Supporting information

Fully Convergent Chemical Synthesis of Ester Insulin: Determination of the High Resolution X-ray Structure by Racemic Protein Crystallography

Michal Avital-Shmilovici, Kalyaneswar Mandal, Zachary P. Gates, Nelson B. Phillips, Michael A. Weiss, and Stephen B.H. Kent

Abbreviations: AcOH = Acetic acid, **Alloc-Cl** = Allyl chloroformate, **Boc** = *N-tert*-butoxycarbonyl, **DCM** = Dichloromethane, **DIEA** = N,N-Diisopropylethylamine, **DMAP** = N,N-Dimethylaminopyridine, **EDC·HCl** = 1-Ethyl-3-(3-dimethyllaminopropyl) carbodiimide hydrochloride, **Et_3N** = Triethylamine, **EtOAc**- Ethylacetate, **Fm** = 9-fluorenylmethyl, **Fmoc** = N-Fluorenyl-9-methyloxycarbonyl, **GnHCl** = Guanidine hydrochloride, **Hex** = Hexane, **MPAA** = (4-carboxymethyl)thiophenol, **TCEP·HCl** = Tris(2-carboxyethyl)phosphine hydrochloride, **THF**= Tetrahydrofuran, **Z(2-Cl)-OSu** = N-2-Chlorobenzyloxycarbonyloxysuccinimide.

Amino Acid Sequence of DKP Insulin

The target amino acid sequence of DKP insulin is:

A Chain:

Gly[1]-Ile-Val-Glu-Gln-Cys-Cys-Thr-Ser—Ile[10]-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cys[20]-Asn

B Chain:

Phe[1]-Val-Asn-Gln-His-Leu-Cys-Gly-Ser-Asp[10]-Leu-Val-Glu-Ala-Leu-Tyr-Leu-Val-Cys-Gly[20]-Glu-Arg-Gly-Phe-Phe-Tyr-Thr-Lys-Pro-Thr[30]

S1: Synthesis of Peptide Segments:

S1.1: Synthesis of peptide Gly^{A1}-Glu^{A4}[OβThr^{B30}-Thz^{B19}]-Cys^{A6}-$^{\alpha}$COSR (3): The peptide was synthesized as discussed in the Experimental Section, and its analytical LC-MS profile is shown in Figure S1.

S1.2: Synthesis of peptide Cys^{A7}-Asn^{A21} (2): The peptide was synthesized on a 0.5 mmol scale on -OCH_2-Pam-resin,[1] by Boc chemistry SPPS using "*in situ* neutralization" protocols[2] and cleaved from the resin using HF–*p*-cresol for 1 h at 0 ^{o}C. The crude product was purified by preparative HPLC and lyophilized to give 223 mg (25% yield based on 0.5 mmol of -OCH_2-Pam-resin) of pure peptide **2**. The product was characterized by LC-MS (ESI); observed: L-Cys^{A7}-Asn^{A21}, 1753.2 ± 0.8 Da; D-Cys^{A7}-Asn^{A21}, 1753.6 ± 0.5 Da; calculated (average isotopes), 1754.0 Da (Figure S2).

S1.3: Synthesis of peptide Phe^{B1}-Val^{B18}- $^{\alpha}$COSR (4): The peptide was prepared on a 0.2 mmol scale on $HSCH_2CH_2CO$-Ala-OCH_2-Pam-resin[3] by manual Boc chemistry stepwise (SPPS) using "*in situ* neutralization" protocols.[2] DNP (His side chain protecting group) was removed by treating the peptide-resin *before* HF cleavage twice with a mixture of 2-mercaptoethanol:DIEA:DMF 2:1:7 for 15 min each at RT. Then the peptide was cleaved from the resin with HF and simultaneously deprotected as described above. The crude product was purified by preparative HPLC and lyophilized to give 132 mg (30% yield based on 0.2 mmol of Boc-Ala-OCH_2-Pam-resin) of pure peptide **4**. The product was characterized by LC-MS (ESI); observed: L-Phe^{B1}-Val^{B18}- $^{\alpha}$COSR, 2179.1± 0.5 Da; D-Phe^{B1}-Val^{B18}- $^{\alpha}$COSR, 2179.3 ± 0.5 Da; calculated (average isotopes), 2179.5 Da (Figure S3).

S2: Chemical synthesis of D-DKP ester insulin: Chemical synthesis of D-DKP ester insulin was performed following the same procedures described in the Experimental Section for the synthesis of L-DKP ester insulin. The analytical data are shown in Figures S4-S5.

