Supplementary Material

S1: Oligomer fits to SAXS data of ωchI in formulation conditions

Data from the concentration series of ωchI_3Zn and ωchI_6Zn were collected at beamline X33 at Deutsches Elektronen Synchrotron (DESY, Hamburg, Germany). The experimental scattering curves were fitted to a linear combination of form factors. The theoretical form factors used for input in Oligomer were calculated from the generated PDB models by the program Crysol as described in Materials and Methods. The fits are shown below.

Table S 1

Sample	Conc µM	Dimer %	Hexamer %	Dihexamer %	R_g Å	MM kDa	X
ωchI_3Zn	1173	14	0	86	26.2	64	2.1
ωchI_3Zn	586	22	8	70	25.8	56	1.43
ωchI_3Zn	327	30	25	45	24.7	45	1.37
ωchI_3Zn	163	35	37	28	23.6	38	1.42
ωchI_6Zn	1154	19	3	78	26.1	60	1.4
ωchI_6Zn	578	23	10	67	25.7	54	1.58

Table S1: The volume fractions (%) of the form factors included in the Oligomer fits (Figure S2) are shown. The R_g and MM obtained from Oligomer are listed along with the X-values of the fits. The fraction of dihexamer increases with rising concentrations, whereas the fractions of hexamer and dihexamer decrease. The form factors of T_2-dimer, R_6-hexamer, and $R_3T_3T_3R_3$-dihexamer (36 Å) were used in the Oligomer run.

