Supplementary Information

Biophysical characterization of a protein for structure comparison: methods for identifying insulin structural changes

Meropi Sklepari,[a] Alison Rodger,*[a] Andrew Reason,[b] Shirin Jamshidi,[a] Ivan Prokes,[a] Claudia A. Blindauer[a]

[a.] Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK. E-mail: a.rodger@warwick.ac.uk
[b.] Biopharmaspec Ltd, Suite 3.1, Lido Medical Centre, St. Saviour, Jersey JE2 7LA, UK.

INTRODUCTION

This section provides material supplementary to what is discussed in the main text. For the DLS temperature experiments, the size of particles in terms of number values is given for all 5 insulin samples, and an additional experiment which establishes the hexamer values is presented. For the CD experiments, the secondary structure analysis of the CD spectra with SSNN and SELCON3 is given and the spectra of the high temperature samples cooled down to 20 °C are discussed, further supporting the previously reported suggestion of a partially folded intermediate prior to the events of unfolding or fibrillization of insulin. For the ^{1}H-NMR experiments, the effect of the previously reported metal coordination on the neutral samples is explained. For the acidic sample, an example is shown where the shift of a specific resonance with the increase of temperature can be used to identify significant changes in the structure of a protein. In addition, the Hα region is discussed in more detail and resonances that are believed to be related with aggregation phenomena are assigned through 2D-NMR experiments. Finally, the molecular modeling simulations for the dimer and hexamer are further discussed, supporting the existence of a partially folded intermediate. A tendency towards dissociation is suggested for the hexamer in accord with our observations in the NMR experiments, while each one of the monomeric/dimeric units composing the dimer and the hexamer respectively, seem to present different unfolding patterns.

MATERIALS AND METHODS

Sample Preparation

DLS hexamer control sample:[1] insulin was initially dissolved in NaOH (0.1 M). Sodium phosphate buffer (100 mM, pH 7.0), NaCl (1 M) and water were added, resulting in an insulin solution of ~8 mg/mL in 10 mM sodium phosphate buffer, 100 mM NaCl and pH 8.0. The refractive index of a blank (identical to control solution omitting the protein) was measured. The viscosity of the dispersant (blank) was corrected[2] (Table S1), by running measurements of a solution consisting of 1 µL of polystyrene latex spheres of known size (100 nm) in 25 µL of the blank solution at each temperature. The sample and the blank were centrifuged with a centrifugal force of 6,000 rpm for 10 minutes and filtered through a 0.2 µm Thermo scientific syringe filter directly into the Quartz cuvette prior to the measurement.

Temperature dependent measurements

CD spectra were recorded at the conclusion of the temperature experiment, 5–8 hours after they had been cooled down to 20 °C. The same parameters described in the main text were used.

The *DLS experiment* was conducted in a temperature range from 20–75 °C with a 5 °C temperature interval, 6 measurements were recorded for each temperature. Each measurement consisted of 10 runs, with 200 s total duration time. 180 s were allowed as the equilibration time after the desired temperature was reached.

Simulated annealing (SA)

Table S2 gives the time and temperature profile of the MD simulations. A total run of 1.68 ns of MD simulated annealing was performed for all insulin systems.

RESULTS AND DISCUSSION

DLS Spectroscopy

Table S3 shows the results from the dynamic light scattering temperature experiments for all 5 samples. The hydrodynamic diameters of insulin particles are shown in terms of number values (in nm). The given percentages refer to the population of each size in the sample, in cases where this was different from 100%.

The hexamer control sample at 1.4 mM protein appeared to be nearly monodisperse in the temperature range from 35–65 °C, showing one main population to be present below 70 °C with hydrodynamic diameter (intensity values) of 5.5–6.1 nm. Larger aggregates appeared at 70 °C, while at 75 °C the sample had probably evaporated as the intensity decreased significantly (data not shown). Table S4 shows the sizes in number values from 20–75 °C.

CD Spectroscopy

Figure S1 shows the secondary structure content for all 5 samples of insulin at 20, 75 and 110 °C. The spectra were analyzed using SSNN[3] and SELCON3.[4] Both programs showed that the acidic and the two neutral samples preserved more helical content than the basic samples at the end of the experiment.

Figure S2A shows the CD spectra for all five samples, 5–8 hours after they had been cooled back to 20 °C. The spectra of the neutral and basic samples suggest that irreversible unfolding takes place in both cases, while the consistent difference in the wavelength of the minima between basic (199.6 nm for both samples) and neutral pH (202.8 nm for the EDTA-containing sample, 205 nm for the EDTA-free sample) is believed to indicate the presence of the completely unfolded monomer at basic pH (minimum below 200 nm is characteristic of random coil), and an equilibrium of unfolded hexamers and dimers that have not completely dissociated and unfolded dissociated monomers at neutral pH[5] (also discussed in the molecular modeling section of supplementary information). The higher minimum wavelength of the neutral EDTA-free sample compared to the EDTA-containing sample further supports this interpretation, as the EDTA-free sample is believed to contain the more stable Zn_2-hexamer as the predominant assembly which is therefore more resistant to dissociation, in contrast to the EDTA-containing sample which is believed to be a mixture of dimers, tetramers and possibly Zn-free hexamers which would dissociate more easily.

The highly helical spectrum at pH 2.8 after cooling back to 20 °C shows that refolding to a partially folded structure occurred, as the spectrum is very similar to that at 45 °C (Figure S2B). This is in agreement with previous studies suggesting the presence of a common partially folded intermediate during fibrillation as well as random coil generation.[6-8] We assume therefore that the key partially folded conformation is shown by the spectrum at 110 °C (with a minimum at 207.6 nm) and suggests that this corresponds to the intermediate that the literature refers to and which leads either to fibrillation or to complete unfolding. Additionally, cooling down to 20 °C causes this intermediate conformation to refold to a structure that resembles the one that is observed at 45 °C, where one of the misfolded helices has probably reformed. A possible explanation for that could be that the dissociation of the dimer (which is the predominant conformation at acidic pH) results in two monomers with different flexibility. This opinion is also supported by our simulation data (see molecular modeling section), which showed two different monomer conformations. Thus, while one monomer probably represents the completely unfolded conformation (due to its resemblance to the final trajectory from the monomer simulation, Figure S11) the other monomer seems to represent the partially folded intermediate, leading to either of the three different pathways mentioned above.