S3: Chemical Conversion of D-DKP Ester Insulin into D-DKP Insulin: Saponification of D-DKP ester insulin **D-1** to give D-DKP insulin **D-17** was performed following the same procedures as described in the Experimental Section for the saponification of L-DKP ester insulin. The analytical data are shown in Figure S6.

S4: Receptor Binding Assay: Affinity for a FLAG-epitope-tagged insulin receptor was determined by a microtiter-plate antibody-capture assay. The insulin receptor was obtained by partial lectin purification following transient transfection of human 293 PEAK rapid fibroblasts as described.[4] Data were analyzed by non-linear regression using a 2-site sequential model.[5] The percentage of tracer bound in absence of competing ligand was in each case <15% to avoid ligand-depletion artifacts. Samples of mirror-image DKP insulin were treated with trypsin as a control to exclude possible contamination by wild-type insulin or L-based insulin analogues. Other methods are as previously described.[6]

S5: Rodent Potency Assay: Male Lewis rats (mean body mass ~300 grams) were rendered diabetic by treatment with streptozotocin (STZ) as described.[7] To test the in vivo potency of insulin analogs in relation to wild-type human insulin, protein solutions containing wild-type human insulin, an analog, or buffer alone (protein-free sterile diluent obtained from Eli Lilly and Co.; composed of 16 mg glycerin, 1.6 mg meta-cresol, 0.65 mg phenol, and 3.8 mg sodium phosphate PH 7.4.) were injected subcutaneously, and resulting changes in blood glucose concentration were monitored by serial measurements using a clinical glucometer (Hypoguard Advance Micro-Draw meter). To ensure uniformity of formulation, insulins were each re-purified by reverse-phase HPLC, dried to powder, dissolved in diluent at the same maximum protein concentration (300 μg/mL) and re-quantified by analytical C4 rp-HPLC; dilutions were made using the above buffer. Rats (N=5 per group) were fasted overnight and were injected subcutaneously at time t = 0 with 20 μg insulin or insulin analog in 100 μl of buffer per 300 g rat. This dose corresponds to ca. 67 μg/kg body weight, which in international units (IU) implies 2 IU/kg body weight. Dose-response studies of wild-type insulin indicated that at this dose a near-maximal rate of glucose disposal during the first hour following injection was achieved. In the experiments shown in panel C of Figure 5

(main text) blood was obtained from the clipped tip of the tail at times 0,10, 20, 30, 40, 50, 60, 70, 80, 100, 120, and 140 min; in panel D blood was obtained at times 0, 10, 20, 30, 40, 50, 60, 80, 100, 120, 150, 180, 240, 300, and 360 min. The efficacy of insulin action to reduce blood glucose concentration was calculated using (a) the change in concentration over time (using least-mean squares and initial region of linear fall) divided by the concentration of insulin injected and (b) the integrated area between the glucose time dependence and a horizontal line at the starting blood glucose concentration. Assessment of statistical significance was performed using a Student's t-test.

References

1. Mitchell, A. R.; Kent, S. B. H.; Engelhard, M.; Merrifield, R. B. *J. Org. Chem.* **1978**, *43*, 2845.
2. Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Res. Ther.* **1992**, *40*, 180.
3. Hackeng, T. M.; Griffin, J. H.; Dawson, P.E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068.
4. Whittaker, J.; Whittaker, L. *J. Biol. Chem.* **2005**, *280*, 20932.
5. Wang, Z. X. *FEBS Lett.* **1995**, *360*, 111.
6. Sohma, Y.; Hua, Q. X.; Liu, M.; Phillips, N. B.; Hu, S. Q.; Whittaker, J.; Whittaker, L. J.; Ng, A.; Roberts Jr, C. T.; Arvan, P.; Kent, S. B. H.; Weiss, M. A. *J. Biol. Chem.* **2010**, *285*, 5040.
7. Yang, Y.; Petkova, A.; Huang, K.; Xu, B.; Hua, Q. X.; Ye, I. J.; Chu, Y. C.; Hu, S. Q.; Phillips, N. B.; Whittaker, J.; Ismail-Beigi, F.; Mackin, R. B.; Katsoyannis, P. G.; Tycko, R.; Weiss, M. A. *J. Biol. Chem.* **2010**, *285*, 10806.

