BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Small Angle X-ray Scattering-Based Elucidation of the Self-Association Mechanism of Human Insulin Analogue Lys^{B29}(N$^{\varepsilon}\omega$-carboxyheptadecanoyl) des(B30)

Malene Hillerup Jensen,[†,‡,§] Per-Olof Wahlund,[§] Katrine Nørgaard Toft,[†] Jes Kristian Jacobsen,[§] Dorte Bjerre Steensgaard,[§] Marco van de Weert,[‡] Svend Havelund,[§] and Bente Vestergaard*[,†]

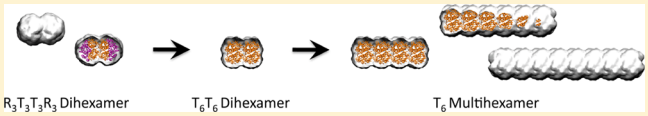
[†]Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark

[‡]Department of Pharmacy, University of Copenhagen, Copenhagen, Denmark

[§]Novo Nordisk A/S, Måløv, Denmark

Supporting Information

ABSTRACT: Lys^{B29}(N$^{\varepsilon}\omega$-carboxyheptadecanoyl) des(B30) human insulin is an insulin analogue belonging to a class of analogues designed to form soluble depots *in subcutis* by self-association, aiming at a protracted action. On the basis of small angle X-ray scattering (SAXS) supplemented by a range of biophysical and structural methods (field flow fractionation, dynamic and multiangle light scattering, circular dichroism, size exclusion chromatography, and crystallography), we propose a mechanism for the self-association expected to occur upon subcutaneous injection of this insulin analogue. SAXS data provide evidence of the in solution structure of the self-associated oligomer, which is a long straight rod composed of "tense" state insulin hexamers (T_6-hexamers) as the smallest repeating unit. The smallest oligomer building block in the process is a T_6T_6-dihexamer. This tense dihexamer is formed by the allosteric change of the initial equilibrium between a proposed "relaxed" state R_6-hexamer and an $R_3T_3T_3R_3$-dihexamer. The allosteric change from relaxed to tense is triggered by removal of phenol, mimicking subcutaneous injection. The data hence provide the first unequivocal evidence of the mechanism of self-association for this type of insulin analogue.



Great effort has been spent on the design and development of long-acting insulin molecules aimed at maintaining a basal level of insulin when administered to diabetics. Ideally, a combination treatment of rapid- and long-acting insulins mimics the natural insulin profile of a healthy person. A modern strategy, which aims to diminish the disadvantages of the first crystalline/amorphous long-acting insulins, is to attach human serum albumin (HSA) binding side chains to insulin, e.g., by acylation.[1−4] Through HSA binding *in subcutis* and in the blood, the circulation time prior to uptake in the liver is prolonged.[1−3,5−7] Some of the acylated insulin analogues modified with cholic acid derivatives, e.g., Lys^{B29}(N$^{\varepsilon}$-lithocholyl) des(B30) human insulin (licI), were shown to have the prolonged effect mainly because of self-association *in subcutis* resulting in a slowly releasing depot[8−11] and hence in a longer duration of action compared to that of binding to HSA alone. Some analogues realize their protracted action primarily by self-association, while others realize their long action because of a combination of self-association and HSA binding.[8] A prolonged action is also expected for Lys^{B29} (N$^{\varepsilon}\omega$-carboxyheptadecanoyl) des(B30) human insulin (ωchI), which is the object of this study. Another example of such an acylated analogue with protracted action is insulin detemir. This analogue, however, has been observed to associate into only dihexamers,[5] in contrast to ωchI, which forms high-molar mass (MM) complexes.[8,12]

Insulin is a small protein of 5.8 kDa, which is stored in the body as crystals in granules in the pancreatic β-cells.[13] Insulin dissolves to monomers, dimers, and hexamers in solution. The human insulin receptor recognizes the monomeric state, resulting in signal transduction.[13] A very well-studied characteristic of zinc-containing insulin is the ligand-induced allosteric change in the hexamer, where the N-terminal part of the B-chain (B1−B8) in the individual protomers adopts an extended ("tense", T) or α-helix ("relaxed", R) conformation. The hexamer can be regarded as a dimer of trimers in which the individual trimers can adopt either allosteric state, meaning that the hexamer can exist in three states, T_6-, T_3R_3-, and R_6-state. The half-site reactivity has been investigated by Dunn and coworkers.[14−16] Binding of phenol and phenolic derivatives in hydrophobic pockets of the hexamer results in a stable R_6-hexamer, whereas the hexamer is found in the T_3R_3- or T_6-state in the absence of phenolic ligands. The allosteric transition in response to various types of allosteric ligands has been studied thoroughly by crystallography,[17−20] nuclear magnetic resonance,[14,21−23] circular dichroism (CD),[15,16,24,25] and other spectroscopic methods.[21,22,26]

The self-association of novel HSA-binding insulin analogues is triggered by removal of phenol, as reported in size exclusion

Received: June 27, 2012
Revised: October 30, 2012
Published: December 20, 2012

chromatography (SEC) studies.[8,10,12] Presumably, the oligomerization is associated with allosteric changes from the R-state to the T-state. Little is known about the underlying mechanism for the self-association *in subcutis*. The first reported crystal structure of an acylated insulin analogue was that of insulin detemir.[4] In the presence of resorcinol, this zinc insulin analogue crystallized as an R_6-dihexamer and the fatty acid side chains form hydrophobic bundles in the interface between the individual hexamers. The only available high-resolution crystal structure of an analogue with a reported higher-order self-association is that of $Lys^{B29}(N^{\varepsilon}$-lithocholyl) des(B30) human insulin (licI, PDB entry 1UZ9). In this structure, licI crystallized with *m*-cresol as an R_6-hexamer with a unique arrangement of the bile acid side chains forming interactions with the side chains from the neighboring hexamers.[9] The structure hints at the involvement of intersubunit contacts between the acylation side chains and specific residues involved in the allosteric changes, and thereby the importance of the shift from the R-state to the T-state of protomers. Until now, all the published crystal structures of acylated insulin analogues are in the R-state, in the presence of phenolic derivatives, and hence do not provide evidence of the molecular mechanism behind self-association.

The aim of this study is to elucidate the molecular mechanism behind self-association of the insulin analogue ωchI. The analogue is formulated under R_6-hexamer stabilizing conditions with Zn(II) and phenol as excipients. In vitro, we mimic the process of subcutaneous self-association upon phenol diffusion by removing phenol chromatographically.[12] The self-association process is studied primarily by applying small angle X-ray scattering (SAXS) and supplemented by a variety of biophysical and structural techniques. Using these methods, we provide direct structural evidence of the self-association mechanism of a modern insulin analogue belonging to the class of long-acting insulin molecules. We show that initially, in the formulation, an equilibrium exists between R_6-hexamers and $R_3T_3T_3R_3$-dihexamers. Subsequently, upon the removal of phenol, the R_6-hexamers and $R_3T_3T_3R_3$-intermediates rearrange to form T_6T_6-dihexamers, which are the building blocks of the process. In the final state, the long straight rodlike multihexamers consist of T_6-hexamers that have been formed via an initial formation of T_6T_6-dihexamers, with the latter further associating into the multihexamer. This mechanism is very similar to that proposed by Steensgaard et al. (DOI 10.1021/bi3008609) for the self-association process of another acylated analogue, insulin degludec, for which marketing approval has recently been filed. This points to common effects of these modifications on the insulin suprahexameric self-association upon acylation with fatty acid side chains.

EXPERIMENTAL PROCEDURES

Preparation and Formulation of Insulin Analogues. The acylated insulin analogue, ωchI, was synthesized and purified as previously described by Havelund.[8,27] The preparation was composed of 600 or 1200 μM insulin analogue, 0, 3, or 6 Zn(II) ions per hexamer, 10 mM Tris-HCl (pH 7.4), and 32 mM phenol, which is similar to the standard composition for clinical use. The samples contain different numbers of Zn(II) ions per hexamer insulin molecule (e.g., ωchI_3Zn). Phenol is used as a preservative in the formulation and to keep the insulin hexamers in the R-state, which inhibits formation of larger oligomers.

Triggering of Self-Association. The self-association process was started by removal of phenol via a short desalting column (NAP5 from GE Healthcare). A sample volume of 300 μL at a protein concentration of 1200 μM was loaded onto the column and was eluted by addition of elution buffer in 100 μL fractions. The concentration of the protein fractions was measured with a NanoDrop ND-1000 instrument using an extinction coefficient of 6200 M^{-1} cm^{-1} and a molar mass (MM) of 5800 Da. Fractions were pooled to obtain the desired concentration. For CD, asymmetrical flow-field flow fractionation (AF4), and small angle X-ray scattering (SAXS) experiments, a final concentration of approximately 750 μM (4.5 mg/mL) was used. For dynamic light scattering (DLS), varying concentrations were used. The elution buffer contained 10 mM Tris (pH 7.4), 140 mM NaCl, 0.01% NaN_3, and 0.6 mM phenol.[12] Buffer exchange was conducted at 25 °C.

The buffer exchange was performed as quickly as possible (applying gravity-induced flow-through); however, the time between the sample being applied to the NAP5 column and the sample being ready for the first measurement (with any of the methods described below) is ∼20 min. Thus, the first 20 min of the process was not monitored.

Dynamic Light Scattering (DLS). DLS was measured using a DynaPro Platereader from Wyatt Technology. For time-resolved DLS (TR-DLS) experiments, ωchI_6Zn was measured at concentrations of 75, 188, 240, 406, and 503 μM (0.45, 0.71, 1.44, 2.44, and 3.02 mg/mL, respectively). Every measurement was obtained by 30 acquisitions of 10−30 s. Data were analyzed using DYNAMICS. The cumulant analysis was used to estimate the average hydrodynamic radius (R_H) of the sample via the Stokes−Einstein equation:

$$R_H = \frac{k_B T}{6\pi\eta D_\tau} \quad (1)$$

where k_B is the Boltzmann constant, T is the temperature, η is the solvent viscosity, and D_τ is the translational diffusion coefficient. D_τ can be calculated from the decay rate ($\Gamma = D_\tau q^2$). Γ is calculated from the autocorrelation function, which is a measure of the intensity fluctuations of the particle in solution:

$$G(\tau) = A + Be^{-2\Gamma\tau} \quad (2)$$

where A and B are instrumental factors and τ is the time delay.