Figure S3 shows the CD spectra for all five samples, at the temperature where it is believed that the partially folded intermediate, with a minimum at 207.0/207.6 nm, or a very similar structure (in terms of secondary structure content), is observed for the first time during the temperature experiment. As we previously described and as expected according to the stability that increases following the order monomer < dimer < hexamer, this structure is observed earlier for the two basic samples where the monomer is predominant (at 65 °C), at 80 °C for the acidic sample (mostly dimeric conformation) and last at 90 °C for both neutral samples where the presence of tetramers/hexamers is favoured.

^{1}H-NMR Spectroscopy

The presence of the peak at 7.43 ppm in the case of the two neutral samples (Figure S4A, B), has been previously found to refer to the Hε1 proton of the Zn-free His(B10) signal in higher aggregates (hexamer and half part of the tetramer), where a shielding effect is caused by the histidine rings from the opposite dimeric units.[9] In the presence of Zn though (EDTA-free sample), the formation of the Zn_2-hexamer is preferred to the Zn-free molecule,[10] thus the 7.43 ppm peak is believed to represent the tetramer, while in the presence of EDTA (complexed Zn) the formation of both tetramers and Zn-free hexamers is favored and also supported by the high association constant of the Zn-free hexamer.[11] In the other half of each tetramer and in dimers, the same proton of His(B10) is exposed and deshielded at 7.71 ppm as previously reported.[9] The broad peak at 7.58 ppm, more obvious in the case of the neutral sample containing Zn (EDTA-free), is believed to represent both the His(B5) Hε1 proton of a high aggregate (hexamer) due to its upfield shift and the Zn-coordinated His(B10) resonance of a hexamer where the downfield effect from the Zn-binding is balanced by the upfield effect from the opposite rings.[9,12,13] The peak at 7.66 ppm in the EDTA-containing neutral sample represents the His(B5) Hε1 proton of

smaller oligomers such as dimers and tetramers. Therefore, the above evidence suggests the presence of tetramers and hexamers as predominant species at neutral pH in the presence of Zn (EDTA-free sample) and the presence of dimers, tetramers and Zn-free hexamers in the case of the EDTA-containing neutral sample. The progressive sharpening of the His(B5) and His(B10) Hε1 proton peaks with the increase of temperature is probably a sign that the Zn is not bound to insulin anymore (e.g. EDTA-free neutral sample at 70-75 °C, Figure S4A), as seems to be the case in acidic conditions (Figure S4C) in which His(B10) is unable to coordinate metal ions.

The sharpening of the Histidine Hε1 peaks was followed with the increase of temperature for all five samples and the line widths measured at half height of the peaks (FWHH) are reported in Table S5. The FWHH values increase from the monomer (basic pH, ~4 Hz) to the dimer (acidic pH, ~7 Hz) and tetramer/hexamer (neutral EDTA-containing, ~27 Hz/neutral EDTA-free sample, ~35 Hz) and decrease to ~4 Hz at higher temperatures for all samples. Thus, indicating the transition to smaller oligomers (probably monomers) that will start aggregating before fibrillation takes place. The line widths of the histidine peak that corresponds to the denatured species is also given for the basic samples and those of the characteristic for the dimeric unit and the monomer peaks (at ~0.40 and 0.15 ppm respectively) for the basic samples and the neutral EDTA-free sample.

At pH 2.5, the upfield shift of a doublet peak from 8.48 ppm (in the 25 °C spectrum) to 8.15 ppm at 65 °C (Figure S4C) is remarkable. This signal is believed to represent the backbone amide proton of Asn(B3) with a coupling constant $^{3}J_{H}{}^{N}{}_{H}{}^{\alpha}$ of 6.7 Hz, and by the –0.37 ppm shift it can be inferred that a gradual shielding effect is caused by the formation of higher aggregates.[14-16] This effect is probably caused by the PheB1 ring, as both residues are positioned on the dimer-dimer interface.[17]

Interesting is the fact that different peaks appear for each sample in the region from 5.0–6.6 ppm (Figure S5), with the only common peak for all samples being the one at 5.04 ppm. The 2D-TOCSY spectrum at pH 2.5 at 25 °C (Figure S6), showed a correlation with resonances at 8.53 ppm and 2.97 ppm. As was previously discussed, the Hε1 proton of the His(B5) residue appears at 8.53 ppm, therefore the 5.04 ppm is likely to represent the Hα proton of His(B5). The peak at 5.37 ppm is also a Hα proton and presents correlations with two spin systems, one of which probably belongs to an aromatic residue as it correlates with two aliphatic peaks at 2.81 and 3.35 ppm and an aromatic peak at 7.22 ppm (Figure S6).[18] This peak is tentatively assigned to PheB24 on the C-terminus of chain B on the monomer-monomer interface, due to its further NOESY correlations (Figure S9). At acidic pH, the peak at 6.28 ppm has a correlation with the peak at 7.49 ppm (Figure S6). Given the combination of chemical shifts, the fact that no cross-peak was present in the ^{13}C-HSQC spectrum (Figure S7), and the exclusive presence at acidic pH, this likely corresponds to an arginine side-chain NH proton. From the 7.49 ppm peak and its through-bond correlations we found that these peaks are related to protons at 2.96 (δ-CH_2), 1.66, 1.41 (γ-CH_2), 1.76, 1.86 (β-CH_2), 4.3 (α-H), 8.26 ppm (backbone ^{N}H) (Figure S8). Thus the 6.28 ppm signal is likely to be one of the side chain NH protons of ArgB22. The peak at 6.6 ppm exhibits a cross-peak that correlates with the peak at 7.09 ppm in the 2D-TOCSY spectrum (Figure S8). The absence of correlation peaks in the ^{13}C-HSQC spectrum (Figure S7) indicates that both peaks also refer to side-chain NH protons, while both the 6.6 ppm and 7.09 ppm signals further correlate with the protons at 3.95 and 2.55 ppm in the 2D-NOESY spectrum (Figure S9). Therefore they are likely to be signals of an asparagine residue. The two neutral samples present peaks at 5.04 and 5.37 ppm, which are also reported for the ^{1}H-NMR spectrum of the native Zn_2-hexamer in the Zn(II)-T_6 state,[19-21] while in the two basic samples only the peak at 5.04 ppm appears.

Molecular Modeling

Figure S10 shows monomer 2 from the dimer simulation at 20, 50, 65, 75 and 110 °C. For both chains A (left) and B (right image), which are depicted separately for clarity, the flexibility of the N-termini is apparent. The displacement of the two chains at 110 °C (in yellow) is indicative of the high tendency for dissociation of the dimer to the two misfolded monomers. The structures from the monomer (white) and the dimer (yellow) simulations at 110 °C are shown in Figure S11. The two monomeric units (1 and 2) that constitute the dimer, present different patterns of unfolding; monomer 1 appears to be significantly similar to the structure of the monomer at 110 °C. The extended C-terminus of chain B leads us to the conclusion that this structure represents the completely unfolded conformation, while monomer 2 probably refers to the partially folded intermediate.