Figures and Schemes

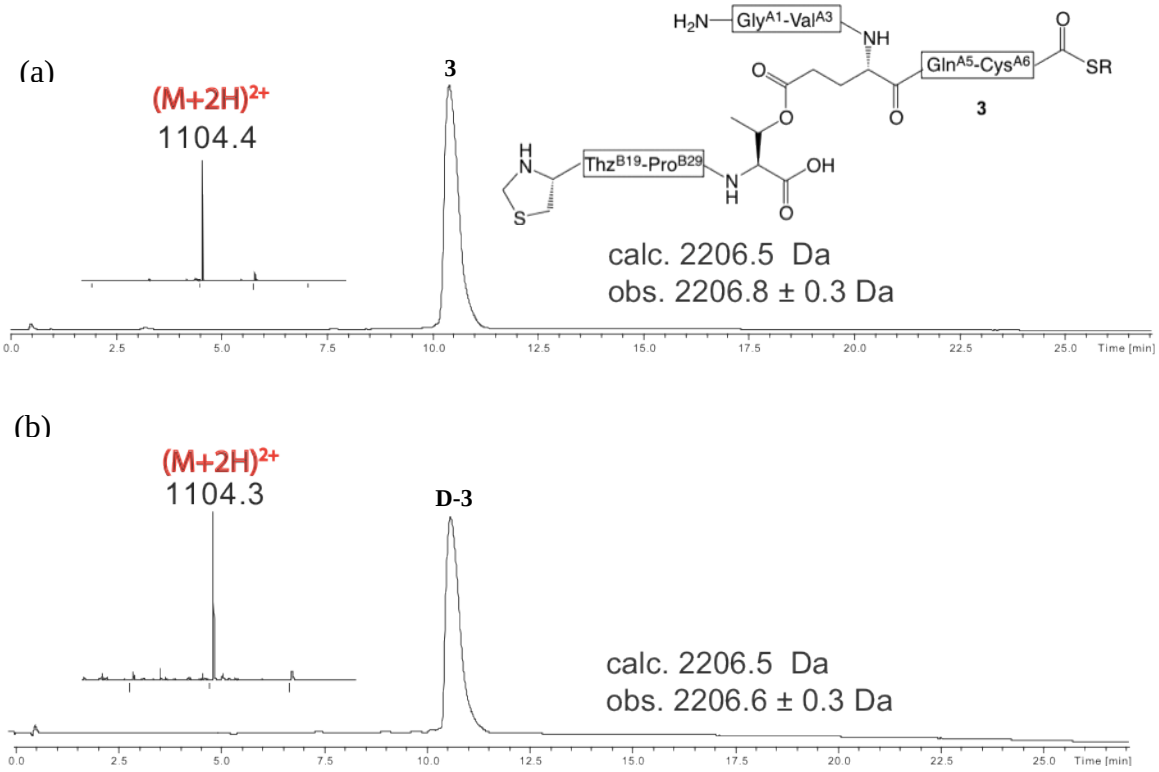


Figure S1: Analytical HPLC profiles together with online ESI MS data (inset) of (a) L-Gly^{A1}-Glu^{A4}[OβThr^{B30}-Thz^{B19}]-Cys^{A6}-$^{\alpha}$COSR (**3**) and (b) D-Gly^{A1}-Glu^{A4}[OβThr^{B30}-Thz^{B19}]-Cys^{A6}-$^{\alpha}$COSR (**D-3**) thioesters (R=–CH_2CH_2CO–Ala–COOH). The chromatographic separations were performed on a C8 column using a linear gradient (9%–53%) of buffer B in buffer A over 22 min (buffer A = 0.1% TFA in water, buffer B = 0.08% TFA in acetonitrile).

