482 nm. All results shown are mean ± standard deviation (SD) of three independent experiments.

The observed fluorescence intensity at 482 nm for different time points was plotted against the time points, and the curve was fitted by Boltzmann equation[72,73] (eq 1)

$$Y = A_2 + \left[\frac{A_2 - A_1}{1 + e^{t_0 - t/\tau}}\right] \quad (1)$$

where A_1 is the initial fluorescence, A_2 is the maximum fluorescence, t_0 stands for the time where the fluorescence has reached to half of the maximum value, and $1/\tau$ is the apparent rate constant of fibril growth and lag time approximated to $t_0 - 2\tau$.

CD Spectroscopy. During insulin amyloid fibrillation, the changes in the secondary structure of insulin samples (BI, HI, and GI) were observed using CD spectroscopy. The incubation condition was kept identical to the ThT assay, prior to recording the spectra; the drawn aliquots were diluted to 25 μM with PBS pH 7.4. The far-UV CD spectra were recorded on a quartz cuvette (0.2 cm) over a range of 200−260 nm with a 1 nm data interval and a scanning speed of 100 nm/min at 298 K using a Jasco J-815 spectrophotometer. Each CD spectrum represents an average accumulation of four scans. The recorded raw spectra data in millidegrees were subtracted from the blank buffer and transformed to molar ellipticity using the following equation (eq 2)

$$\text{Molar ellipticity } (\theta) = m_0 M/10 \times L \times C \quad (2)$$

where m_0 represents millidegrees, M stands for molecular weight (g mol^{-1}), L is the path length of quartz cuvette used (cm), and C is the molar concentration of protein.

Atomic Force Microscopy. For AFM, the insulin samples were diluted to a concentration of 50 μM with ultrapure water. Aliquots of 10 μL were cast on freshly cleaved mica and left for 10 min to adsorb. The samples were rinsed with ultrapure water and dried under N_2 gas. The images were taken using a Veeco di Innova atomic force microscope in a tapping mode under ambient conditions using a triangular silicon nitride cantilever SNL-10 (Bruker AFM Probes) with a tip radius of 2 nm, a typical resonance frequency of 65 kHz, and a spring constant of 0.35 N/m. All images are unfiltered. Analysis of AFM images was performed using Gwyddion software (for height and width) and FiberApp software (for length of fibrils).

NMR Experiments. The NMR experiments were performed at 298 K on an AVANCE 600 MHz spectrophotometer equipped with a QCI Cryoprobe. For data acquisition and processing, TopSpin v3.1 was utilized. For all the experiments, Zn-free insulin was used.[69]

The sample was dissolved in 20% acetic acid-d_4, 70% water, and 10% D_2O, pH 1.9, at a final concentration of 350 μM. Spectra were recorded immediately in a Shigemi tube. 2D homonuclear total correlation spectroscopy (TOCSY) and NOESY were performed at 298 K with mixing times of 40 and 200 ms, respectively. The NOESY spectrum was calibrated with respect to the residual methyl group peak (2.03 ppm) of d_4-acetic acid. The NOESY spectra of insulin were recorded with 64 scans and 512 increments in t_1 and 2048 data points in t_2. The spectral width was 12 for both dimensions. For quadrature detection in t_1 dimension, states TPPI was used and water suppression was done using excitation sculpting. The insulin TOCSY and NOESY spectra were analyzed with Sparky software (https://www.cgl.ucsf.edu/home/sparky).

Raman Spectroscopy. The insulin samples (BI, HI, and GI) were incubated in 20% acetic acid (pH 1.9) at 335 K. The aliquots were drawn at 0, 38 (before entering into elongation phase), and 72 h (mature fibril). The collected aliquots were drop-cast on a fresh clean aluminum foil mounted on the glass slide.[56] Raman spectra were obtained in backscattering geometry using a LabRAM HR—Jobin Yvon (Horiba, Kyoto, Japan) spectrometer equipped with a Peltier-cooled CCD. A diode laser of wavelength 785 nm was used as an excitation source, and the light was focused on the sample using a 50× objective.[56] All data were recorded within a wavenumber range from 800 to 1800 cm^{-1}.