The autocorrelation function was used to separate the contributions from the different species present in the sample and obtain the R_H, polydispersity (%Pd), and mass (%Mass) of these species. The size range of particles was set to 1−1000 nm. The quality of the data was ensured by setting limits for the sum of squares error for cumulant fit (SOS), baseline, and amplitude to <10, ±0.01, and 1, respectively. The self-association process was followed until it had reached a plateau or until the signal was out of the range of the instrument. The experiments were performed at 25 °C.

Small Angle X-ray Scattering Data Collection and Primary Data Analysis. Three kinds of ωchI SAXS data were collected: (i) from samples in the pre-self-assembly formulation, (ii) from ωchI dihexamer, purified online during SAXS data collection, and (iii) from fully self-associated samples. Data were collected at the X33 beamline (Deutsches Elektronen-Synchrotron, Hamburg, Germany) and at the SWING beamline (Synchrotron Soleil, Paris, France).

Under pharmaceutically relevant conditions (i), SAXS data of ωchI_3Zn and ωchI_6Zn were collected in concentration

series with concentrations of 327−1173 μM (1.96−7.04 mg/mL) and 578−1154 μM (3.47−6.93 mg/mL), respectively. The data were collected at X33, at a wavelength of 1.5418 Å with an exposure time of 120 s in a q range of 0.0073−0.4945 Å^{-1} [where the scattering vector $q = 4\pi^* \sin(\theta)/\lambda$, and θ is half the scattering angle]. The data were collected on an image plate detector. Fully self-associated samples (iii) after 24, 32, and 40 h at a concentration of 1250 μM (7.5 mg/mL) were measured at X33 with a 4 × 30 s exposure time on a Pilatus detector and a q range of 0.0077−0.6237 Å^{-1}. The data were collected according to the standard procedures at X33 with one buffer measurement on either side of the protein measurements. The two-dimensional images were radial averaged to give the one-dimensional scattering function. The intensity was normalized by the intensity of the incoming beam and the protein concentration and was also corrected for the detector response (image plate detector only). Subsequently, the protein curves were subtracted by the averaged buffer data curve. The Guinier approximation was used to estimate the radius of gyration (R_g) and the intensity at zero angle (I_0). For high-MM assemblies in the fully associated samples, the Guinier assumption is compromised within the range of collected data; hence, indirect Fourier transformation was used to obtain estimates of R_g and I_0. The MM of the sample, which is proportional to I_0, was estimated via comparison to extrapolated I_0 values from standard measurements of BSA. The data of the formulation samples and of the fully self-associated samples are shown in sections S1 and S3 of the Supporting Information.

The online HPLC/SAXS setup at the SWING beamline[28] permitted collection of pure ωchI dihexamer data (ii) after elution on a Superdex 200 column in the system described in section S2 of the Supporting Information. Scattering data from the fully self-associated state (after 48 h) (iii) were also collected at SWING as described for the dihexameric sample, but without using a SEC column (section S3 of the Supporting Information).

Data analysis was conducted using appropriate programs from the ATSAS suite[29] using the default program input parameters if nothing else is stated. PRIMUS was used to perform background subtractions (however, MatLab and local software were used for samples collected at SWING) and Guinier analysis;[30] GNOM to obtain pair distance distribution functions [$P(r)$] by applying the indirect Fourier transform, which also yields the maximal distance in the particle (D_{max});[31] Crysol to evaluate solution scattering from model PDB structures and fit to experimental SAXS data;[32] Oligomer to fit an experimental scattering curve from a multicomponent mixture with a linear combination of form factors calculated from model protein structures also providing estimated volume fractions of the respective structures used;[30] and Dammin to perform ab initio modeling on selected data curves.[33]

Ab Initio Modeling from SAXS Data. Ab initio modeling with Dammin[33] was performed on the SAXS data of the pure ωchI_3Zn dihexamer (section S2 of the Supporting Information) and the fully self-associated sample after 48 h (section S3 of the Supporting Information). For modeling of the pure dihexamer, the $P(r)$ function from a q range of 0.0124−0.3003 Å^{-1} with a D_{max} of 78 Å was calculated using GNOM[31] and subsequentely used for input in Dammin version 51. The initial search volume was a cylinder with a radius of 60 Å and a height of 90 Å. Ten individual models were calculated using the basic simulated annealing protocol in Dammin. Damaver[34] was used to compare the individual models, and all models had normalized spatial discrepancy (NSD) values of <0.6 and were consequently included in the averaging and filtering procedures.

For modeling of the fully self-associated sample after 48 h, the $P(r)$ function from a q range of 0.0054−0.1325 Å^{-1} with a D_{max} of 705 Å was calculated using GNOM[31] and used for input in Dammin. All the points in the low q range were included, but the high q range containing the Bragg peak was cut away. The initial search volume was evaluated from the $P(r)$ function, where the characteristic shape suggests that the scatterer can roughly be described as a cylinder with a 100 Å radius and a 900 Å height. Ten individual models were calculated using the basic simulated annealing protocol in Dammin. Damaver was used to compare the individual models, and all models had NSD values of <0.7 and were included in the averaged and filtered models.

PDB Models for Crysol and Oligomer. PDB entries 1EV3, 1TRZ, and 1MSO were used to create models resembling ωchI in the three conformational states, R_6, T_3R_3, and T_6, respectively. "Dummy" residues were built in to mimic the fatty acid side chain (and phenol in R_6) and to obtain a MM close to $MM_{\omega chI}$, including phenol and Zn(II) when appropriate.

For pre-self-association SAXS data (i) and pure dihexamer data (ii), PDB models of hexamers and dihexamers in the T_6-state, the T_3R_3-state, and the R_6-state were built (Figure 1A,B). Models of monomers, dimers, trimers, tetramers, and pentamers were only prepared in the T_6-state and the R_6-state. For fully self-associated samples (iii), rod-shaped T_6-oligomers were made from one hexamer up to 30 hexamers by translation along the central axis in the insulin hexamer with a distance of 35.1 Å (Figure 1C). A search of the PDB[35] revealed that insulin hexamers have varying heights along the 3-fold Zn(II) axis ranging from 33.5 to 39.6 Å, which would result in likewise varying interhexameric distances. All insulin structures in the rhombohedral $R3$ space group deposited at present were investigated. The interhexameric distances fall into three groups, which corresponded to a grouping of the three different hexameric states of insulin, T_6, T_3R_3, and R_6, with corresponding average distances of 33.9 ± 0.2 Å (17 crystal structures), 37.2 ± 0.8 Å (26 crystal structures), and 39.3 ± 0.3 Å (4 crystal structures), respectively. Hence T_6T_6-dihexamers were created with interhexameric distances of 33−35 Å, $R_3T_3T_3R_3$- and $T_3R_3R_3T_3$-dihexamers with interhexameric distances of 33−39 Å, and R_6R_6-dihexamers with interhexameric distances of 38−41 Å; finally, one T_6-tetrahexamer was created with distances of 35.1 Å (the distance derived from the Bragg peak in the fully self-associated sample). The dihexamers were created by rotation and translation along the 3-fold symmetry axis in the hexamer (Figure 1B). All manipulations of structures were performed in PyMol.[36]

Crysol and Oligomer Analysis of SAXS Data. Theoretical scattering from the prepared PDB models was calculated with Crysol[32] using the default parameters, but using 256 points in the reciprocal space grid. The calculated form factors from the models were used as input for Oligomer.[29] Oligomer analysis was conducted with appropriate sets of insulin oligomers (Figure 1A−C). Models in all three allosteric states were included in the analysis, thus avoiding making subjective choices prior to the tests. The form factor sets were gradually reduced to narrow the possible models and to compare the fits from the T_6-, T_3R_3-, and R_6-states. For the pre-self-association samples (i) and the pure dihexamer (ii), several combinations

dx.doi.org/10.1021/bi3008615 | Biochemistry 2013, 52, 282−294

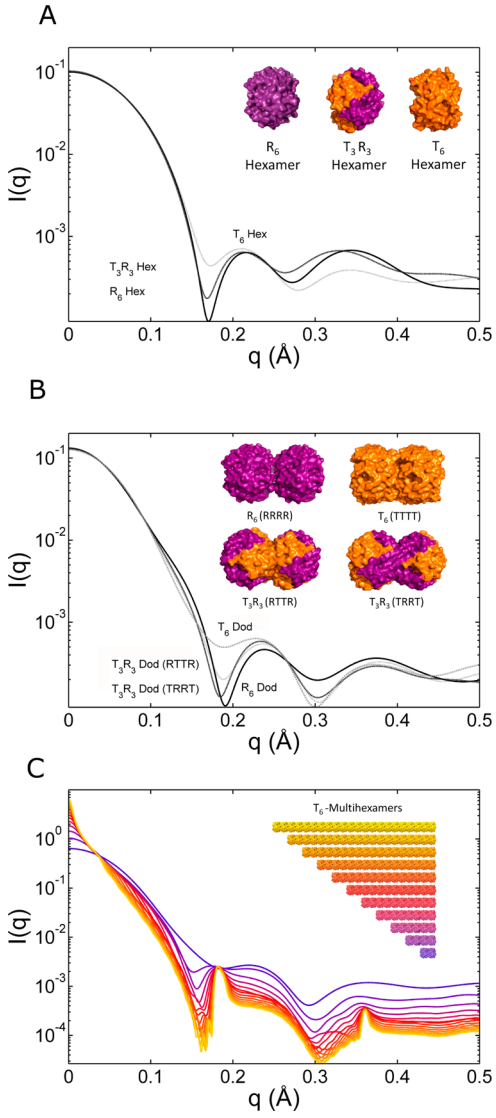


Figure 1. Examples of the protein models and their corresponding calculated scattering curves used as input in Oligomer. Insulin hexamers in the R_6-, T_3R_3-, and T_6-states (A), dihexamers in the R_6R_6-, $R_3T_3T_3R_3$-, $T_3R_3R_3T_3$-, and T_6T_6-states (B), and finally a series of even-numbered multihexamers in the T_6-state (C) are shown. The insulin monomers are colored according to their allosteric state: purple for the R-state and orange for the T-state.