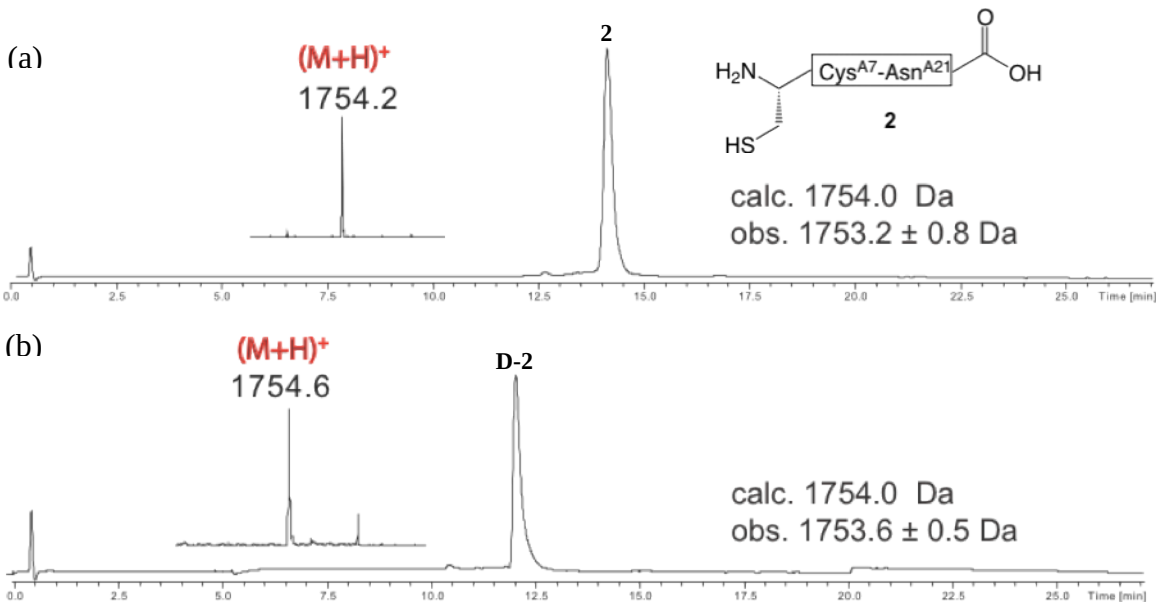


Figure S2: Analytical HPLC profiles together with online ESI MS data (inset) of (a) L-Cys^{A7}-Asn^{A21} (**2**) and (b) D-Cys^{A7}-Asn^{A21} (**D-2**) peptides. The chromatographic separations were performed on (a) C18 column and (b) C8 column using a linear gradient (9%–53%) of buffer B in buffer A over 22 min (buffer A = 0.1% TFA in water, buffer B = 0.08% TFA in acetonitrile).

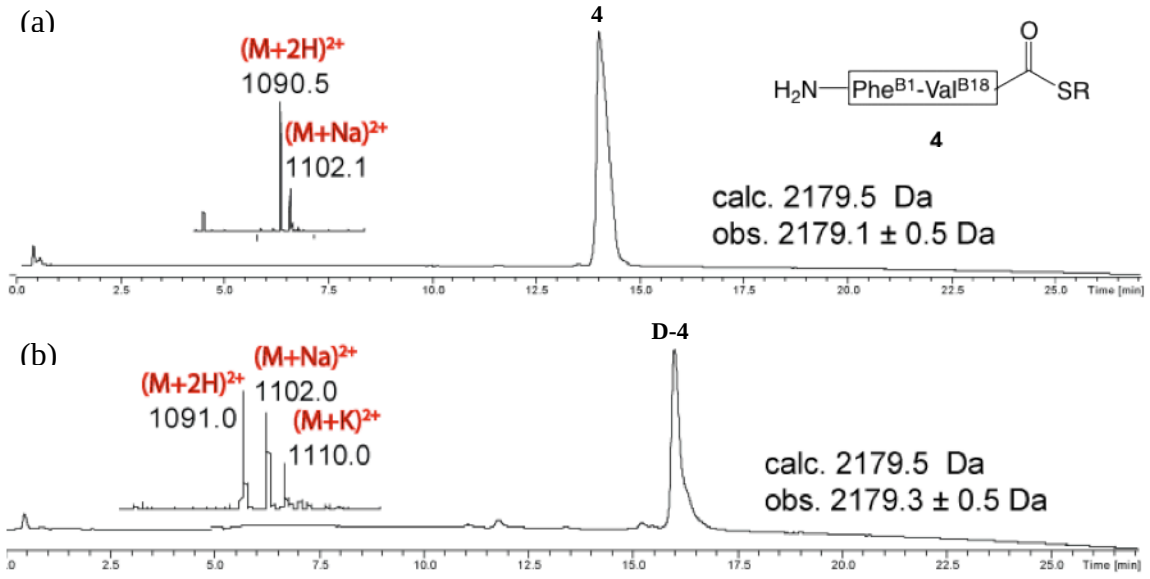


Figure S3: Analytical HPLC profiles together with online ESI MS data (inset) of (a) L-Phe^{B1}-Val^{B18}-$^{\alpha}$COSR (**4**) and (b) D-Phe^{B1}-Val^{B18}-$^{\alpha}$COSR (**D-4**) thioesters (R=–CH_2CH_2CO–Ala–COOH). The chromatographic separations were performed on (a) C8 column and (b) C18 column using a linear gradient (9%–53%) of buffer B in buffer A over 22 min (buffer A = 0.1% TFA in water, buffer B = 0.08% TFA in acetonitrile).