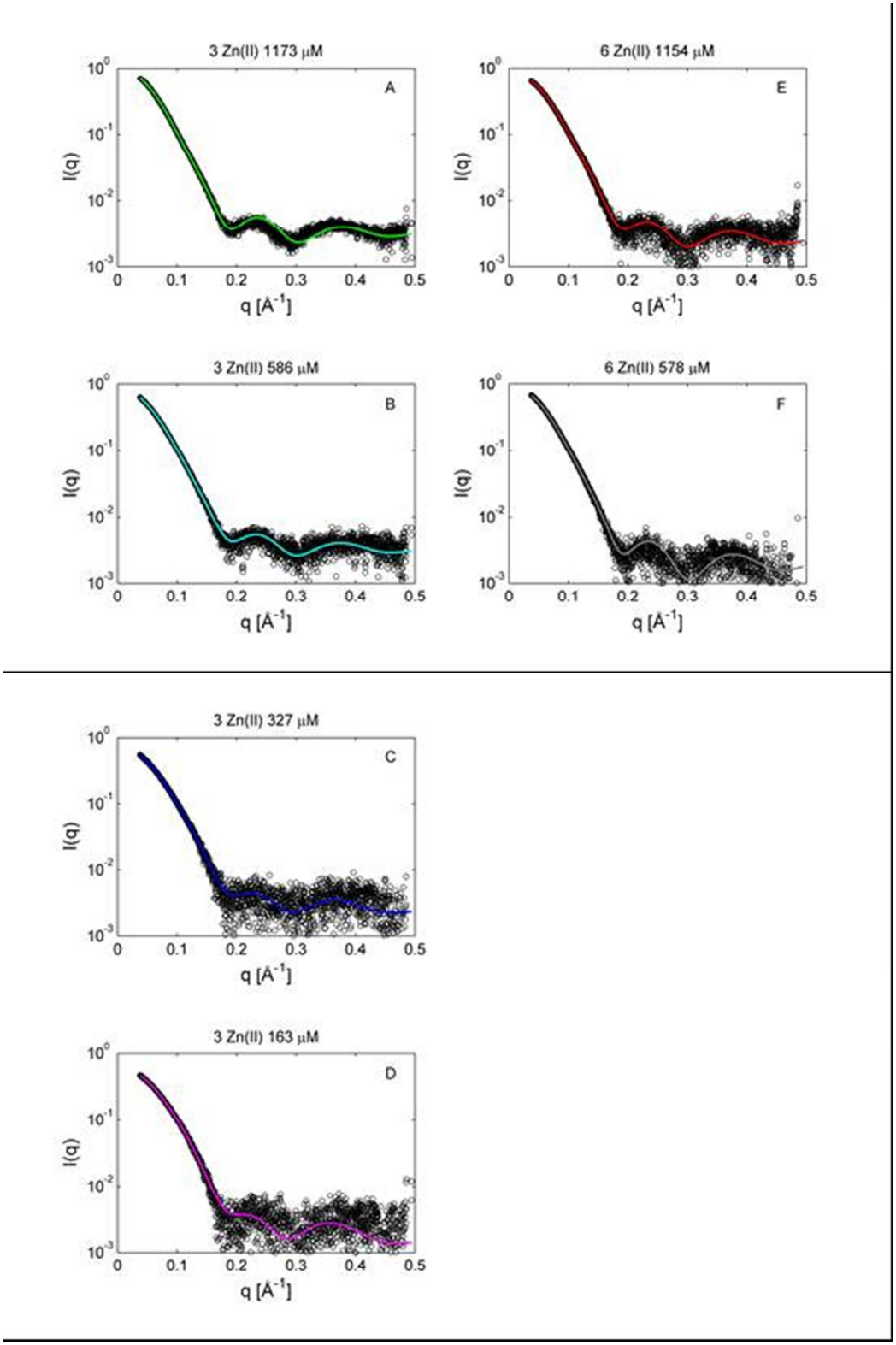


Figure S 1 Oligomer fits to experimental SAXS data of samples ωchI_3Zn and ωchI_6Zn. Experimental data are shown as black circles. The concentration series of ωchI_3Zn includes the concentrations (A

to D): 1173 (green), 586 (cyan), 327 (blue), and 163 µM (magenta); and ωchI_6Zn (E and F): the concentrations 1154 (red) and 578 (grey).

S2: Collection of pure ωchI dihexamer by SEC/SAXS

The experimental conditions described above were furthermore used to collect SAXS data from solutions containing the pure ωchI_3Zn dihexamer.

An HPLC was connected online to the SAXS sample cell at the SWING beamline at Synchrotron Soleil, France [28]. Figure S 2A shows the SEC elution profile of the insulin dihexamer. Here, 199 SAXS dataframes were collected over the elution peak with an individual exposure time of two seconds. 90 frames of the SAXS data are shown in Figure S 2B. After the peak, 20 frames of the elution buffer were collected (when a stable baseline was restored), merged, and used for background subtraction. After background subtraction, Guinier analysis was carried out with AutoRg [40] for each data curve through the peak. In Figure S 2A, it can be seen that the R_g forms a plateau through the chromatographic peak. The R_g values are slightly higher in the front of the peak and slightly lower in the last half of the peak. Frames 56 to 65 was merged (R_g = 26.8 Å) and used for analysis. The best Crysol fit of the experimental data was of a $R_3T_3T_3R_3$-dihexamer with an interhexameric distance of 36 Å (see the main paper **Error! Reference source not found.**). The R_g of 26.7 Å obtained from the Crysol fit corresponds nicely to a dihexamer. A second merged curve was created from the front of the peak of frames 43 to 54. Oligomer analysis showed that a combination of 96% T_3R_3-dihexamer and 4% T_6-tetrahexamer gave the best fit to the experimental data (Figure S 2C). The MM and R_g of 74 kDa and 27.9 Å, respectively, obtained from Oligomer analysis corresponds to a dihexamer with a fraction of a larger species.