of form factors were tested: all hexamers and dihexamers (a); subsequently, the sets were reduced to contain (b) only R_6R_6-dihexamers, (c_i) only internal $R_3T_3T_3R_3$-dihexamers, (c_e) only external $T_3R_3R_3T_3$-dihexamers, or (d) only T_6T_6-dihexamers. In all sets, an R_6-hexamer and a T_6-tetrahexamer were included. The $R_3T_3T_3R_3$-dihexamers were furthermore combined with monomers, dimers, trimers, tetramers, pentamers, and hexamers of the T- and R-states (e), which was subsequently reduced to include only $R_3T_3T_3R_3$-dihexamers, R_6-hexamers, and T_2-dimers (f). For analysis of the fully self-associated samples (iii), the following sets were tested: (g) R_6-hexamer, $R_3T_3T_3R_3$-dihexamers, and T_6-multihexamers of 1−30 hexamers, (h) $R_3T_3T_3R_3$-dihexamers and only even-numbered T_6-multihexamers of 2−30 hexamers, (i) only even-numbered T_6-multihexamers of 2−30 hexamers, and (j) only even-numbered T_6-multihexamers of 2−14 plus 30 hexamers. A constant (1000) was always included as an additional component in the Oligomer analysis.

Asymmetrical Flow-Field Flow Fractionation. Asymmetrical flow-field flow fractionation (AF4) was conducted with the analogue samples prior to (see section S5 of the Supporting Information) and during self-association. AF4 was performed on an Eclipse AF4 instrument (Wyatt Technology) integrated with an Agilent 1200 series HPLC system further coupled to a multiangle light scattering (MALS) detector. The channel length was 145 mm with a spacer height of 490 μm. Polyethersulfone (PES) membranes with a cutoff of 3 kDa were used to ensure that all states of insulin, from monomers to larger oligomers, would be separated and to prevent molecules from passing through or sticking to the membrane. The channel flow (F_c) rate was 1.0 mL/min. The focusing step was performed for 4 min with a focusing rate of 1.5 mL/min for all experiments. For analysis of ωchI_6Zn during self-association (commenced as described above), the eluent consisted of 0.6 mM phenol, 10 mM Tris-HCl (pH 7.4), 140 mM NaCl, and 0.01% NaN_3. The protein concentration was 770 μM (4.6 mg/mL). F_c was set to 1 mL/min, and F_x was a linear gradient of the cross-flow from 3 to 0 mL/min for 30 min. A blank sample was measured and used for baseline subtraction prior to analysis. Cobalt(III) insulin [Co(III), hexameric insulin, 35 kDa] was used as a marker.[37]

MALS in Connection with SEC and AF4. MALS was performed in connection with size exclusion chromatography (SEC) and AF4 runs of ωchI to obtain the weight-average molecular mass (M_W) of the species being separated. A DAWN HELEOS MALS instrument with 18 photodetectors was connected to the Agilent HPLC and the Eclipse system, along with a differential refractometer (Optilab rEX) and a DLS detector (WyattQELS) (all instruments except the HPLC system were from Wyatt Technology). The Wyatt COMET was turned on prior to all experiments to clean the flow cell before light scattering experiments were performed. The signal from UV_{276}, refractive index (RI), DLS, and MALS was collected continuously.

In the presence of high concentrations of phenol, RI was used instead of UV to calculate the concentration of the sample.

Calculation of M_W was performed with ASTRA (Wyatt Technology) using a refractive index increment (dn/dc) of 0.185 mL g^{-1} and Zimm's equation:

$$K^* \frac{c}{R_\theta} = \frac{1}{M_W P_\theta} + 2A_2 c \tag{3}$$

where c is the concentration of the solute (grams per milliliter), M_W is the weight-average molecular mass, R_θ is the Rayleigh ratio, A_2 is the second virial coefficient, and P_θ is the scattering function. The optical constant, K^*, can be calculated as

$$K^* = 4\pi^2 n_0^2 \left(\frac{dn}{dc}\right)^2 N_A^{-1} \lambda_0^{-4} \tag{4}$$

where n_0 is the refractive index of the solvent, dn/dc is the refractive index increment, N_A is Avogadro's number, and λ_0 is the wavelength of the light in vacuum. M_W was determined only in the chromatographic regions where both UV_{276} (or RI) and the MALS signal were sufficiently high. An extinction coefficient of 1037.84 g^{-1} mL cm^{-1} was used for calculation of the concentration of ωchI. The recovery of the sample was calculated from the injected and recovered mass. The recovery of the samples was always >90%.

dx.doi.org/10.1021/bi3008615 | *Biochemistry* 2013, 52, 282−294

The radius of gyration (R_g), also known as the root-mean-square radius, was calculated through the scattering function, P_θ:

$$P_\theta = 1 - \frac{q^2 R_g^{\ 2}}{3} \quad (5)$$

where q is the scattering vector. Finally, the DLS signal was used to determine the R_H of the sample by the Stokes−Einstein equation, which relates to spherical particles (see eq 1). It was not possible to calculate the R_H values for all the samples. Samples with molecular masses of >1000 kDa did not give a reliable DLS signal.

Near-UV Circular Dichroism (NUV-CD). Dimerization of the insulin analogue and interconversion among the allosteric states of the insulin hexamer (R_6 to T_6 through T_3R_3) were measured by CD in the near-UV range of 245−350 nm. The dimerization was followed at the tyrosine band at 276 nm, while the change from the R_6-state to the T_6-state was followed by a change in the disulfide band at 251 nm.[10] The ellipticity constant, $\Delta\varepsilon$, was normalized to the molar concentration of the protein. CD measurements were taken on a Jasco J-715 spectrometer calibrated with (+)-10-camphorsulfonic acid. The path length of the cuvette was 0.05 cm. A concentration series of ωchI was measured in a Tris-HCl-buffered solution at pH 7.4 with concentrations from 4 to 1800 μM. A zinc titration was performed at a protein concentration of 600 μM, where Zn(II), as zinc acetate, was titrated in prior to every measurement at concentrations of 0−0.69 mM. Finally, a phenol scan of ωchI_6Zn was performed in protein preparations [600 μM, 6 Zn(II) ions per 6 insulin monomers] containing phenol (0−30 mM). Time resolved CD (TR-CD) data were collected in intervals of 20 min over 17 h. The concentration before self-association was 1070 μM and during self-association was ∼800 μM. During self-association, the process was monitored in two buffers, containing either 0 or 0.6 mM phenol.

RESULTS

The acylated insulin analogue, ωchI, has been analyzed in solution immediately prior to self-association, in its fully self-associated state, and during self-association. The analogue has been investigated under experimental conditions similar to the formulations used for clinical trials for this class of analogues, i.e., including Zn(II) and using phenol as a preservative, which prevent self-association. In the majority of our experiments, a ratio of six Zn(II) ions per six insulin monomers (ωchI_6Zn) was used in the formulation. In the characterization of ωchI in the formulation prior to self-association, ratios of zero and three Zn(II) per six insulin monomers (ωchI_0Zn and ωchI_3Zn, respectively) were also used.

SAXS Solution Analysis Prior to Self-Association. When attempting to understand the molecular mechanism behind self-association, we need to characterize the pre-self-association state. In accordance, we collected SAXS data from concentration series of ωchI_3Zn and ωchI_6Zn in the presence of phenol. Zinc is necessary for the assembly into insulin hexamers and thus is required for further assembly into the multihexamers, reported here. The SAXS data of ωchI_3Zn and ωchI_6Zn showed a concentration-dependent increase in scattering intensity at low scattering angles, which suggests the presence of several oligomers in a concentration-dependent equilibrium. The behavior of ωchI_6Zn was qualitatively the same as that of ωchI_3Zn, and here we specify the observations for ωchI_3Zn only. The data sets were thus fit to a linear combination of a set of theoretical scattering curves using Oligomer[29] as explained for the self-associated state (vide infra). For the pre-self-association samples (i), form factor sets a−f were used (Figures 1B and 3C and Materials and Methods for details about the tested set of models). No individual model fit the data better than a combination of the form factors. This may be an effect of including more degrees of freedom, which will always give a better fit to the data. Hence, the Occam's razor approach was applied in the Oligomer analysis, where more and more form factors were removed from the input. Theoretical scattering from $R_3T_3T_3R_3$-dihexamers fit significantly better than those from T_6- or R_6-dihexamers. The best fits to the entire concentration series with the fewest form factors were a combination of form factors for dimers, hexamers, and $R_3T_3T_3R_3$-dihexamers (with an interhexameric distance of 36 Å) (f). There was no difference between the use of T_2-/R_2-dimers or T_6-/R_6-hexamers (evaluated by visual inspection and equal χ values), and hence, the analysis is not conclusive with regard to the allosteric state of individual dimers and hexamers. Within the concentration series of ωchI_3Zn, the volume fractions of $R_3T_3T_3R_3$-dihexamers increased from 28 to 86%, that of R_6-hexamers (or T_6-hexamers) decreased from 37 to 0%, and that of T_6-dimers decreased from 35 to 14% with an increasing protein concentration. All individual fits (section S1) and a table of calculated volume fractions from all samples (Table S1) are given in the Supporting Information.

