Electronic Supplementary Information

Zinc-coordination and C-peptide complexation: a potential mechanism for the endogenous inhibition of IAPP aggregation

Xinwei Ge,[1] Aleksandr Kakinen,[2] Esteban N. Gurzov,[3,4] Wen Yang,[5] Lokman Pang,[3,4] Emily H. Pilkington,[2] Praveen Govindan-Nedumpully,[1] Pengyu Chen,[5] Frances Separovic,[6] Thomas P. Davis,[2,7] Pu Chun Ke,[2*] and Feng Ding[1*]*

[1]Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA
[2]ARC Centre of Excellence in Convergent Bio-Nano Science and Technology, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia
[3]St Vincent's Institute of Medical Research, 9 Princes Street, Fitzroy, VIC 3065, Australia
[4]Department of Medicine, St. Vincent's Hospital, University of Melbourne, Melbourne, Australia
[5]Department of Material Engineering, Auburn, AL 36849, USA
[6]School of Chemistry, Bio21 Institute, University of Melbourne, Melbourne, Victoria 3010, Australia
[7]Department of Chemistry, Warwick University, Gibbet Hill, Coventry, CV4 7AL, United Kingdom

Supplementary Materials and Methods

Materials

Human islet amyloid polypeptide (IAPP) (37-residue sequence: KCNTATCATQRLANFLVHSSNNFGAILSSTNVGSNTY; disulfide bridge: 2-7; MW: 3,906; purity: 98%) and C-peptide (31-residue sequence: EAEDLQVGQVELGGGPGAGSLQPLALEGSLQ; MW: 3,020; purity: 96%) were obtained as lyophilized powders from AnaSpec. Zinc chloride and ThT kit were acquired from Sigma-Aldrich. Milli-Q water was used for all dilutions.

ThT spectroscopic assay. To study IAPP amyloid fibrillation, IAPP (16.9 μM final) was mixed with ThT (16.9 μM final) and at varied molar ratios with C-peptide and/or Zn (0.3, 0.7, 1.3, 2.7 and 5.3 of IAPP to 1 of C-peptide and/or Zn). The IAPP-C-peptide-Zn molar ratio of 2.7:1:1 was slightly off the 2:1:1 ratio of interest due to sample weighing, but the range of chosen concentrations was deemed optimal based on the ThT and TEM observations. 100 μL of the each sample was added to a 96-well plate and changes in ThT fluorescence were recorded every 5 min over 14 h at a constant temperature of 25 °C using a FlexStation 3 Multi-Mode Microplate Reader (Molecular Devices; excitation: 440 nm, emission: 485 nm).

High-resolution transmission electron microscopy. IAPP fibrils were visualized by high-resolution TEM imaging. 5 μL of 24 h pre-incubated IAPP or IAPP with C-peptide and/or Zn at different molar ratios were pipetted onto glow discharged (15 s) 400 mesh copper grids (Formvar film, ProSciTech) and allowed for 60 s of adsorption. The final concentration of IAPP was 16.9 μM in all samples. Excess samples were then drawn off using filter paper and the grids were washed twice using 10 μL Milli-Q water, with excess drawn off. Grids were stained with 5 μL 1% uranyl

acetate for 30 s with excess stain drawn off and air-dried as needed. Imaging was performed on a Tecnai G^2 F20 Transmission Electron Microscope (FEI, Eindhoven, The Netherlands), operating at a voltage of 200 kV. Images were recorded using a Gatan UltraScan 1000 P 2k CCD camera (Gatan, California, USA) and Gatand Digital Micrograph 3.9.5 software.

Circular Dichroism Spectroscopy. To access the cooperative effect of Zinc and C-peptide with different molar ratios on the aggregation of IAPP, we collected the spectra of control protein samples and IAPP/C-peptide/Zn^{2+} mixtures using a J-810 circular dichroism spectrometer (JASCO) with a quartz cuvette of 1 mm path length over a wavelength range of 190-260 nm at room temperature (Fig. 2). Data were collected every 0.1 nm at a scanning speed of 50 nm/min with a bandwidth of 1 nm and averaged over eight measurements. The IAPP samples before and after 24 h incubation were measured as the control group and the mixture of IAPP-C-peptide and/or Zn^{2+} at different molar ratios were measured after 24 h incubation, respectively. The final spectra were baseline-corrected and the data were measured in mean residue ellipticity (θ) and converted to standard unit of deg·cm^2 $dmol^{-1}$ using equation $[\theta]=(\theta\times M_0)/(10000\times C_{soln}\times L)$, where M_0 is the mean residue molecular weight (3904.5g/mol for IAPP and 3020.5g/mol for C-peptide), C_{solnis} is the protein concentration (g/mL), and L is the path length through the buffer (cm). The percentages of the secondary structure content – i.e., α-helix, β-sheet, turn, and coil – were calculated by the non-negative least square (nnls) fitting method based on the Brahms&Brahms reference spectra.[32]

IAPP fibrils tracking and physical characterizations. Fibril tracking and analysis were performed with software FiberApp[28] to determine the morphology and mesoscopic parameters of persistence length (λ) and contour length (l) of IAPP fibrils in the presence and absence of C-peptide and/or Zn. The FiberApp open source code was developed based on statistical polymer physics and enables structural analysis of fiber-like, filamentous, and macromolecular objects. The persistence length λ reflects an intrinsic property of a polymer, denoting its rigidity and is mathematically defined via the bond correlation function (BCF) in 3D or 2D as the length over which angular correlations in the tangential direction decrease by a factor of e. Here the λ values of IAPP fibrils were estimated using the BCF, mean-squared end-to-end distance (MSED) and mean-squared midpoint displacement (MSMD) methods and presented as average values determined by the three methods[28]. The contour length corresponds to the end-to-end length of a polymer along its contour. The values of persistence length and contour length were obtained based on statistical analysis of 1,869 fibrils.

DMD simulations. DMD is a special type of molecular dynamics approach in which discrete step functions instead of continuous functions are used to mimic the constraints, and more details could be found elsewhere.[21] In DMD, we used an implicit solvent model to make the simulations faster. An adapted force filed from Medusa was used to model inter-atomic interactions, including van der Waals, solvation, hydrogen bonds, and electrostatic interactions. DMD also makes use of Debye—Hückel approximation to model screened electrostatic interactions between charged atoms, with the Debye length assigned ~10 Å and water electric susceptibility 80. The temperature was maintained in all simulations at 300 K using the Anderson's thermostat.[33] Each production run after relaxation lasted 50 ns.

