Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.
This journal is © The Royal Society of Chemistry 2016

Supplementary Information

Synthesis of fluorescent IGF-II analogues for FRET-based investigations into the binding of IGF-II to the IGF-1R

J. M. Cottam Jones,[a] P. W. R. Harris,[c,d,e] D. B. Scanlon,[a]

B. E. Forbes,[f,g] M. A. Brimble[c,d,e] and A. D. Abell[a,b]

[a.] *School of Chemistry and Physics, The University of Adelaide, Adelaide 5001, SA, Australia.*
[b.] *ARC Centre of Excellence for Nanoscale BioPhotonics (CNBP), The University of Adelaide, South Australia, Australia 5005*
[c.] *School of Chemical Sciences, The University of Auckland, 23 Symonds Street, Auckland 1010, New Zealand.*
[d.] *Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Private Bag 92019, Auckland 1010, New Zealand.*
[e.] *Institute for Innovation in Biotechnology, The University of Auckland, 3A Symonds Street, Auckland 1010, New Zealand.*
[f.] *School of Molecular and Biomedical Sciences, The University of Adelaide, Adelaide 5005, SA, Australia.*
[g.] *School of Medicine, Flinders University of South Australia, Bedford Park, 5042, SA, Australia.*

Table of Contents

1 Reagents and materials 3

2 Experimental techniques 3

3 Experimental Methods 4

3.1 Synthesis of coumaryl amino acid building blocks 4

3.2 Peptide Synthesis 5

4 Procedure for oxidative folding/disulphide bond formation 7

5 Synthesis of peptide fragments 7

5.1 Synthesis of N-terminal IGF-II (1-46) thioester 7

5.2 Synthesis of C-terminal IGF-II (47-67) fragment 8

5.3 Synthesis of IGF-II (1-20) thioester 9

5.4 Synthesis of the F19Cou IGF-II (1-20) thioester 10

5.5 Synthesis of IGF-II (Thz-46) thioester 11

5.6 F28Cou IGF-II (Thz-46) thioester 12

5.7 Two fragment synthesis of the native IGF-II protein (IGF-II (3 Fragment)) by native chemical ligation 13

5.8 Three fragment synthesis of the native IGF-II protein (IGF-II (3 Fragment)) 14

5.9 Three fragment synthesis of the F19Cou IGF-II protein 16

5.10 Three fragment synthesis of the F28Cou IGF-II protein using native chemical ligation 17

6 Competition binding assays 20

6.1 Materials 20

6.2 Procedure for competition binding assays 20

7 Experimental details for Fluorescence Resonance Energy Transfer (FRET) 21

7.1 Experimental materials 21

7.2 Procedure for FRET experiments (native IGF-II, F19Cou IGF-II, F28Cou IGF-II in *presence* of sIGF-1R) 21

7.3 Procedure for control experiments (native IGF-II, F19Cou IGF-II, F28Cou IGF-II in *absence* of sIGF-1R) 21

7.4 Summary of results from FRET experiments 22

8 Schematic of naturally occurring tryptophan residues located adjacent to the putative IGF-II binding sites 32

9 References 33

1 Reagents and materials

Solvents and reagents were of reagent grade or higher and used as supplied unless otherwise stated. Solvents used for HPLC were of HPLC grade and used without further purification. Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Cys(trt)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH), Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH) were purchased from CEM Corporation or GL Biochem (Shanghai, China). Boc-Ala-OH, Boc-Asp(OcHex)-OH, Boc-Arg(Tos)-OH, Boc-Gly-OH, Boc-Ile-OH, Boc-Leu-OH, Boc-Phe-OH, Boc-Ser(Bzl)-OH, Boc-Thr(Bzl)-OH, Boc-Tyr(Br-Bzl)-OH, Boc-Val-OH) and Boc-L-Ala-OCH_2-Phi-CH_2-COOH (Boc-Ala-PAM-COOH) linker were purchased from Polypeptide Group (Strasbourg, France), Boc-Pro-OH and Boc-Gln-OH were purchased from AusPep (Tullamarine, Victoria, Australia), and Boc-Cys(4-MeBzl)-OH, Boc-Glu(OcHex)-OH, Boc-Thz-OH) and Fmoc-PAL-tentagel resin (0.21 mmol/g) were purchased from ChemImpex (Illinois, U.S.A). 