SEC- and SAXS-Based Visualization of Isolated ωchI Dihexamers in the Pre-Self-Associated State. The online HPLC at the SWING beamline at Synchrotron Soleil[28] facilitated the collection of SAXS data from the pure ωchI_3Zn dihexamer (ii) during a SEC run. The R_g values were calculated via the Guinier approximation, resulting in two levels (27.2 and 26.8 Å) through the chromatogram of the eluting sample, indicating larger particles being present at the front of the chromatographic peak (section S2 of the Supporting Information). The data curves resulting in the lower R_g values were merged, and the resulting data curve was fit to several theoretical curves from R_6R_6-, T_6T_6-, $T_3R_3R_3T_3$-, and $R_3T_3T_3R_3$-dihexamer PDB models (for details, see Experimental Procedures). The best fit from a single model was obtained for an $R_3T_3T_3R_3$-dihexamer with an interhexameric distance of 36 Å (Figure 2A), in accordance with the result from the non-gel-filtered samples also revealing the presence of the same dihexamers in equilibrium with dimers and hexamers. Also, the theoretical and experimentally derived R_g values agree (Tables 1 and 2), and the calculated molecular mass (MM) corresponded to 72.1 kDa as expected. The SAXS curves collected at earlier elution times (average R_g of 27.2 Å) could be fit only with a combination of form factors of different models. Different combinations of form factors (a−d) were tested. The best fit showed that a combination of 96% $R_3T_3T_3R_3$-dihexamer (again with an interhexameric distance of 36 Å) and 4% T_6-tetrahexamer fit the data well (section S2 of the Supporting Information), also in accordance with the observed slightly larger average dimensions and the experimentally observed small increase in MM (74 kDa). The $P(r)$ of the ωchI_3Zn dihexamer is shown in Figure 2C, with a D_{max} of 78 Å and a typical distance of ∼32 Å.

The merged SAXS dihexamer data curve was used for ab initio modeling of the dihexamer by applying Dammin.[33] The averaged and filtered model is shown in Figure 3A. The model

dx.doi.org/10.1021/bi3008615 | Biochemistry 2013, 52, 282−294

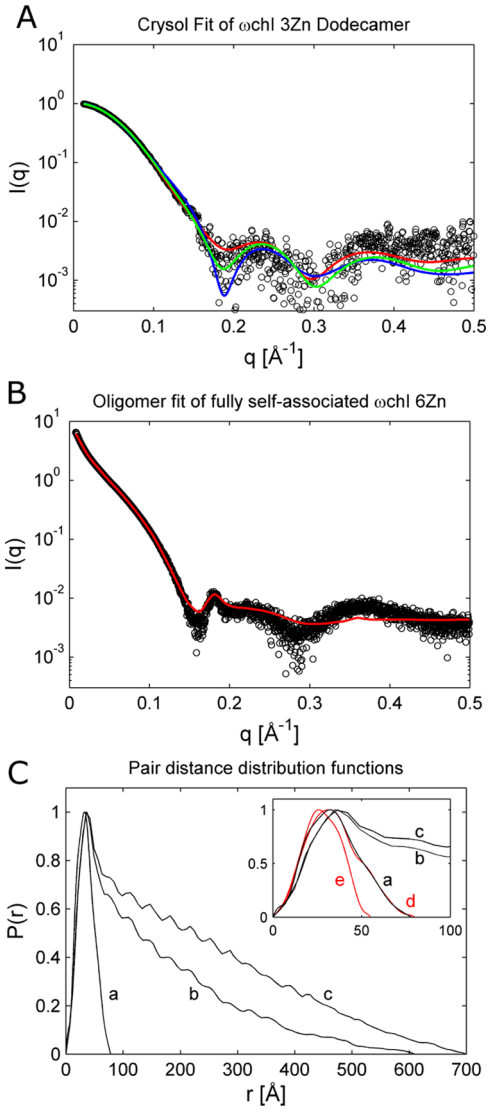


Figure 2. Experimental SAXS data collected from ωchI samples. (A) Crysol fits to the pure ωchI_3Zn dihexamer are shown. Solid lines are theoretical scattering curves from the T_6T_6-dihexamer (red; χ^2 = 3.5), the $R_3T_3T_3R_3$-dihexamer (green; χ^2 = 5.9), and the R_6R_6-dihexamer (blue; χ^2 = 9.4). (B) An Oligomer fit to data from a fully self-associated ωchI_6Zn sample after 24 h is shown as a solid red line (χ^2 = 1.9), and black circles are experimental data points. (C) Pair distance distribution functions of the pure ωchI_3Zn dihexamer (a) and two self-associated samples after 24 h (b) and 48 h (c) with a close-up of the first 100 Å as an inset. Theoretical pair distance distribution functions for an R_6-hexamer (d) and an $R_3T_3T_3R_3$-dihexamer (e) are included for comparison (red). The pair distance distribution functions are normalized to the same maximal value for better comparison of the shapes.

has the characteristic shape of a dihexamer built by two, almost spherical hexamers. The model has a diameter of ~50 Å and a length of ~80 Å, which fit with the dimensions of a dihexamer and correspond well with the calculated D_{max} of 78 Å. Two cavities are observed in the model, one in each hexameric unit.

Complementary Biophysical Solution Analysis Prior to Self-Association. To supplement and substantiate the conclusions based on our SAXS data, we analyzed the pre-self-associated state of ωchI with DLS, SEC, AF4, MALS, and CD. Batch DLS measurements resulted in estimates of the hydrodynamic radius (R_H) of 16, 26, and 26 Å for ωchI_0Zn, ωchI_3Zn, and ωchI_6Zn, respectively (Table 1), with corresponding polydispersities of 24, 16, and 15%, respectively. In Table 2, experimentally and theoretically derived values of R_H and R_g are presented for hexamers and dihexamers. Evidently, the R_H of the two samples in the presence of Zn(II) corresponded to a hexamer, and the sample without Zn(II) was on average smaller than a hexamer.

In addition, we applied the two separation techniques SEC and AF4, both in connection with MALS. The results are summarized in Table 1, and the chromatograms and fractograms with the corresponding calculated weight-average MM (M_W) can be seen in sections S4 and S5 of the Supporting Information. The zinc-free samples eluted as an asymmetrical peak with M_W distributions from 10 to 22 kDa and from 10 to 63 kDa for SEC and AF4, respectively. Zinc-containing samples eluted as single peaks with an M_W close to 72 kDa (Table 1). Clearly, the experimentally determined R_H and R_g from SEC and AF4 fit very well with the calculated values of insulin dihexamers determined by HYDROPRO[38] and Crysol[32] in Table 2.

Finally, the oligomerization and allosteric state of ωchI were investigated with NUV-CD (section S6 of the Supporting Information). We conclude that under the conditions used for the experiments in this paper ωchI is in a fully dimerized/hexamerized state; thus, all ωchI molecules are involved in dimer interactions, and no monomers are present in solution. Furthermore, a screen of phenol concentrations from 0 to 30 mM (section S6 of the Supporting Information) shows that in the absence of phenol, ωchI_6Zn is in the T_6-state, as previously published.[10,39] The ellipticity constant at 251 nm ($\Delta\varepsilon_{251}$) rapidly reaches saturation at 3 mM phenol at a value of −4.5 M^{-1} cm^{-1}. This value does not correspond to the full R_6-transition seen for human insulin (−6.5 M^{-1} cm^{-1}). This indicates that for all phenol-containing samples a partial allosteric transition has taken place, thus rendering the hexamers in the T_3R_3-state or in equilibrium between the T_3R_3-state and the R_6-state.

Crystal Structure in the Non-Self-Associative R-State. The crystal structure of ωchI can be found in the PDB to a resolution of 1.6 Å (entry 3ZU1). The ωchI dimer in the crystal forms essentially the same structure as human insulin in its R_6-complex with resorcinol. This is seen in Figure 4, where a ribbon representation of ωchI is superimposed on human insulin in the R_6-state (PDB entry 1EV3) and in the T_6-state (PDB entry 1MSO). The N-terminus of the B-chain adopts an α-helix in the R_6-structure. The fatty acid chain was not present in the PDB model, hence suggesting that the fatty acid does not adopt to a well-defined conformation in the structural state present in the crystals. Two Zn(II) ions with a distance of 15.9 Å were coordinated in a tetrahedral geometry by nitrogen in the three usual His^{B10} residues and one chloride ion along the symmetry axis of the hexamer. Resorcinol was bound the same way as in the human insulin structures in its allosteric binding pocket. This structure is not in the state necessary for self-association, which is also the case for all previously deposited crystal structures of acylated human insulin analogues.

SAXS Analysis of the Fully Self-Associated Sample. Self-association was initiated on samples of ωchI_6Zn via column-mediated buffer exchange, resulting in a final phenol concentration of 0.6 mM. Maintaining a low final phenol concentration was necessary to slow the rate of the process to study the process over time as discussed in a subsequent section.[12] The samples were subsequently stored at room

dx.doi.org/10.1021/bi3008615 | *Biochemistry* 2013, 52, 282−294

Table 1. Experimentally Derived Biophysical Parameters (R_H, R_g, and MM) Obtained from DLS, MALS, AF4, and SAXS Experiments[a]

	R_H (Å)				MM				
sample	batch DLS	SEC/MALS	AF4/MALS	R_g (Å), SAXS	theory	SAXS (I_0)	SAXS (Oligomer)	SEC−MALS	AF4−MALS
ωchI_0Zn	16	10−25	na	na	35.7	na	na	10−22	10−63
ωchI_3Zn	26	35	36	27	35.9	50	69[c]	72	74
ωchI_6Zn	26	35	67	26	36.1	46	64[d]	68	70
ωchI_3Zn[b]	na	na	na	26.8[e]	71.8	na	72[e]	na	na

[a]The given theoretical MM is that of a ωchI hexamer, including the appropriate number of zinc ions and six phenol molecules. The R_H measured with batch DLS is from a formulation of 1200 μM insulin analogue and represents the species with a mass that is >99% of that of the entire sample. na means not available. [b]Dihexameric state of ωchI_3Zn purified via SEC. [c]Equilibrium between hexamer and dihexamer shown by concentration series. The value represents a concentration of 1173 μM, with a population of 11% hexamer and 87% dihexamer according to Oligomer. [d]Equilibrium between hexamer and dihexamer shown by concentration series. The value represents a concentration of 1154 μM, with a population of 22% hexamer and 78% dihexamer according to Oligomer. [e]SAXS data of the ωchI_6Zn dihexamer were obtained by SAXS with online SEC at the SWING beamline at the Synchrotron Soleil. R_g is estimated from Guinier, and MM is calculated from Oligomer.