Modeling zinc-coordination. The binding between zinc and histidine on IAPP, aspartic acid and glutamic acid on C-peptide was modeled by a number of constraints with a series of square well functions. The interaction parameters were obtained by statistical analysis of all available

structures containing Zn^{2+} in protein databank (PDB).[34] To capture the distance and angular dependence of zinc coordination, we computed the interatomic distances between zinc and all atoms in the associated chemical groups, including imidazole of histidine and carboxyl of aspartic (Asp) and glumatic (Glu) acids (Fig. S9). Since the adjacent carbons (i.e., the CB atom in aspartic acid and the CG atom in glutamic acid) were coplanar to carboxyl groups, we also included their distances to the coordinated zinc. Based on the histograms of interatomic distances, we determined the lower bound and upper bound values to assign the coordination interactions.

For histidine, zinc could bind either NE2 or ND1 atoms of the imidazole, giving two peaks. The zinc-NE2 binding was more common with a higher peak at the shorter distance. In addition, NE2 was un-protonated in our model. Therefore, we considered the binding of zinc with NE2 in our simulations. Since the coordination between zinc and imidazole was coplanar, we used additional constraints with square-well potential functions between zinc ion and all the other four imidazole atoms to ensure the planarity (Fig. S10A). The step parameters of these square-well functions were listed in Table S1. To determine the interaction strength, we varied ε_{HIS} from 0 to 40 kcal/mol with the increment every 4 kcal/mol. For each value assigned, the constraints were tested by 10 independent simulations under 300 K with each simulation performed for 50 ns. The number of contact between Zn^{2+} and targeted peptide was examined (Fig. S11A). We found that ε_{HIS} = 20 kcal/mol corresponds to the minimal value with which Zn^{2+} could fully stay in contact with histidine. Therefore, we assigned ε_{HIS} = 20 kcal/mol for zinc-histine coordination in our simulations.

Since the two oxygen atoms (i.e., OD1, OD2 of Asp and OE1, OE2 for Glu) in the carboxyl group was chemically identical and zinc could bind to both of them, there were two peaks in the histogram of zinc-oxygen distances (Fig. S9). For simplicity, we assumed the direct coordinate bond of zinc with OD1 of Asp and OE1 of Glu, whose bonding parameters were obtained by fitting the first peak with shorter distances in Fig. S5. Similarly to the zinc-imidazole coordination interactions, we assigned additional constraints between zinc and other three co-planar atoms (Fig. S6B). The interaction parameters between zinc and the other oxygen atoms (i.e., OD2 of Asp and OE2 of Glu) were obtained by fitting the second peak at larger distances in Fig. S5. In DMD simulations, we assigned the formation charge (-1e) to the centers of carboxyl group, i.e., the CG atom of Asp and the CD atom of Glu). We included in the interatomic interactions between zinc and the charge centers both the screened electrostatic and coordination interactions (Fig. S10B). The interaction parameters for zinc-carboxyl coordination are listed in Table S2. To determine the interaction strength, we also varied the values of $\varepsilon_{GLU/ASP}$ from 0 to 40 kcal/mol. With each value tested, we performed 10 independent simulations and examined the number of Asp/Glu residues for coordination bonds between Zn^{2+} and targeted peptide as well as the protein's secondary structure content (Fig. S11B). We found that $\varepsilon_{GLU/ASP}$ = 15 kcal/mol corresponded to the minimal value, where the secondary structure of C-peptide became stabilized and the number of coordinated Asp/Glu residues started to saturate. Interestingly, the maximum number of acidic residues forming the coordination bonds with zinc was 4 out of the total 5 Asp/Glu residues in the C-peptide, because the electrostatic repulsion was too strong for the last acidic residue to bind the zinc center.

Simulation setup. In all simulations we used the IAPP (PDB ID: 2L86) and C-peptide (PDB ID: 1T0C) structures obtained from PDB. To maintain the same IAPP peptide concentration, a cubic box with the dimensions of 100 Å, 126 Å and 144 Å was used to model 1, 2, and 3 IAPPs,

respectively. The periodic boundary conditions were used in all simulations. For simulations to obtain the binding kinetics (such as modeling the effect of C-peptide alone on IAPP aggregation, and the binding of zinc/C-peptide/IAPP at 1:1:1 molecular ratio) and the modeling of zinc-coordinated IAPP and C-peptide complexation, 100 independent simulations with different initial conditions were performed to acquire sufficient samplings. For the nucleation test, 10 independent simulations for each representative molecular complexes were performed in parallel. Counter ions including Cl^- and Na^+ were added to neutralize the systems if necessary. In addition, no coordination interaction was assigned between Zn^{2+} and the histidine of incoming IAPP in the nucleation tests for both the heterodimer and heterotrimer.

Reconstruction of the structural ensemble of zinc-coordinated complexes. From simulation trajectories, we selected all the complex structures coordinated by zinc, where zinc was bound to His18 of IAPP and at least one acidic residue of C-peptide. For each independent simulation (lasted 50 ns), 1,000 snapshot structures were sampled every 50 ps. We counted a coordination bond formation by the interatomic distance cutoff corresponding to the upper bound of the coordination interactions of zinc with NE1 of histidine, OD1 of aspartate, or OE1 for glutamate. The potential energy histogram of these zinc-coordinated heterodimer and heterotrimer featured a Gaussian-like distribution (Fig. 3A). For both the heterodimer and heterotrimer, we reconstructed their *low energy* structural ensembles by selecting those with their potential energies one standard deviation below the average. We obtained 17,237 low energy structures for the heterodimer and 33,182 for the heterotrimer.

Clustering Analysis. For the structural ensembles of both the heterodimer and heterotrimer, we performed the clustering analysis using a hierarchical clustering program, oc (www.compbio.dundee.ac.uk/downloads/oc). The clustering analysis was based on the calculation of root-mean-square-deviation (RMSD) between two structures. Given the two IAPP were indistinguishable in the heterotrimer, we performed structural alignment twice by alternating the index of the two IAPPs and used the small RMSD value for any two structures. For the clustering of the heterodimer, we set the cutoff value as 6 Å and all low energy structures were grouped into 10 larger clusters (shown in Fig. S7) and 256 smaller ones with one or two structures. We used the centroid node of the top 10 large clusters for the nucleation test. Given the larger heterotrimer, we found a cutoff of 8 Å was able to effectively distinguish different structures, ending with 7 clusters.

Compared to the heterodimer, the heterotrimer had significantly smaller structural variations. To further improve our structural sampling of the heterotrimer, we applied 7 replica exchange DMD simulations[30] with each of the centroid structure as the starting configuration. In all simulations, the coordination interactions were maintained. Based on simulation trajectories from all replica runs, we selected again the low energy states, with potential energy one standard deviation bellow the lowest Gaussian-like peak in the histogram. Similar clustering analysis resulted into only 3 clusters with similar over-all structural features as in Fig. 3C.