Comparison of the 20 °C and 110 °C structures of the hexamer (Figure S12) shows that the interactions between the antiparallel β-strands of the B-chain C-terminal portions are weakened, and that the B chain N-terminus is completely extended. A more careful look into the structures from the hexamer simulation (Figure S13) suggests that each one of the three dimers can potentially follow a different unfolding pattern. Both chains A (left) and B (right) of dimer 1 appear to be more stable and compact in contrast to the increased flexibility of dimers 2 and 3, which gradually diverge leading to more expanded structures and to a complete extension of the N-termini of both chains B in the case of dimer 3. However, dimer 3 appears to preserve its antiparallel β-strand character in the same way as dimer 1, while dimer 2 appears prone to further dissociation to monomers as the antiparallel β-strand character has disappeared. Therefore, in the timescale that our CD, NMR and DLS experiments were conducted (which was significantly longer than that of the simulations), the hexameric form of insulin can result in a mixture of misfolded dimers and monomers at increased temperatures and that also explains its stability, as the dissociation and unfolding process needs to evolve through several stages.

Videos S1-S3 show the simulations for the monomer, dimer and hexamer respectively as a function of increasing temperature.

REFERENCES

1. W. Kadima, L. Øgendal, R. Bauer, N. Kaarsholm, K. Brodersen, J. F. Hansen and P. Porting, *Biopolymers*, 1993, **33**, 1643-1657.
2. E. L. Gilroy, M. R. Hicks, D. J. Smith and A. Rodger, *Analyst*, 2011, **136**, 4159-4163.
3. V. Hall, A. Nash and A. Rodger, *Anal. Methods*, 2014, **6**, 6721-6726.
4. N. Sreerama and R. W. Woody, *Anal Biochem*, 1993, **209**, 32-44.
5. V. N. Uversky and A. L. Fink, *Biochim Biophys Acta*, 2004, **1698**, 131-153.
6. M. R. H. Krebs, E. H. C. Bromley, S. S. Rogers and A. M. Donald, *Biophys. J.*, 2005, **88**, 2013-2021.
7. A. Arora, C. Ha and C. B. Park, *Protein Sci.*, 2004, **13**, 2429-2436.
8. Q.-x. Hua and M. A. Weiss, *J. Biol. Chem.*, 2004, **279**, 21449-21460.
9. W. Kadima, M. Roy, R. W. Lee, N. C. Kaarsholm and M. F. Dunn, *J. Biol. Chem.*, 1992, **267**, 8963-8970.
10. T. Blundell, G. Dodson, D. Hodgkin and D. Mercola, in *Advances in Protein Chemistry*, eds. J. T. E. C.B. Anfinsen and M. R. Frederic, Academic Press, 1972, vol. Volume 26, pp. 279-402.
11. B. Skelbaek-Pedersen, J. Brange, L. Langkjaer, U. Damgaard, H. Ege, S. Havelund, L. G. Heding, K. H. Joergensen, J. Lykkeberg and J. Markussen, *Galenics of Insulin: The Physico-chemical and Pharmaceutical Aspects of Insulin and Insulin Preparations*, Springer Berlin Heidelberg, 2012.
12. J. H. Bradbury, V. Ramesh and G. Dodson, *J. Mol. Biol.*, 1981, **150**, 609-613.
13. R. Palmieri, R. W. K. Lee and M. F. Dunn, *Biochemistry*, 1988, **27**, 3387-3397.
14. D. S. Wishart, B. D. Sykes and F. M. Richards, *J. Mol. Biol.*, 1991, **222**, 311-333.
15. S. Ludvigsen, M. Roy, H. Thoegersen and N. C. Kaarsholm, *Biochemistry*, 1994, **33**, 7998-8006.
16. H. B. Olsen, S. Ludvigsen and N. C. Kaarsholm, *Biochemistry*, 1996, **35**, 8836-8845.
17. E. Jacoby, Q. X. Hua, A. S. Stern, B. H. Frank and M. A. Weiss, *J. Mol. Biol.*, 1996, **258**, 136-157.
18. M. Roy, R. W. K. Lee, N. C. Kaarsholm, H. Thøgersen, J. Brange and M. F. Dunn, *Biochim. Biophys. Acta, Mol. Cell Res.*, 1990, **1053**, 63-73.
19. M. Roy, M. L. Brader, R. W. Lee, N. C. Kaarsholm, J. F. Hansen and M. F. Dunn, *J. Biol. Chem.*, 1989, **264**, 19081-19085.
20. M. L. Brader, N. C. Kaarsholm, R. W. K. Lee and M. F. Dunn, *Biochemistry*, 1991, **30**, 6636-6645.
21. P. S. Brzovic, W. E. Choi, D. Borchardt, N. C. Kaarsholm and M. F. Dunn, *Biochemistry*, 1994, **33**, 13057-13069.

Table S1: Corrected values for viscosity at each temperature and refractive indices for each one of the blank solutions, obtained as described in the main text and in the supplementary information text. The values presented for the blank at pH 2.5 are the same as water values.

Temperature (°C)	Blank pH 2.5	Blank pH 7.7	Blank pH 7.6 EDTA	Blank pH 10.9	Blank pH 10.5 EDTA	Blank NaCl 100 mM
20	1.003	1.018	0.991	1.075	1.038	0.988
25	0.887	0.936	0.906	0.945	0.939	0.874
30	0.792	0.852	0.776	0.790	0.831	0.785
35	0.714	0.759	0.683	0.724	0.710	0.701
40	0.649	0.675	0.648	0.673	0.683	0.637
45	0.595	0.616	0.592	0.621	0.631	0.593
50	0.548	0.580	0.533	0.566	0.556	0.552
55	0.508	0.522	0.491	0.520	0.527	0.514
60	0.471	0.533	0.455	0.476	0.479	0.479
65	0.438	0.477	0.433	0.478	0.461	0.440
70	0.407	0.453	0.398	0.444	0.410	0.414
75	0.378	0.456	0.375	0.425	0.400	0.385
Refractive Index	1.333	1.3334	1.3332	1.3345	1.3335	1.334

Table S2: Time and temperature profile of the SA-MD simulations for all insulin systems.