Table 2. Calculated R_H and R_g Values of Insulin Hexamers and Dihexamers in the R_6-, T_3R_3-, and T_6-States Using HYDROPRO[29] and Crysol[30]

initial PDB ID prior to modification	oligomeric state	allosteric state	interhexameric distance (Å)	HYDROPRO R_H (Å)	HYDROPRO R_g (Å)	Crysol R_g (Å)
1EV3	hexamer	R_6		25.6	19.4	19.5
1EV3	dihexamer	R_6	39.5	33.8	27.8	27.9
1TRZ	dihexamer	T_3R_3	35	32.9	26.5	26.4
1TRZ	dihexamer	T_3R_3	37	33.2	26.9	27.1
1MSO	hexamer	T_6		26.4	20	19.8
1MSO	dihexamer	T_6	35.1	34.3	26.9	26.8

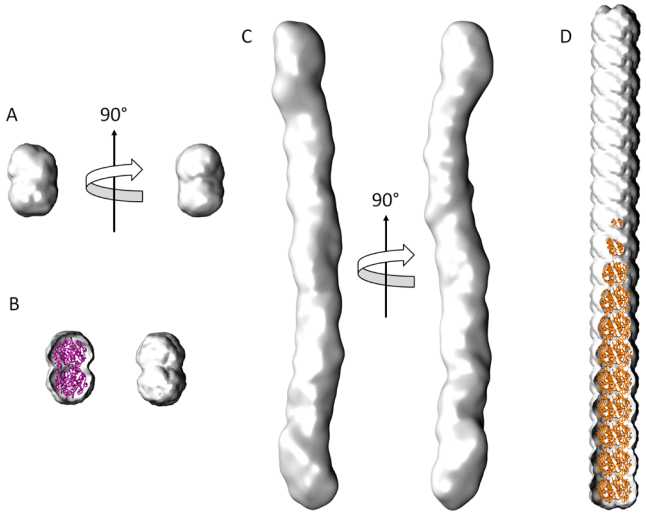


Figure 3. Ab initio models of the SEC-purified dihexamer (A) and the fully self-associated sample after 48 h (C). The models are averaged and filtered from 10 individual models calculated with Dammin.[33] Two models of an $R_3T_3T_3R_3$-dihexamer (B) created from crystal structures of PDB entries 1MS0 and 1TRZ and a T_6-multihexamer (18 hexamers) (D) are shown for comparison. This figure was prepared with VMD.

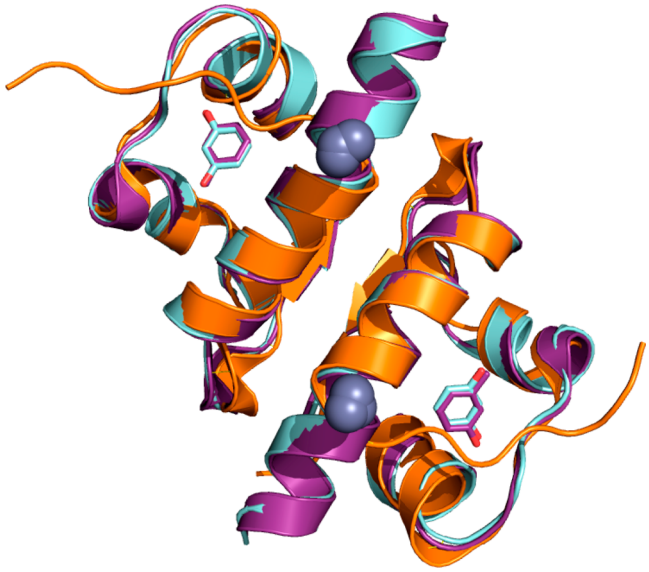


Figure 4. X-ray structure of ωchI in the R_6-conformation (cyan, PDB entry 3ZU1). The dimer in the asymmetrical unit is superimposed with the R_6-crystal structure (purple, PDB entry 1EV3, rmsd of 0.316 Å) and the T_6-crystal structure of HI (orange, PDB entry 1MSO, rmsd of 1.056 Å). Two Zn(II) ions (gray) are located on the 3-fold symmetry axis, which is oriented vertically. The two resorcinol molecules are bound in the hydrophobic pocket as in PDB entry 1EV3. The N-terminal part of the B-chain (B1−B8) adopts an extended conformation in the T_6-state.

temperature, and SAXS data were collected 24−48 h after the initiation of association. Scattering data were measured at two different beamlines in two different sample cell types but resulted in the same distinct scattering curves; i.e., the sample environment during data collection did not influence the data. The curves revealed a steep increase at low angles signifying the formation of very long species in solution (Figure 2B and section S3 of the Supporting Information). Starting around 0.18 $Å^{-1}$ and toward higher scattering angles, another characteristic "round" feature is evident with two minima around 0.2 and 0.3 $Å^{-1}$. These features are also found in scattering curves from pre-self-associated samples, suggesting the preservation of several structural features during self-association (Figure 2A and section S1 of the Supporting

dx.doi.org/10.1021/bi3008615 | Biochemistry 2013, 52, 282−294

Information). Notably, a Bragg peak is present at a q value of 0.185 Å^{-1} in all samples, arising from a repeating distance within the scatterer of 35.1 ± 0.2 Å. Only the first-order Bragg peak is observed. The steep increase at low angles obstructs Guinier analysis of the data;[29] instead, we calculated R_g and I_0 (Table 1) values using indirect Fourier transformation performed with GNOM[31] and determined a pair distance distribution function [$P(r)$] to obtain an approximate maximal distance within the scatterers (D_{max}). D_{max} was estimated to be 610 and 705 Å for the 24−40 and 48 h samples, respectively. The resulting $P(r)$ of the self-associated samples, in particular, the sample after 48 h, has a characteristic shape with a peak at 35 Å continuing in a slope to 705 Å with ripples equally spaced by 35.1 Å (Figure 2C). The ripples directly result from the Bragg peak in the SAXS data curves. Such a scattering function and $P(r)$ function arise from rod-shaped particles with axial inhomogeneity[40] formed by units with a 35.1 Å repeated distance. At the same time, the round features in the data curve with the two minima suggested that the self-associated form consists of hexameric units.[11,41] The data collected after 48 h result from particles slightly longer than the 24−40 h samples but otherwise with the same dimensions and Bragg distance as earlier samples (24−40 h).

Ab initio modeling directly from the SAXS data using Dammin[33] resulted in a low-resolution model of the sample in the fully self-associated state (Figure 3C). The averaged and filtered model showed a long straight rod-shaped envelope. The average diameter of the particle was ∼50 Å and the length ∼620 Å, which are comparable to those of the constructed T_6-multihexamer PDB model (Figure 3D). The Dammin models appear slightly flexible around the central axis. These overall dimensions hence confirm the notion based on the raw data and the $P(r)$ function, suggesting that the self-associated state is formed by stacks of hexamer-sized units.

Given the observation of a repeating distance of 35.1 Å within the self-associated molecule, the allosteric form of the hexamer in the self-associated state could be determined. An examination of the insulin crystal structures deposited in the PDB revealed a clear relationship between the allosteric state and the maximal dimension of the hexamer along the crystallographic 3-fold axis. The interhexameric distances fell into three groups with average distances of 33.9 ± 0.2, 37.2 ± 0.8, and 39.3 ± 0.3 Å corresponding to the T_6-, T_3R_3-, and R_6-states of insulin, respectively. Interestingly, the distance corresponding to the scattering angle of the Bragg peak was 35.1 Å, which falls between the values of the T_6- and T_3R_3-states but most closely resembles the typical dimensions of T_6-hexamers. This thus strongly suggests that the self-associated species is a T_6-multihexamer.

As an alternative approach to reconstructing a model of the self-associated insulin analogue, the data sets were fit using Oligomer[29] to a linear combination of theoretical scattering curves calculated with Crysol[32] from modeled PDB structures of different T_6-multihexamers (Figure 1C and section S3 of the Supporting Information). The insulin crystal structures were modified to resemble ωchI, and the typical distance of 35.1 Å was used in construction of the linear multihexamers. Several sets of form factors (g−j) were used for input. As described above, the number of form factors was reduced to find the best fit, including the fewest possible form factors. The minimal form factor set that fit adequately to the data from the 24 h sample included only even-numbered T_6-multihexamers, resulting in 6% dimers, 51% octamers, 29% decamers, 4% 12-hexamers, 5% 14-hexamers, and 5% 30-hexamers (Figure 2B). Because fitting of scattering curves from individual single multihexamers to the data was not possible, this strongly suggests that the samples are heterogeneous in length but apparently consist of different lengths of only the same type of multihexamers. There is a slight discrepancy between the theoretical scattering and the experimental data, which may be ascribed to some flexibility around the longest axis in the T_6-multihexamers, as suggested by the ab initio models (Figure 3C).

It should be noted that the SAXS data were collected with 0.6 mM phenol present in the buffer, which may restrict the extent of multihexamerization. In a phenol-free environment, the multihexamers may thus grow to larger sizes.

Solution Analysis during Self-Association. After the investigation of ωchI prior to and after self-association, elucidation of the dynamic process during self-association was undertaken. The process was initiated by buffer exchange as described above.