For curve fitting, Origin program was used. First, the Raman spectra were processed by multipoint baseline correction. To derive the content of secondary structure, the amide I region (1590−1700 cm^{-1}) of the spectra was selected for curve fitting analysis. For the band fitting, we first selected, by peak picking option, six peaks in the broad amide I region obtained from experimental data, and to define the shape of the fitted curves, a mixed Gaussian and Lorentzian function embedded in the Origin program was used. Four of the peaks are assigned to different protein conformation/secondary structures and two of them were assigned to aromatic residues. The peaks are assigned values, which are most probable for the particular conformation/residue based on several references.[50] For α-helical conformation, it was selected at 1658 cm^{-1}; for β-sheet component, it was assigned at 1673 cm^{-1}; for PPII and disordered structure, it was assigned at 1685 cm^{-1}; and for other disordered and extended structures, it was assigned at 1644 cm^{-1}. For aromatic residues, the peaks were assigned at ∼1605 and ∼1615 cm^{-1}. In the fitting protocol, we also provide a window of bandwidth range between 15 and 35 cm^{-1} and allowed a ±5 cm^{-1} window for the band maxima. After incorporating these parameters, the band fitting analysis was performed, and the best fitted results are taken. The total area under the four conformation-linked fitted curves represents the total secondary structure, and the area under each curve represented the corresponding fraction (%) of the protein secondary structure.

Solution Structure of GI. Three-dimensional (3D) solution structure of GI was calculated based on the experimentally measured NOEs and ff14SB force fields in AMBER.[74] The protonation states of titratable residues were determined using the H++ server.[75] The interproton distance from the NOE along with the torsional information was used as the restraint for structural model building. The torsional angle restraints were calculated from the PREDITOR[76] web server with the help of a ^{1}H chemical shift deviation, and ±20 to 25° from the calculated torsional angles were used for structure calculation. The resolved ring protons of F24^B (viz., F24-2H and F24-3H) or Y26^B (viz. Y26-2H and Y26-3H) were used to calculate the upper bound distance constraints from the NOE intensities. The 2.0 Å distance was fixed for lower distance constraints. A strong (2.5 Å), medium (2.6−3.5 Å), and weak (3.6−5.0 Å) distances were categorized from NOE intensities for further structure calculation using AMBER. Hydrogen bonds were included in the structure calculation. A total of 319 short range, 35 medium-range ($2 < |i - j| < 5$), and 25 long-range ($|i - j| \geq 5$) NOEs (382 total) and 101 torsion angles were used in this procedure. A starting coordinate of GI with an additional "Thr30ArgArg32" motif at the C-ter of B-

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

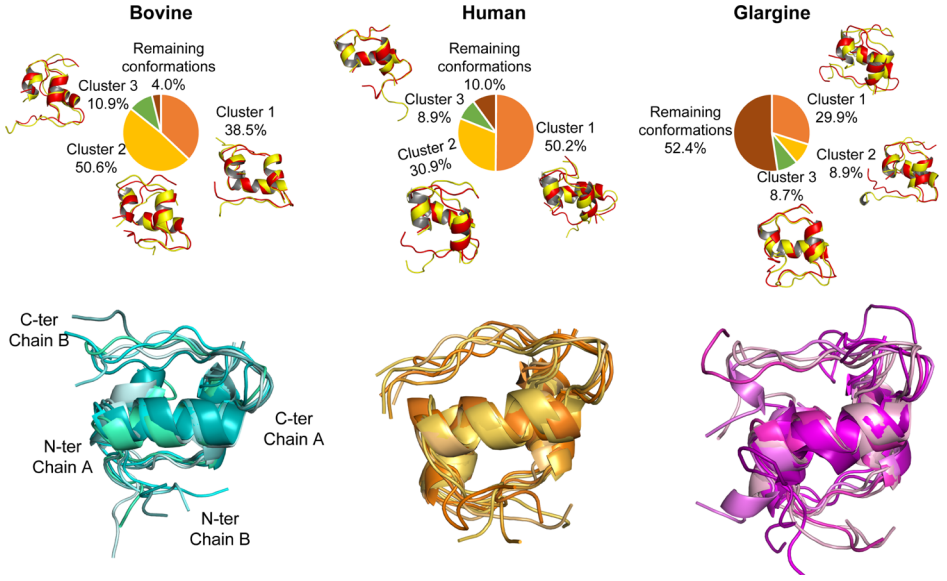


Figure 7. Cluster representation of MD simulation of insulin variants: dominant conformations from MD simulations of each variant are shown based on cluster analysis. The upper panel shows the representative structure from each cluster with minimum (yellow) and maximum (red) deviation from the cluster center. The lower panel shows the structural overlay from the MD simulation trajectory.