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU), N-hydroxybenzotriazole hydrate (HOBt•H_2O) were purchased from Advanced Chemtech (Louisville, KY). Diisopropylcarbodiimide (DIC) and 3-(tritylthio)propionic acid were purchased from GL Biochem (Shanghai, China). Aminomethyl polystyrene resin was made 'in house' according to the procedure described by Harris *et al.*[1, 2] Trifluoroacetic acid (TFA) was purchased from Halocarbon (New Jersey). Diisopropylethylamine (DIPEA), N-methylpyrrolidine (NMP), piperidine, 3,6-dioxa-1,8-octanedithiol (DODT), triisopropylsilane (TIPS), tris(2-carboxyethyl)phosphine hydrochloride (TCEP), mercaptopropanoic acid (MPAA), methoxyamine hydrochloride ($MeONH_2$•HCl) were purchased from Sigma Aldrich. Dimethylformamide (DMF) (AR grade), acetonitrile (HPLC grade), guanidine hydrochloride (GnHCl) were purchased from Scharlau (Sentmenat, Barcelona, Spain). Diethyl ether and disodium hydrogen phosphate (Na_2HPO_4) were purchased from Ajax chemicals. HF gas was purchased from Matheson Tri-Gas (Basking Ridge, NJ). Buffers were degassed by sparging with Ar for 30 min prior to use and pH measurements were carried out using an ISFET pocket pH meter with an accuracy of ± 0.01 pH units.

2 Experimental techniques

NMR

^{1}H and ^{13}C NMR spectra were acquired on a Varian Gemini 2000 spectrophotometer (operating at 300 and 75 MHz respectively) or Varian Inova 600 Spectrophotometer (operating at 600 and 150 MHz respectively, with a delay (D1) of 1 s). All spectra were obtained at 23 °C and chemical shifts (δ) are reported in parts per million (ppm) relative to a residual solvent peak ($CDCl_3$ (ppm): δH = 7.26, δC = 77.0; D_2O (ppm): δH = 4.79, δC: sr = -32.10 Hz; $DMSO_{-d6}$ (ppm): δH = 2.50, δC = 39.5).

Mass spectrometry

LCMS analysis was carried out on an Agilent 1120 LC compact equipped with a variable wavelength detector and coupled to a Hewlett Packard 1100 MSD mass spectrometer, using ESI in the positive mode. Gradient analysis was conducted on an Agilent Zorbax C3 (3.5 µm, 3 mm x 150 mm) column using a linear gradient of 5-65% D over 20 min with a flow rate of 0.3 mL/min, where the solvent system used was: A (MilliQ water + 0.1% formic acid) and D (acetonitrile + 0.1% formic acid). Elution of product was monitored at 220 nm and additional wavelengths (215 nm and 280 nm) were used as required.

Peptide masses were confirmed by MS analysis on a Micromass Q-TOF 2 spectrometer using Nano-ESI in the positive mode. Accurate high resolution mass spectrometry (HRMS) data was collected on an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) equipped with a Nano-ESI source in the positive mode.

RP-HPLC

Analytical RP-HPLC was conducted on an Agilent 1200 Series HPLC equipped with a diode array detector and a Supelco Discovery Biowide Pore C5 column (5µm, 4.6 x 250mm). Product was eluted using a linear gradient of 0-70% D over 30 min and a flow rate of 1 mL/min, where the solvent system used was: A (MilliQ water + 0.1% TFA) and D (acetonitrile + 0.08% TFA). Elution of the peptide was monitored at 220nm, with the use of additional wavelengths (215 nm, 280 nm and 320 nm) as required.