Cell culture and viability: The insulin producing NIT-1 cell line was cultured in DMEM (Invitrogen, UK) supplemented with 10% fetal calf serum. The percentage cell death of NIT-1 cells was determined by one observer, who was unaware of sample identity, for ≥2,000 cells in each experimental condition and four independent repeats of the assay. The observation was performed using an inverted fluorescence microscope after staining samples with the DNA dyes

Hoechst-33342 (10 μg/mL) and propidium iodide (5 μg/mL). [35] The cells were cultured in Opti-MEM (Life Technologies) for treatment with IAPP (16.9 μM final concentration), C-peptide (6.3 μM), Zn (6.3 μM) or their combinations. The cells were treated for 24 h before the viability assessment.

Statistical analysis: Viability data are represented as means ± SEM. Given the paired nature of the experimental design, comparisons between treated groups were made by analysis of variance (ANOVA) with Tukey correction. A p-value <0.05 was considered statistically significant.

Table S1. Zinc-imidazole binding parameters. The interaction parameters were obtained from statistical analysis of zinc-histidine binding in PDB (Fig. S6). The values of d_0 and d_1 were obtained by fitting the highest peaks in the histograms of interatomic distances in Fig. S6. The secondary peaks were simply due to the binding of zinc with ND1, which was not included in our model. The additional steps for zinc-NE1binding were assigned to capture their van der Waals attractions. The unit is in Å.

Atom pairs	d_0	d_1	d_2	d_3	d_4
Zinc-CG	4.12	4.28			
Zinc-ND1	4.08	4.30			
Zinc-CD2	2.98	3.16			
Zinc-CE1	2.96	3.16			
Zinc-NE1	1.99	2.21	2.65	3.31	4.19

Table S2. Zinc-carboxyl binding parameters. As for zinc-histidine binding, the interaction parameters were obtained from statistical analysis of zinc-Asp/Glu binding in PDB (Fig. S6). The values of d_0 and d_1 were obtained by fitting the highest peaks in the histograms of interatomic distances in Fig. S6. The unit is in Å.

Atom pairs (Asp/Glu)	d_0	d_1
Zinc-OD1/OE1	1.88	2.12
Zinc-OD2/OE2	2.56	3.20
Zinc-CG/CD	2.61	3.11
Zinc-CD/CG	4.05	4.37

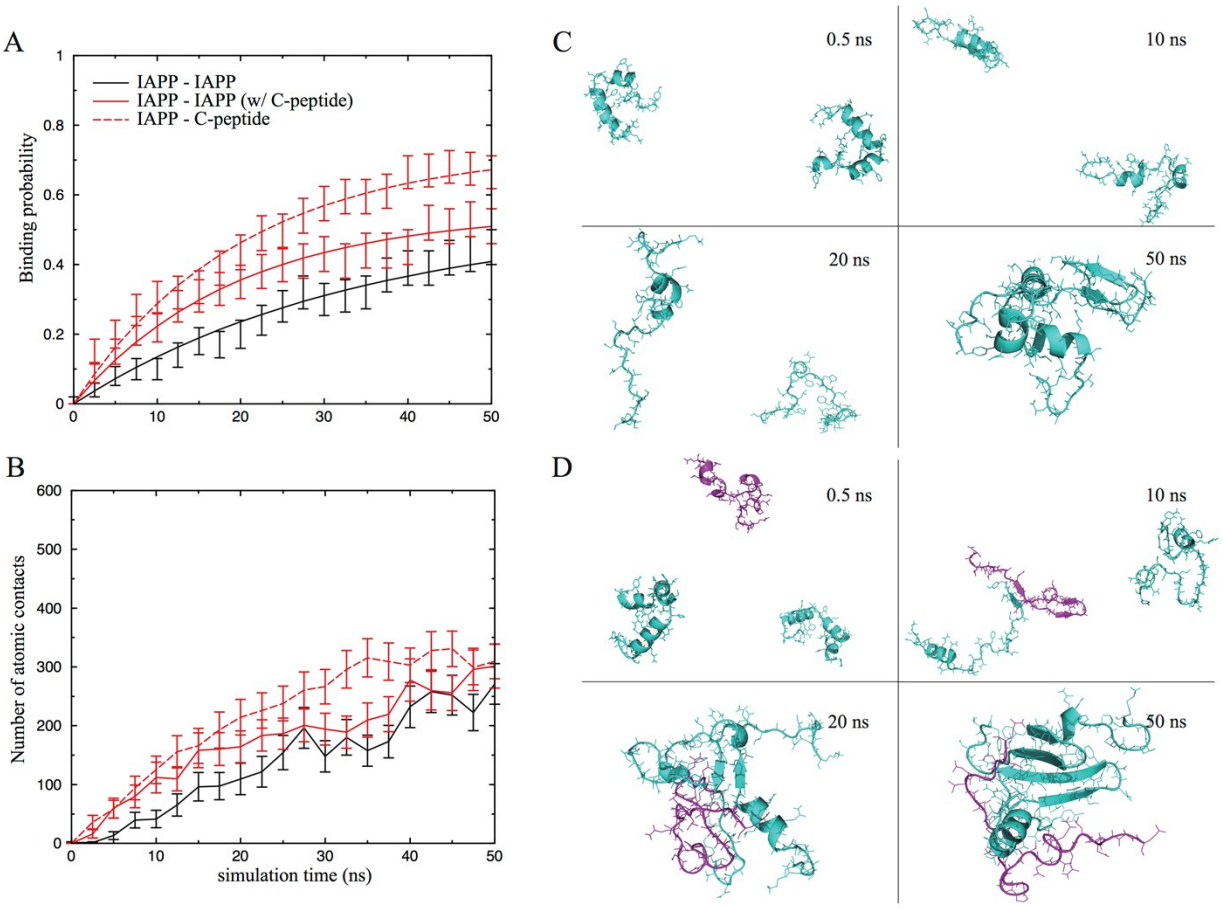

Figure S1. DMD simulations of IAPP dimerization with and without the presence of C-peptide. (A) Trajectories of binding probability and (B) trajectories of number of atomic contacts between different peptides, including IAPP-IAPP binding without the presence of C-peptide (solid black), and IAPP - IAPP binding with the presence of C-peptide (solid red), and IAPP - C-peptide binding (dashed red). The results were averaged over 100 independent simulations and the error bars were derived from standard errors. The trajectories of binding probability were also fitted exponentially. By binding both IAPPs with the C-peptide, the rate of self-association of IAPP dimers was increased from 0.029 ns^{-1} to 0.052 ns^{-1}. (C) Snapshot structures taken from a typical DMD simulations of IAPPs alone and (D) IAPPs with C-peptide. All peptides are shown in cartoon illustrations, while IAPPs are coloured in cyan and C-peptide coloured magenta. The formation of inter-peptide beta sheets indicates potential initiation of IAPP fibrillization.