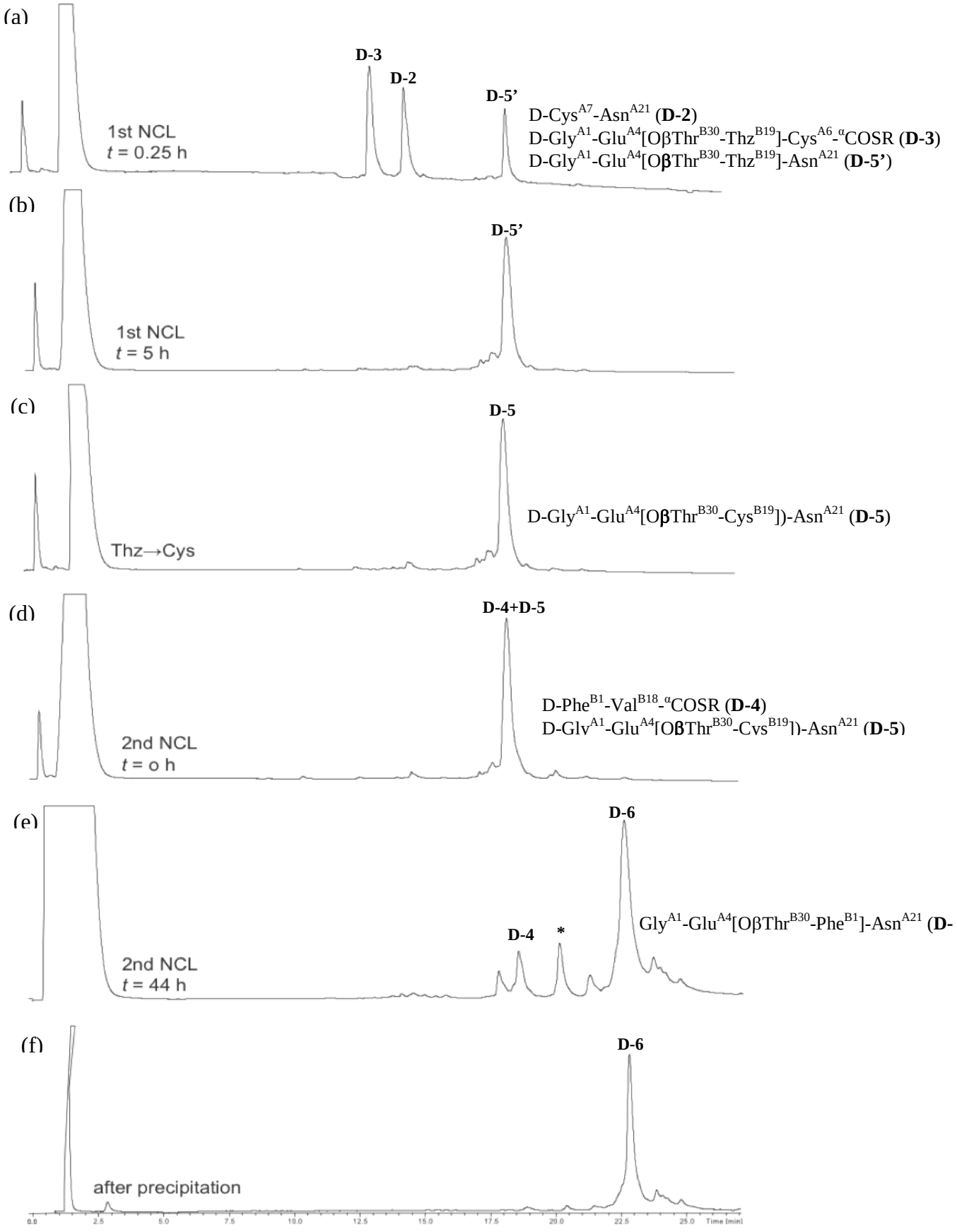

(a)
D-3
D-2
D-5'
1st NCL
t = 0.25 h
D-Cys^A7-Asn^A21 (D-2)
D-Gly^A1-Glu^A4[OβThr^B30-Thz^B19]-Cys^A6-αCOSR (D-3)
D-Gly^A1-Glu^A4[OβThr^B30-Thz^B19]-Asn^A21 (D-5')
(b)
D-5'
1st NCL
t = 5 h
(c)
D-5
Thz→Cys
D-Gly^A1-Glu^A4[OβThr^B30-Cys^B19])-Asn^A21 (D-5)
(d)
D-4+D-5
2nd NCL
t = o h
D-Phe^B1-Val^B18-αCOSR (D-4)
D-Gly^A1-Glu^A4[OβThr^B30-Cys^B19])-Asn^A21 (D-5)
(e)
D-6
D-4
*
2nd NCL
t = 44 h
Gly^A1-Glu^A4[OβThr^B30-Phe^B1]-Asn^A21 (D-
(f)
D-6
after precipitation
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5
25.0
Time (min)