chain was prepared using homology modeling approach, taken from the previous experimental structure of GI from the Protein Data Bank (PDB) (accession code: 5VIZ) as a template. For the homology modeling, automated SWISS-MODEL protocol was used.[77] Using the restraint information, simulated annealing was performed for 20 ps with a generalized born model. The force constant values for the distance restraint varied from 0 to 20 kcal/mol Å^2 for the lower and upper bound values. System temperature was regulated using the Berendsen coupling algorithm, and the nonbonded interactions were accounted with a cutoff value of 15.0 Å. The lowest energy conformation of GI representing the experimental average structure was presented as a structural ensemble. Structural analysis was performed using the SuperPose v1.0 server.[78] The electrostatic surface potential for GI was calculated using the APBS program.[79]

Molecular Dynamics Simulation. A comparative structural analysis for GI with similar variants, viz., BI and HI, were studied using molecular dynamics (MD) simulation (Figure 7). The coordinates of BI and HI were obtained from PDB (accession code: 2ZP6 and 3W7Y, respectively). The 3D structural coordinates of insulin variants were processed in a protein preparation wizard of Schrödinger Suite 2017-4.[80] The objective here was to correct the bond order, addition of missing hydrogen atoms, and removing the crystallographic water molecules.[80] The structures were finally refined at acidic pH 1.9, thereby changing the protonation states of titratable residues. Solvation of the protein was modeled using the TIP3P water model in a octahedral box with a 9.5 Å edge distance using the H++ server.[75] The system was neutralized by addition of appropriate number of Cl^- ions. For calculation of pK_a values, internal and external dielectric constants were set to values 10 and 80 as recommended by Grimsley et al.[81]

MD simulation was performed using the ff14SB force field using the AMBER program.[74] The adapted force field is a modified version of ff99SB, containing adjustment of φ and ψ parameters, side-chain rotamers, and alternate protonation states of ionizable residues. The force field has been tested against the secondary structure content of short peptides and is recommended for proteins and peptides. The SHAKE algorithm[82] was applied to constrain the bonds involving hydrogen. Gradual heating of the system to 335 K was controlled using Langevin dynamics[83] with a restraint force constant (100 kcal/mol Å^2) on solute atoms. The particle mesh Ewald method[84] was used to demonstrate the non-bonded interaction with a cutoff value of 8 Å. System equilibration was performed with gradual release of restraint force constant, from backbone atoms and then from side-chain atoms, at an interval of 10 ns. Equilibration steps of all systems were performed for a total of 35 ns, in which a final 20 ns time step was performed without any positional restraints. Production trajectories were calculated for a timescale of 500 ns, where the snapshots were saved at an interval of 20 ps, for analysis. Cluster analysis was performed using *K*-means algorithm with a radius of 4 Å, using MMTSB tools.[85]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.9b10349.

AFM images of fibrils of BI, HI, and GI; 2D NMR spectra of insulin in 20% acetic acid; chemical shift deviation; fibril morphological parameters; chemical shift assignments of BI, HI, and GI; structural statistics of GI; and pK_a analysis of BI, HI, and GI at pH 1.9 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Anirban Bhunia − *Department of Biophysics, Bose Institute, Kolkata 700054, India;* orcid.org/0000-0002-8752-2842; Phone: +91-33-25693336; Email: anirbanbhunia@gmail.com, bhunia@jcbose.ac.in

Authors

Bhisma N. Ratha − *Department of Biophysics, Bose Institute, Kolkata 700054, India*

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

Rajiv K. Kar − *Department of Biophysics, Bose Institute, Kolkata 700054, India;* orcid.org/0000-0003-4629-5863
Zuzana Bednarikova − *Department of Biophysics, Institute of Experimental Physics Slovak Academy of Sciences, Kosice 040 01, Slovakia*
Zuzana Gazova − *Department of Biophysics, Institute of Experimental Physics Slovak Academy of Sciences, Kosice 040 01, Slovakia;* orcid.org/0000-0002-0670-3431
Samuel A. Kotler − *National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, Maryland 20850, United States*
Sreyan Raha − *Department of Physics, Bose Institute, Kolkata 700009, India*
Soumya De − *School of Bioscience, IIT Kharagpur, Kharagpur 721302, India;* orcid.org/0000-0002-4347-3137
Nakul C. Maiti − *Division Structural Biology and Bioinformatics, CSIR-Indian Institute of Chemical Biology, Kolkata 700032, India;* orcid.org/0000-0002-8498-6502