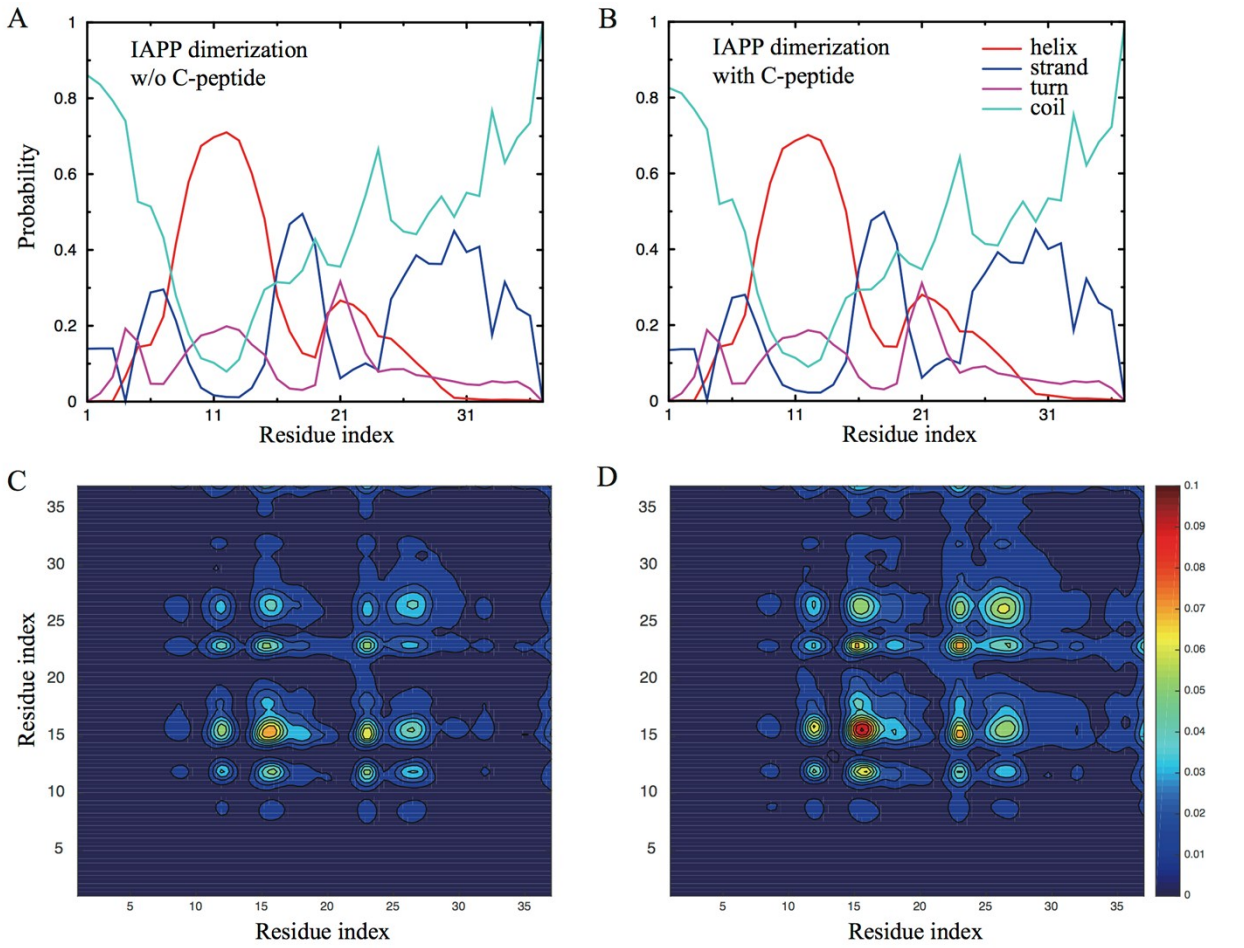

Figure S2. Secondary structure information and pair-wise contact frequency maps of IAPPs with and without C-peptide of 1:1 ratio. (A, B) For each residue on IAPP, its probability to be in a helix (red), strand (blue), turn (magenta), or coil (cyan) was calculated and plotted. The statistic work was over the second half of 100 independent simulations for both simulation systems. (C, D) Contour plots illustrating the pair-wise contact frequency information for IAPP – IAPP association, for both simulation systems respectively. The colour bar on right shows the scale and absolute values of contact frequency among residues on IAPPs. With or without C-peptide, we found residues Leu 12, Phe 15, Leu 16, and Phe 23 were highly engaged in IAPP self-association. The calculations were averaged for 100 independent simulations and 50 ns for each simulation, and data are illustrated as contour plots.

