DTU Library

Metal ion controlled self-assembly of a chemically reengineered protein drug studied by small-angle X-ray scattering

Jesper, Nygaard; Munch, Henrik K. ; Thulstrup, Peter W. ; Christensen, Niels Johan; Hoeg-Jensen, Thomas; Jensen, Knud J. ; Arleth, Lise

[Link back to DTU Orbit]

Metal Ion Controlled Self-Assembly of a Chemically Reengineered Protein Drug Studied by Small-Angle X-ray Scattering

Jesper Nygaard,[†,‡] Henrik K. Munch,[§] Peter W. Thulstrup,[§] Niels J. Christensen,[∥] Thomas Hoeg-Jensen,[⊥] Knud J. Jensen,[§] and Lise Arleth*[†]

[†]Niels Bohr Institute, Faculty of Science, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark

[‡]MAX-lab, Lund University, Ole Römers väg 1, SE-223 63 Lund, Sweden

[§]Department of Chemistry, Faculty of Science, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark

[∥]DTU Chemistry, Department of Chemistry, Technical University of Denmark, Kemitorvet, Building 206, 2800 Kgs. Lyngby, Denmark

[⊥]Diabetes Protein and Peptide Chemistry, Novo Nordisk Park D6.1.152, 2760 Maaloev, Denmark

(S) Supporting Information

ABSTRACT: Precise control of the oligomeric state of proteins is of central importance for biological function and for the properties of biopharmaceutical drugs. Here, the self-assembly of 2,2′-bipyridine conjugated monomeric insulin analogues, induced through coordination to divalent metal ions, was studied. This protein drug system was designed to form non-native homo-oligomers through selective coordination of two divalent metal ions, Fe(II) and Zn(II), respectively. The insulin type chosen for this study is a variant designed for a reduced tendency toward native dimer formation at physiological concentrations. A small-angle X-ray scattering analysis of the bipyridine-modified insulin system confirmed an organization into a novel well-ordered structure based on insulin trimers, as induced by the addition of Fe(II). In contrast, unmodified monomeric insulin formed larger and more randomly structured assemblies upon addition of Fe(II). The addition of Zn(II), on the other hand, led to the formation of small quantities of insulin hexamers for both the bipyridine-modified and the unmodified monomeric insulin. Interestingly, the location of the bipyridine-modification significantly affects the tendency to hexamer formation as compared to the unmodified insulin. Our study shows how combining a structural study and chemical design can be used to obtain molecular understanding and control of the self-assembly of a protein drug. This knowledge may eventually be employed to develop an optimized *in vivo* drug release profile.
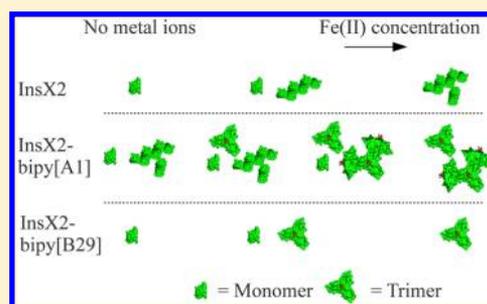
INTRODUCTION

Protein-based drugs constitute a growing share of the pharmaceutical market. Unfortunately, proteins are generally more prone to degradation, as compared to most small-molecule drugs.[1] However, the properties of proteins can be modulated by the addition of appendages, such as lipid chains and PEG.[2] A central research and development activity associated with protein drug development concentrates on optimizing the protein drug release profile following injection. Discovering new approaches to control the self-assembly of proteins, for example, to modulate their function or to prolong their *in vivo* half-life, has extensive research focus, both in academia and in the pharmaceutical industry.[3] The present study investigates different strategies to obtain better control of the nanoscale structural properties of protein-based drugs using insulin as the model protein.

Insulin is a small protein of ∼5.7 kDa. In its native form, it has a preference for forming dimers through interactions between hydrophobic surfaces. These dimers are the foundation for the insulin hexamer that is built up from three dimers. In the hexamer, the three dimers are held together by hydrophobic surface contacts and by the coordination around two Zn(II)-

ions.[4,5] Insulin is stored as hexamers in the pancreas and is released into the bloodstream where its monomeric form is part of the regulation of the glucose level in the blood.[6] Diabetes, both type 1 and 2, is associated with faults in the regulation of or sensitivity to insulin levels. Current estimates are that 346 million people worldwide suffer from diabetes and that in 2004 3.4 million died due to effects related to high blood sugar level. The number of diabetic patients is expected to double during the next 20 years.[7]

In diabetes treatment, administration of insulin is often required, and different long-acting insulin drugs have been developed to deliver the constant insulin level present in healthy individuals and without which the diabetic patient is at risk of hyperglycemia.[8] In addition, several fast-acting insulin analogues have been developed to establish the insulin peaks required shortly after the intake of a meal.[9,10] These fast-acting insulin analogues generally have a tendency to resist oligomerization as
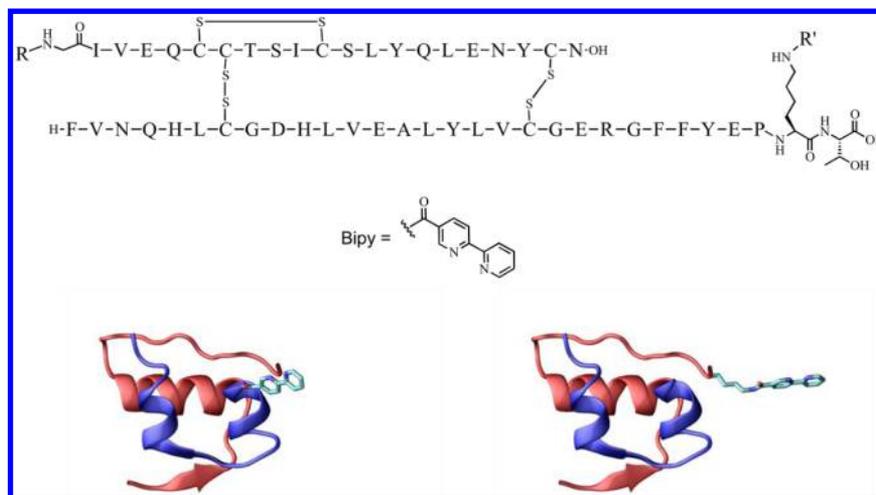
Figure 1. Top: Primary structure of bipy and the InsX2 analogues. InsX2: R,R′ = H; InsX2-bipy[A1]: R = bipy, R′ = H; InsX2-bipy[B29]: R = H, R′ = bipy. Bottom left: Tertiary structure of InsX2-bipy[A1]. Bottom right: Tertiary structure of InsX2-bipy[B29]. The tertiary structures are modified from pdb structure 2JV1; not energetically optimized.

compared to the long-acting analogues. In the present study, we use one of these monomeric insulin types as a model protein. We have selected the insulin analogue InsX2, which is produced at Novo Nordisk A/S for research purposes.[10] Compared to human insulin, InsX2 is modified at positions 9 and 27 on the B chain where substitutions have been made to, respectively, Asp and Glu (B9D, B27E; Chain, Position, Substituted with). Both substitutions are made on the hydrophobic surface involved in the formation of the insulin dimer. The alterations decrease the propensity for dimerization and consequently also for hexamer formation, resulting in a more monomeric insulin variant.[10]

It has previously been shown that 2,2′-bipyridine (bipy) linked to GalNAc to form bipy-GalNAc can result in bipy-GalNAc trimers by ligand binding between bipy and Fe(II),[11] relying on the strong propensity of Fe(II) for forming a *tris* complex with the bidentate bipy ligand. Recently, we demonstrated that bipy modified InsX2 variants would form first published insulin trimer upon the addition of Fe(II), as studied by chromatographic and optical spectroscopic methods. The circular dichroism (CD) data clearly proved formation of trimers, proving that the geometry of the artificial metal ion site can drive protein oligomerization, interestingly with orientations which depended on the local chiral environment provided by B29 vs A1 positions of InsX2.[12] However, the study was limited, as the formation of higher-order assemblies could not be studied by these methods.

Here, we report the structure and nanoscale properties of these two InsX2-bipy variants upon the addition of Fe(II) or Zn(II) ions, respectively, as studied by small angle X-ray scattering (SAXS). The present SAXS analysis does confirm that trimers are formed from both InsX2-bipy variants upon addition of Fe(II). However, we find here that the monomer–trimer equilibrium distribution is significantly affected by the bipy binding site and that the InsX2-bipy[B29] system requires a higher excess of Fe(II) to form trimers than the InsX2-bipy[A1]. We also find that these trimers do not appear when adding Zn(II) instead of Fe(II). Instead, upon addition of Zn(II), the modified insulin self-organizes into the native hexamer structure, however, to a lesser extent than human insulin. Again, the equilibrium is significantly shifted toward monomers for the InsX2-bipy[B29] system. Finally, we observe that the bipy binding site has a significant impact on the formation of larger molecular assemblies. In both the original InsX2 and the InsX2-bipy[A1]

systems, addition of excess concentrations of Fe(II) and Zn(II), respectively, leads to the formation of larger fractal-like molecular assemblies based on monomers, trimers, or hexamers. However, in the case of InsX2-bipy[B29] only trimers or hexamers are observed depending on whether Fe(II) and Zn(II) is added. The larger fractal-like assemblies of these are not observed. This suggests that the functionalization with bipy at position [B29] inhibits the formation of larger molecular assemblies and in this way stabilizes the smaller and more well-ordered assemblies.