Figure S4: Analytical HPLC data for the total chemical synthesis of D-DKP ester insulin. (a,b) 1st NCL: reaction of peptide segments D-Gly^{A1}-Glu^{A4}[OβThr^{B30}-Thz^{B19}]-Cys^{A6}-$^{\alpha}$COSR (**D-3**) (R=–CH_2CH_2CO–Ala–COOH) and D-Cys^{A7}-Asn^{A21}(**D-2**). (c) The crude product obtained after treatment with 0.15 M methoxylamine·HCl at pH 4 to give peptide **D-5**. (d,e) 2nd NCL: reaction of D-Gly^{A1}-Glu^{A4}[OβThr^{B30}-Cys^{B19}])-Asn^{A21} (**D-5**) and D-Phe^{B1}-Val^{B18}-$^{\alpha}$COSR (**D-4**). Both peptide segments have the same retention time, as it was confirmed by LC-MS. Upon completion, the excess D-Phe^{B1}-Val^{B18}-$^{\alpha}$COSR (**D-4**) was left. *: D-Phe^{B1}-Val^{B18}-$^{\alpha}$CO-MPAA thioester. f) Crude polypeptide (**D-6**) after precipitation from water. The chromatographic separations were performed on a C18 column using a linear gradient (9%–53%) of buffer B in buffer A over 22 min (buffer A = 0.1% TFA in water, buffer B = 0.08% TFA in acetonitrile). The times shown refer to overall elapsed times for the synthesis.

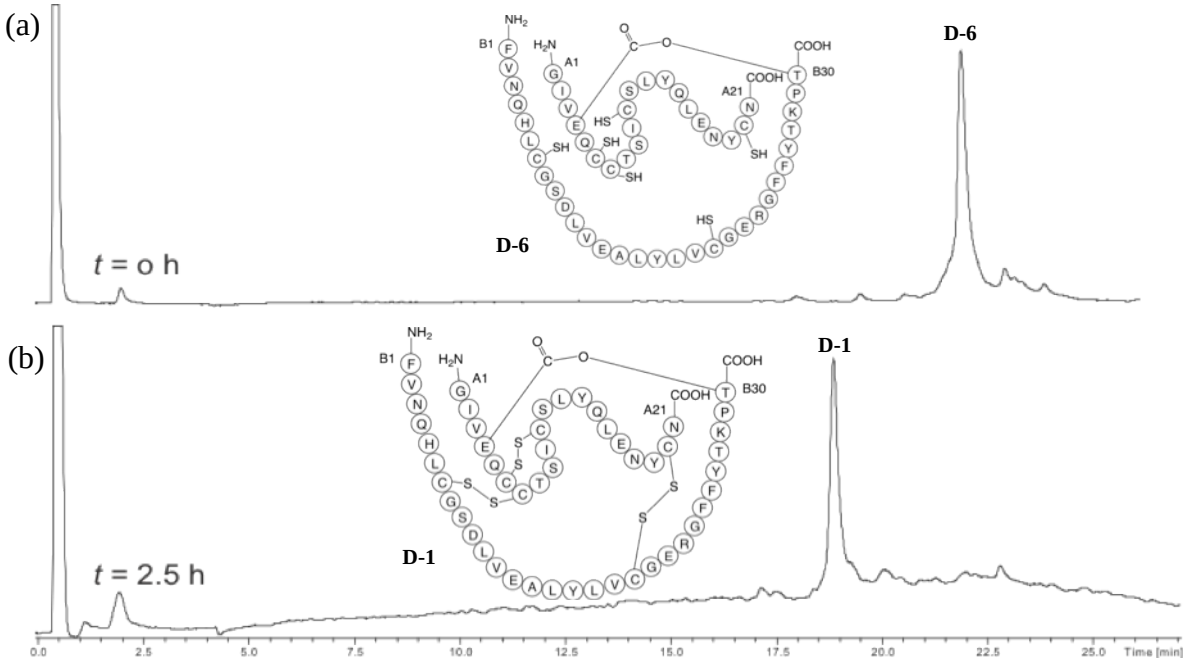


Figure S5: HPLC characterization of the folding of D-DKP ester insulin. 1a) Crude D-DKP ester insulin linear polypeptide (**D-6**) at *t* = 0 h. 1b) Folding reaction after 2.5 h. The chromatographic separations were performed on a C18 column using a linear gradient (9%–53%) of buffer B in buffer A over 22 min (buffer A = 0.1% TFA in water, buffer B = 0.08% TFA in acetonitrile).