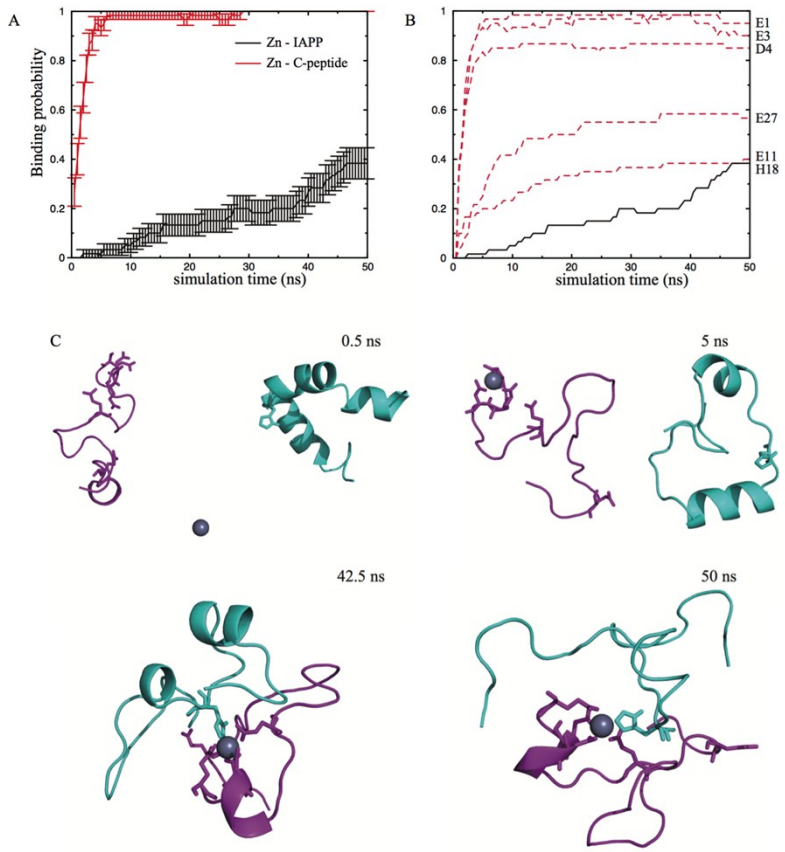

Figure S3. DMD simulations of the binding between Zn^{2+}, IAPP, and C-peptide at 1:1:1 molecular ratios. (A) Trajectories of the binding probabilities of Zn^{2+}-IAPP and Zn^{2+}-C-peptide as functions of simulation time, coloured as red and black, respectively. The data were averaged over 100 independent simulations and standard deviations were denoted as error bars. (B) Trajectories illustrating the averaged binding probabilities of Zn^{2+} to His18 (H18) on IAPP, as well as to Glu1 (E1), Glu3 (E3), Asp (D4), Glu11 (E11), and Glu27 (E27) on C-peptide. The binding with H18 is presented as the solid line, while interactions with C-peptide residues are presented as dashed lines. (C) Snapshot structures taken from a typical DMD simulation of Zn^{2+} coordinating IAPP and C-peptide. The peptides are shown in cartoon with IAPP coloured cyan and C-peptide magenta. The binding residues on both peptides are highlighted as sticks. The zinc ion is illustrated as a grey sphere. All the structures in this and following figures are prepared with PyMol (www.pymol.org).

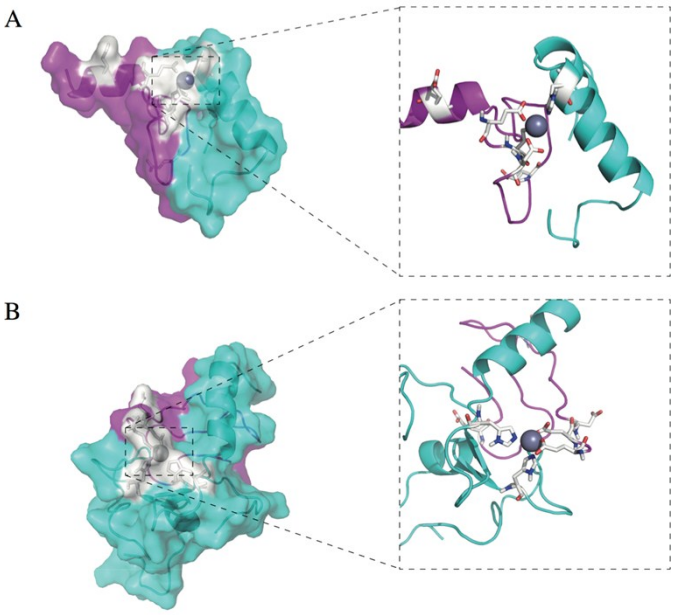

Figure S4. Surface structures of the zinc-coordinated IAPP and C-peptide heterodimer and heterotrimer. (A) The zinc ion inside the heterodimer is partially solvent exposed, and can possibly bind more incoming IAPP. (B) The zinc ion inside the heterotrimer is fully buried, preventing further coordination with IAPP. In both panels, IAPPs (and surfaces) are shown in cartoon representation and coloured in cyan, C-peptide (and surfaces) also shown in cartoon and coloured in magenta, and zinc ion as a grey sphere. All binding residues (and solvent accessible surfaces) are highlighted in enlargements, with stick illustrations and coloured in white.

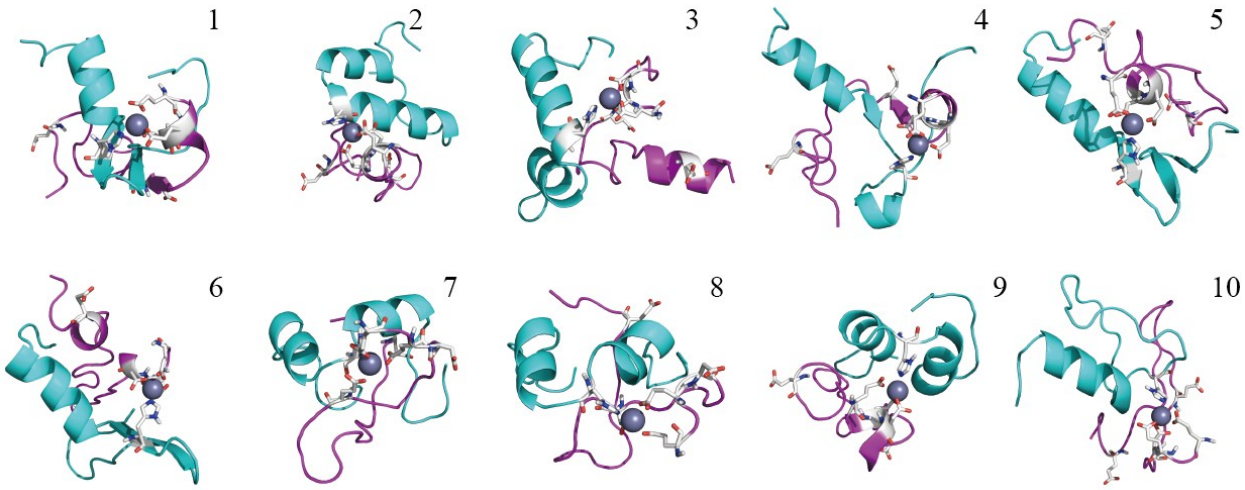

Figure S5. The centroid heterodimer structures of the top 10 clusters, with relative percentage of 1.57%, 2.36%, 1.57%, 1.57%, 1.77%, 3.74%, 2.16%, 1.97%, 2.56% and 1.18% of total structural ensemble respectively. IAPPs are shown in cartoon illustrations and coloured in cyan, C-peptide also shown in cartoon illustrations and coloured in magenta, and zinc ion as a grey sphere. All binding residues (and surfaces) are highlighted in enlargements, with stick illustrations and coloured in white.