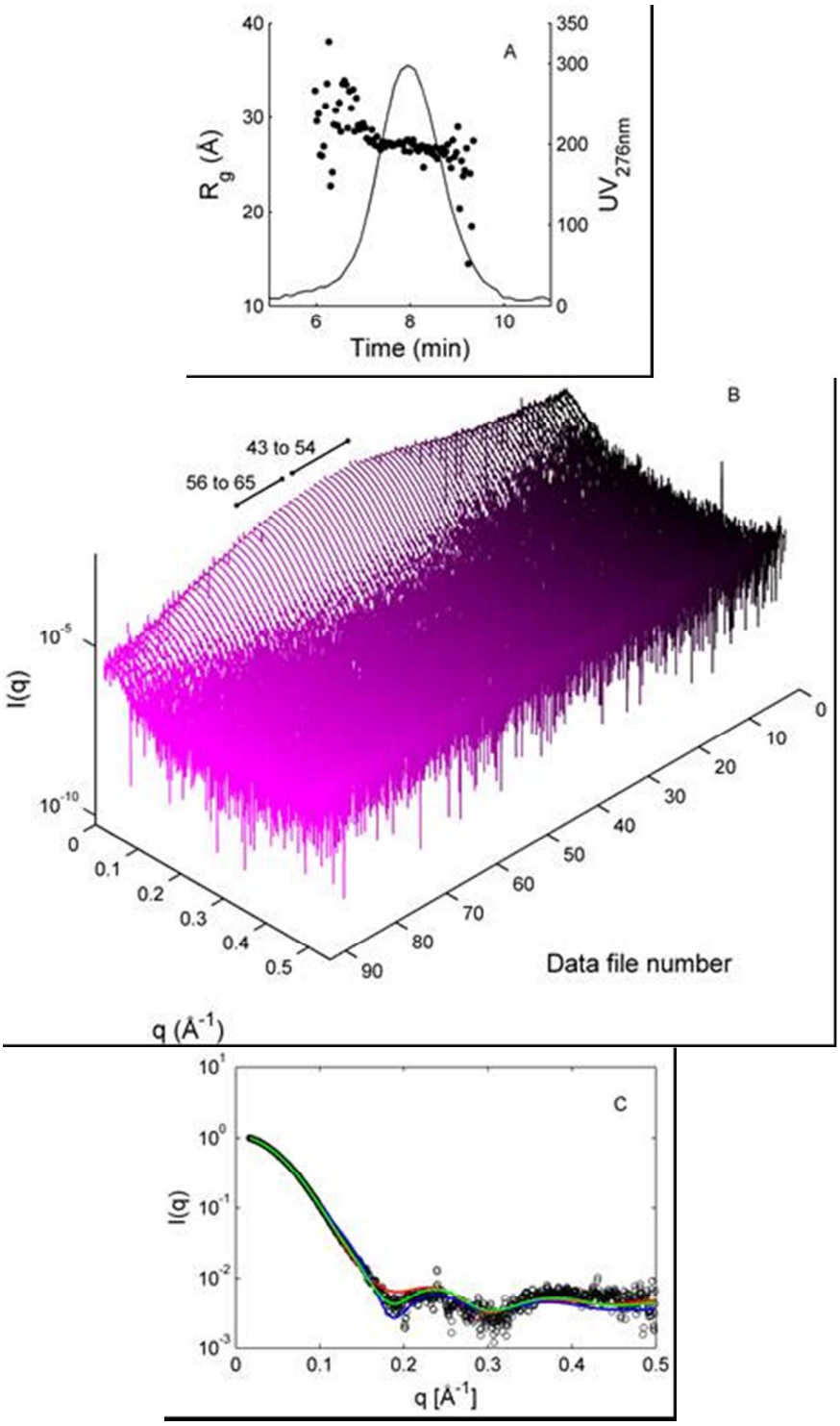


Figure S 2 SAXS data was collected of ωchI_3Zn dihexamer after online SEC separation. (A) shows the UV_{276nm} signal from the SAXS sample cell and the R_g-values derived from the SAXS curves through the peak (AutoRg) [27]. The data curves measured through the dihexameric peak are shown in three dimensions (B). The two lines indicate the curves used for merging and further analysis. (C) The Oligomer fit (green line: 96% $R_3T_3T_3R_3$-dihexamer and 4% T_6-tetrahexamer) of the merged curve

(frames 43 to 54) is shown. Also shown are the theoretical scattering curves for 96% R_6R_6-dihexamer and 4% T_6-tetrahexamer (blue line), and for 96% T_6T_6-dihexamer and 4% T_6-tetrahexamer (red line).