The transition between the allosteric states during self-association was monitored by time-resolved CD (see Figure 5)

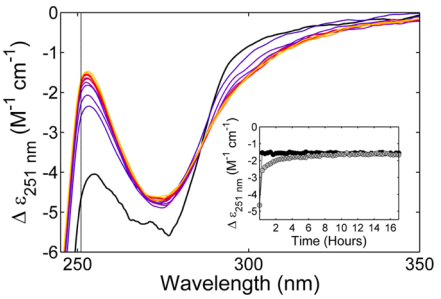


Figure 5. Self-association of ωchI_6Zn followed by TR-CD. The full NUV-CD spectra are shown. The allosteric change from R to T can be followed at 251 nm (marked with a vertical line). The measurements prior to self-assembly are shown in black (1070 μM). After buffer exchange over a NAP5 column, the conformational change can be followed in the sample (800 μM) (from purple to yellow). Here the samples from 0 to 8 and 16 h are shown, corresponding to the fractograms in Figure 7. The inset shows the development of $\Delta\varepsilon_{251}$ during the next 17 h in a buffer containing 0 (black circles) and 0.6 mM phenol (gray circles).

for ωchI_6Zn in buffers containing 0 and 0.6 mM phenol. For the cysteinyl band, a $\Delta\varepsilon_{251}$ of −4.6 M^{-1} cm^{-1} is found before phenol removal (1070 μM insulin, equal to 6.7 mg/mL). This value changes essentially immediately to a constant level at −1.6 M^{-1} cm^{-1} at an ωchI_6Zn concentration of 785 μM (4.9 mg/mL) in the absence of phenol. A different picture is seen when using the buffer with 0.6 mM phenol. The cysteinyl band is shown to gradually change from −2.6 to a stable level at −1.6 after approximately 10 h at a protein concentration of 833 μM (5.2 mg/mL) compared to the preassociation state of −4.6 M^{-1} cm^{-1} before phenol removal. This shows that the transition from T_3R_3 to T_6 in dihexamers is slowed in the presence of even small amounts of phenol. In both buffer systems, the initial $\Delta\varepsilon_{276}$ of approximately −4.8 M^{-1} cm^{-1} indicates full dimerization as expected (section S6 of the Supporting Information). The 260−280 nm region is disturbed by the absorbance of phenol in the sample prior to self-association and can thus not be used to investigate any changes around the tyrosine upon self-association.

dx.doi.org/10.1021/bi3008615 | Biochemistry 2013, 52, 282−294

The self-association kinetics of ωchI were also monitored with time-resolved DLS (TR-DLS). For ωchI_6Zn in a buffer containing 0.6 mM phenol (as in the TR-CD experiments), the increase in R_H during self-association could be followed over a period of ≥1 day, whereas preliminary tests showed that samples with a lower zinc content per hexamer reached a high-MM plateau very quickly, before any DLS measurements could be performed. Hence, a concentration series was analyzed for only ωchI_6Zn. In Figure 6, the growth of the cumulant R_H is shown. The five concentrations (75, 118, 240, 406, and 503 μM) reached an increasing size plateau at 18, 21, 27, 52, and 63 nm, respectively. The growth started immediately, and no lag phase was observed in the TR-DLS data. The initial rate of the cumulant R_H was found to be directly proportional to the concentration of the sample (inset of Figure 6). Regularization analysis showed the existence of differently sized species within the sample (data not shown).

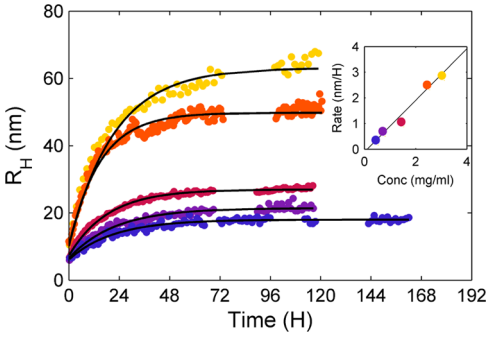


Figure 6. TR-DLS measurements of ωchI_6Zn are shown where the growth of R_H is followed over several days. The development of the cumulant R_H of analogue ωchI is shown at the following concentrations: 75 (a, blue), 118 (b, violet), 240 (c, red), 406 (d, orange), and 503 μM (e, yellow). The growth of R_H was fit to the function $f(t) = y_0 + a[1 - \exp(-bt)]$. y_0 corresponds to the intercept with the y-axis; a is the increase in R_H at the plateau (minus y_0), and b is the initial slope. The initial rate is plotted vs concentration in the inset.

We also used AF4 in connection with MALS to follow the development during self-association in a time-resolved manner. After self-association had commenced, ωchI_6Zn was injected once every hour for 17 h to follow the change in the sample composition. The fractograms can be seen in Figure 7A. The first sample (0 h) eluted with a pronounced peak with a MM of 72 kDa and a broad slope toward higher MM where at least three individual peaks can be distinguished. In the subsequent samples, the area under the 72 kDa peak decreased while the high-MM portion increased. Finally after 17 h, the majority of the sample eluted at the end of the cross-flow gradient at 36 min with an average MM of ∼1200 kDa. In Figure 7B, the decrease in UV_{276} of the dihexameric species and the increase of the high-MM species are plotted as a function of time. The amount of dihexameric species decreases as a decaying function, while the growth of the high-MM associates follows a sigmoidal shape. The lag phase of the high-MM associates, which is not observed in the TR-DLS experiments, is formed presumably because of the buildup of smaller oligomeric species before large high-MM associates are formed. Notably, according to the AF4 data, no hexamers are present after initiation of the self-association, which leaves the dihexamer as the smallest species present in solution.

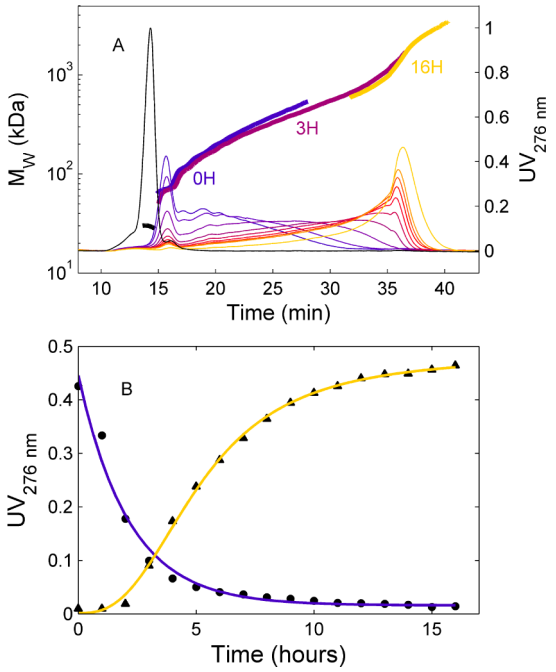


Figure 7. (A) Fractograms show the developing self-association of ωchI_6Zn (770 μM). The sample was injected once every hour 17 times. The first curve at 0 h (blue) is the first measurement 20 min after the buffer exchange on the NAP5 buffer exchange column had commenced. Injections from 0 to 8 and 16 h (from blue to yellow) corresponding to the CD spectra in Figure 5 are shown. The Co(III) insulin[37] is used as reference for the hexameric state of insulin. M_W is shown as a thicker line in the same color as the respective UV_{276} signal. (B) Decay of the dihexameric species over time fit to exponential decaying function $f(t) = y_0 + a \exp(-bt)$ (blue) and growth of high-MM associates fit to logistic sigmoidal function $f(t) = y_0 + a/[1 + (t/t_0)b]$ (yellow).

DISCUSSION

It is of great and evident interest to understand the molecular mechanism behind the self-association observed for the novel classes of chemically modified insulin analogues. Previous studies have shown that the self-association of ωchI is triggered by the removal of phenol,[12,8] but it has remained unclear whether the self-assembly is initiated by the expected associated conformational change of the hexamer from the R_6-state to the T_6-state or another effect caused by phenol removal. Our experiments reveal structural and biophysical evidence regarding one particular human insulin analogue acylated with ω-carboxyheptadecanoyl, ωchI. In the following paper (DOI 10.1021/bi3008609), similar findings are reported for insulin degludec, a clinically relevant analogue with a different fatty acid modification.

Solution Structure of the ωchI Dihexamer Revealed by SAXS. Extensive SAXS analysis was applied to ωchI prior to self-association in the presence of phenol. We concluded that in formulation with zinc, ωchI exists in equilibrium between hexamers and dihexamers (Table 1), with small amounts of dimers at low concentrations. A similar equilibrium is investigated thoroughly for insulin degludec in the following paper by Steensgaard et al. (DOI 10.1021/bi3008609), where the dependence on zinc, phenol, resorcinol, NaCl, and imidazole is studied.

It is evident from our analysis that the application of individual biophysical methods alone without the SAXS analysis

would have resulted in nonconclusive observations. Batch DLS measurements of ωchI_3Zn and ωchI_6Zn show an R_H of 2.6 nm, corresponding to the expected radius of a hexamer. In contrast, the two separation techniques, SEC and AF4, strongly indicate the formulations to be completely in a dihexameric state, where we see a single peak with an M_W very close to 72 kDa. The ab initio model built from the SAXS data of the online SEC-purified species conclusively confirms the presence of a dihexamer, revealing a structure of two hexamers stacking on top of each other. This is the first ever direct visualization of a dihexameric insulin analogue structure. Furthermore, the internal allosteric structure of the dihexamer was found to be $R_3T_3T_3R_3$, based on fits of theoretical scattering curves to the experimental data.

Finally, SAXS concentration series collected on zinc-containing ωchI clearly reveal a concentration-dependent equilibrium among dimers, hexamers, and $R_3T_3T_3R_3$-dihexamers, as also reported for insulin degludec (DOI 10.1021/bi3008609).