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.9b10349

Author Contributions

A.B. designed the research; B.N.R., R.K.K., Z.B., Z.G., S.R., S.D., N.C.M., and A.B. performed the experiments; B.N.R., R.K.K., Z.G., S.A.K., N.C.M., and A.B. analyzed the results; and B.N.R., R.K.K., S.A.K., and N.C.M. wrote the manuscript; all authors reviewed the manuscript; and A.B. arranged funding for this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was partly supported by the Council of Scientific and Industrial Research (02(0292)/17/EMR-II to A.B.), partly by the Department of Biotechnology (BT/PR29978/MED/30/2037/2018 to A.B.), Govt. of India, and partly by the Intramural External Fund of Bose Institute (R/16/19/1615 to A.B.). On the Slovak side the work was supported by the projects VEGA 2/0145/17 and APVV-18-0284. B.N.R. thanks CSIR-UGC, Govt. of India, for providing the fellowship. The structure of GI has been deposited to the PDB with the accession code of 6K59.

■ ABBREVIATIONS

BI, bovine insulin; HI, human insulin; GI, glargine insulin; ThT, thioflavin T; CD, circular dichroism; h, hour/s; t_{lag}, lag time; $t_{1/2}$, half-time; AFM, atomic force microscopy; NMR, nuclear magnetic resonance; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; MD simulation, molecular dynamics simulation; d_4-acetic acid, tetradeutero acetic acid