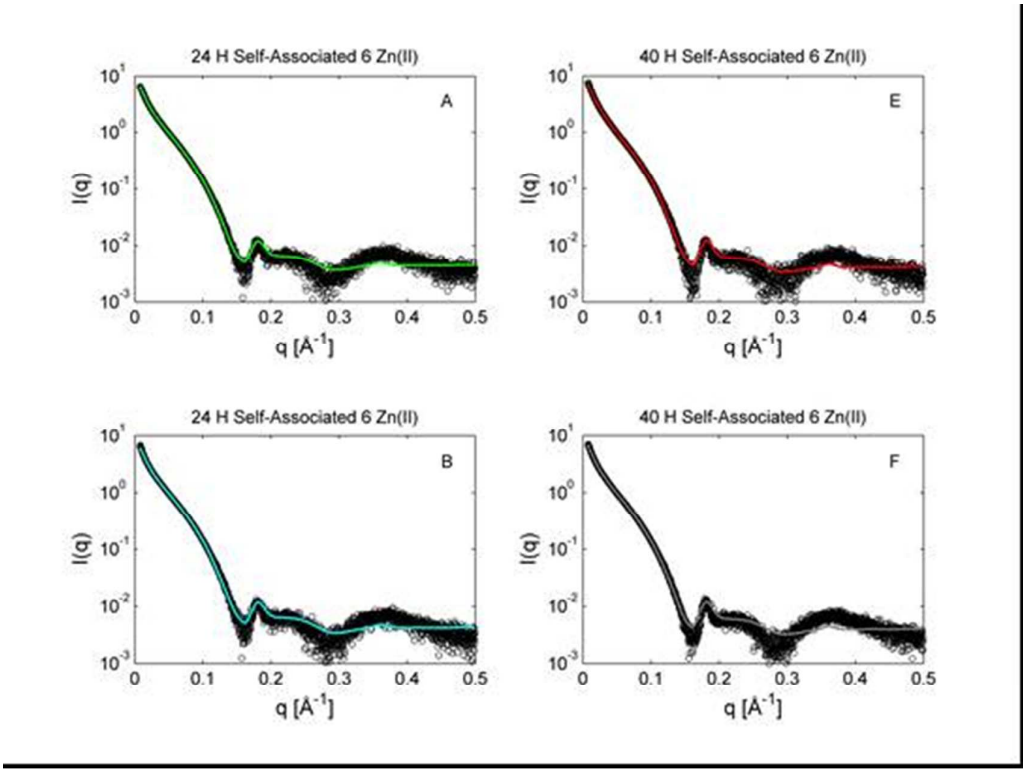
S3: Oligomer fits to fully self-associated SAXS data

SAXS data of the fully self-associated ωchI_6Zn were collected at beamline X33 at Deutsches Elektronen Synchrotron (DESY, Hamburg, Germany) and at the SWING beam line at Synchrotron Soleil (Paris, France). The time points from 24 to 40 hours were collected at X33. The 48 hour sample was collected at SWING using the HPLC as auto sampler as described in S2, however without a column, and is a merge of 12 frames of 2 seconds exposure. The experimental scattering curves were fitted to a linear combination of form factors using the program Oligomer [27]. The theoretical form factors used for input in Oligomer were calculated from the generated PDB models by the program Crysol [31] as described in Materials and Methods. The fits are shown below.

Table S 2

Sample	dihex %	Tetrahex %	hexahex %	octahex %	decahex %	12-hex %	14-hex %	30-hex %	R_g Å	MM kDa	X
24 H	6			51	29	4	5	5	114,7	350	1,85
24 H	9			42	36	5	3	5	113,8	342	1,96
32 H				29	52			19	145,6	474	3,97
32 H				47	41	3		9	126	394	2,82
40 H				37	47	6		10	128,8	410	2,33
40 H				54	28	8	5	5	116,3	370	2,18
48 H					25	54	10	11	138,2	492	20,63

Table S2: The volume fractions (%) of the form factors included in the Oligomer fits (figure S1) are shown. The R_g and MM obtained from Oligomer are listed along with the X-values of the fits. The form factors of T_6-multihexamers (35.1 Å) composed of 2 to 14 plus 30 hexamers were used in the Oligomer run.


24 H Self-Associated 6 Zn(II)
A
40 H Self-Associated 6 Zn(II)
E
24 H Self-Associated 6 Zn(II)
B
40 H Self-Associated 6 Zn(II)
F
I(q)
q [Å^{-1}]
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
0
0.1
0.2
0.3
0.4
0.5