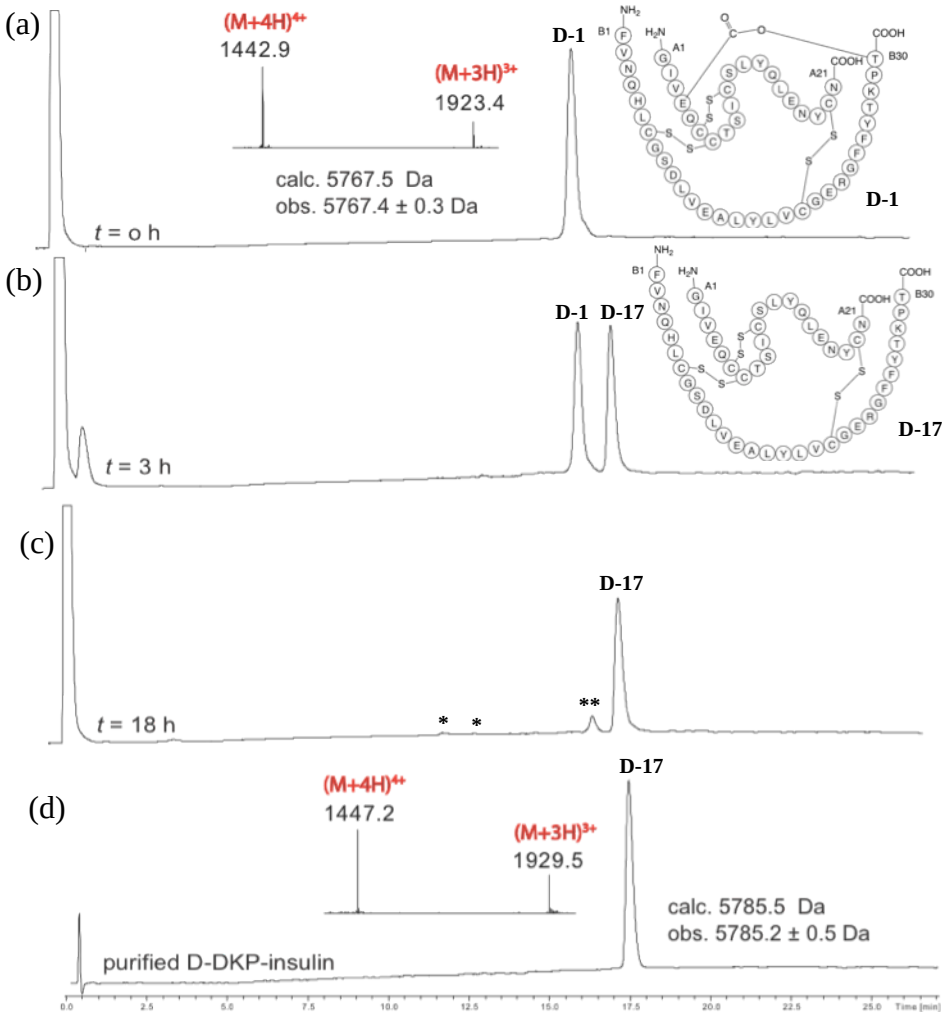


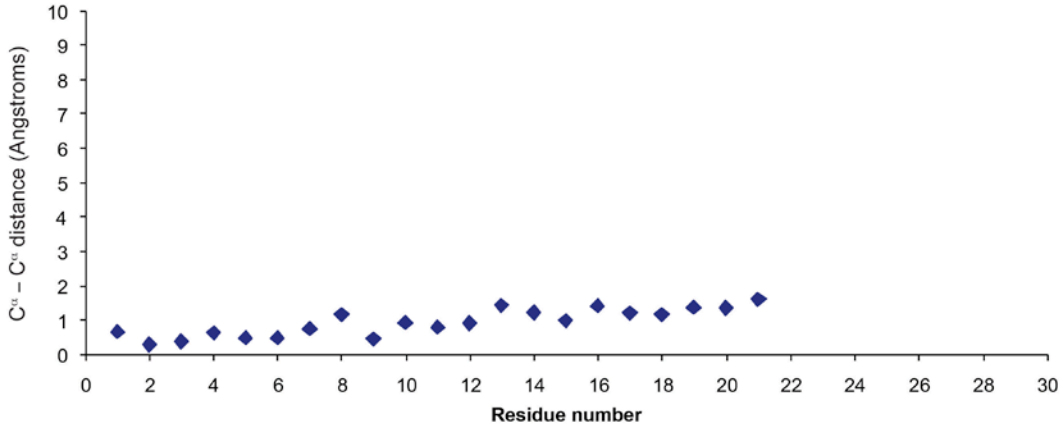
Figure S6: HPLC and MS characterization of the conversion of D-DKP ester insulin into D-DKP insulin. Saponification of D-DKP ester insulin **D-1** to give D-DKP insulin **D-17**: (a) D-DKP ester insulin **D-1** at *t* = o h (Inset: on-line ESI-MS spectra of the main peak). (b,c) Reaction mixture at (b) *t* = 3 h and (c) *t* = 18 h. *: derived from A chain. **: folded B chain. (d) Purified D-DKP insulin (Inset: on-line ESI-MS spectra of the main peak). The observed increase of 18 Da for the hydrolyzed D-DKP insulin confirms the conversion of D-DKP ester insulin to D-DKP insulin by addition of water element. The chromatographic separations were performed on a C18 column using a linear gradient (9%–53%) of buffer B in buffer A over 22 min.

Data collection	
Space group	P $\bar{1}$