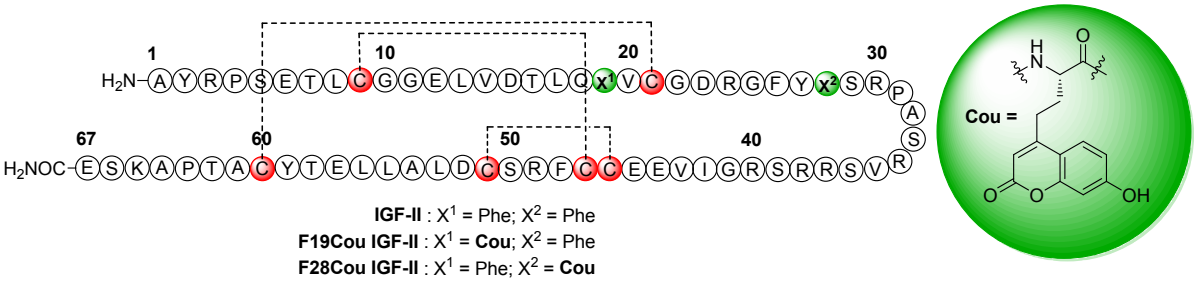
Semi-preparative RP-HPLC was carried out on a Dionex Ultimate U3000 system equipped with a multiple wavelength detector and using a Phenomenex Gemini C18 column (5 µm, 250 mm x 10 mm) or on a HP 1100

series HPLC equipped with a diode array detector using a Supelco Discovery C18 column (5 µm, 250 mm x 10 mm). The product was eluted using a linear gradient of 1% D/min, with a flow rate of 5 mL/min, where the solvent system used was: A (MilliQ water + 0.01% TFA) and D (acetonitrile + 0.08% TFA). Typically, a 0-40% acetonitrile over 40 min gradient was used; however gradient systems were adjusted according to the elution profiles and peak profiles obtained from the analytical RP-HPLC chromatograms. Fractions were analysed by LCMS. Pure fractions were pooled, and lyophilised.

Protein Quantification

IGF-II analogues were quantified by comparing analytical RP-HPLC C4 profiles with profiles of standard Long™Arg3IGF-I preparations, on an Agilent 1100 series RP-HPLC using the procedure described by Denley *et al.*[3]

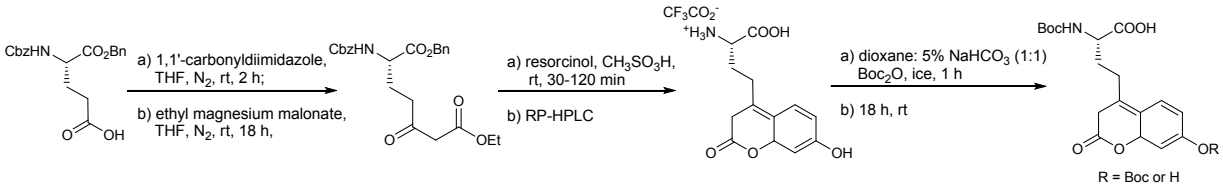
IGF-II Sequence and Analogues



3 Experimental Methods

3.1 Synthesis of coumaryl amino acid building blocks

Coumaryl amino acid building block (Cou) shown below was synthesised following the methods outlined by Brun *et al.*[4] and Wang *et al.*[5].



Ethyl magnesium malonate

An aqueous solution of magnesium chloride hexahydrate ($MgCl_2 \bullet 6H_2O$) (3.05 g, 10 ml, 1.5 M) was added to an aqueous solution of ethyl potassium malonate (5.10 g, 10 ml, 3 M). The resulting solution was stirred at rt for 30 min. The reaction mixture was diluted with isopropanol (200 ml) and the mixture was stirred at rt for a further 30 min. The mixture was then filtered and the filtrate concentrated *in vacuo* to give ethyl magnesium malonate (4.21 g, 98%) as a white crystalline solid. The product dried under vacuum and used without further purification. ^{1}H NMR (300 MHz, D_2O, δ) ppm: 4.15 (app qd, *J* = 7.2, 0.9 Hz, C*H*$_2$CH$_3$, 4H), 3.26 (app d, *J* = 0.9 Hz, $CH_3CH_2OCOCH_2COO$, 4H), 1.22 (app td, *J* = 7.2, 1.0 Hz, CH_2CH_3, 6H).