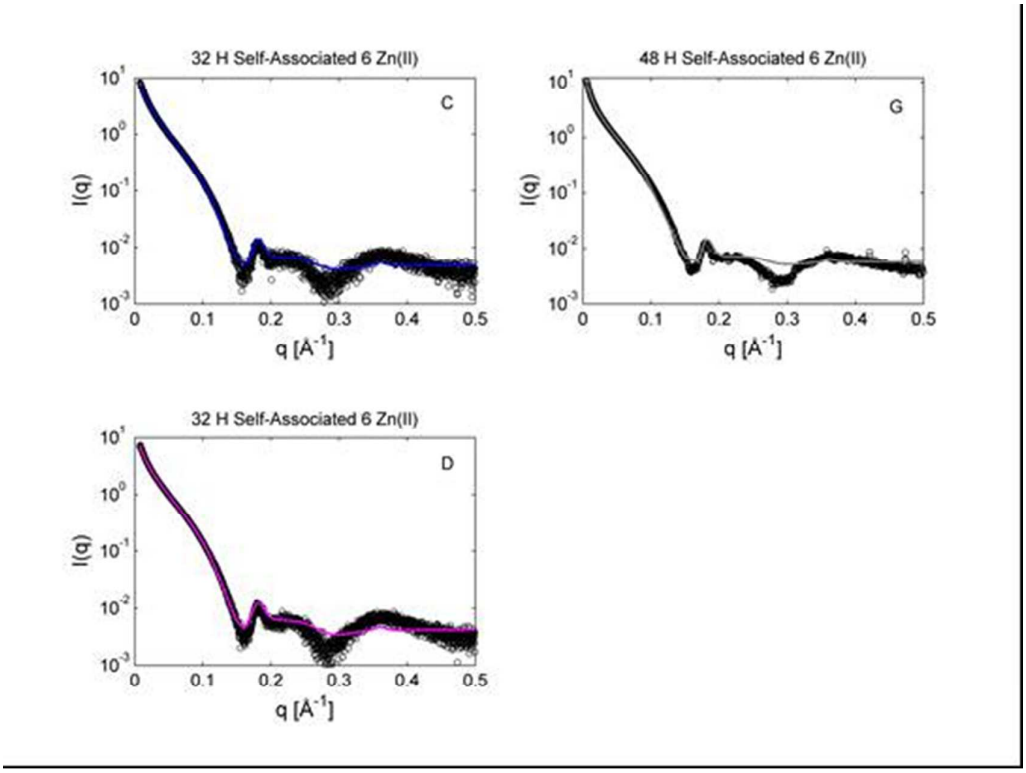


Figure S 3 SAXS data of the self-associated state of ωchI_6Zn (24 to 48 hours (A-G)) is presented with fits from Oligomer [27] analysis. Experimental data are shown as black circles. Panel H shows the volume fractions of the models included in the oligomer fits.

S4: Size Exclusion Chromatography (SEC)

Prior to self-assembly (in presence of phenol) SEC in connection with MALS was used to characterise ωchI in conditions with 0, 3, and 6 of Zn(II) ions per 6 insulin monomers. SEC was performed on an Agilent 1200 series HPLC system coupled to a MALS DAWN® HELEOS™ and an OptiLab® rEX refractive index detector from Wyatt Technology.

A Superdex 200 HR 10/30 column with an eluent consisting of 10 mM Tris/HCl pH 7.4, 140 mM NaCl, and 16 mM phenol was used for analysis of the three samples with a flow of 0.5 ml/min and an injection volume of 200 µL. Due to the large absorbance of phenol at 276 nm and thereby obstruction of the UV_{276nm} signal, the separation was followed with the

refractive index (RI) which was also used for concentration determination of the sample. Further details about MALS can be found in Materials and Methods.

A high salt content, which is not present in the formulation, is necessary in the SEC separation in order to maintain the chromatographic performance. In conclusion, ωchI_0Zn elutes in a peak with a M_W distribution between 10 and 22 kDa. ωchI_3Zn and ωchI_6Zn eluted in peaks with average M_W's of 73 and 69 kDa, respectively, although the M_W also had a slight slope indicating a distribution of sizes. Thus ωchI_0Zn can be regarded as existing as equilibrium between smaller oligomers, such as monomers, dimers, trimers, and tetramers, while ωchI_3Zn and ωchI_6Zn primarily exist in the dihexamer form.