■ EXPERIMENTAL METHODS

Samples. Three insulin analogues were examined in this work: InsX2, which is the unmodified monomeric insulin analogue, was kindly provided by Novo Nordisk A/S; InsX2-bipy[B29], which is InsX2 modified with 2,2′-bipyridine at position B29; and InsX2-bipy[A1], which is InsX2 modified with 2,2′-bipyridine attached at position A1. InsX2-bipy[A1] and InsX2-bipy[B29] were synthesized as previously described.[12] The primary and tertiary structures of InsX2-bipy[B29] and InsX2-bipy[A1] are illustrated in Figure 1. The systems were formulated with and without Fe(II) or Zn(II) metal ions at different molar ratios of metal ion to InsX2.

For the SAXS measurements, InsX2, InsX2-bipy[A1], or InsX2-bipy[B29] was dissolved in Milli-Q water, and the pH was lowered to between pH 2 and pH 3 by addition of HCl. The protein concentration of the stock solution was monitored with a Nanodrop 1000 spectrophotometer from Thermo Scientific.

Samples with No Added Metal Ions. InsX2, InsX2-bipy[B29], and InsX2-bipy[A1] without Fe(II) or Zn(II) were used as reference systems for the analysis of the systems formulated *with* metal ions. For the preparation of the metal ion-free samples, the pH was raised, from pH 2–3, to pH 7.5 and a concentrated HEPES buffer at pH 7.5 was added to obtain a final concentration of 10 mM Hepes. The samples were adjusted with Milli-Q water, and the following final concentrations were obtained: InsX2, 2.4 mg/mL (422 μM); InsX2-bipy[B29], 3.0 mg/mL (608 μM); and InsX2-bipy[A1], 2.7 mg/mL (542 μM).

Samples with Fe(II). To prepare the Fe(II)-containing samples, different amounts of Fe(Cl)$_2$, dissolved in 0.1 M HCl, to prevent oxidation, were added to vials of concentrated insulin solutions while these were still at pH 2–3. The pH was then raised to pH 7.5, HEPES buffer at pH 7.5 was added, and the volume was adjusted with Milli-Q water to obtain a buffer concentration of 10 mM HEPES and the same insulin concentrations as for the metal ion-free samples. The prepared molar stoichiometries of Fe(II) to the various InsX2-analogues, [Fe(II)]/[InsX2], are listed in Table 1. The solid Fe(Cl)$_2$ was kept dry and unoxidized through storage in a desiccator, and no sign of

oxidation of Fe(II) was observed in the solutions with the applied procedures.

Table 1. Fe(II) and Zn(II) Series Measured on the InsX2 Analogues

sample	[Fe(II)]/[InsX2]				
InsX2	0	0.13	0.25	0.51	0.75
InsX2-bipy[B29]	0	0.1	0.2	0.3	0.7
InsX2-bipy[A1]	0	0.1		0.3	0.7
sample	[Zn(II)]/[InsX2]				
InsX2	0	0.2	0.3	0.6	0.09
InsX2-bipy[B29]	0	0.1	0.3	0.7	
InsX2-bipy[A1]	0	0.1	0.3	0.7	

Samples with Zn(II). The samples with Zn(II) were prepared using the same procedure as described for Fe(II) except that $Zn(OAc)_2$ was added at low pH instead of $Fe(Cl)_2$. The pH was then raised to pH 7.5, the Hepes buffer was added, and the volume was adjusted with Milli-Q water to obtain a final Hepes concentration of 10 mM. The prepared molar stoichiometries of Zn(II) to the various InsX2-analogues, [Zn(II)]/[InsX2], are listed in Table 1.

After the preparation, all samples were set to equilibrate for a minimum of 8 h at 5 °C.

Small Angle X-ray Scattering Experiments. The SAXS measurements were performed partly at the SAXS beamline I711 at MAX-lab, Lund, Sweden,[13] and partly at the ID14−3 beamline at ESRF, Grenoble, France.[14] Buffer backgrounds were measured before and after each sample, and the averaged background was subtracted from the sample measurement. For the data sets from I711, both the initial data reduction including azimuthal averaging, q-calibration, absolute scattering intensity calibration, and background subtraction were performed by use of the *BioXTAS RAW* software.[15] For the data sets from ID14−3, the data reduction was performed by means of the fully automatized version of the *ATSAS* package.[16] As usual, the scattering data are represented as a function of the momentum transfer $q = 4\pi \sin \theta / \lambda$, where λ is the X-ray wavelength (1.1 Å at the I711 setup used and 0.931 Å at ID14−3) and θ is the half scattering angle. By using the scattering from H_2O and bovine serum albumin, respectively, as external standards, the scattering intensities, $I(q)$, could be obtained in terms of the absolute scattering cross sections per unit volume in units of cm^2/cm^3.

CD and Absorbance Spectrum Measurements. The CD measurements were performed on a Jasco J-815 CD spectrometer with a Jasco CDF-426 S/15 Peltier-element temperature controller, attached to a Julabo water bath. Corresponding UV−vis absorbance measurements were carried out on a Shimadzu UV2401PC instrument. For the optical spectroscopy, the sample temperature was maintained at 25 °C, and Hellma quartz cuvettes of either 10 mm or 1 mm were used. Blank spectra of reference solutions were recorded and subsequently subtracted, and finally, the recorded data were converted to molar units.

■ THEORY AND MATHEMATICAL MODELING

The small-angle scattering intensity from a general sample of particles in solution can be described by the product of a form factor, $P(q)$, describing the *intra*particle scattering, a structure factor, $S(q)$, describing the *inter*particle interactions and a prefactor, c, which may be calculated by the product of the concentration of the particles, their molar mass and the second power of their excess scattering length per unit mass[17]

$$I(q) = c \cdot P(q) \cdot S(q) \tag{1}$$

The samples investigated in the present project consisted of a mixture of different oligomeric forms of InsX2-bipy. The overall sample concentration was low and kept at ∼2.5 mg/mL, implying that the structure factor, $S(q)$, could be assumed to be

unity. Consequently, the following general expression for the small-angle scattering could be used:

$$I(q) = c \sum_i X_i \cdot P_i(q) \tag{2}$$

where X_i denotes the weight fraction of the proteins that are in a given oligomeric form, and $P_i(q)$ describes the corresponding form factor. In order to ensure mass-conservation, $\sum_i X_i = 1$. Oligomeric particles, as those in focus in the present work, are composed of small subunits, and the form factor $P(q)$ may be decomposed into the scattering from the single subunit, $p(q)$, and a function $f(q)$ describing the effect of the oligomeric organization on the scattering:

$$P(q) = p(q) \cdot f(q) \tag{3}$$

The function $f(q)$ may be a function describing the effect of organizing the monomeric subunit in a well-defined oligomer such as, e.g., a dimer or a hexamer. Alternatively, it may be a function that describes a less-ordered structure, such as, e.g., the structure of protein subunits aggregated into a fractal-like structure.

Well-Ordered Oligomers. In the present work, it was relevant to model the ordered structures of insulin monomers, trimers, and hexamers, as well as fractal-like structures composed of these subunits. Here, we here take advantage of the fact that most of the ordered oligomeric subunits can be found in the protein data bank, thus the $f(q)$'s corresponding to the ordered oligomeric substructures were, in fact, not explicitly calculated. Instead, $P(q)$ values for the entire insulin monomer, trimer, or hexamer were calculated using the program *CRYSOL*[18] combined with information about the insulin crystal structure from the protein data bank. For monomeric insulin, the PDB structure 3I3Z.pdb was used.[19] The hexamer structures were generated from 1EV3.pdb by use of *PYMOL*.[20] Putative 3D structure models for the previously unknown quaternary trimer structure for InsX2-bipy[B29] trimer and InsX2-bipy[A1] trimer were generated *in silico*. To this end, the NMR solution structure of the human insulin monomer[21] (PDB ID: 2JV1) was obtained from the protein data bank,[22] and the first structure of the ensemble was extracted for further work. The B9D and B27E mutations characteristic of InsX2 were introduced into the NMR structure with the mutation functionality in *Maestro*.[23] Subsequently, the bipyridine moiety was introduced at position B29 or A1, yielding the InsX2-bipy[B29] monomer or InsX2-bipy[A1] monomer, respectively. To generate the corresponding trimers, the superposition function in Maestro was employed to overlay the bipyridine moieties of three duplicates of the bipyridine functionalized InsX2 monomer with the bipyridine moieties in the crystal structure of a tris-(2,2′-bipyridine) cobalt(III) complex.[24] After substituting Fe(II) for Co(III), the resulting trimer models were subjected to energy minimization in *MacroModel*[25] using the OPLS_2005[26] force field and relaxed convergence criteria (gradient <1.0 kJ Å^{-1} mol^{-1}). This allowed the removal of steric clashes while largely conserving the native insulin geometry. Relevant PDB files were generated by means of the software *Visual Molecular Dynamics* (VMD).[27] In the process of determining the trimer structure, different geometry optimization criteria were used and outputs were subsequently converted to reciprocal space using *CRYSOL* and compared to the experimental SAXS data. The structure that gave the best fit to the experimental data for the trimer was used for the subsequent SAXS analysis. The applied form factor scattering
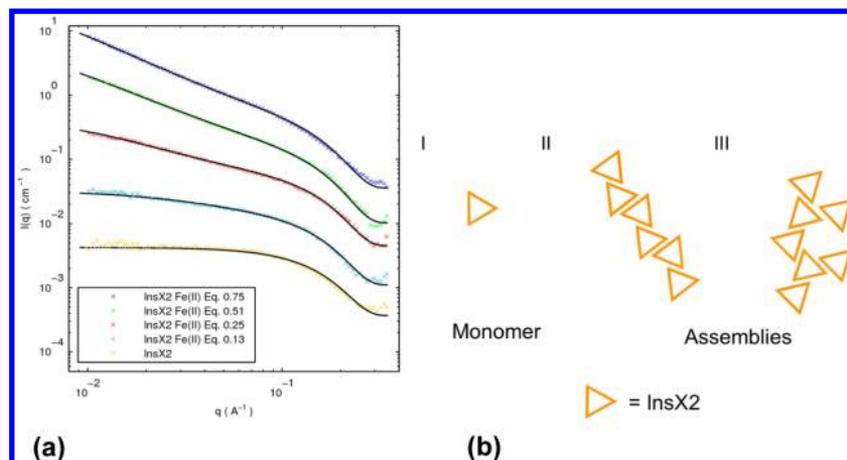
Figure 2. (a) SAXS data (colored points) and model fits (full black lines). The InsX2 data (yellow points) are on absolute scale. The above data are rescaled by 3^n, where n runs from 1 to 4. (b) (I) monomeric insulins. (II) Insulin assemblies built from monomeric insulin subunits with a fractal dimension close to 1. (III) Branched insulin assemblies built from monomeric insulin subunits.

intensities for the insulin monomer, trimer, and hexamer are plotted in Figure S1 of Supporting Information.