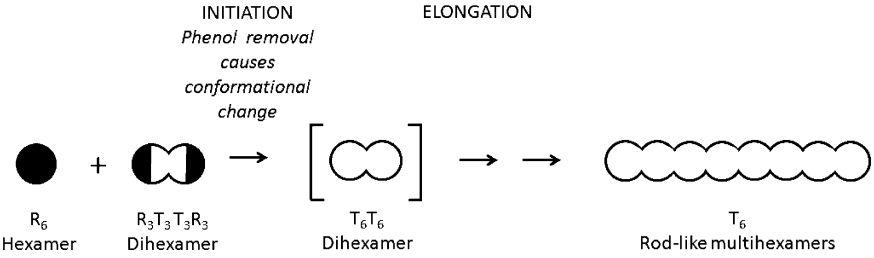
The $R_3T_3T_3R_3$-Dihexamer Exists in the Formulation in Equilibrium with Smaller Oligomers (in the presence of phenol). An important element in the self-association of ωchI is the conformational change that occurs within the individual hexamers. NUV-CD analysis with varying concentrations of protein and phenol revealed that even at the high phenol concentrations (32 mM) the analogue never reaches the $\Delta\varepsilon_{251}$ level of the R_6-state as previously reported for human insulin by Olsen et al.[39] At 30 mM phenol, the fraction of the R-state is ∼60% of the level of human insulin. Because the T-state is assumed to be the self-association prone state, which is confirmed by our SAXS analysis of the fully self-associated state, the most obvious interpretation of the data is that the sample is in the T_3R_3-state or in equilibrium between an R_6-hexamer and a dihexamer in the $R_3T_3T_3R_3$-form [R-ends are facing the solvent (Figure 1)]. This is strongly supported by the SAXS data, fitting well with theoretically calculated scattering curves from this type of oligomer, both for concentration series prior to self-association and for data from the SEC-purified dihexamer. From inspection of the theoretical scattering curves of all three kinds of hexamers (T_6, T_3R_3, and R_6) and four kinds of dihexamers (T_6T_6, $T_3R_3R_3T_3$, $R_3T_3T_3R_3$, and R_6R_6) in Figure 1, there is a clear difference between the T- and R-states of both oligomers at the minimum at 0.2 $Å^{-1}$. The R_6-state has a deep minimum, whereas the T_6-state has a shallow minimum; the T_3R_3-state is found between the R_6- and T_6-states. These structural features reflected in the scattering curves help in the identification of the conformational state of the insulin hexamer. Taken together, the characterization of ωchI in the formulation prior to stronger self-association shows that ωchI in the presence of three or more Zn(II) ions per six monomers primarily exists in the $R_3T_3T_3R_3$-dihexameric form in equilibrium with the hexamer in the R_6-form. In the following paper by Steensgaard et al. on insulin degludec (DOI 10.1021/bi3008609), the existence of both the $R_3T_3T_3R_3$-dihexamer and the R_6-hexamer was shown with SAXS, suggesting some minor difference in the self-association behavior between the two analogues.

The Self-Associated State Is a Rod-Shaped T_6-Multihexamer. The SAXS data of the fully self-associated sample show that this state consists of long rod-shaped particles with axial inhomogeneity. Importantly, the Bragg peak derived from frequently occurring real-space distances of 35.1 Å directly reveals a repeating distance within these rods. We can therefore conclude that the rods are built up by hexamers as the asymmetrical unit with an interhexameric distance of 35.1 Å, which also accounts for the inhomogeneity along the length axis. In accord with these data, we fit the SAXS data to combinations of rods of different lengths, built by such units. Extensive analysis of the available crystal structures of different allosteric forms of insulin hexamers suggests that the rods are built from T_6-multihexamers, which is confirmed by fitting to sets of modeled multihexamers applying variants of allosteric forms. We can thus, on the basis of SAXS solution data alone, conclude that the high-MM associates of ωchI are linear assemblies of T_6-multihexamers. Although this has been suggested previously, on the basis of the observation that phenol diffusion preceded oligomerization, this is the first ever direct evidence of the presence of such an oligomeric species. Another interesting finding is that SAXS data recorded at different late time points (24, 32, and 40 h) reveal no differences in the final association state. This suggests that the depotlike high-MM structures are stable and intact in vitro over the time range analyzed.

The individual hexamers in the high-MM associates are in the T_6-state, with slightly more space (35.1 Å) between the individual hexamers than in the insulin T_6-hexamers found in previously published crystal structures of human insulin (33.9 Å). One may speculate that the ω-carboxyheptadecanoyl side chains occupy a part of the space between the hexamers, which is very open in the T_6-state with direct access to the zinc binding site. An example of this type of structural arrangement is found in the R_6-crystal structure of insulin detemir (PDB entry 1XDA) where the fatty acid acylation side chains form a hydrophobic bundle in the interface between hexamers;[4] here the interhexameric distance is 39.6 Å. This structure, however, cannot be directly compared to that of self-associated ωchI because the detemir crystal structure is in the R_6-state and the fatty acid side chains do not contain the terminal carboxyl group. If the side chains occupy the interface between two T_6-hexamers, the carboxyl groups must be located very close to the center with the exposed zinc site.

The Crystal Structure (R_6) of ωchI Is Not Comparable to the Self-Associated State. A crystal structure of ωchI in the T_6-state would offer valuable insight into the type of interactions involved in self-association; however, crystallization of ωchI in the T_6-state, like that of other acylated self-associating insulin analogues, has not yet been possible. We can speculate that crystallization is complicated because of solution equilibria reported here. However, the crystal structure of ωchI in the R_6-state has been determined to a resolution of 1.6 Å (PDB entry 3ZU1). The interhexameric distance along the crystallographic 3-fold axis was 41 Å in this structure and hence does not correspond to the interhexameric distance in the self-associated state. Regrettably, no electron density of the acylated side chains was seen in the structure. In the R_6-structure of LicI,[9] the lithocholyl groups (100% occupancy) are evenly distributed around the hexamer and form interactions with Phe^{B1} from its own monomer and the lithocholyl group from the neighboring hexamer. We speculated that the interaction of Phe^{B1} with the lithocholyl group in the R_6-hexamer prevents self-association and thus allows crystallization; however, upon the transition to T_6, the displacement of Phe^{B1} allows interaction of the side chain with neighboring hexamers in the solution, thus allowing interactions possibly responsible for self-association.[9] It is not possible to predict a similar scheme

Scheme 1. Proposed Polymerization Mechanism for ωchI Self-Association[a]



[a]Prior to self-association in the presence of zinc, ωchI exists in equilibrium between primarily R_6-hexamers and $R_3T_3T_3R_3$-dihexamers (the R-state is represented in black and the T-state in white). Upon removal of phenol, a conformational change (initiation) renders ωchI in the T_6-state, which is the self-association prone state. T_6-dihexamers act as the smallest building block in the self-association into multihexamers (elongation). The fully self-associated multihexamers are long rods composed of T_6-hexamers as the repeating unit.

for ωchI, because of the lacking electron density of the side chains in the current crystal structure.

A Shift from $R_3T_3T_3R_3$-Dihexamers to T_6T_6-Dihexamers Triggers Self-Association. The self-association process of ωchI_6Zn was monitored via TR-DLS, TR-CD, and TR-AF4. The TR-CD experiment monitors the conformational change, which triggers the self-association process, whereas TR-DLS and TR-AF4 monitor the growth of the high-MM associates.

The TR-CD measurements confirm that ωchI adopts the tense state after phenol removal, which further confirms our previous conclusions that the high-MM associates are in the T_6-state, which is also the case for individual building blocks. Clearly, the addition of a small amount of phenol to the eluent slows the rate of self-association (Figure 5), presumably because of the resulting slower release of phenol from the ωchI hexamers. Previously reported SEC results confirm that in an isotonic eluent (without phenol) high-MM associates are formed essentially immediately.[12]

The TR-DLS measurements show that the rate of self-association and the final size of the high-MM assemblies are dependent on the concentration (see Figure 6).

Finally, TR-AF4 fractograms visualize the development of high-MM assemblies during self-association (Figure 7), and it is evident that a dihexamer is the smallest species present under these experimental conditions. Furthermore, the relative amount of dihexamers decreases over time, in accordance with a model in which dihexamers are incorporated into the growing self-associates. Thus, when phenol is removed from the environment, the equilibrium of R_6-hexamers and $R_3T_3T_3R_3$-dihexamers in the formulation shifts toward the formation of T_6T_6-dihexamers, which are then able to self-assemble into high-MM complexes consisting of T_6-multihexamers (Scheme 1).

Proposed Mechanism of ωchI Self-Association. The proposed mechanism of self-association for ωchI in the presence of Zn(II) is a polymerization of dihexamers, which elongate into multihexamers (Scheme 1). The mechanism may be comparable to downhill polymerization (irreversible polymerization), which consists of two steps, initiation and elongation, with different rate constants.[42] Prior to self-association, under formulation conditions, our data suggest that ωchI exists in equilibrium between an R_6-hexamer and $R_3T_3T_3R_3$-dihexamers with the R-state facing the environment. Removal of phenol triggers the R to T conversion of the ωchI oligomers into the T_6T_6-dihexamer (initiation), which is thought to be the "nucleus" of the process and the smallest building block. Self-association (elongation) follows, as seen in DLS and SAXS data, and long rod-shaped multihexamers are built with T_6-hexamers as the repeating unit.

In conclusion, it is proposed that the conformational change from the R-state to the T-state in the dihexamer is the driving force of the process. Only the trimer faces in the T-state are able to self-associate, whereas trimer faces in the R-state hinder higher self-association. Thus, pre-self-associating R_6-hexamers and $R_3T_3T_3R_3$-dihexamers are indisposed to self-association, while T_6T_6-dihexamers are prone to self-association.

■ ASSOCIATED CONTENT

Supporting Information

SAXS data from ωchI under formulation conditions (S1), in the pure dihexameric state (S2), and in the fully self-associated state (S3) and SEC-MALS (S4), AF4-MALS (S5), and CD (S6) data under the formulation conditions. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Universitetsparken 2, DK-2100 Copenhagen, Denmark. E-mail: bente.vestergaard@sund.ku.dk. Phone: +45 35 336 403. Fax: +45 35 336 041.

Funding

We are very thankful for funding from the Drug Research Academy at the Faculty of Health and Medical Sciences, University of Copenhagen (M.H.J.), Novo Nordisk A/S (M.H.J., P.-O.W., J.K.J., D.B.S., and S.H.), the Carlsberg Foundation (K.N.T.), the Danish Council for Independent Research, Medical Sciences (B.V.), and DANSCATT (M.H.J., K.N.T., and B.V.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge expert assistance in the laboratory from Michael Docherty, Ulrich Ehrbar, Birgit Dræby Spoon, and Lene Grønlund Andersen (Novo Nordisk A/S). We thank group leader Dmitri I. Svergun (X33, Deutsches Elektronen-Synchrotron), principal beamline scientist Javier Perez (SWING, Synchrotron Soleil), and Charlotte Rode Mosbæk, Annette Eva Langkilde, and Magda Møller (Department of Drug Design and Pharmacology, University of Copenhagen) for support during beamtime. The crystal structure was kindly provided by Dr. Gerd Schlückebier (Novo Nordisk A/S).