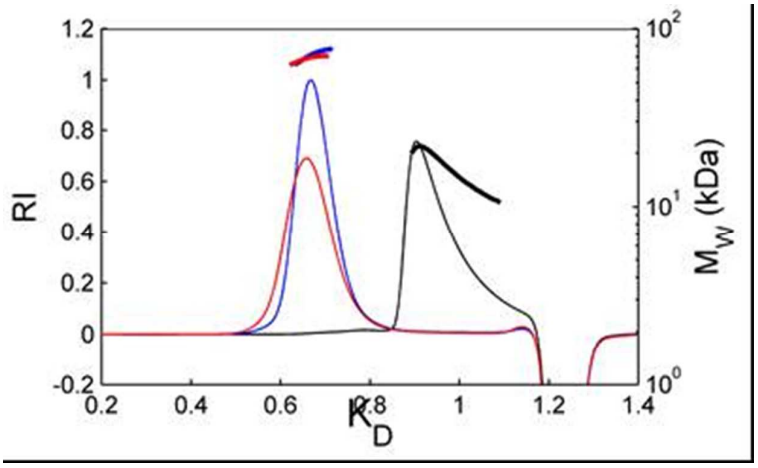


Figure S 4 Chromatograms showing SEC/MALS separation of ωchI_0Zn (black), ωchI_3Zn (blue), and ωchI_6Zn (red) in formulation conditions. The eluent is composed by 10 mM Tris/HCl pH 7.4, 140 mM NaCl, and 16 mM phenol. The SEC analysis needs to be run in presence of NaCl, in order to maintain the separation of the chromatographic system.

S5: Asymmetrical Flow-Field Flow Fractionation (AF4)

Like SEC, AF4 was used to analyse the sample prior to self-association. The channel length was 145 mm with a spacer height of 490 µm. Polyethersulfone (PES) membranes (from Wyatt Technology) with a cut-off of 3 kDa were used. Such a low cut-off was chosen in order to ensure that all states of insulin – from monomers (M_W ~ 6 kDa) to larger oligomers (dimers, hexamers, dihexamers and higher order associates) – would be separated and to avoid molecules passing through the membrane. The channel flow (F_c) rate was 1.0 ml/min. The focussing step was performed for 4 min with a focussing rate of 1.5 ml/min for all experiments. For separation of the insulin analogues in formulation conditions, the eluent consisted of 10 mM Tris/HCl pH 7.4, 10 or 140 mM NaCl, and 16 mM phenol. 10 mM NaCl was added in order to avoid insulin from sticking to the membrane. A cross flows (F_x) of 0.8 ml/min was used for ωchI_0Zn, ωchI_3Zn, and ωchI_6Zn. All samples had an elution period of 65 min and a succeeding elution period with F_x of 0 ml/min for 25 min. AF4 was coupled to MALS and further details can be found in the paper.

The SEC analysis was run with 140 mM NaCl, which was not present in the formulation. In order to investigate if the samples are affected by 'high salt' and 'low salt' conditions in the separation, AF4 was run in conditions with 10 mM NaCl ('low salt'). The samples containing zinc were not affected by the salt content. ωchI_3Zn and ωchI_6Zn eluted in peaks with a M_W's of 73 and 69 kDa; the M_W had a slight slope indicating a distribution of sizes. ωchI_0Zn eluted in a peak with a M_W distribution between 10 and 63 kDa. Thus, the two techniques gave slightly varying results without zinc, where AF4 indicated a broader equilibrium from monomers to dihexamers.

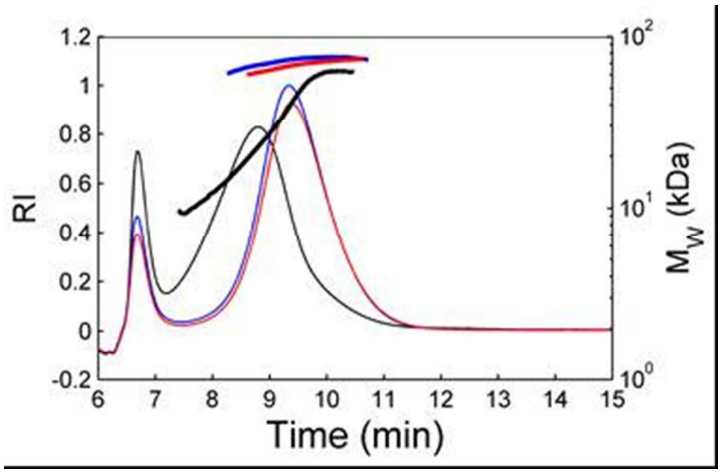


Figure S 5 Fractograms showing AF4/MALS runs of ωchI_0Zn (black), ωchI_3Zn (blue), and ωchI_6Zn (red) in formulation conditions. The channel flow was 1 ml/min for all samples. The protein was eluted with a constant cross flow of 0.8 ml/min. The eluent is composed by 10 mM Tris/HCl pH 7.4, 10 mM NaCl, and 16 mM phenol.