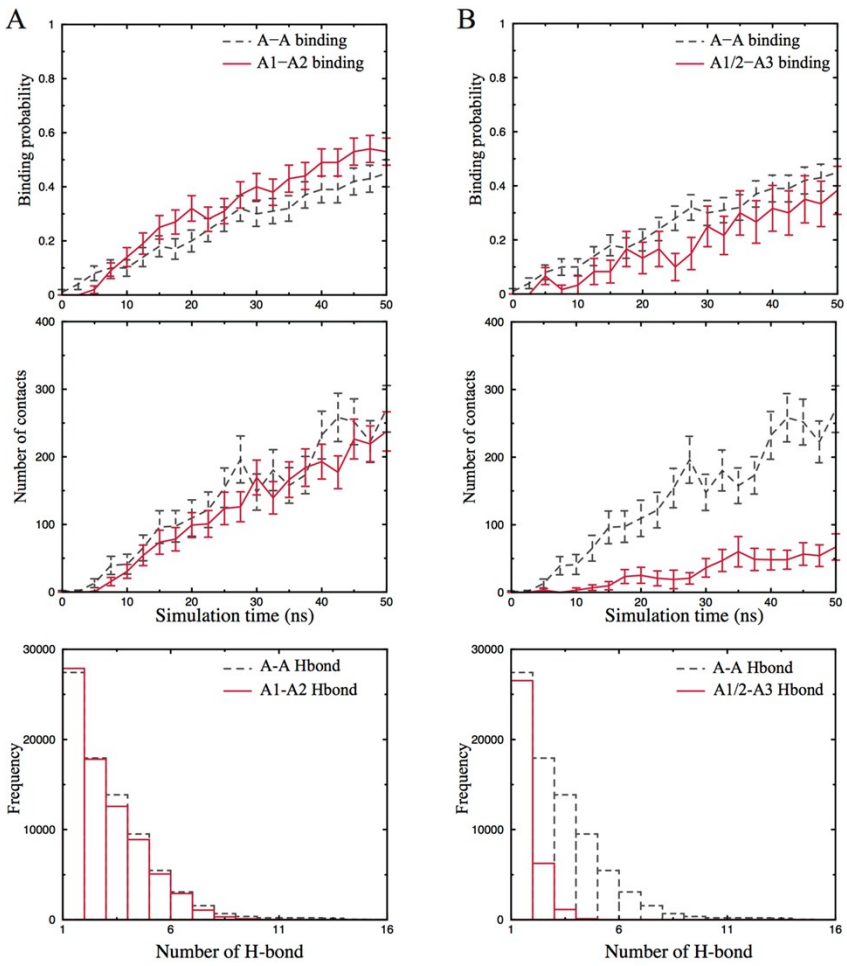

Figure 6. Nucleation test of zinc-coordinated heterodimer and heterotrimer. (A) The binding of IAPP in the heterodimer with an incoming IAPP are estimated in terms of IAPP-IAPP binding probability, number of atomic contact, and inter-chain hydrogen bond. Compared to the dimer simulations of two IAPPs alone (A-A binding, coloured in grey), the IAPP in the heterodimer (A1) could still bind the incoming IAPP (A2) and form extensive contacts and hydrogen bonds. (B) The IAPPs in the heterotrimer (A1 and A2) did not form many contacts and hydrogen bonds with the incoming IAPP (A3), even though they may be associated with each other via dynamic binding. Data were averaged for both IAPPs in the heterotrimer.

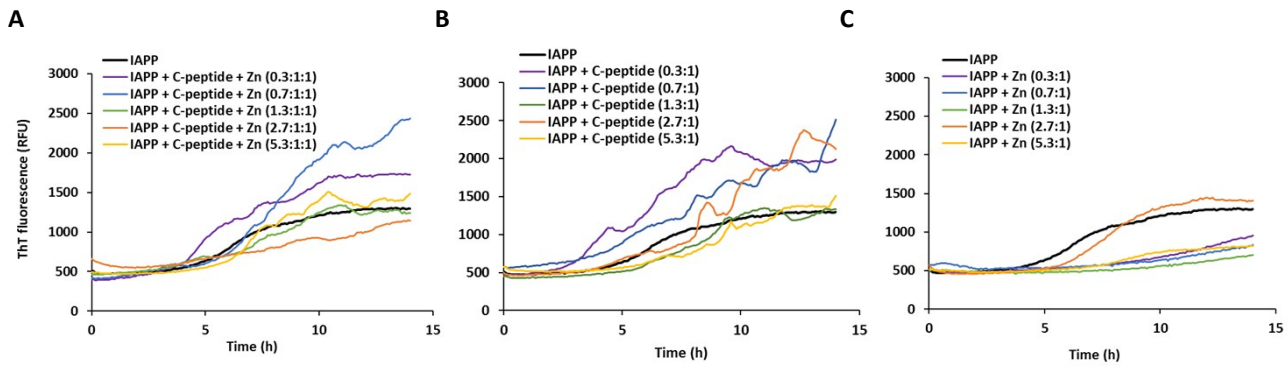

Figure S7. ThT fluorescence assay data on zinc-coordinated IAPP-C-peptide complexation for 14 h of sample incubation. (A) ThT fluorescence assay of IAPP with C-peptide and zinc of various molar ratios. (B) ThT fluorescence assay of IAPP with C-peptide of various molar ratios. (C) ThT fluorescence assay of IAPP with zinc of various molar ratios. The final concentration of IAPP was 16.9 μM in all samples.