Steps	Temperature (K)	Time (fs)
Annealing Step 1	0–293	500
Annealing Step 2	293	60000
Annealing Step 3	293–323	60000
Annealing Step 4	323	240000
Annealing Step 5	323–338	30000
Annealing Step 6	338	120000
Annealing Step 7	338–348	20000
Annealing Step 8	348	80000
Annealing Step 9	348–383	70000
Annealing Step 10	383	800000
Annealing Step 11	383–293	100000
Annealing Step 12	293	100000
Total Time		1685000 fs = 168 ps = 1.68 ns

Table S3: Dynamic light scattering temperature experiment showing the size (hydrodynamic diameter) of insulin particles in terms of number values (in nm). Insulin solutions are 1 mM at acidic, neutral (EDTA-free, EDTA-containing) and basic (EDTA-free, EDTA-containing) pH in 10 mM sodium phosphate buffer. The percentages, where mentioned, describe the population of each size in the sample. 100% population is implied where no percentages are mentioned.

	20°C	25°C	30°C	35°C	40°C	45°C	50°C	55°C	60°C	65°C	70°C	75°C
pH 2.5	2.3	2.4	2.3	2.3	2.3	2.4	2.1	2.0	2.1	2.1	2.0	2.2
pH 7.7	3.7	3.7	3.4	3.4	3.5	3.4	3.2	3.4	3.2	3.7 (87%) 1.7 (13%)	5.3	2.6 (50%) 8.7 (12%)
pH 7.7 EDTA	3.4	3.2	3.2	3.3	3.0	2.9	2.9	2.7	2.6	2.6	3.3	2.5 (62%) 8.5 (38%)
pH 10.9	2.5	2.3	2.8	2.6	2.5 (73%) 1.2 (27%)	2.9	3.4 (74%) 1.6 (26%)	3.1 (59%) 1.6 (41%)	2.4	2.8	2.7	2.8 (51%) 1.1 (49%)
pH 10.5 EDTA	2.8	2.3	2.4	2.4	2.4	2.4	2.6	2.5	3.3 (83%) 0.9 (17%)	2.2 (83%) 6.4 (17%)	5.1	2.0 (14%) 3.1 (19%) 7.6 (67%)

Table S4: Dynamic light scattering temperature experiment showing the size (hydrodynamic diameter) of insulin particles in terms of number values (in nm) in 1.4 mM protein solution of pH 8.0 in 100 mM NaCl, 10 mM sodium phosphate buffer. The percentages, where mentioned, describe the population of each size in the sample. 100% population is implied where no percentage is mentioned.

	20°C	25°C	30°C	35°C	40°C	45°C	50°C	55°C	60°C	65°C	70°C	75°C
pH 8.0, 100 mM NaCl	4.1	4.0	3.7	3.8	3.8	3.7	3.7	3.6	3.4	3.5	3.7	2.9 (33%)

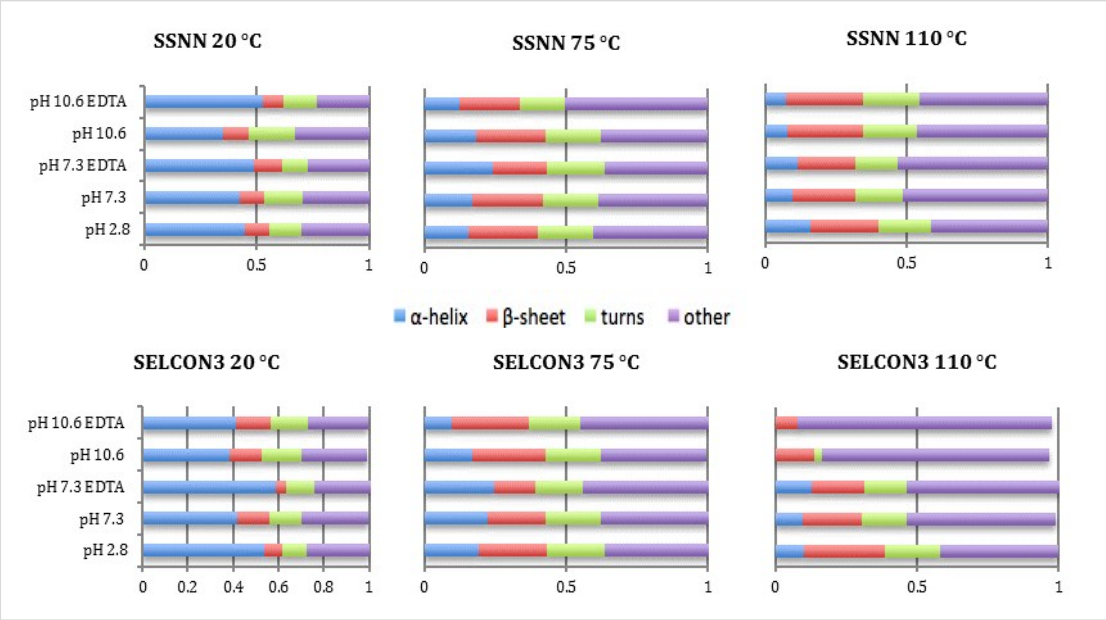

Figure S1: Secondary structure content of the 5 different samples of insulin at 20 °C, 75 °C and 110 °C. The CD spectra were analyzed with SSNN and SELCON3. Insulin solutions are 0.1 mg/mL at acidic, neutral (EDTA-free, EDTA-containing) and basic (EDTA-free, EDTA-containing) pH in 10 mM sodium phosphate buffer.