Fractal-Like Oligomers. In the case of the larger oligomeric structures also observed in this project, the fractal model developed by Teixeira[28] was used. The model (eq 4) describes the scattering of aggregates from a subunit characterized by an interaction radius r_0, fractal dimension D, and correlation length ξ:

$$f_{\text{frac}}(q) = 1 + \frac{1}{(q \cdot r_0)^D} \frac{D \cdot \Gamma(D-1)}{\left[1 + 1/(q^2 \cdot \xi^2)\right]^{(D-1)/2}}$$
$$\cdot \sin\left[(D-1) \cdot \tan^{-1}(q \cdot \xi)\right] \tag{4}$$

In eq 4, $\Gamma(x)$ denotes the gamma function. In the present work, the fractal model is combined with a numerical form factor for an insulin monomer, hexamer, or trimer using a previously developed strategy.[29,30] By combining the form factor of the insulin subunit with the fractal model through the so-called decoupling approximation,[31] it is taken into account that the insulin subunit is not centrosymmetric. The fractal model in combination with the form factor for subunit i is denoted $F_{\text{frac}}[q, P_i(q)]$. The value for r_0 was determined from the half D_{max} values based on the PDB structures and was fixed to $r_{0 \text{ monomer}} = 19$ Å, $r_{0 \text{ trimer}} = 24$ Å, and $r_{0 \text{ hexamer}} = 24$ Å. Examples of the fractal-like structured assemblies of the insulin monomer, trimer and hexamer are plotted in figure S1 of Supporting Information.

Mixtures of Different Oligomers. A scenario where there is equilibrium between different oligomeric states is readily described by combination of eqs 2 and 3:

$$I(q) = c \sum_i X_i \cdot p_i(q) \cdot f_i(q) \tag{5}$$

In the case of a sample composed of a mixture of, e.g., monomers, trimers, and fractals composed from monomers eq 5 becomes

$$I(q) = c\{X_{\text{mon}} P_{\text{mon}}(q) + X_{\text{tri}} P_{\text{tri}}(q) + X_{\text{SA.mon}} F_{\text{frac}}$$
$$[q, P_{\text{mon}}(q)]\} \tag{6}$$

where $f_{\text{mon}}(q) = 1$ such that $p_{\text{mon}}(q) = P_{\text{mon}}(q)$. Further details about the specific models used in the different experimental cases are provided in the Results section along with the obtained model fit parameters.

Analysis Software. All models were implemented in an in-house made Fortran77 program containing a least squares fitting (LSF) routine that allowed fitting the mathematical models to the experimental data. As described above, *CRYSOL* was used to determine the numerical scattering intensities corresponding to the well-ordered oligomeric states, and as described in more detail elsewhere,[29] the information about the form factor amplitudes was also obtained from the *CRYSOL* output files.

■ RESULTS

In the following, the results of the SAXS analysis of the three different insulin analogues prepared with, respectively, Fe(II) or Zn(II) are presented. The data are rescaled in all plots to allow for evaluation of the quality of the model fits. An unscaled data representation of the data is plotted in Figure S2 of the Supporting Information.

Structural Rearrangements Introduced by Titration of Insulin Analogues with Fe(II). *InsX2 with Fe(II).* In this experiment, samples of unmodified InsX2 were prepared with increasing amounts of Fe(II) as described in Experimental Methods. SAXS data were obtained at the SAXS beamline I711 at MAX-lab, Sweden. The aim of the series was to investigate the behavior of InsX2 in the presence of Fe(II) as a reference for the InsX2-bipy measurements. The obtained SAXS data are plotted in Figure 2a. It is readily seen that the intensity in the low q part of the data increases with increasing Fe(II) concentration. This clearly indicates that, although InsX2 has no bipy moiety that can bind Fe(II), the addition of Fe(II) still leads to the formation of larger but less ordered molecular assemblies.

To obtain a better and more quantitative description of the formed structures, the SAXS data were simulated using the following special case of eq 5 and keeping $1 = X_{\text{mon}} + X_{\text{SA,mon}}$:

$$I(q) = c\{X_{\text{mon}} \cdot P_{\text{mon}}(q) + X_{\text{SA,mon}} F_{\text{frac}}[q, P_{\text{mon}}(q)]\} \tag{7}$$

This analysis showed that the pure InsX2 data were consistent with a model for monomeric insulin, thus having $X_{\text{mon}} = 1$ and $X_{\text{SA,mon}} = 0$. Titration of more and more Fe(II) into the sample gave rise to structures that could best be described as long, slightly branched insulin aggregates built from monomers. A decreasing fraction of monomers, X_{mon}, and a correspondingly increasing fraction, $X_{\text{SA,mon}}$, of these fractal-like aggregates were observed. Furthermore, an increase of both the fractal dimension, D, and the correlation length, ξ, was observed. The growth of the

fractal dimension corresponds to the transition from a needle-like to a more branched structure.[32] Had the correlation length been relatively constant and independent of the Fe(II) concentration, that would have indicated an ordered repetitive substructure in the molecular assemblies. The observation of an increasing correlation length, however, suggests that the molecular assemblies do not possess such a repetitive substructure and they are not well-ordered. In the process of analyzing the data, a more simple model that assumed all the insulin took part in the molecular assemblies, i.e., a model that assumed $X_{SA,mon} = 1$, was evaluated for the samples with Fe(II) concentrations above zero. However, this model failed to describe the experimental data at the three intermediate Fe(II) concentrations. The simulation results are summarized in Table 2. The corresponding model fits are shown in Figure 2a and the formed structures are illustrated Figure 2b.

Table 2. InsX2 Formulated with Fe(II)—Model Fit Results[a]

[Fe(II)]/[InsX2]	X_{mon}[c] [%]	$X_{SA,mon}$ [%]	ξ [Å]	D
0	100	0	n.a.	n.a.
0.13	57.4	42.6 ± 0.9	58 ± 6	1.13 ± 0.04
0.25	63.8	36.2 ± 1.7	110 ± 12	1.50 ± 0.04
0.51	12.0	88.0 ± 0.3	261 ± 17	1.50 ± 0.07
0.75	0	100[b]	280 ± 18	1.57 ± 0.01

[a]X_{mon}: The fraction of InsX2 that is in monomeric form. $X_{SA,mon}$: The fraction of InsX2 that participates in larger assemblies. ξ is the correlation length and D the fractal dimension of the larger monomer-based assemblies. [b]Parameter fixed manually. [c]X_{mon} is calculated using $X_{mon} = 1 - X_{SA,mon}$.

InsX2-bipy[B29] with Fe(II). The goal of modifying insulin X2 with bipy[B29] was, vide supra, to develop a protein which formed trimers upon addition of Fe(II). Thus, a Fe(II) titration scan of InsX2-bipy[B29] was performed and studied by SAXS at the ID14−3 beamline at ESRF, France. The obtained data are shown in Figure 3a. A sharp upturn in the intensity at low q-values is observed in all data sets. The upturn shows that aggregates were present in the sample even before Fe(II) was added. However, we observe that aggregation does not increase with titration of Fe(II) suggesting that the aggregated part was already present in the stock solution and is unaffected by the Fe(II) addition. The data analysis shows that the SAXS data from InsX2-bipy[B29] in the presence of varying concentrations of Fe(II) can be described very well as a combination of insulin monomers, insulin trimers, and a constant fraction of preaggregated insulin. While the titration of Fe(II) leads to a gradual decrease of the monomer fraction and a corresponding increase of the trimer fraction, the preaggregated insulin fraction remained constant. Thus, the experimental data could be fitted directly by eq 6. The presence of the preaggregated insulin is unfortunate and could most likely have been dealt with by careful centrifugation prior to the experiments. However, the volume fraction of the preaggregated insulin is only minor and its effect is negligible on the high-q part of the data, which contains the information about the monomer to trimer transition.

In order to determine the structure of the trimers, we exploited our previous findings[12] that indicated that the InsX2-bipy[B29] trimers formed with Fe(II) are primarily of facial (*fac*) stereochemistry. Several hypothetical PDB structures for the InsX2-bipy[B29] trimer were generated for both the meridional (*mer*) and *fac* isomers of the trimer and tested against the SAXS data. The model fitting to the very high-quality SAXS data clearly confirmed that the *fac* isomeric geometry was the dominating species.