■ REFERENCES

(1) Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 384−396.
(2) Dobson, C. M. The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc., B* **2001**, *356*, 133−145.
(3) Zimmet, P. Z. Diabetes and its drivers: the largest epidemic in human history? *Clin. Diabetes Endocrinol.* **2017**, *3*, 1.
(4) Hua, Q.-x.; Weiss, M. A. Mechanism of insulin fibrillation: the structure of insulin under amyloidogenic conditions resembles a protein-folding intermediate. *J. Biol. Chem.* **2004**, *279*, 21449−21460.
(5) Groenning, M.; Frokjaer, S.; Vestergaard, B. Formation Mechanism of Insulin Fibrils and Structural Aspects of the Insulin Fibrillation Process. *Curr. Protein Pept. Sci.* **2009**, *10*, 509−528.
(6) Ivanova, M. I.; Sievers, S. A.; Sawaya, M. R.; Wall, J. S.; Eisenberg, D. Molecular basis for insulin fibril assembly. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18990−18995.
(7) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. Ø.; Riekel, C.; Eisenberg, D. Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature* **2007**, *447*, 453−457.
(8) Krebs, M. R. H.; Macphee, C. E.; Miller, A. F.; Dunlop, I. E.; Dobson, C. M.; Donald, A. M. The formation of spherulites by amyloid fibrils of bovine insulin. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14420−14424.
(9) Röder, P. V.; Wu, B.; Liu, Y.; Han, W. Pancreatic regulation of glucose homeostasis. *Exp. Mol. Med.* **2016**, *48*, No. e219.
(10) Bouchard, M.; Zurdo, J.; Nettleton, E. J.; Dobson, C. M.; Robinson, C. V. Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy. *Protein Sci.* **2000**, *9*, 1960−1967.
(11) Smith, M. I.; Foderà, V.; Sharp, J. S.; Roberts, C. J.; Donald, A. M. Factors affecting the formation of insulin amyloid spherulites. *Colloids Surf., B* **2012**, *89*, 216−222.
(12) Buchanan, L. E.; Dunkelberger, E. B.; Tran, H. Q.; Cheng, P.-N.; Chiu, C.-C.; Cao, P.; Raleigh, D. P.; de Pablo, J. J.; Nowick, J. S.; Zanni, M. T. Mechanism of IAPP amyloid fibril formation involves an intermediate with a transient β-sheet. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 19285−19290.
(13) Lemkau, L. R.; Comellas, G.; Kloepper, K. D.; Woods, W. S.; George, J. M.; Rienstra, C. M. Mutant Protein A30P α-Synuclein Adopts Wild-type Fibril Structure, Despite Slower Fibrillation Kinetics. *J. Biol. Chem.* **2012**, *287*, 11526−11532.
(14) Koo, H.-J.; Lee, H.-J.; Im, H. Sequence determinants regulating fibrillation of human α-synuclein. *Biochem. Biophys. Res. Commun.* **2008**, *368*, 772−778.
(15) Brender, J. R.; Hartman, K.; Reid, K. R.; Kennedy, R. T.; Ramamoorthy, A. A Single Mutation in the Nonamyloidogenic Region of Islet Amyloid Polypeptide Greatly Reduces Toxicity. *Biochemistry* **2008**, *47*, 12680−12688.
(16) Ke, P. C.; Sani, M.-A.; Ding, F.; Kakinen, A.; Javed, I.; Separovic, F.; Davis, T. P.; Mezzenga, R. Implications of peptide assemblies in amyloid diseases. *Chem. Soc. Rev.* **2017**, *46*, 6492−6531.
(17) Nettleton, E. J.; Tito, P.; Sunde, M.; Bouchard, M.; Dobson, C. M.; Robinson, C. V. Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. *Biophys. J.* **2000**, *79*, 1053−1065.
(18) Tito, P.; Nettleton, E. J.; Robinson, C. V. Dissecting the hydrogen exchange properties of insulin under amyloid fibril forming conditions: a site-specific investigation by mass spectrometry. *J. Mol. Biol.* **2000**, *303*, 267−278.
(19) Dzwolak, W.; Ravindra, R.; Lendermann, J.; Winter, R. Aggregation of bovine insulin probed by DSC/PPC calorimetry and FTIR spectroscopy. *Biochemistry* **2003**, *42*, 11347−11355.
(20) Yamamoto, S.; Watarai, H. Raman optical activity study on insulin amyloid- and prefibril intermediate. *Chirality* **2012**, *24*, 97−103.
(21) Arora, A.; Ha, C.; Park, C. B. Insulin amyloid fibrillation at above 100 °C: New insights into protein folding under extreme temperatures. *Protein Sci.* **2004**, *13*, 2429−2436.
(22) Ahmad, A.; Millett, I. S.; Doniach, S.; Uversky, V. N.; Fink, A. L. Partially Folded Intermediates in Insulin Fibrillation. *Biochemistry* **2003**, *42*, 11404−11416.
(23) Vestergaard, B.; Groenning, M.; Roessle, M.; Kastrup, J. S.; van de Weert, M.; Flink, J. M.; Frokjaer, S.; Gajhede, M.; Svergun, D. I. A

helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *PLoS Biol.* **2007**, *5*, No. e134.
(24) Buse, J. Insulin glargine (HOE901). *Diabetes Care* **2000**, *23*, 576.
(25) Strub, C.; Alies, C.; Lougarre, A.; Ladurantie, C.; Czaplicki, J.; Fournier, D. Mutation of exposed hydrophobic amino acids to arginine to increase protein stability. *BMC Biochem.* **2004**, *5*, 9.
(26) Taylor, T. J.; Vaisman, I. I. Discrimination of thermophilic and mesophilic proteins. *BMC Struct. Biol.* **2010**, *10*, S5.
(27) Haney, P. J.; Badger, J. H.; Buldak, G. L.; Reich, C. I.; Woese, C. R.; Olsen, G. J. Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic Methanococcus species. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3578−3583.
(28) Nick Pace, C.; Alston, R. W.; Shaw, K. L. Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci.* **2000**, *9*, 1395−1398.
(29) Kumar, S.; Nussinov, R. How do thermophilic proteins deal with heat? *Cell. Mol. Life Sci.* **2001**, *58*, 1216−1233.
(30) Biancalana, M.; Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta* **2010**, *1804*, 1405−1412.
(31) Groenning, M.; Norrman, M.; Flink, J. M.; van de Weert, M.; Bukrinsky, J. T.; Schluckebier, G.; Frokjaer, S. Binding mode of Thioflavin T in insulin amyloid fibrils. *J. Struct. Biol.* **2007**, *159*, 483−497.
(32) Schlein, M. Insulin Formulation Characterization-the Thioflavin T Assays. *AAPS J.* **2017**, *19*, 397−408.
(33) Banerjee, V.; Das, K. P. Modulation of pathway of insulin fibrillation by a small molecule helix inducer 2,2,2-trifluoroethanol. *Colloids Surf., B* **2012**, *92*, 142−150.
(34) Hua, Q.; Weiss, M. A. Comparative 2D NMR studies of human insulin and despentapeptide insulin: sequential resonance assignment and implications for protein dynamics and receptor recognition. *Biochemistry* **1991**, *30*, 5505−5515.
(35) Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An analytical solution to the kinetics of breakable filament assembly. *Science* **2009**, *326*, 1533−1537.
(36) Benseny-Cases, N.; Cócera, M.; Cladera, J. Conversion of non-fibrillar β-sheet oligomers into amyloid fibrils in Alzheimer's disease amyloid peptide aggregation. *Biochem. Biophys. Res. Commun.* **2007**, *361*, 916−921.
(37) Arosio, P.; Knowles, T. P. J.; Linse, S. On the lag phase in amyloid fibril formation. *Phys. Chem. Chem. Phys.* **2015**, *17*, 7606−7618.
(38) Dolui, S.; Roy, A.; Pal, U.; Saha, A.; Maiti, N. C. Structural Insight of Amyloidogenic Intermediates of Human Insulin. *ACS Omega* **2018**, *3*, 2452−2462.
(39) Das, A. K.; Rawat, A.; Bhowmik, D.; Pandit, R.; Huster, D.; Maiti, S. An Early Folding Contact between Phe19 and Leu34 is Critical for Amyloid-β Oligomer Toxicity. *ACS Chem. Neurosci.* **2015**, *6*, 1290−1295.
(40) Marshall, K. E.; Vadukul, D. M.; Dahal, L.; Theisen, A.; Fowler, M. W.; Al-Hilaly, Y.; Ford, L.; Kemenes, G.; Day, I. J.; Staras, K.; Serpell, L. C. A critical role for the self-assembly of Amyloid-beta1-42 in neurodegeneration. *Sci. Rep.* **2016**, *6*, 30182.
(41) Canet, D.; Last, A. M.; Tito, P.; Sunde, M.; Spencer, A.; Archer, D. B.; Redfield, C.; Robinson, C. V.; Dobson, C. M. Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme. *Nat. Struct. Biol.* **2002**, *9*, 308−315.
(42) Nielsen, S. B.; Macchi, F.; Raccosta, S.; Langkilde, A. E.; Giehm, L.; Kyrsting, A.; Svane, A. S. P.; Manno, M.; Christiansen, G.; Nielsen, N. C.; Oddershede, L.; Vestergaard, B.; Otzen, D. E. Wildtype and A30P mutant alpha-synuclein form different fibril structures. *PLoS One* **2013**, *8*, No. e67713.
(43) Smith, R. A. S.; Nabok, A.; Blakeman, B. J. F.; Xue, W. F.; Abell, B.; Smith, D. P. Analysis of Toxic Amyloid Fibril Interactions at Natively Derived Membranes by Ellipsometry. *PLoS One* **2015**, *10*, No. e0132309.