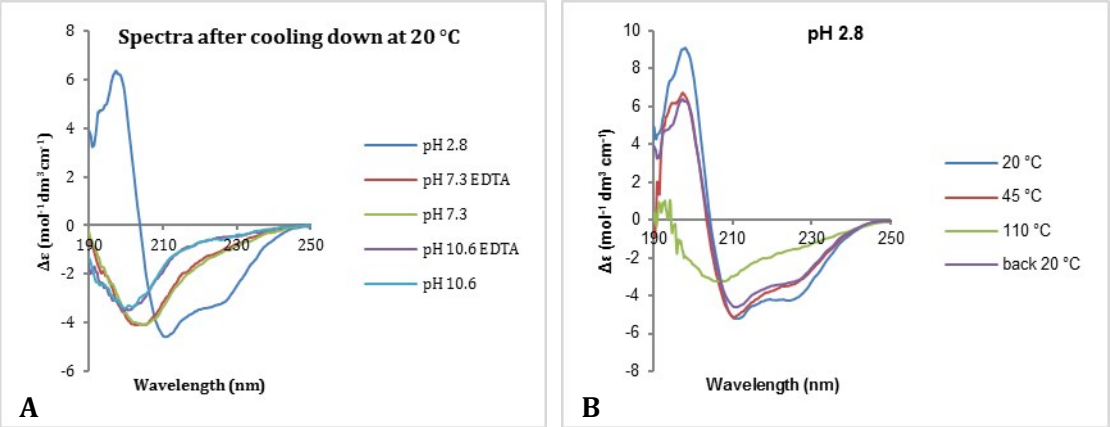

Figure S2: *A*. CD spectra for Insulin 0.1 mg/mL at acidic, neutral (EDTA-containing, EDTA-free) and basic (EDTA-containing, EDTA-free) pH in 10 mM sodium phosphate buffer after having been heated and cooled down to 20 °C. Unfolding of the protein at neutral and basic pH seems to be irreversible, while refolding to a partially folded conformation is apparent in acidic conditions. *B*. Comparison of CD spectra for Insulin at pH 2.8 at 20 °C, 45 °C, 110 °C and after being cooled down to 20 °C. Upon cooling, the partially folded intermediate appearing at 110 °C seems to refold to a structure very close to that observed at 45 °C.

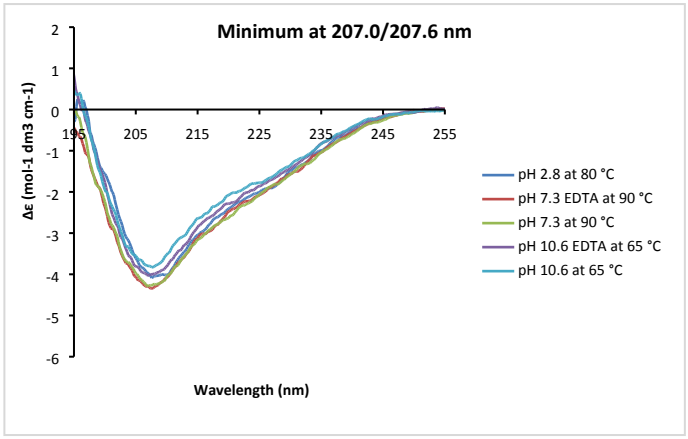

Figure S3: CD spectra for each one of the five samples at the temperature where they first present a minimum at 207.0 or 207.6 nm. This structure is believed to represent that of the partially folded intermediate and is observed at an earlier temperature for the two basic samples (65 °C), at 80 °C for the acidic sample and at 90 °C for the two neutral samples, following their expected stability order as mentioned in the text.

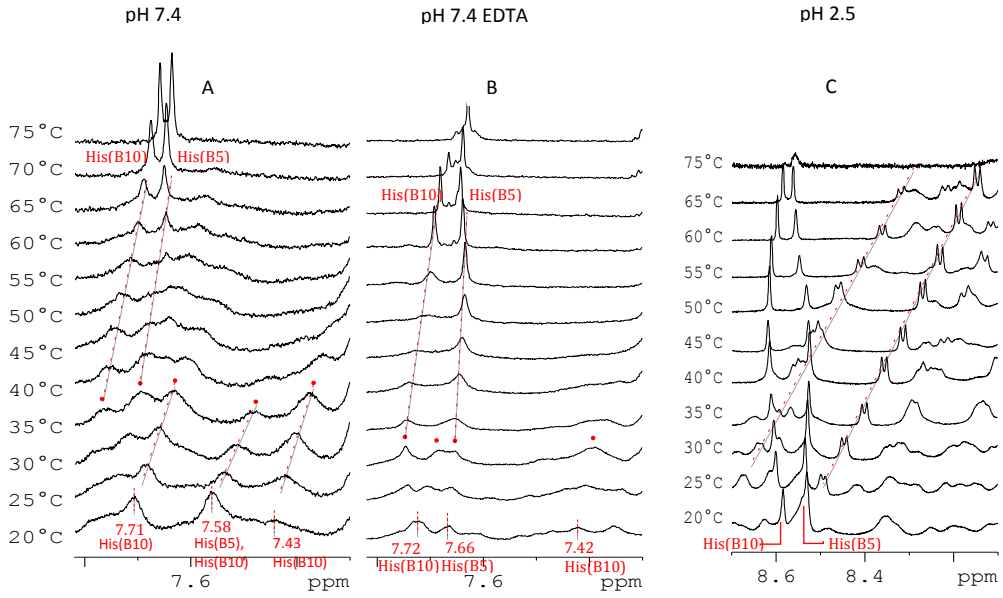

Figure S4: The respective histidine region from the ^{1}H-NMR spectra of insulin at neutral (A, B) and acidic (C) pH is shown as a function of increasing temperature. *A*. EDTA-free neutral sample (Zn-containing), *B*. EDTA-containing neutral sample. The Hε1 proton signals for His(B10) and His(B5) are indicated at low and increased temperatures as broad (high aggregates) and sharp (smaller conformations, not coordinated to Zn) peaks respectively. *C*. acidic sample. The increasing shielding effect on the two indicated doublet peaks with the increase of temperature (dashed lines), is suggested to be a sign of the formation of higher aggregates.

Table S5: The changes on line widths (in Hz) with temperature are shown. The line widths were measured at half height of the peaks. The selected peaks are: the Histidine Hε1 peaks of the five samples, the Leu(B15) Hδ methyl peaks for the dimeric unit (~0.40 ppm) and the monomer (~0.15 ppm) for the basic samples and the neutral EDTA-free sample (dimeric unit only), and the His peak for the denatured species for the basic samples. The measured line widths are indicative of the oligomerisation state and the predominant species in each sample.

		Width (Hz) at T (°C)											
Characteristic peaks	peaks (ppm)	20°C	25°C	30°C	35°C	40°C	45°C	50°C	55°C	60°C	65°C	70°C	75°C
	pH 2.5												
His(B10)	8.58	6.40	7.97	8.13	9.58	6.13	3.75	3.47	3.16	2.91	3.00		
His(B5)	8.53	7.24	6.61	5.27	4.59	5.85		5.76	5.01	3.95	3.15		
	pH 7.4												
His(B10)	7.71	33.54	33.50	32.03	35.19					20.25	12.99	6.20	3.76
His(B5)	7.56	38.72	42.31	41.59	30.71					17.05	8.51	5.39	4.19
Leu(B15) Hδ methyl (dimeric unit)	0.42	47.60	55.03	59.22	43.93	42.44	42.45	43.72	41.95	41.54	41.19	40.36	