In order to describe the fractal-like aggregates visible in the scattering data as the sharp upturn at low q, it was found that the fractal dimension, D, had to be fixed to a value of 3, corresponding to a rather compact aggregate. For the volume fraction of the large insulin aggregates, $X_{SA,mon}$, and the correlation length of the aggregates, ξ, the solution space was much larger, and many pairs of the two variables gave good
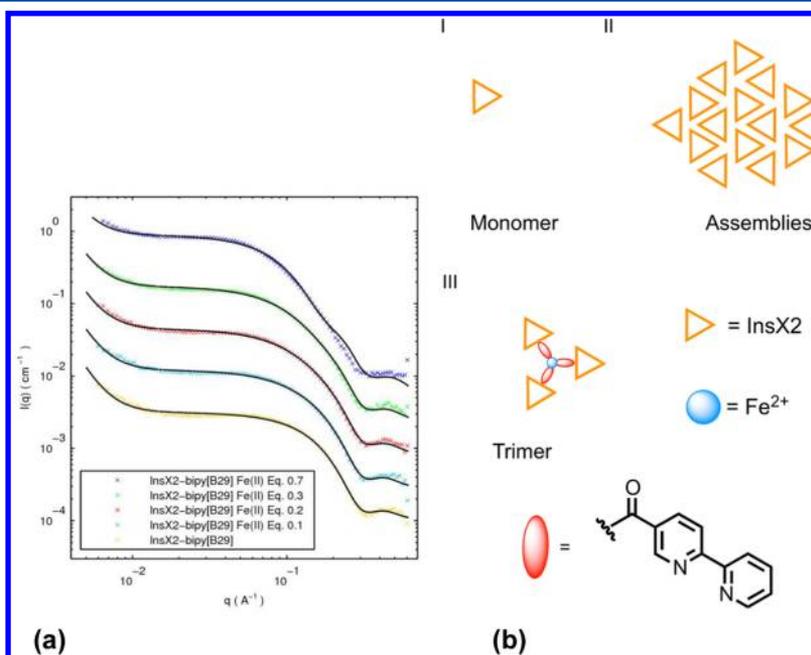

Figure 3. (a) SAXS data (colored points) and model fits (full black lines). The InsX2-bipy[B29] data (yellow points) are on absolute scale. The above data have been rescaled by 3^n where n runs from 1 to 4. (b) (I) Free monomeric insulin. (II) Assemblies built from monomeric insulin. (III) Insulin trimer coordinating around a Fe(II) ion.

solutions. The values $X_{SA,mon}$ = 1% and a ξ = 4200 Å were used. This gave a good description of the upturn in all cases and allowed for focusing on the Fe(II) induced transition from monomeric to trimeric insulin visible at high q.

The resulting model fits are plotted in Figure 3a, the fit parameters are reported in Table 3, and the corresponding

Table 3. InsX2-bipy[B29] Formulated with Fe(II)—Model Fit Results[a]

[Fe(II)]/[InsX2-bipy[B29]]	X_{mon} [%][c]	X_{tri} [%]	$X_{SA,mon}$ [%]
0	99	0[b]	1[b]
0.1	88.7	10.3 ± 0.2	1[b]
0.2	77.2	21.8 ± 0.2	1[b]
0.3	61.1	37.9 ± 0.3	1[b]
0.7	0	99[b]	1[b]

[a]X_{mon}: The fraction of insulin that is in monomeric form. X_{tri}: The fraction of insulin that is in trimeric form. $X_{SA,mon}$: The fraction of insulin that participates in larger assemblies. [b]Parameter fixed manually. [c]X_{mon} is calculated using $X_{mon} = 1 - (X_{tri} + X_{SA,mon})$.

structures are illustrated in Figure 3b. It is clear from the modeling outputs that the amount of trimers formed increased with the Fe(II) concentration from 0% to a maximum close to 100%.

InsX2-bipy[A1] with Fe(II). InsX2-bipy[A1] was the third insulin analogue investigated; it had also been designed to form trimers upon addition of Fe(II). SAXS data were obtained at beamline I711 at MAX-lab, Sweden. As for InsX2-bipy[B29], the aim of the InsX2-bipy[A1] series was to investigate the InsX2-bipy[A1] trimer formation upon addition of Fe(II). The obtained data are plotted in Figure 4a.

When inspecting the data set for InsX2-bipy[A1] in Figure 4a, it can be seen from the lack of a Guinier region at low q that there is already a high degree of molecular assembly present in the sample even before Fe(II) is added to the sample. However, the

titration with Fe(II) clearly leads to further growth of the size of the molecular assemblies in the sample as observed from the increase in the low-q scattering with increasing Fe(II) concentration. The data from this series were modeled with a further generalized version of eq 6 that included the possibility of having fractal-like aggregates from both monomeric and trimeric subunits:

$$I(q) = c\{X_{mon}\cdot P_{mon}(q) + X_{tri}\cdot P_{tri}(q) + X_{SA,mon}\cdot$$
$$F_{frac}[q, P_{mon}(q)] + X_{SA,tri}\cdot F_{frac}[q, P_{tri}(q)]\} \quad (8)$$

where, analogous to the more simple cases, $X_{mon} + X_{tri} + X_{SA,mon} + X_{SA,tri}$ = 1. In practice, none of the data sets contained enough information to allow for fitting all four terms of eq 8 at the same time. The results reported in the following are hence the best compromise in the need to minimize the number of fitting parameters while still obtaining good model fits.

The InsX2-bipy[A1] sample without Fe(II) was modeled as a combination of monomers and assemblies of monomers, i.e., by having $X_{tri} = X_{SA,tri} = 0$ in eq 8 such that it reduces to eq 7. This suggested that 72% of the sample was still in monomeric form, while the remaining 28% of the insulin formed larger assemblies that could be described as slightly branched structures, with a fractal dimension D of 1.4. A correlation length, ξ, of ~200 Å was found. Since this type of structure is not observed for InsX2, it is clear that the molecular assemblies must be mediated by bipy, as expected. If the molecular assemblies were due to simple pairs of bipy–bipy interactions, the size of the assemblies would be expected to be limited by geometry, as the InsX2 is much larger than the bipy moiety. This is not the case, which suggests an interaction between the bipy group of one molecule and the InsX2 of another molecule, e.g., hydrophobic, ionic, and so forth. This would allow the formation of chains and branched chains, in accordance with the modeling results.
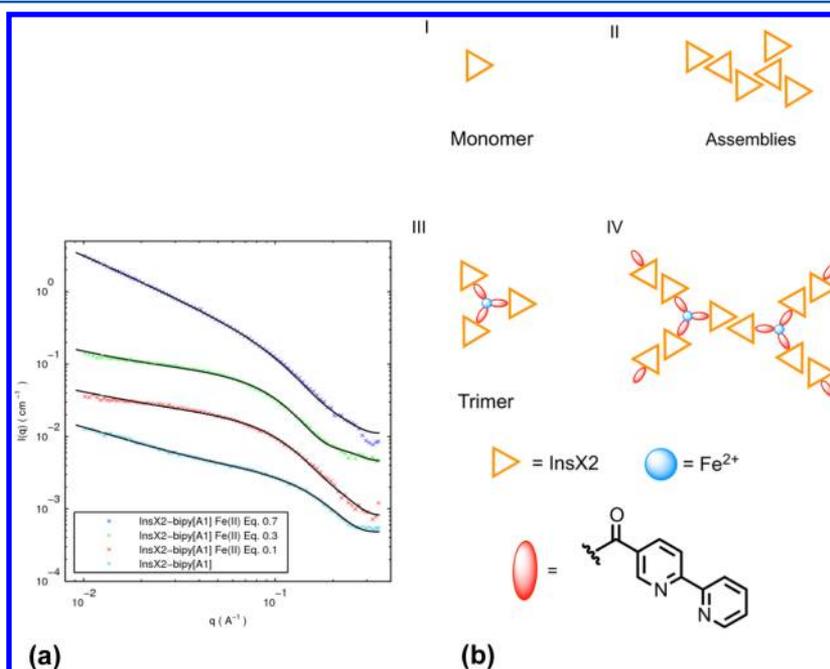

Figure 4. (a) SAXS data (colored points) and model fits (black lines). The InsX2-bipy[A1] data (light blue points) are on absolute scale. The above data are rescaled by 3^n where n runs from 1 to 3. (b) (I) Monomeric insulin. (II) Assemblies built from monomeric insulin subunits. (III) Trimer built from monomeric insulin around a Fe(II) ion. (IV) Assemblies built from trimer insulin subunits.