(44) Jakhria, T.; Hellewell, A. L.; Porter, M. Y.; Jackson, M. P.; Tipping, K. W.; Xue, W.-F.; Radford, S. E.; Hewitt, E. W. β2-Microglobulin Amyloid Fibrils Are Nanoparticles That Disrupt Lysosomal Membrane Protein Trafficking and Inhibit Protein Degradation by Lysosomes. *J. Biol. Chem.* **2014**, *289*, 35781−35794.
(45) Lee, Y. J.; Savtchenko, R.; Ostapchenko, V. G.; Makarava, N.; Baskakov, I. V. Molecular Structure of Amyloid Fibrils Controls the Relationship between Fibrillar Size and Toxicity. *PLoS One* **2011**, *6*, No. e20244.
(46) Xue, W.-F.; Hellewell, A. L.; Hewitt, E. W.; Radford, S. E. Fibril fragmentation in amyloid assembly and cytotoxicity. *Prion* **2010**, *4*, 20−25.
(47) Kar, R. K.; Gazova, Z.; Bednarikova, Z.; Mroue, K. H.; Ghosh, A.; Zhang, R.; Ulicna, K.; Siebert, H. C.; Nifantiev, N. E.; Bhunia, A. Evidence for Inhibition of Lysozyme Amyloid Fibrillization by Peptide Fragments from Human Lysozyme: A Combined Spectroscopy, Microscopy, and Docking Study. *Biomacromolecules* **2016**, *17*, 1998−2009.
(48) Apetri, M. M.; Maiti, N. C.; Zagorski, M. G.; Carey, P. R.; Anderson, V. E. Secondary Structure of α-Synuclein Oligomers: Characterization by Raman and Atomic Force Microscopy. *J. Mol. Biol.* **2006**, *355*, 63−71.
(49) Kurouski, D.; van Duyne, R. P.; Lednev, I. K. Exploring the structure and formation mechanism of amyloid fibrils by Raman spectroscopy: a review. *Analyst* **2015**, *140*, 4967−4980.
(50) Huang, K.; Maiti, N. C.; Phillips, N. B.; Carey, P. R.; Weiss, M. A. Structure-Specific Effects of Protein Topology on Cross-β Assembly: Studies of Insulin Fibrillation. *Biochemistry* **2006**, *45*, 10278−10293.
(51) Zako, T.; Sakono, M.; Hashimoto, N.; Ihara, M.; Maeda, M. Bovine insulin filaments induced by reducing disulfide bonds show a different morphology, secondary structure, and cell toxicity from intact insulin amyloid fibrils. *Biophys. J.* **2009**, *96*, 3331−3340.
(52) Maiti, N. C.; Apetri, M. M.; Zagorski, M. G.; Carey, P. R.; Anderson, V. E. Raman Spectroscopic Characterization of Secondary Structure in Natively Unfolded Proteins: α-Synuclein. *J. Am. Chem. Soc.* **2004**, *126*, 2399−2408.
(53) Lippert, J. L.; Tyminski, D.; Desmeules, P. J. Determination of the secondary structure of proteins by laser Raman spectroscopy. *J. Am. Chem. Soc.* **1976**, *98*, 7075−7080.
(54) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A. Elucidating Peptide and Protein Structure and Dynamics: UV Resonance Raman Spectroscopy. *J. Phys. Chem. Lett.* **2011**, *2*, 334−344.
(55) Dong, J.; Wan, Z.; Popov, M.; Carey, P. R.; Weiss, M. A. Insulin assembly damps conformational fluctuations: Raman analysis of amide I linewidths in native states and fibrils. *J. Mol. Biol.* **2003**, *330*, 431−442.
(56) Sereda, V.; Sawaya, M. R.; Lednev, I. K. Structural Organization of Insulin Fibrils Based on Polarized Raman Spectroscopy: Evaluation of Existing Models. *J. Am. Chem. Soc.* **2015**, *137*, 11312−11320.
(57) Selvaratnam, R.; Chowdhury, S.; VanSchouwen, B.; Melacini, G. Mapping allostery through the covariance analysis of NMR chemical shifts. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6133−6138.
(58) Bhattacharyya, D.; Kumar, R.; Mehra, S.; Ghosh, A.; Maji, S. K.; Bhunia, A. Multitude NMR studies of α-synuclein familial mutants: probing their differential aggregation propensities. *Chem. Commun.* **2018**, *54*, 3605−3608.
(59) Rafferty, B.; Flohr, Z. C.; Martini, A. Protein Contact Maps. https://nanohub.org/resources/contactmaps (accessed March 25, 2019).
(60) Sneideris, T.; Darguzis, D.; Botyriute, A.; Grigaliunas, M.; Winter, R.; Smirnovas, V. pH-Driven Polymorphism of Insulin Amyloid-Like Fibrils. *PLoS One* **2015**, *10*, No. e0136602.
(61) Gibson, T. J.; Murphy, R. M. Inhibition of insulin fibrillogenesis with targeted peptides. *Protein Sci.* **2006**, *15*, 1133−1141.