	pH 7.3 EDTA												
His(B10)	7.70							30.65	16.18	4.46	3.12	4.05	
His(B5)	7.66	27.25	26.54	27.92	36.71	34.23	20.10	9.02	4.99	4.01	3.98	3.96	4.36
	pH 10.8												
His(B10)	7.67	3.85	3.88	3.95	3.91	3.51	3.43						
His (denatured species)	7.66						11.36	10.54	10.63	9.58	8.22	8.82	10.29
His(B5)	7.53	6.12	6.39	5.91	5.25	4.09	3.76						
Leu(B15) Hδ methyl (dimeric unit)	0.42	33.75	29.25	19.33									
Leu(B15) Hδ methyl (monomer)	0.16	20.12	19.31	18.06	18.40	15.60	14.03						
	pH 10.7 EDTA												
His(B10)	7.65	4.33	4.01	3.78	3.36	3.22	3.04						
His (denatured species)	7.64									7.28	6.65	6.93	7.67
His(B5)	7.51	5.12	4.24	3.88	3.52	4.48	3.60						
Leu(B15) Hδ methyl (dimeric unit)	0.39	30.65	25.07	20.41									
Leu(B15) Hδ methyl (monomer)	0.13	19.73	19.00	17.91	17.61	15.81	16.36						

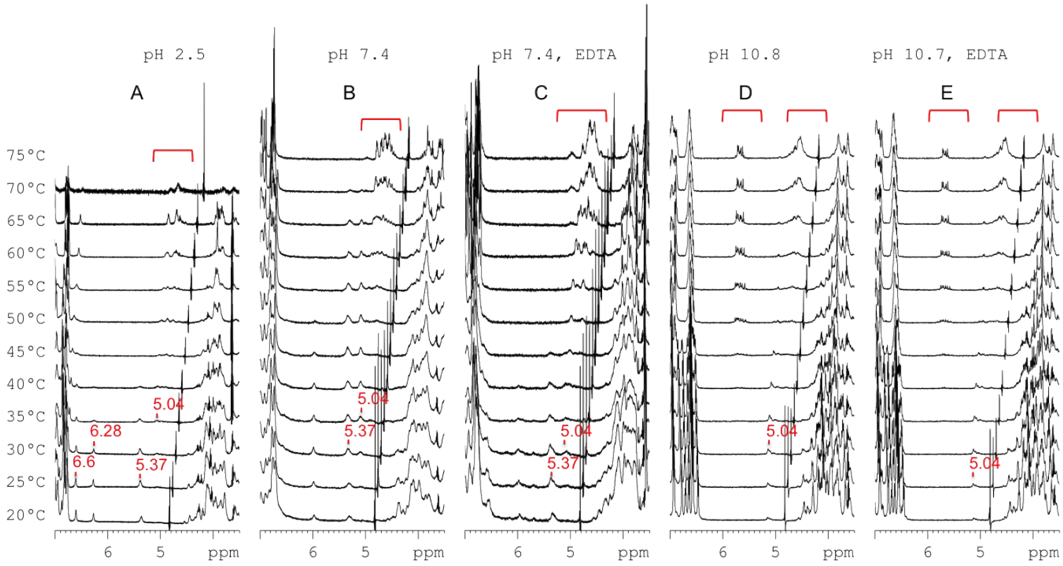

Figure S5: The Hα region (3.5-7.0 ppm) of ^{1}H-NMR spectra of insulin at acidic (A), neutral EDTA-free (B) and EDTA-containing (C), basic EDTA-free (D) and EDTA-containing (E) pH at increasing temperature is shown. The peaks at 5.04 and 5.37 ppm belong to Hα protons, while the peaks at 6.28 and 6.6 ppm are side chain NH protons.

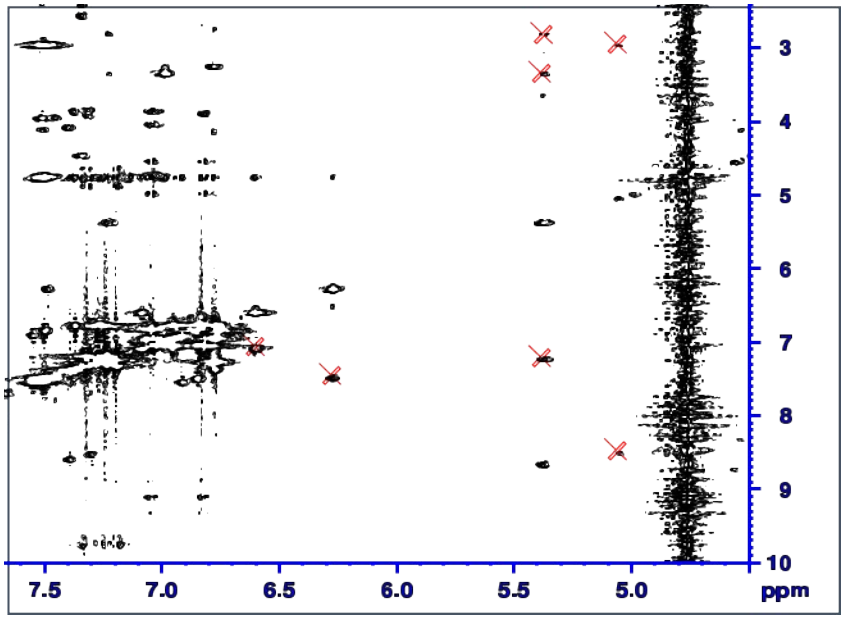

Figure S6: 2D-TOCSY spectrum at pH 2.5 25 °C, showing the correlations of the protons at 5.04, 5.37, 6.28 and 6.6 ppm.

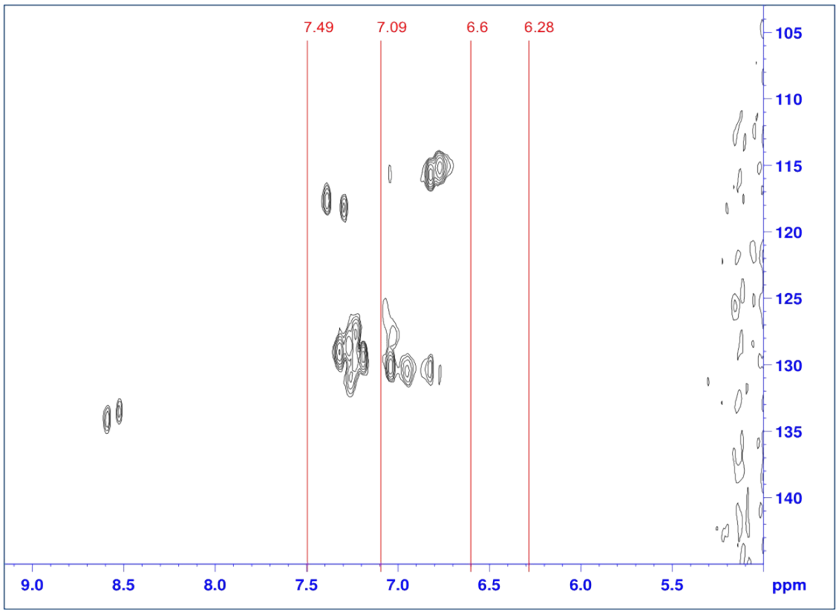

Figure S7: 2D-HSQC (^{13}C and ^{1}H) spectrum at pH 2.5 25 °C, showing the absence of cross-peaks for the peaks at 6.28, 6.6, 7.09 and 7.49 ppm, therefore indicating that these peaks refer to side chain NH protons.