Table 4. InsX2-bipy[A1] Formulated with Fe(II)—Model Fit Results[a]

[Fe(II)]/[InsX2-bipy[A1]]	X_{mon}[c] [%]	X_{tri} [%]	$X_{SA,mon}$ [%]	$X_{SA,tri}$ [%]	ξ_{mon} [Å]	D_{mon}	ξ_{tri} [Å]	D_{tri}
0	72.3	0[b]	27.7 ± 2.1	n.a.	198 ± 56	1.39 ± 0.07	n.a.	n.a.
0.1	40.8	43.1 ± 1.1	16.1 ± 0.5	0[b]	198[b]	1.39[b]	n.a.	n.a.
0.3	25.74	72.1 ± 1.1	0[b]	2.16 ± 0.04	n.a.	n.a.	309[b]	1.43[b]
0.7	0[b]	52.8[d]	0[b]	47.2 ± 0.7	n.a.	n.a.	309 ± 25	1.43 ± 0.01

[a]X_{mon}: The fraction of insulin that is in monomeric form. X_{tri}: The fraction of insulin that is in trimeric form. $X_{SA,mon}$: The fraction of monomeric insulin that participates in larger assemblies. X_{SAtri}: The fraction of trimeric insulin that participates in larger assemblies. ξ_{mon} is the correlation length and D_{mon} the fractal dimension of the monomer assemblies. ξ_{tri} is the correlation length and D_{tri} the fractal dimension of the trimer assemblies. [b]Parameter fixed manually. [c]X_{mon} is calculated using $X_{mon} = 1 - (X_{tri} + X_{SA,mon} + X_{SA,tri})$. [d]Case where X_{tri} is calculated using $X_{tri} = 1 - X_{SA,tri}$.
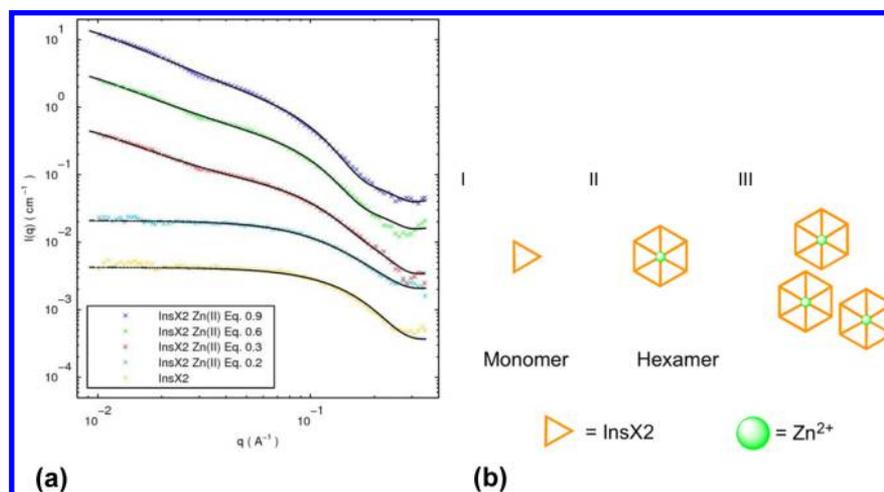

Figure 5. (a) SAXS data (colored points) and model fits (full black lines). The InsX2 data (yellow points) are on absolute scale. The above data are rescaled by 3^n where n runs from 1 to 4. (b) (I) Monomeric insulin. (II) Cartoon of hexameric insulin. (III) Assemblies built from hexameric insulin subunits.

Table 5. InsX2 Formulated with Zn(II)—Model Fit Results[a]

[Zn(II)]/[InsX2]	X_{mon} [%]	X_{hex} [%]	$X_{SA,hex}$ [%]	ξ [Å]	D
0	100[c]	0[b]	n.a.	n.a.	n.a.
0.2	91.3[c]	8.7 ± 0.3	0[b]	n.a.	n.a.
0.3	78.2 ± 0.7	19.1[d]	2.7 ± 0.1	100 ± 13	2.04 ± 0.06
0.6	58 ± 1	36[d]	6.2 ± 0.1	140 ± 15	1.84 ± 0.05
0.9	53 ± 2	35[d]	11.8 ± 0.3	142 ± 17	1.80 ± 0.04

[a]X_{mon}: The fraction of insulin that is in monomeric form. X_{hex}: The fraction of insulin that is in hexameric form. $X_{SA,hex}$: The fraction of insulin that participates in larger assemblies with hexameric subunits. ξ is the correlation length of the assemblies and D is their fractal dimension. [b]Parameter fixed manually. [c]X_{mon} is calculated using $X_{mon} = 1 - (X_{hex} + X_{SA,hex})$. [d]$X_{hex}$ is calculated using $X_{hex} = 1 - (X_{mon} + X_{SA,hex})$.

Our recent CD data revealed an opposite handedness of the bipy-Fe(II) complexes.[12] Here, the *fac* isomeric trimer structure found for InsX2-bipy[B29] was tested as input for the modeling of the InsX2-bipy[A1] samples with Fe(II) and turned out to be in good agreement with the experimental data.

During the modeling of InsX2-bipy[A1] with Fe(II), it was found that, depending on the concentration of Fe(II), the system could be described as either (1) a mixture of monomeric insulin, trimeric insulin, and assemblies of monomer insulin, which was observed at low Fe(II) concentrations, or (2) a combination of monomeric insulin, trimeric insulin, and assemblies of trimers as was observed at high Fe(II) concentrations. While it is most likely that some of the samples contain all four species in eq 8, the experimental SAXS data do not allow for such a detailed decomposition.

For the samples with high concentrations of Fe(II), the data suggest that, since almost all the bipy in the sample coordinated to Fe(II) and formed trimers, the aggregation of trimers was due to Ins−Ins binding mediated by the Fe(II).

The obtained model fits are plotted in Figure 4a, the corresponding fit parameters are listed in Table 4, and the obtained structures are illustrated in Figure 4b.

Structural Rearrangements Introduced by Titration of Insulin Analogues with Zn(II). *InsX2 with Zn(II).* The SAXS data from InsX2 titrated with Zn(II) were obtained at the I711 beamline at Max-Lab, Sweden. As discussed above, the InsX2 without metal ions is primarily in monomeric form. As Zn(II) was titrated into the InsX2 sample, an increase of the scattering at low q was observed (Figure 5a) and a kink of the $I(q)$ at approximately $q = 0.03$ Å^{-1} appeared. The analysis showed that the data could be described as a mixture of monomers, hexamers, and assemblies of hexamers, according to eq 9:

$$I(q) = c\{X_{mon}{\cdot}P_{mon}(q) + X_{hex}{\cdot}P_{hex}(q) + X_{SA,hex}$$
$${\cdot}F_{frac}[q, P_{hex}(q)]\} \qquad (9)$$

When inspecting the obtained values for X_{mon}, X_{hex}, and $X_{SA,hex}$ in Table 5, it is seen that, with increasing Zn(II) concentration,

the hexamers started to form larger assemblies before all the InsX2 had formed hexamers. This suggests that Zn(II), besides inducing hexamer formation, can also to a lesser extent mediate an attraction between already formed hexamers. Whether this is due to the existence of a specific ligand binding site at the water-exposed side of the insulin remains to be investigated. It can be noted, however, that, in a Zn_4 T_3R_3 insulin hexamer crystal structure, there is evidence for the presence of a Zn-binding site at the hexamer surface, with Zn coordination by HisB5 from different hexamers.[33] Here, we find that addition of Zn(II) to InsX2 leads to the formation of both InsX2 hexamers and to branched assemblies of InsX2 hexamers, with a fractal dimension D of ~1.8 and a correlation length ξ of ~140 Å.

For the two samples with the highest amounts of Zn(II), we were unable to obtain fully satisfactory fits in the q-region from ~0.02 Å^{-1} to ~0.05 Å^{-1}. A marked kink is observed in the experimental data at ~0.03 Å^{-1} as an indication of a well-defined characteristic length, but the kink is not reproduced by the model. This indicates that the description provided by eq 9 of a mixture of separated monomers, hexamers, and fractal-like assemblies of hexamers is a slightly too simplistic picture of the sample; and most likely, mixed aggregates will be present, e.g., fractal-like aggregates of monomers, dimers, and hexamers. An attempt to fit the data and in particular the observed kink with a model approach previously developed to describe casein micelles was also tried.[34] In this approach, the particles are described as larger, e.g., spherically shaped aggregates with a protein or other appropriate substructure depending on the actual sample. While this approach gave a decent description of the kink-region, it failed to give good model fits to the remaining part of the data, and the approach was not applied further.

The obtained model fits using eq 9 are plotted in Figure 5a, the corresponding fit parameters are listed in Table 5 and a cartoon of the systems is provided in Figure 5b.

CD Analysis of Ins-bipy Titrated with Zn(II). The trimer formation in the InsX2-bipy[A1] and InsX2[B29] with Fe(II) systems was previously studied by CD spectroscopy.[12] In order to complement the SAXS analysis of the InsX2-bipy in the presence of Zn(II), CD and near-UV absorption measurements were performed on the InsX2-bipy[A1] under similar conditions to investigate if insulin trimers formed when adding Zn(II). For the InsX2-bipy[A1], optical spectroscopy in the near-UV spectra showed absorption from the $\pi-\pi^*$ transition of the bipy chromophore overlapping the weaker signal from the Tyr side-chains of the insulin. Thus, the absorption spectrum in the absence of Zn(II) displayed a ~280 nm peak, and in the corresponding CD spectrum, a weak positive signal was observed in the same region. Gradual addition of Zn(II) red-shifts the absorbance, and a new peak was observed with maxima at 305 and 318 nm and with isosbestic points at 275 and 300 nm in agreement with metal ion coordination by bipy. The CD spectrum similarly displayed a red-shifted signal with two local maxima above 300 nm. The observed positive CD signal above 300 nm is likely to correspond to a single bipy unit from InsX2-bipy[A1] binding Zn(II) under the influence of the chiral environment of the protein. It is important to note that the observed bipy CD signals in Figure 6 do not correspond to the characteristic exciton signature of a *tris* bipy coordination environment, thus suggesting that trimers are not formed. It must be said, however, that, in theory, equal amounts of Δ and Λ optical isomer trimers might have been formed such that the CD signals would cancel each other out.
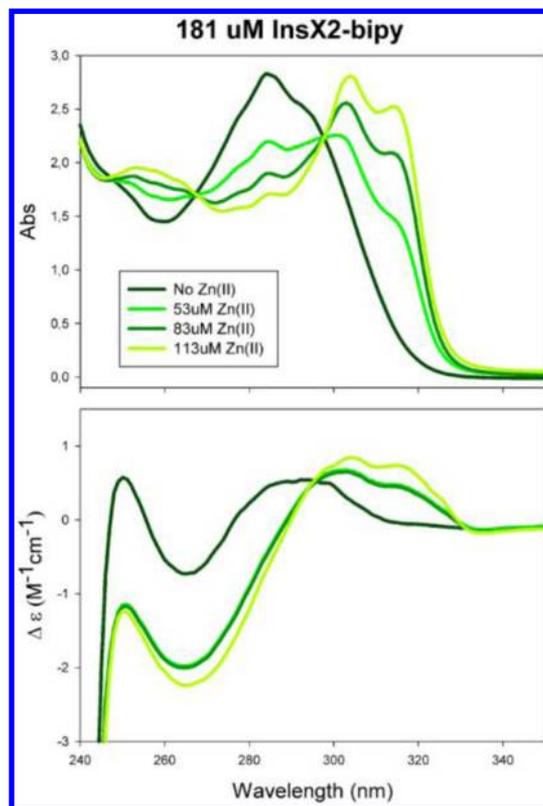

Figure 6. Near-UV absorption (upper panel) and CD (lower panel) of 181 μM InsX2-bipy[A1]. The following [Zn(II)]/[InsX2] stoichiometries were employed: 0, 0.29, 0.45, and 0.62.