Wavelength (Å)	0.97921
Cell dimensions	
a, *b*, *c* (Å)	20.5, 27.8, 34.8
α, β, γ (°)	96.4, 98.6, 104.5
Mol/asymmetric unit	1
Mol/unit cell	2
Resolution (Å)	50.0 - 1.6 (1.63 – 1.60)
R_{merge}	0.048 (0.487)
$I/\sigma I$	34.5 (3.5)
Redundancy	4.0 (4.1)
Refinement*	
Resolution (Å)	34.0 - 1.6 (1.64 – 1.60)
No. reflections	8146
Completeness (%)	89.7 (84.5)
$R_{work/}$ R_{free}	0.221 / 0.250
No. non-hydrogen atoms	459
Solvent	14
Average B-factor (Å^2)	29.4
R.m.s deviations	
Bond lengths (Å)	0.023
Bond angles (°)	2.036

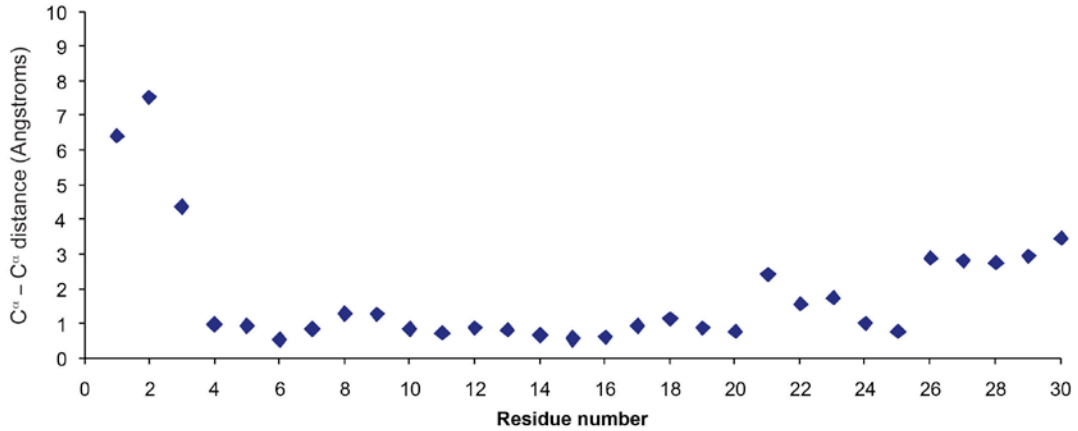
*Highest resolution shell is shown in parenthesis.

Table S1: X-ray data collection and refinement statistics for racemic DKP ester insulin PDB ID code 4IUZ.

(a)



(b)



Graph S1: Distances between Cα atoms of the (a) A chain and (b) B chain of the DKP ester insulin and KP insulin (PDB ID: 1LPH) crystal structures.