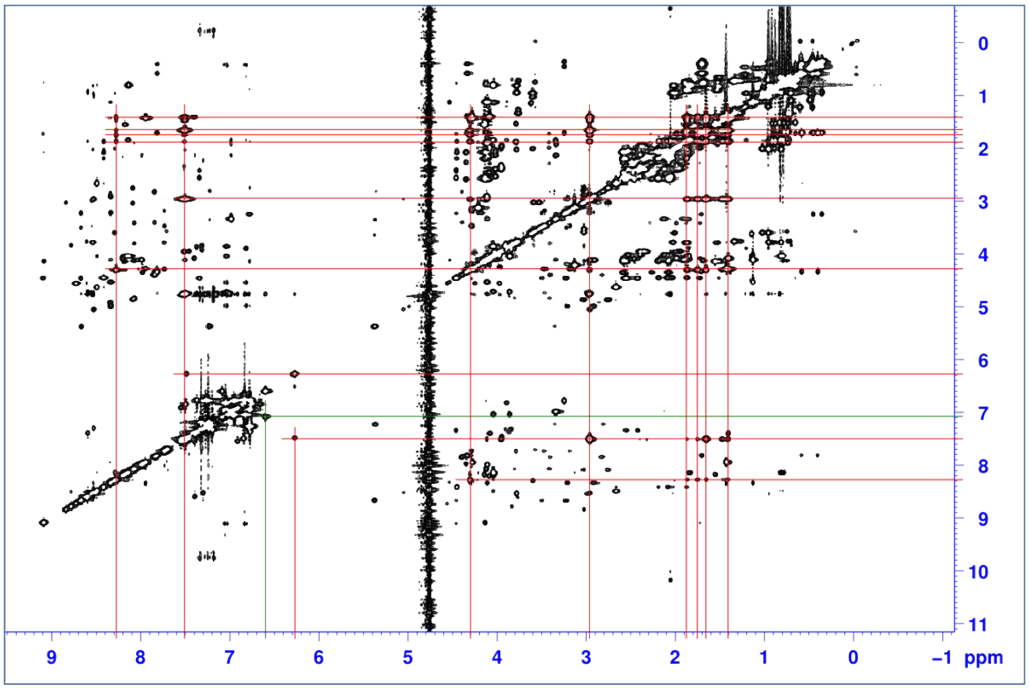

Figure S8: 2D-TOCSY spectrum at pH 2.5 25 °C, showing the correlation of the peak at 6.28 ppm with the peak at 7.49 ppm and the through-bond correlations of the peak at 7.49 ppm with the peaks at 1.41, 1.66 (γ-CH_2), 1.76, 1.86 (β-CH_2), 2.96 (δ-CH_2), 4.3 (α-H) and 8.26 ppm (backbone ^{N}H) in red, probably matching an arginine residue. The correlation of the peak at 6.6 ppm with the peak at 7.09 ppm is shown in green.

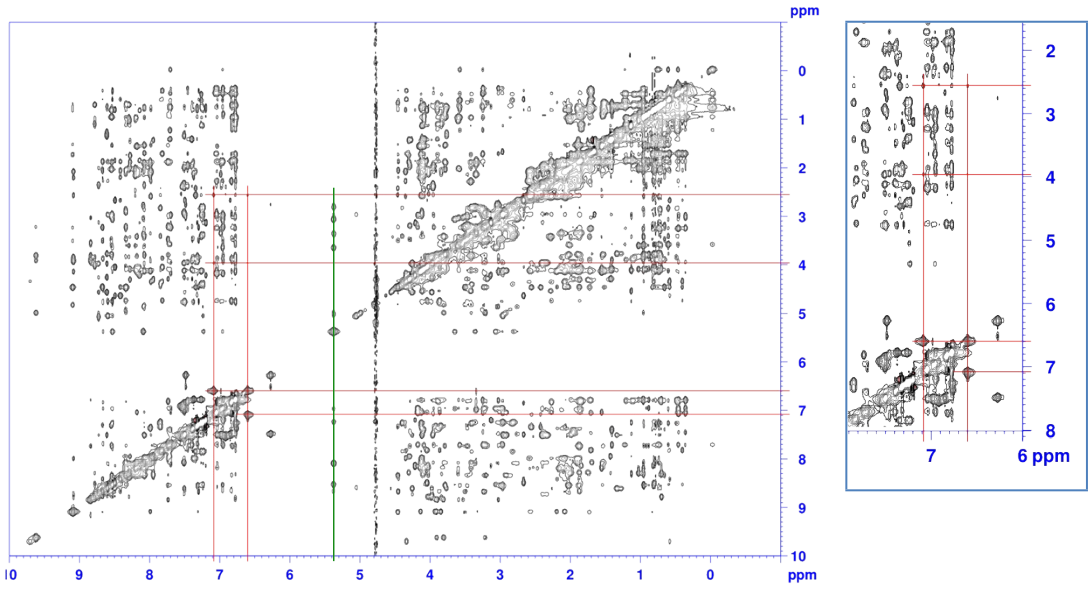

Figure S9: *Left*: 2D-NOESY spectrum at pH 2.5 25 °C, showing the correlation of the peaks at 6.6 and 7.09 ppm with the protons at 2.55 and 3.95 ppm, probably referring to the signals of an asparagine residue. *Right*: zoom in image of the selected area. The green line shows the correlation peaks of the Hα proton at 5.37 ppm. Their through correlation peaks (not indicated) led us to tentatively assign the Hα proton to PheB24.