InsX2-bipy[B29] with Zn(II). SAXS measurements of the InsX2-bipy[B29] samples titrated with increasing concentrations of Zn(II) were performed at the ID14−3 beamline at ESRF, France. As explained above, the CD analysis did not indicate the formation of bipy-induced trimers upon addition of Zn(II) in the InsX2-bipy[A1]-system, and trimers were therefore also not expected in the InsX2-bipy[B29]-system. This expectation was confirmed in the SAXS data analysis where we found the best agreement between experimental data and model fits by using a combination of monomers, hexamers, and monomer-based fractal-like aggregates:

$$I(q) = c\{X_{mon} \cdot P_{mon}(q) + X_{hex} \cdot P_{hex}(q) + X_{SA,mon}$$
$$\cdot F_{frac}[q, P_{mon}(q)]\} \qquad (10)$$

An upturn at low q similar to the one observed in the preparations of InsX2-bipy[B29] with Fe(II) is present in the data. The two sets of samples were prepared from the same stock, and as in the preparation with Fe(II), no significant dependence was observed between the upturn and the amount of titrated metal ion. The upturn could be sufficiently well accounted for by the same approach and model parameters as in the analysis of the InsX2-bipy[B29] with Fe(II). This allowed for focusing on the information present in the intermediate to high q data. As can be seen in Figure 7a, the high q part of the SAXS data only changes slightly upon the addition of Zn(II), thus indicating only minor structural changes. This observation is supported by the model fit parameters reported in Table 6, which show that the sample only reacts weakly upon the addition of Zn(II) as compared to the above-described ordinary InsX2 sample with Zn(II). Our analysis suggests that when [InsX2-bipy[B29]]/[Zn(II)] = 0.7
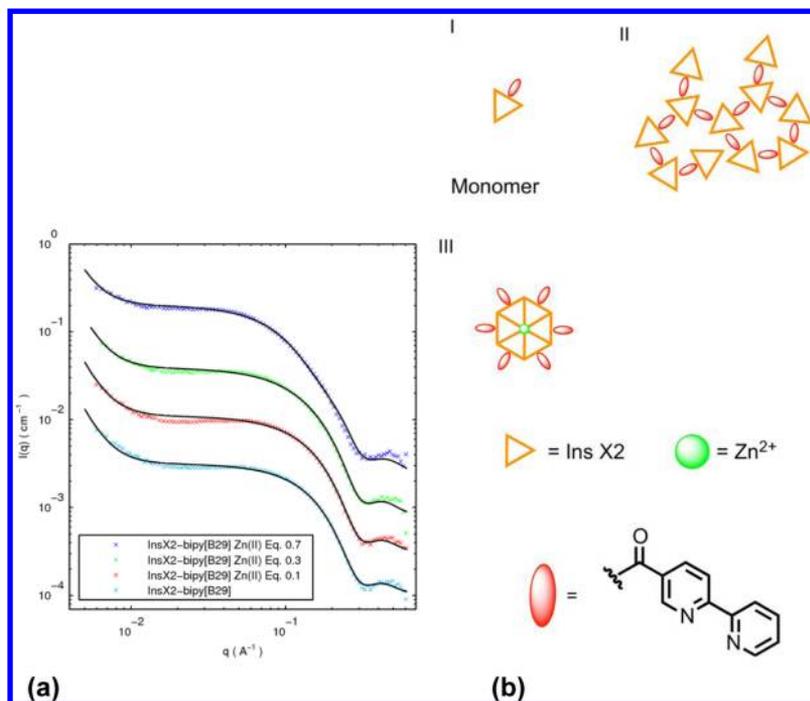
Figure 7. (a) SAXS data (colored points) and model fits (full black lines). InsX2-bipy[B29] data (light blue points) are on absolute scale. The above data are rescaled by 3^n where n runs from 1 to 3. (b) (I) Monomeric insulin. (II) Assemblies of monomeric insulin. (III) Insulin hexamers.

Table 6. InsX2-bipy[B29] with Zn(II)—Model Fit Results[a]

[Zn(II)]/[InsX2-bipy[B29]]	X_{mon}[c] [%]	X_{hex} [%]	$X_{SA,mon}$ [%]
0	99	0[b]	1[b]
0.1	99	0[b]	1[b]
0.3	95.3	3.7 ± 0.8	1[b]
0.7	77.6	21.4 ± 0.2	1[b]

[a] X_{mon}: The fraction of insulin that is in monomeric form. X_{hex}: The fraction of insulin that is in hexameric form. $X_{SA,mon}$: The fraction of insulin that participates in larger assemblies with monomeric subunits. [b] Parameter fixed manually. [c] X_{mon} is calculated using $X_{mon} = 1 - (X_{hex} + X_{SA,mon})$.

is fitted with model described by eq 10 then only 21% of the InsX2-bipy[B29] are in hexamer form, while there are no indications of the more unspecific Zn(II)-mediated Ins—Ins interactions that gave rise to the large fractal-like structures observed in the pure InsX2 system. Thus, InsX2-bipy[B29] seem to have both a lower tendency to form hexamers and a lower tendency to form hexamer-based larger aggregates, and the monomeric insulin is by a large margin the dominating species.

A weak depression around $q = 0.02$ Å^{-1} followed by a slightly higher scattering intensity around $q = 0.05$ Å^{-1} is also observed. For protein samples, which have a nearly homogeneous scattering length density, the most obvious explanation to this effect is a weak repulsive interaction between the single proteins. Indeed, using a previously developed strategy[29] for the incorporation of a hard-spheres structure factor in the modeling of the free monomer and hexamer insulins allowed for a superior model fit to the scattering data at intermediate q-values (results not shown). However, as the structure factor effect is weak and mainly affects data in the intermediate q-range, while the information about the monomer to hexamer transition is mainly present at high q-values, the two types of information do to a good approximation not interfere. Hence, in practice the same results for the monomer to hexamer fractions were obtained with

and without the structure factor incorporated. For simplicity, we have therefore chosen only to report the fit results obtained without structure factor, i.e., eq 10. The obtained model fits are plotted in Figure 7b, a cartoon of the structures is provided in Figure 7b, and the model fit parameters are reported in Table 6.

InsX2-bipy[A1] with Zn(II). SAXS data from InsX2-bipy[A1] with Zn(II) were obtained at the I711 beamline at Max-lab, Sweden. The obtained data were best described as a combination of monomers and hexamers in combination with fractal-like aggregates of, respectively, monomers and hexamers:

$$I(q) = c\{X_{mon} \cdot P_{mon}(q) + X_{hex} \cdot P_{hex}(q) + X_{SA,mon} \cdot$$
$$F_{frac}[q, P_{mon}(q)] + X_{SA,hex} \cdot F_{frac}[q, P_{hex}(q)]\} \quad (11)$$

where, as usual, $X_{mon} + X_{hex} + X_{SA,mon} + X_{SA,hex} = 1$. As in the case of InsX2-bipy[A1] with Fe(II), not all the terms of the model were fitted in any of the data sets. As explained above, the InsX2-bipy[A1] with no added metal ions consists of monomers and chain-like assemblies of monomers. As more and more Zn(II) was added to the system, a population of hexamers first appeared, and then a population of aggregated hexamers appeared at the highest Zn(II) concentration. As in the InsX2-bipy[A1] with Fe(II) system, it is most likely that the samples will contain populations of monomer-based and hexamer-based aggregates, as well as even mixed monomer—hexamer aggregates. However, the information contained in our experimental data did not allow for introducing these more advanced descriptions while at the same time obtaining reliable estimates of the fit parameters, and the model was restrained to the more simple approximation provided above. The model-fits are shown in Figure 8a, the corresponding structures are illustrated in Figure 8b, and the obtained fit parameters are reported in Table 7.
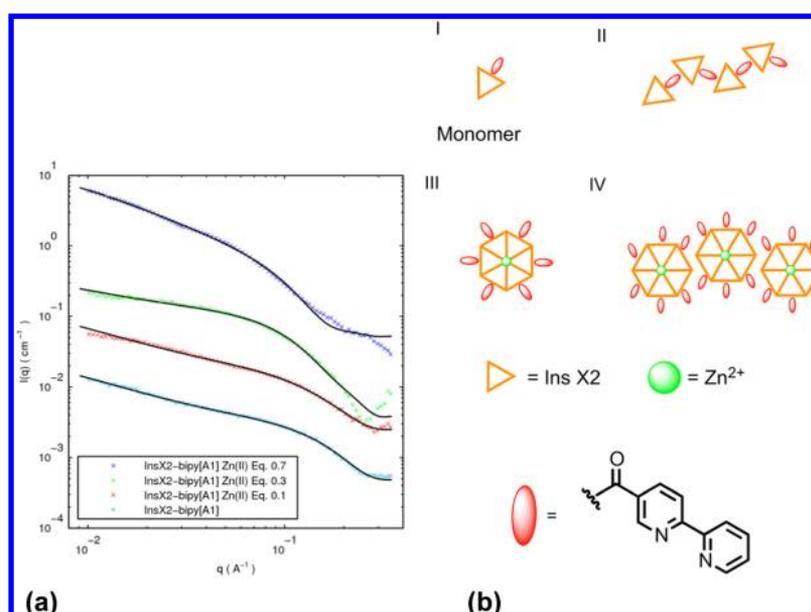
Figure 8. (a) SAXS data (colored points) and model fits (full black lines). InsX2-bipy[A1] data (light blue points) are on absolute scale. The above data are rescaled for readability by 4n where n runs from 1 to 3. (b) (I) monomeric insulin. (II) Assemblies of monomeric insulin. (III) Insulin hexamer. (IV) Assemblies built from hexameric insulin.