S6: NUV-CD measurements of ωchI_6Zn

To fully understand the oligomerisation state and conformational state of ωchI in the formulation conditions before self-association, the samples were studied with CD in the near UV range (245-350 nm). Here information about the insulin dimerisation can be extracted and also about the allosteric state of the zinc hexamer from R_6 over T_3R_3 to T_6. The NUV-CD spectrum has a distinct negative band with a minimum at 276 nm; this band originates from the tyrosine chromophore, which is located in the monomer-monomer interface of the dimer, and is therefore affected by the change in environment due to dimer formation. Furthermore, the negative band with a maximum at 251 nm is sensitive towards change in environment of the cysteinyl chromophore in insulin and has earlier been used to be descriptive of the transition between the R and T states.

The monomer-dimer equilibrium was investigated with a concentration series of ωchI in absence of zinc and phenol see Figure S 6A. The 276 nm band shows a $\Delta\varepsilon_{276}$ of -1.5 M^{-1} cm^{-1}

at 4-10 µM ωchI, and furthermore that dimerisation reaches a saturated level of about -3.5 M^{-1} cm^{-1} at concentrations ≥ 1000 µM; the concentration used in the formulation here (1200 µM) was in the saturated area with a $\Delta\varepsilon_{276}$ of -3.0 M^{-1} cm^{-1}. Subsequently, a zinc titration of ωchI (600 µM) in absence of phenol was performed and showed that the dimer equilibrium is pushed further with only a small change in $\Delta\varepsilon_{276}$ from -3.0 M^{-1} cm^{-1} to -4.3 M^{-1} cm^{-1}, where the endpoint corresponds with full dimerisation of zinc-free human insulin reported earlier [32]. This suggests that the equilibrium is not pushed entirely to the dimeric state at the saturated level in the concentration series, and that in presence of zinc all the monomer-monomer interfaces are formed due to zinc hexamer formation.

Finally, the allosteric state was investigated by a titration of ωchI_6Zn (600 µM, ~4 mg/ml) with phenol from 0 to 30 mM. It shows a large change in the 251 nm band from -1 to -4.5 M^{-1} cm^{-1} and a rapid saturation of $\Delta\varepsilon_{251}$ at 3 mM phenol. Previously, phenol titrations of human insulin with two Zn(II) per six ins have been reported [10, 32], where $\Delta\varepsilon_{251}$ develops from -1.5 to -6.5 M^{-1} cm^{-1} over the same range of phenol concentrations. This may indicate that for ωchI_6Zn with 32 mM phenol only a partial allosteric transition has taken place, thus rendering the hexamer in the T_3R_3 state or in equilibrium between T_3R_3 and R_6.

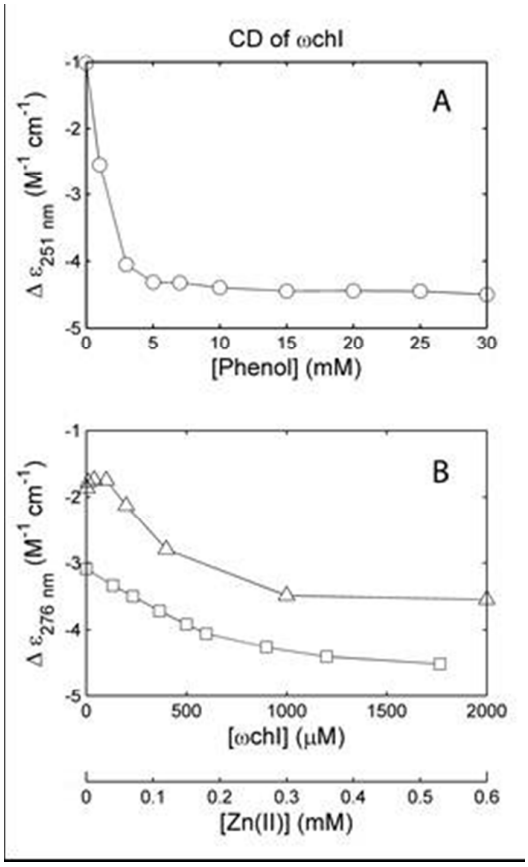


Figure S 6 NUV-CD measurements of ωchI. The ellipticity constant, Δε, at a wavelength of 251 and 276 nm is shown in the plots. The allosteric state of ωchI_6Zn (600 μM) is monitored at phenol concentrations 0-30 mM (A). The dimerisation was measured in a concentration series of ωchI from 0 to 2000 μM (triangles) in absence of Zn(II) and phenol; and secondly, in a Zn(II) titration experiment (squares) from 0 to 0.53 mM (600 μM ωchI) (B).