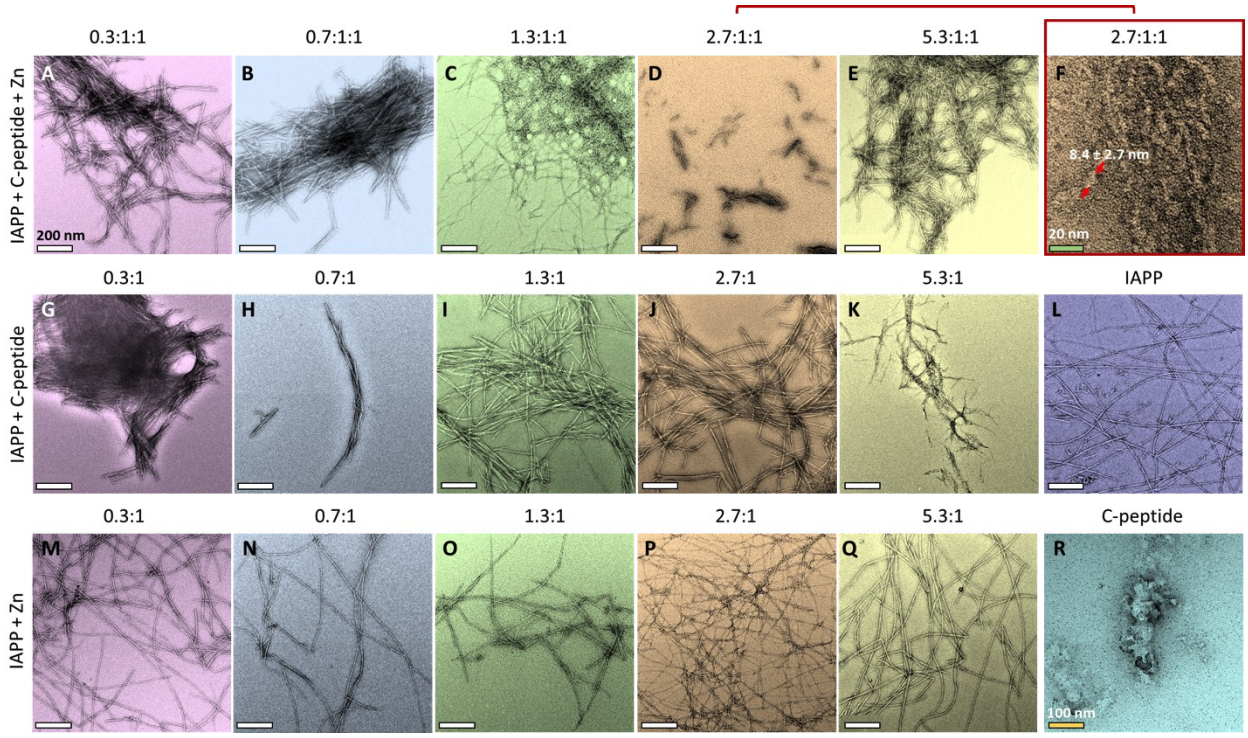

Figure S8. High-resolution TEM imaging of zinc-coordinated IAPP-C-peptide complexation for 24 h of sample incubation. A-E: IAPP with C-peptide and zinc of various molar ratios. F: magnified view of an aggregate in D. G-K: IAPP with C-peptide of various molar ratios. L: IAPP control. M-Q: IAPP with zinc of various molar ratios. R: C-peptide control. The final concentration of IAPP was 16.9 μM in all samples. Scale bars: 200 nm unless specified otherwise.

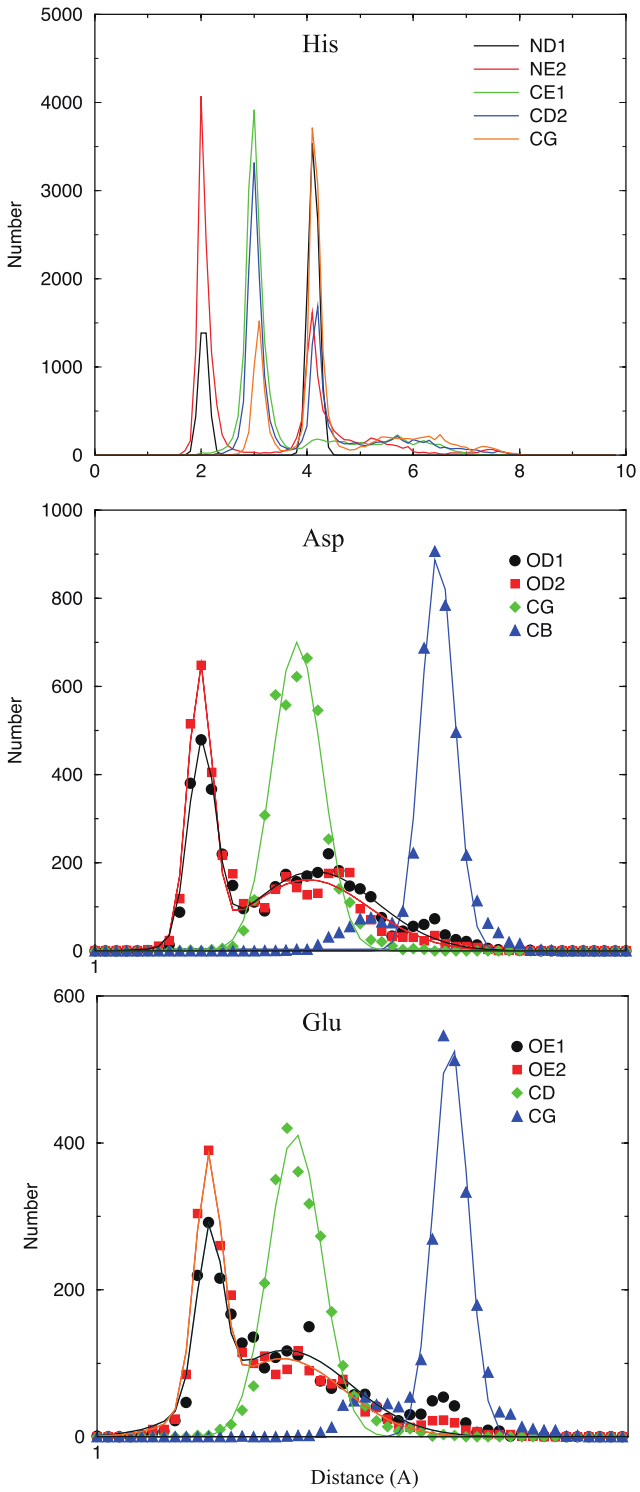

Figure S9. The distributions of inter-atomic distances for coordination between Zn^{2+} and histidine (His), aspartic acid (Asp), and glutamic (Glu) acid respectively. Five atoms of the histidine imidazole (CG, ND1, CD2, CD2 and NE2), our atoms of aspartic (CG, OD1, OD2 of the carboxyl group and the attached CB) and glumatic (CD, OE1, OE2 of the carboxyl group and the attached CG) acids were included for the calculation.