Table 7. InsX2-bipy[A1] Formulated with Zn(II)—Model Fit Results[a]

[Zn(II)]/[InsX2-bipy[A1]]	X_{mon} [%]	X_{hex} [%]	$X_{SA,mon}$ [%]	$X_{SA,hex}$ [%]	ξ_{mon} [Å]	D_{mon}	ξ_{hex} [Å]	D_{hex}
0	72.3^c	0^b	27.7 ± 2.1	n.a.	198 ± 56	1.39 ± 0.07	n.a.	n.a.
0.1	62.0^c	5.2 ± 0.1	32.8 ± 0.5	0^b	198^b	1.39^b	n.a.	n.a.
0.3	54.7^c	27.0 ± 0.7	18.3 ± 0.1	0^b	198^b	1.39^b	n.a.	n.a.
0.7	54.5^b	21.8^d	0^b	23.7 ± 0.8	n.a.	n.a.	180 ± 18	1.37 ± 0.03

[a]X_{mon}: The fraction of insulin that is in monomeric form. X_{hex}: The fraction of insulin that is in hexameric form. $X_{SA,mon}$ and $X_{SA,hex}$: The fraction of insulin that participates in larger assemblies with, respectively, monomeric and hexameric subunits. ξ is the correlation length of the assemblies in question and D is their fractal dimension. [b]Parameter fixed manually. [c]X_{mon} is calculated using $X_{mon} = 1 - (X_{hex} + X_{SA,mon} + X_{SA,hex})$. [d]$X_{hex}$ is calculated using $X_{hex} = 1 - (X_{mon} + X_{SA,mon} + X_{SA,hex})$.

■ DISCUSSION

Our recently reported chromatographic and optically based study demonstrated that both bipy-functionalized InsX2 variants form trimers in the presence of Fe(II).[12] However, the CD data did not provide information on the higher-order self-assembly of the InsX2-bipy variants. The present SAXS-based structural study indeed not only confirmed the presence of these trimers, but also suggests that there were other interaction sites for divalent metal ions in this insulin-based system, which gave rise to the formation of larger and less ordered structured assemblies with the trimers or the hexamers as the basic "building blocks". Binding sites for metal ions, e.g., Co^{3+} and Ca^{2+} in insulin have previously been found inside the insulin hexamer[35,36] and have been shown to stabilize this. The present study and the observation of large metal ion-mediated assemblies, however, indicate that some type of interaction sites are present at the water-exposed part of the X2 insulin surface, including on the X2 insulin hexamer.[33] The SAXS analysis furthermore showed that both location of the bipy and the type of metal ion, Fe(II) or Zn(II), significantly affected the self-assembly behavior and thus the nanoscale structures of the formed molecular assemblies. An overview of the rich structural behavior of the system as determined from the SAXS analysis is provided in Figure 9.

In summary, we find that insulin trimers form when adding Fe(II) to both the InsX2-bipy[A1] and the InsX2-bipy[B29] system. When Fe(II) is added to insulin at a molar ratio of 0.7,

both systems are saturated and are composed of ∼100% insulin trimers. However, the analysis clearly shows that, in the InsX2-bipy[A1] system, the trimers form and become predominant at much lower Fe(II) concentrations than in the InsX2-bipy[B29] system.

The analysis previously performed by Munch et al.[12] suggested that the trimers formed in the InsX2-bipy systems were most likely *fac* isomeric. In the present SAXS analysis on the same system, both the *fac* and *mer* conformations were tested as structural inputs for the modeling routines. The *mer* conformations were unable to fit the high q part of the SAXS data, whereas the *fac* conformation fitted the data sets very well, clearly showing that the insulin trimers formed had a preference for the *fac*. Munch et al.[12] showed that InsX2-bipy[B29] and InsX2-bipy[A1] have opposite enantiomeric excess (Δ/Λ). The resolution of the SAXS data does, however, not allow for a distinction between the two enantiomers, and the same structure was used for both the InsX2-bipy[B29] and InsX2-bipy[A1] trimers.

The CD analysis of the InsX2-bipy[A1] system with Zn(II) strongly indicated that Zn(II) does not promote trimer formation in these systems, as no exciton signal was observed. However, there was clear evidence from CD and absorbance that some Zn(II) ions coordinate to single bipy units. The SAXS analysis of the same systems showed that part of the Zn(II) ions give rise to the formation of insulin hexamers, even though the
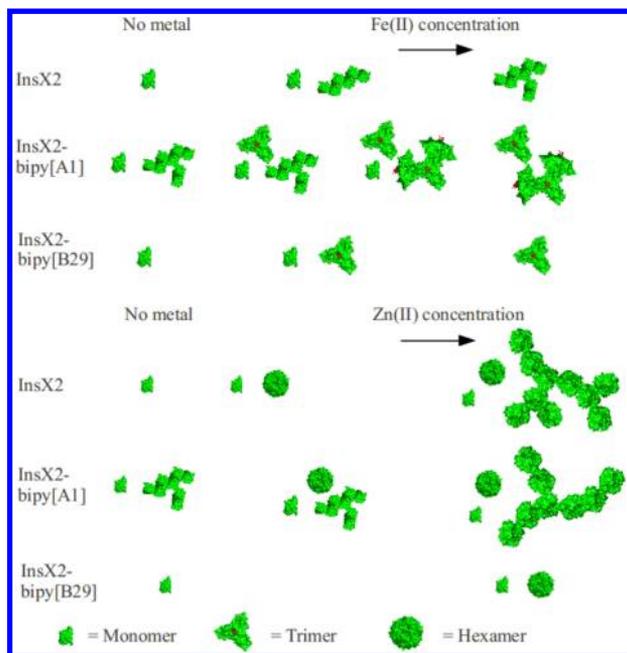
Figure 9. Different types of insulin assemblies found as the metal concentration in the sample is increased. Only the most significant assembly types are shown. The binding sites for the larger assemblies are only tentative and should not be seen as a true depiction of the assemblies. The bipy are only shown for the trimer (in red) and assemblies of trimers.

monomer−hexamer equilibrium is significantly shifted toward monomers as compared to native human insulin. It is noteworthy that the Zn(II) concentration dependence of the hexamer formation and the formation of larger hexamer assemblies were comparable in the InsX2 and the InsX2-bipy[A1] systems, while the propensity of the Zn(II) induced hexamer formation was significantly lower in the InsX2-bipy[B29]-system. Hence, we observe that the bipy modification of the insulin in position B29 impairs the formation of both Fe(II) induced trimers and Zn(II) induced hexamers.

It is also very interesting to note that the large fractal-like molecular assemblies were much more abundant in the InsX2 and InsX2-bipy[A1] systems than in the InsX2-bipy[B29] system. Thus, the bipy modification in position B29 not only impairs the specific oligomer formation, but also the unspecific higher oligomer formation, whether this has monomers, trimers, or hexamers as the basic building blocks.

The pharmaceutical implication of this finding is apparent. InsX2 has been designed as monomeric insulin with limited propensity for dimerization. If insulin drugs are injected subcutaneously, larger oligomers yield slower diffusion into the bloodstream leading to protracted drug delivery. Adding a synthetic side-chain to the B29 site can potentially impair the hexamer formation of InsX2 and consequently prevent assembly of larger hexamer-based aggregates. Another important aspect is that the receptor binding is not impaired by the synthetic side-chain. Receptor binding studies previously conducted for InsX2[10] by use of the free fat cell activity (FFC) and mouse blood glucose assay (MBGA) gave values of 31% (FFC) and 93% (MBGA). Receptor binding studies have also been performed on InsX2-bipy[A1] and InsX2-bipy[B29][12] using a human insulin receptor affinity scintillation proximity assay (HIRAspa).[37] For InsX2-bipy[A1], the receptor binding was found to be 5.2%, while for InsX2-bipy[B29], it was 29.4%,[12] hence still

subnanomolar. Although the different methods are not directly comparable, the results indicate that the receptor binding is not impaired by adding bipy to the B29 site. Further studies with other synthetic side chains could potentially lead to faster-acting insulin analogues applicable in the treatment of diabetes mellitus.