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136

(62) Nielsen, L.; Frokjaer, S.; Brange, J.; Uversky, V. N.; Fink, A. L. Probing the Mechanism of Insulin Fibril Formation with Insulin Mutants. *Biochemistry* **2001**, *40*, 8397−8409.
(63) Brange, J.; Andersen, L.; Laursen, E. D.; Meyn, G.; Rasmussen, E. Toward understanding insulin fibrillation. *J. Pharm. Sci.* **1997**, *86*, 517−525.
(64) Brange, J.; Dodson, G. G.; Edwards, D. J.; Holden, P. H.; Whittingham, J. L. A model of insulin fibrils derived from the x-ray crystal structure of a monomeric insulin (despentapeptide insulin). *Proteins* **1997**, *27*, 507−516.
(65) Markussen, J.; Diers, I.; Engesgaard, A.; Hansen, M. T.; Hougaard, P.; Langkjaer, L.; Norris, K.; Ribel, U.; Sørensen, A. R.; Sørensen, E.; et al. Soluble, prolonged-acting insulin derivatives. II. Degree of protraction and crystallizability of insulins substituted in positions A17, B8, B13, B27 and B30. *Protein Eng.* **1987**, *1*, 215−223.
(66) Hilgenfeld, R.; Seipke, G.; Berchtold, H.; Owens, D. R. The evolution of insulin glargine and its continuing contribution to diabetes care. *Drugs* **2014**, *74*, 911−927.
(67) Guo, M.; Gorman, P. M.; Rico, M.; Chakrabartty, A.; Laurents, D. V. Charge substitution shows that repulsive electrostatic interactions impede the oligomerization of Alzheimer amyloid peptides. *FEBS Lett.* **2005**, *579*, 3574−3578.
(68) Mompeán, M.; Chakrabartty, A.; Buratti, E.; Laurents, D. V. Electrostatic Repulsion Governs TDP-43 C-terminal Domain Aggregation. *PLoS Biol.* **2016**, *14*, No. e1002447.
(69) Ratha, B. N.; Ghosh, A.; Brender, J. R.; Gayen, N.; Ilyas, H.; Neeraja, C.; Das, K. P.; Mandal, A. K.; Bhunia, A. Inhibition of Insulin Amyloid Fibrillation by a Novel Amphipathic Heptapeptide: MECHANISTIC DETAILS STUDIED BY SPECTROSCOPY IN COMBINATION WITH MICROSCOPY. *J. Biol. Chem.* **2016**, *291*, 23545−23556.
(70) Banerjee, V.; Kar, R. K.; Datta, A.; Parthasarathi, K.; Chatterjee, S.; Das, K. P.; Bhunia, A. Use of a small peptide fragment as an inhibitor of insulin fibrillation process: a study by high and low resolution spectroscopy. *PLoS One* **2013**, *8*, No. e72318.
(71) Hackl, E. V.; Darkwah, J.; Smith, G.; Ermolina, I. Effect of acidic and basic pH on Thioflavin T absorbance and fluorescence. *Eur. Biophys. J.* **2015**, *44*, 249−261.
(72) Ratha, B. N.; Kar, R. K.; Kalita, S.; Kalita, S.; Raha, S.; Singha, A.; Garai, K.; Mandal, B.; Bhunia, A. Sequence specificity of amylin-insulin interaction: a fragment-based insulin fibrillation inhibition study. *Biochim. Biophys. Acta, Proteins Proteomics* **2019**, *1867*, 405−415.
(73) Morris, A. M.; Watzky, M. A.; Finke, R. G. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochim. Biophys. Acta, Proteins Proteomics* **2009**, *1794*, 375−397.
(74) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(75) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537−W541.
(76) Berjanskii, M. V.; Neal, S.; Wishart, D. S. PREDITOR: a web server for predicting protein torsion angle restraints. *Nucleic Acids Res.* **2006**, *34*, W63−W69.
(77) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296−W303.
(78) Maiti, R.; van Domselaar, G. H.; Zhang, H.; Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **2004**, *32*, W590−W594.
(79) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037−10041.
(80) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(81) Grimsley, G. R.; Martin Scholtz, J.; Nick Pace, C. A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* **2009**, *18*, 247−251.
(82) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, *22*, 501−508.
(83) Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin thermostat for rigid body dynamics. *J. Chem. Phys.* **2009**, *130*, 234101.
(84) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(85) Feig, M.; Karanicolas, J.; Brooks, C. L., 3rd MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.

https://dx.doi.org/10.1021/acs.jpcb.9b10349
J. Phys. Chem. B 2020, 124, 1125−1136