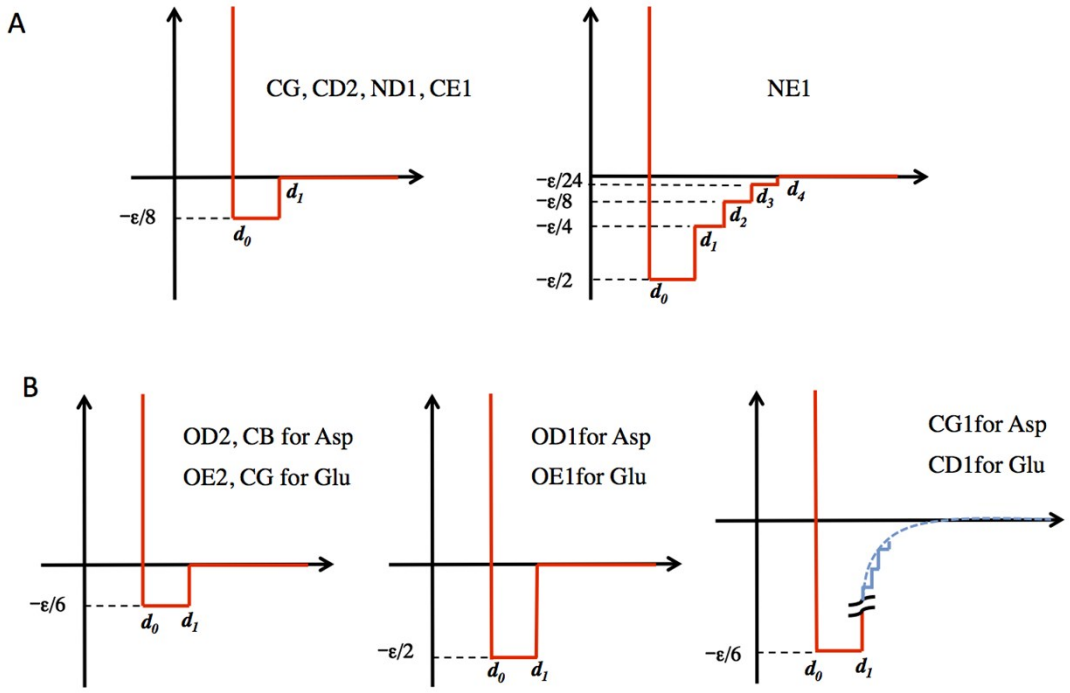

Figure S10. The schematic potential functions used to model the coordination interactions of zinc with (A) imidazole of histidine and (B) carboxyl of aspartate and glutamate. The parameters d_0 and d_1, corresponding to the lower and upper bounds of interatomic distances in the zinc-bound conformation, were obtained from Gaussian-fit of peaks in Fig. S5. The values for zinc- imidazole and zinc-carboxyl were listed in Tables S1 and S2, respectively. The additional interaction steps between zinc and NE1 were assigned to capture the corresponding van der Waals attractions. Similarly, the additional potential steps between zinc and CG1 of Asp and between zinc and CD1 of Glu correspond to the electrostatic interactions, because CG1 of Asp and CD1 of Glu were assumed as charge centers of their carboxyl groups.

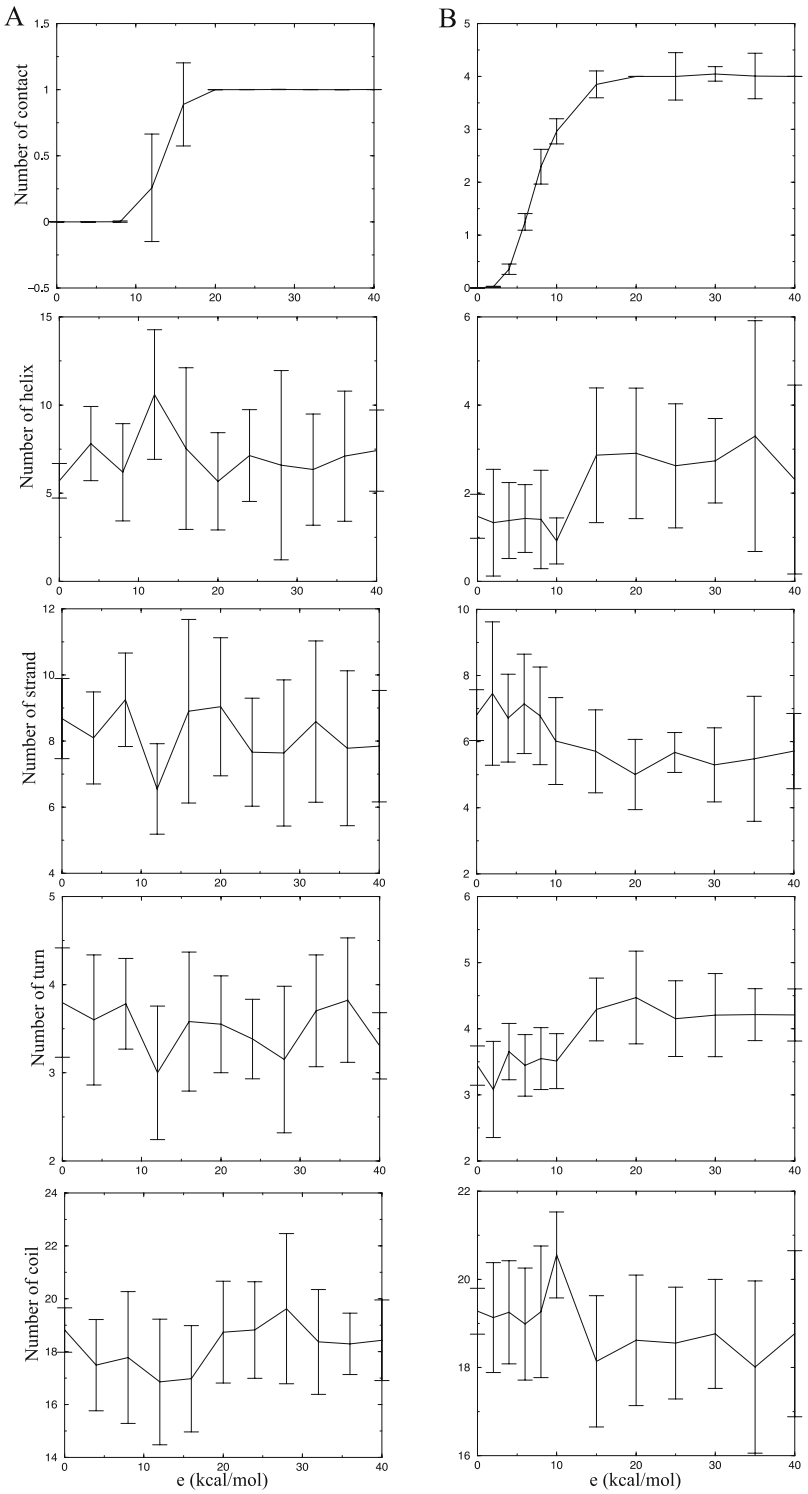

Figure S11. Simulations to determine bond strengths. Parameters including number of zinc-coordinated residues and various secondary structure contents (helix, strand, turn and coil) for (A) the coordination between Zn^{2+} and IAPP and (B) the coordination between Zn^{2+} and C-peptide. Error bars were derived from 10 independent simulations for each tested value of ε.