■ ABBREVIATIONS

AF4, asymmetrical flow-field flow fractionation; DLS, dynamic light scattering; HI, human insulin; I_0, scattering intensity at zero angle; MALS, multiangle light scattering; MM, molar mass; M_W, weight-average MM; NSD, normalized spatial discrepancy; NUV-CD, near-UV circular dichroism; ωchI, Lys^{B29}(N^{ε}ω-carboxyheptadecanoyl) des(B30) human insulin; $P(r)$, pair distance distribution function; PDB, Protein Data Bank; R_H, hydrodynamic radius; R_g, radius of gyration; rmsd, root-mean-square deviation; SAXS, small angle X-ray scattering; SEC, size exclusion chromatography; sc, subcutaneous.

■ REFERENCES

(1) Kurtzhals, P., Havelund, S., Jonassen, I., Kiehr, B., Larsen, U. D., Ribel, U., and Markussen, J. (1995) Albumin binding of insulins acylated with fatty acids: Characterization of the ligand protein interaction and correlation between binding affinity and timing of the insulin effect in vivo. *Biochem. J. 312*, 725−731.

(2) Markussen, J., Havelund, S., Kurtzhals, P., Andersen, A. S., Halstrom, J., Hasselager, E., Larsen, U. D., Ribel, U., Schaffer, L., Vad, K., and Jonassen, I. (1996) Soluble, fatty acid acylated insulins bind to albumin and show protracted action in pigs. *Diabetologia 39*, 281−288.

(3) Kurtzhals, P., Havelund, S., Jonassen, I., Kiehr, B., Ribel, U., and Markussen, J. (1996) Albumin binding and time action of acylated insulins in various species. *J. Pharm. Sci. 85*, 304−308.

(4) Whittingham, J. L., Havelund, S., and Jonassen, I. (1997) Crystal structure of a prolonged-acting insulin with albumin-binding properties. *Biochemistry 36*, 2826−2831.

(5) Havelund, S., Plum, A., Ribel, U., Jonassen, I., Volund, A., Markussen, J., and Kurtzhals, P. (2004) The mechanism of protraction of insulin detemir, a long-acting, acylated analog of human insulin. *Pharm. Res. 21*, 1498−1504.

(6) Hamilton-Wessler, M., Ader, M., Dea, M., Moore, D., Jorgensen, P. N., Markussen, J., and Bergman, R. N. (1999) Mechanism of protracted metabolic effects of fatty acid acylated insulin, NN304, in dogs: Retention of NN304 by albumin. *Diabetologia 42*, 1254−1263.

(7) Radziuk, J., Pye, S., Bradley, B., Braaten, J., Vignati, L., Roach, P., Bowsher, R., DiMarchi, R., and Chance, R. (1998) Basal activity profiles of NPH and [Nε-palmitoyl Lys (B29)] human insulins in subjects with IDDM. *Diabetologia 41*, 116−120.

(8) Havelund, S., Balschmidt, P., Jonassen, I., and Hoeg-Jensen, T. (1999) Aggregates of human insulin derivatives. Patent WO 99/21888.

(9) Whittingham, J. L., Jonassen, I., Havelund, S., Roberts, S. M., Dodson, E. J., Verma, C. S., Wilkinson, A. J., and Dodson, G. G. (2004) Crystallographic and solution studies of N-lithocholyl insulin: A new generation of prolonged-acting human insulins. *Biochemistry 43*, 5987−5995.

(10) Jonassen, I., Havelund, S., Ribel, U., Plum, A., Loftager, M., Hoeg-Jensen, T., Volund, A., and Markussen, J. (2006) Biochemical and physiological properties of a novel series of long-acting insulin analogs obtained by acylation with cholic acid derivatives. *Pharm. Res. 23*, 49−55.

(11) Jonassen, I., Havelund, S., Hoeg-Jensen, T., Steensgaard, D. B., Wahlund, P. O., and Ribel, U. (2012) Design of the novel protraction mechanism of insulin degludec, an ultra-long acting basal insulin. *Pharm. Res. 29*, 2104−2114.

(12) Jensen, M. H., Wahlund, P.-O., Jakobsen, J. K., Vestergaard, B., van de Weert, M., and Havelund, S. (2011) Self-association of long-acting insulin analogues studied by size exclusion chromatography coupled to multi angle light scattering. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci. 879*, 2945−2951.

(13) Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., and Mercola, D. A. (1972) Three-dimensional atomic structure of insulin and its relationship to activity. *Diabetes 21*, 492−505.

(14) Brzovic, P. S., Choi, W. E., Borchardt, D., Kaarsholm, N. C., and Dunn, M. F. (1994) Structural asymmetry and half-site reactivity in the T to R allosteric transition of the insulin hexamer. *Biochemistry 33*, 13057−13069.

(15) Bloom, C. R., Choi, W. E., Brzovic, P. S., Ha, J. J., Huang, S. T., Kaarsholm, N. C., and Dunn, M. F. (1995) Ligand binding to wild-type and E-B13Q mutant insulins: A three-state allosteric model system showing half-site reactivity. *J. Mol. Biol. 245*, 324−330.

(16) Bloom, C. R., Wu, N., Dunn, A., Kaarsholm, N. C., and Dunn, M. F. (1998) Comparison of the allosteric properties of the Co(II)- and Zn(II)-substituted insulin hexamers. *Biochemistry 37*, 10937−10944.

(17) Smith, G. D., Swenson, D. C., Dodson, E. J., Dodson, G. G., and Reynolds, C. D. (1984) Structural stability in the 4-zinc human insulin hexamer. *Proc. Natl. Acad. Sci. U.S.A. 81*, 7093−7097.

(18) Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philos. Trans. R. Soc. London, Ser. B 319*, 369−456.

(19) Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., and Swenson, D. (1989) Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. *Nature 338*, 594−596.

(20) Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A., and Blessing, R. H. (2000) R6 hexameric insulin complexed with m-cresol or resorcinol. *Acta Crystallogr. D56*, 1541−1548.

(21) Roy, M., Brader, M. L., Lee, R. W., Kaarsholm, N. C., Hansen, J. F., and Dunn, M. F. (1989) Spectroscopic signatures of the T to R conformational transition in the insulin hexamer. *J. Biol. Chem. 264*, 19081−19085.

(22) Kaarsholm, N. C., Ko, H. C., and Dunn, M. F. (1989) Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin, and miniproinsulin. *Biochemistry 28*, 4427−4435.

(23) Brader, M. L., Kaarsholm, N. C., Lee, R. W., and Dunn, M. F. (1991) Characterization of the R-state insulin hexamer and its derivatives. The hexamer is stabilized by heterotropic ligand binding interactions. *Biochemistry 30*, 6636−6645.

(24) Renscheidt, H., Strassburger, W., Glatter, U., Wollmer, A., Dodson, G. G., and Mercola, D. A. (1984) A solution equivalent of the 2Zn→4Zn transformation of insulin in the crystal. *Eur. J. Biochem. 142*, 7−14.

(25) Shneine, J., Voswinkel, M., Federwisch, M., and Wollmer, A. (2000) Enhancing the T→R transition of insulin by helix-promoting sequence modifications at the N-terminal B-chain. *Biol. Chem. 381*, 127−133.

(26) Huang, S. T., Choi, W. E., Bloom, C., Leuenberger, M., and Dunn, M. F. (1997) Carboxylate ions are strong allosteric ligands for the HisB10 sites of the R-state insulin hexamer. *Biochemistry 36*, 9878−9888.

(27) Schäffer, L., and Balschmidt, P. (1997) Insulin derivatives and their use. Patent WO 9731022.

(28) David, G., and Perez, J. (2009) Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr. 42*, 892−900.

(29) Konarev, P. V., Petoukhov, M. V., Volkov, V. V., and Svergun, D. I. (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr. 39*, 277−286.

(30) Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr. 36*, 1277−1282.

(31) Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr. 25*, 495−503.

(32) Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr. 28*, 768−773.

(33) Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J. 76*, 2879−2886.
(34) Volkov, V. V., and Svergun, D. I. (2011) Uniqueness of ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr. 36*, 860−864.
(35) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235−242.
(36) DeLano, W. L. (2002) *The PyMOL molecular graphics system*, version 1.2r3pre, Schrödinger, LLC, New York.
(37) Kurtzhals, P., and Ribel, U. (1995) Action profile of cobalt(III)-insulin: A novel principle of protraction of potential use for basal insulin delivery. *Diabetes 44*, 1381−1385.
(38) de la Garcia, T., Huertas, M. L., and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J. 78*, 719−730.
(39) Olsen, H. B., and Kaarsholm, N. C. (2000) Structural effects of protein lipidation as revealed by LysB29-myristoyl, des(B30) insulin. *Biochemistry 39*, 11893−11900.
(40) Glatter, O. (1991) Small-angle scattering and light scattering. In *Neutron, X-ray and light scattering: Introduction to an investigative tool for colloidal and polymeric systems* (Lindner, P., and Zemb, Th., Eds.) pp 33−82, North-Holland Delta, Amsterdam.
(41) Jonassen, I., Havelund, S., Ribel, U., Hoeg-Jensen, T., Steensgaard, D. B., Johansen, T., Haahr, H., Nishimura, E., and Kurtzhals, P. (2010) Insulin degludec: Multi-hexamer formation is the underlying basis for this new generation ultra-long acting basal insulin. *Diabetologia 53*, S388.
(42) Cohen, R. J. (1991) Determining rate constants for irreversible polymerization where the initial step and propagation steps have different rate constants: Consideration of polyadenylate polymerase. *J. Theor. Biol. 150*, 529−537.