CONCLUSION

The nanoscale-scale structures formed by self-assembly of bipyridine modified insulin variants have been investigated by SAXS. The bipy modification led to the formation of well-defined trimers when adding Fe(II). The addition of Zn(II) led to the formation of the same type of hexamers as those present in human insulin, however, with a monomer−hexamer equilibrium significantly shifted toward the monomer side. The choice of binding site position (A1 or B29) for bipy on the insulin molecule played a significant role in the propensity of the system to form well-ordered trimers and hexamers, respectively. The choice of bipy binding site also affected the propensity of the system to form larger and less well-structured molecular assemblies. Insulin modified with bipy at position A1 easily formed both well-ordered oligomers and larger aggregates, whereas bipy modification at position B29 partly suppressed the formation of well-ordered oligomers and apparently fully suppressed the formation of larger aggregates. These structural findings suggest that the studied insulin analogues display interesting self-assembly behavior, with potential relevance for future development of insulin and other biopharmaceutical drugs. However, detailed pharmaceutical studies would obviously have to be performed in order to fully elucidate this potential. An advantage of our strategy for control of protein self-assembly is the use of non-native ligands. We believe that it can be extended to other proteins for a wide range of applications in biopharmaceuticals, chemical biology, and material science.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional figures as described. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: arleth@nbi.ku.dk.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Wang, W. Instability, stabilization, and formulation of liquid protein pharmaceuticals. *Int. J. Pharm.* **1999**, *185*, 129−88.

(2) Veronese, M., F.; Pasut, G. PEGylation, successful approach to drug delivery. *Drug Discovery Today* **2005**, *10*, 1451−1458.

(3) Hostrup, S.; Huus, K.; Parshad, H. Modification of Peptides and Proteins. In *Delivery Technologies for Biopharmaceuticals: Peptides, Proteins, Nucleic Acids and Vaccines*, Jorgensen, L.; Nielsen, H. M., Eds.; John Wiley & Sons Ltd: Chichester, West Sussex, UK, 2009; pp 171−191.

(4) Dodson, E. J.; Dodson, G. G.; Hubbard, R. E.; Moody, P. C. E.; Turkenburg, J.; Whittingham, J.; Xioa, B.; Brange, J.; Kaarsholm, N.; Thogersen, H. Insulin assembly: its modification by protein engineering and ligand binding. *Phil. Trans. R. Soc. London, Ser. A* **1993**, *345*, 153−164.

(5) Adams, M. J.; Baker, E. N.; Blundell, T. L.; Harding, M. M.; Dodson, E. J.; Hodgkin, D. C.; Dodson, G. G.; Rimmer, B.; Vijayan, M.; Sheat, S. Structure of Rhombohedral 2 Zinc Insulin Crystals. *Nature* **1969**, *224*, 491−495.

(6) Brange, J. The new era of biotech insulin analogues. *Diabetologia* **1997**, *40*, 48−53.

(7) WHO Fact sheet No. 312; http://www.who.int/mediacentre/factsheets/fs312/en/index.html 2011−10−06.

(8) Mayer, J. P.; Zhang, F.; Dimarchi, R. D. Insulin Structure and Function Insulin Structure and Function. *Biopolymers* **2007**, *88*.

(9) Bhatnagar, S.; Srivastava, D.; Jayadev, M. S. K.; Dubey, A. K. Molecular variants and derivatives of insulin for improved glycemic control in diabetes. *Prog. Biophys. Mol. Biol.* **2006**, *91*, 199−228.

(10) Brange, J.; Ribel, U.; Hansen, J. F.; Dodson, G.; Hansen, M. T.; Havelund, S.; Melberg, S. G.; Norris, F.; Norris, K.; Snel, L; Sørensen, A. R.; Voigt, H. O. Monomeric insulins obtained by protein engineering and their medical implications. *Nature* **1988**, *333*, 679−682.

(11) Sakai, S.; Sasaki, T. Multivalent Carbohydrate Ligands Assembled on a Metal Template. *J. Am. Chem. Soc.* **1994**, *116*, 1587−1588.

(12) Munch, H. K.; Heide, S. T.; Christensen, N. J.; Hoeg-Jensen, T.; Thulstrup, P. W.; Jensen, K. J. Controlled self-assembly of re-engineered insulin by Fe(II). *Chem.—Eur. J.* **2011**, *17*, 7198−7204.

(13) Knaapila, M.; Svensson, C.; Barauskas, J.; Zackrisson, M.; Nielsen, S., S.; Toft, K. N.; Vestergaard, B.; Arleth, L.; Olsson, U.; Pedersen, J. S.; Cerenius, Y. A new small-angle X-ray scattering set-up on the crystallography beamline I711 at MAX-lab. *J. Synchrotron Rad.* **2009**, *16*, 498−504.

(14) Pernot, P.; Theveneau, P.; Giraud, T; Fernandes, R., N.; Nurizzo, D.; Spruce, D.; Surr, J.; McSweeney, S.; Round, A.; Felisaz, F.; Foedinger, L.; Gobbo, A. .; Huet, J.; Villard, C.; Cipriani, F. New beamline dedicated to solution scattering from biological macromolecules at the ESRF. *JPCS* **2010**, *247*, 012009.

(15) Nielsen, S., S.; Toft, K., N.; Snakenborg, D.; Jeppesen, M., G.; Jacobsen, J., K.; Vestergaard, B.; Kutter, J., P.; Arleth, L. BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *J. Appl. Crystallogr.* **2009**, *42*, 959−964.

(16) Petoukhov, M., V.; Konarev, P., V.; Kikhney, A., G.; Svergun, D., I. Supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **2007**, *40*, 223−228.

(17) Pedersen, J. S. Modelling of Small-Angle Scattering Data from Colloids and Polymer Systems. In *Neutrons, X-ray and Light: Scattering Methods Applied to Soft Condensed Matter*, Lindner, P.; Zemb, T., Eds.; North-Holland, 2002; pp 391−420.

(18) Svergun, D.; Barberato, C.; Koch, M. H. J. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **1995**, 768−773.

(19) Protein Data Bank http://www.pdb.org/ 2012−07−25.

(20) Delano, W. L. The PyMOL molecular graphics system, version 1.2rl; Schrödinger, LLC, 2009.

(21) Bocian, W.; Sitkowski, J.; Bednarek, E.; Tarnowska, A.; Kawecki., R.; Kozerski, L. Structure of human insulin monomer in water/acetonitrile solution. *J. Biomol. NMR* **2008**, *40*, 55−64.

(22) Bernstein, F. C.; Koetzle, T., F.; Williams, G., J., B.; Meyer, E., F., Jr.; Brice, M., D.; Rodgers, J., R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank A Computer-Based Archival File for Macromolecular Structures. *Eur. J. Biochem.* **1977**, *80*, 319−324.

(23) Maestro, version 9.2; Schrödinger, LLC; New York, 2011.

(24) Sieber, R.; Decurtins, S.; Stoeckli-evans, H.; Wilson, C.; Yufit, D.; Howard, J.; A., K.; Capelli, S., C.; Hauser, A. A Thermal Spin Transition in [Co(bpy)3][LiCr(ox)3] (ox = C2O42; bpy = 2,2′-bipyridine). *Chem.—Eur. J.* **2000**, *6*, 361−368.

(25) MacroModel, version 9.9; Schrödinger, LLC; New York, 2011.

(26) Banks, J., L.; Beard, H., S.; Cao, Y.; Cho, A., E.; Damm, W.; Farid, R.; Felts, A., K.; Halgren, T., A.; Mainz, D., T.; Maple, J., R.; Murphy, R.; Philipp, D., M.; Repasky, M., P.; Zhang, L., Y.; Berne, B., J.; Friesner, R., A.; Gallicchio, E.; Levy, R., M. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26*, 1752−1780.

(27) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(28) Teixeira, J. Small-angle scattering by fractal systems. *J. Appl. Crystallogr.* **1988**, *21*, 781−785.

(29) Høiberg-Nielsen, R.; Westh, P.; Skov, L., K.; Arleth, L. Interrelationship of steric stabilization and self-crowding of a glycosylated protein. *Biophys. J.* **2009**, *97*, 1445−1453.

(30) Malik, L.; Nygaard, J.; Hoiberg-Nielsen, R.; Arleth, L.; Hoeg-Jensen, T.; Jensen, K., J. Perfluoroalkyl chains direct novel self-assembly of insulin. *Langmuir* **2012**, *28*, 593−603.

(31) Kotlarchyk, M.; Chen, S.-H. Analysis of small angle neutron scattering spectra from polydisperse interacting colloids. *J. Chem. Phys.* **1983**, *79*, 2461−2469.

(32) Bushell, G. C.; Yan, Y. D.; Woodfield, D.; Raper, J.; Amal, R. On techniques for the measurement of the mass fractal dimension of aggregates. *Adv. Colloid Interface Sci.* **2002**, *95*, 1−50.

(33) Smith, D., G.; Pangborn, W., A.; Blessing, R., H. Phase changes in T3R3f human insulin: temperature or pressure induced? *Acta Crystallogr.* **2001**, *57*, 1091−1100.

(34) Hansen, S.; Bauer, R.; Lomholt, S., B.; Quist, K., B.; Pedersen, J., S.; Mortensen, K. Structure of casein micelles studied by small-angle neutron scattering. *Eur. Biophys. J.* **1996**, *24*, 143−147.

(35) Kurtzhals, P.; Kiehr, B.; Sørensen, R. A. Cobalt(lll)-Insulin Hexamer is a Prolonged-Acting Insulin Prodrug. *J. Pharm. Sci.* **1995**, *84*, 1164−1168.

(36) Coffmant, F, D; Dunn, M., F. Insulin-Metal Ion Interactions: The Binding of Divalent Cations to Insulin Hexamers and Tetramers and the Assembly of Insulin Hexamerst. *Biochemistry* **1988**, *27*, 6179−6187.

(37) Glendorf, T.; Stidsen, C., E.; Norrman, M.; Nishimura, E.; Sørensen, A, R.; Kjeldsen, T. Engineering of insulin receptor isoform-selective insulin analogues. *PloS One* **2011**, *6*, 1−6.