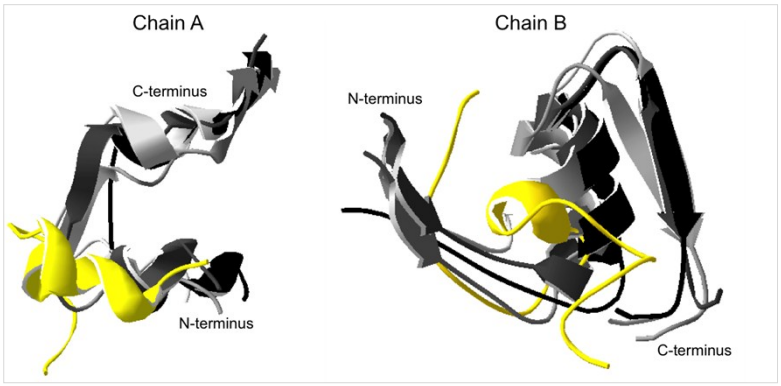

Figure S10: Overlaid structures from the dimer simulation as a function of temperature for chain A (*Left*) and chain B (*Right*). The second monomer is depicted. From black to light gray with decreasing brightness, the frames are at 20, 50, 65 and 75 °C. The frame at 110 °C is shown in yellow.

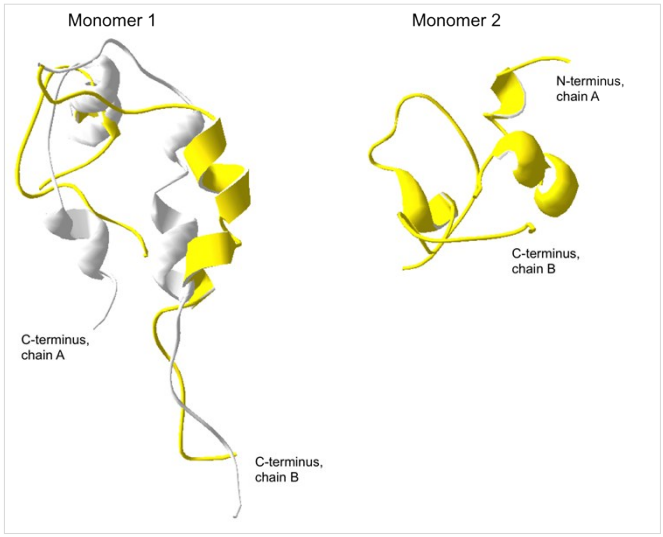

Figure S11: Structures from the monomer (white) and dimer (yellow) simulations at 110 °C. The dimer presents the tendency to dissociate to monomers 1 and 2, leading to two differently misfolded conformations. The structure of monomer 1 is similar to that of the monomer at the end of the simulation, and is believed to represent the completely unfolded conformation.

Figure S12: Overlaid structures from the hexamer simulation at 20 °C (black) and 110 °C (yellow), showing the weakened interactions between the antiparallel β-strands and the extended N-terminus of chain B.

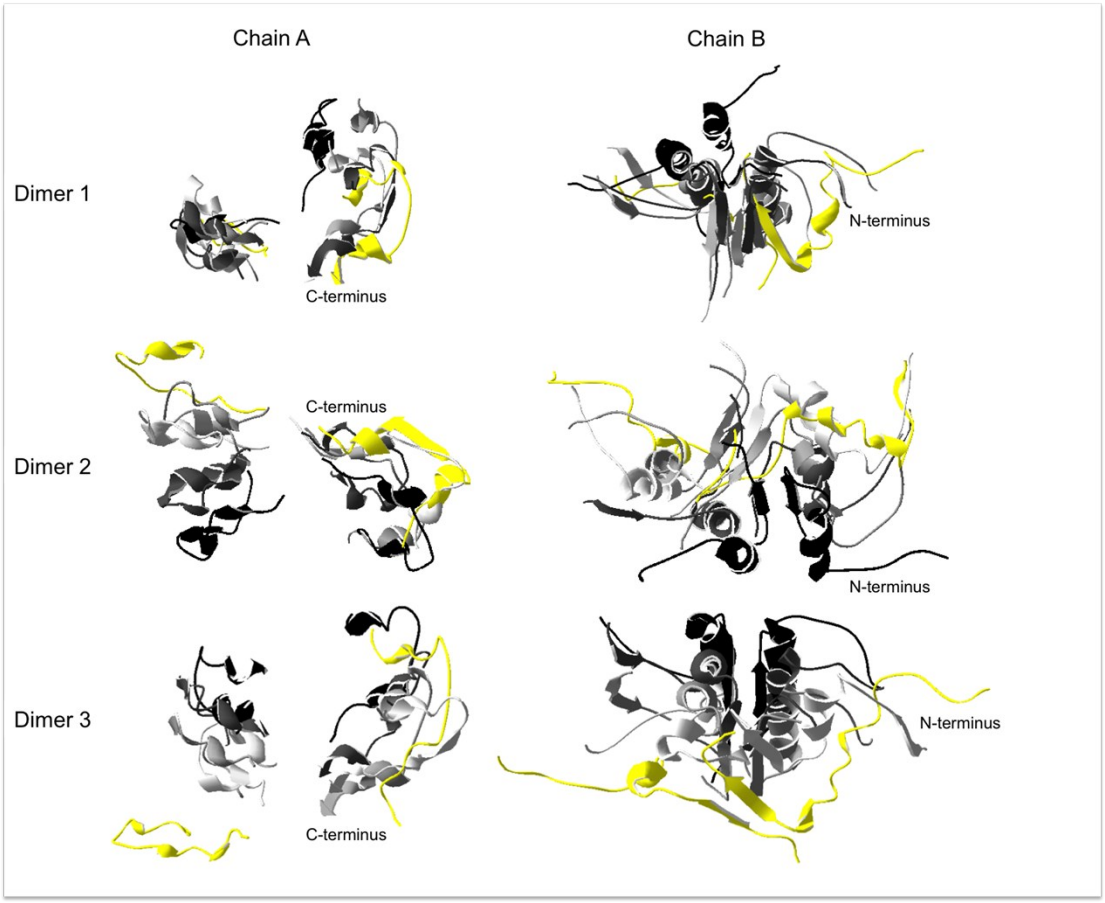

Figure S13: Structures from the hexamer simulation as a function of increasing temperature. Chains A (left) and B (right) are depicted separately for each one of the three dimers. From black to light grey with decreasing brightness, the frames are at 20, 50, 65 and 75 °C. The frame at 110 °C is shown in yellow. The three dimers present differences in their unfolding patterns, therefore contributing to the increased stability of the hexamer, which needs significantly longer time to dissociate and unfold compared to the dimer.

Video S2: Simulated annealing molecular dynamics (SA-MD) of insulin dimer.

Video S3: Simulated annealing molecular dynamics (SA-MD) of insulin hexamer.