1-Benzyl 6-ethyl 2-(((benzyloxy)carbonyl)amino)-4-oxoheptanedioate

Cbz-Glu-OBn (10.98 g, 30.0 mmol) was dissolved in anhydrous THF (4 mL/mmol) and CDI (5.27 g, 32.5 mmol) was slowly added, the solution was stirred at rt for 2 h under a nitrogen environment. Ethyl magnesium malonate (4.7 g, 16.0 mmol) was added and reaction was stirred under a nitrogen environment at rt for a further 16 h. The reaction mixture was extracted with diethyl ether (3 x 50 mL) and washed with saturated aqueous $NaHCO_3$ (3 x 50 mL), water (2 x 50 mL) and brine (3 x 50 mL), dried over Na_2SO_4 and concentrated *in vacuo*. The resulting oil was purified by silica column chromatography (1:1 PE:EtOAc) to give the β-ketoester (12.4 g, 95%) as a white waxy solid. mp 52-59 °C. ^{1}H NMR (300 MHz, $CDCl_3$) δH ppm: 7.42-7.28 (m, Ar*H*, 10H), 5.38 (d, *J* = 7.7 Hz, N*H*CHCH$_2$CH$_2$, 1H), 5.16 (s, -COOC*H*$_2$Ph, 2H), 5.10 (s, PhC*H*$_2$OCONH, 2H), 4.40 (dd, *J* = 13.6, 9.2 Hz, NHC*H*CH$_2$CH$_2$,1H), 4.17 (q, *J* = 7.1 Hz, OC*H*$_2$CH$_3$, 2H), 2.71 – 2.39 (m, NHCHCH$_2$C*H*$_2$CO, 2H), 2.38 – 1.80 (m, NHCHC*H*$_2$CH$_2$CO, 2H), 1.57 (s, COC*H*$_2$CO, 2H), 1.25 (t, *J* = 7.1 Hz, -OCH$_2$C*H*$_3$, 3H). ^{13}C NMR (75 MHz, $CDCl_3$, δ)

ppm: 201.65, 171.81, 167.04, 156.07, 136.21, 135.20, 128.72, 128.62, 128.61, 128.45, 128.29, 128.18, 67.44, 67.15, 61.50, 53.27, 49.24, 38.65, 26.29, 14.15.
Spectral data matches literature:[5]

L-(7-Hydroxycoumarin-4-yl) ethylglycine trifluoroacetate salt
The β-ketoester (2.4 g, 5.5 mmol) was slowly added to a stirring solution of resorcinol (3.0 g, 27.0 mmol) in methanesulfonic acid (9.0 mL, 0.14 mol). The solution was stirred vigorously at rt for 120 min an then diluted with cold ether (200 mL). The mixture was cooled in a dry ice/acetone bath for 30 min and the precipitate was isolated by filtration, dissolved in water, and the resulting solution was filtered and lyophilised. The resulting residue was purified by semi-preparative RP-HPLC to give the coumaryl amino acid (517 mg, 28%) as an off-white solid. mp > 220 °C. $[\alpha]^{23}_{D}$ = 24.18° (c. 1.0, 1 M HCl). ^{1}H NMR (600 MHz, DMSO-*d6*, δ) ppm: 9.23 (br s, O*H*, 1H), 7.66 (d, *J* = 8.7 Hz, Ar*H*, 1H), 6.85 (dd, *J* = 8.6, 1.7 Hz, Ar*H*, 1H), 6.72 (d, *J* = 1.8 Hz, Ar*H*, 1H), 6.07 (s, CH_2CCHCO, 1H), 3.56 – 3.12 (m, $NHCHCH_2CH_2$, 1H), 3.01 – 2.67 (m, CH_2, 2H), 2.15 – 1.85 (m, CH_2, 2H). ^{13}C NMR (75 MHz, DMSO-*d6*, δ) ppm: 169.70, 161.60, 160.35, 156.24, 155.19, 126.13, 113.01, 110.79, 109.19, 102.49, 102.46, 29.90, 27.40. MS (ESI+) Calculated mass for $C_{13}H_{13}NO_5$: 263.08Da (average isotopes); observed mass: 264.13 (M + H).
Spectral data matches literature:[5]

(S)-2-((tert-Butoxycarbonyl)amino)-4-(7-hydroxy-2-oxo-2H-chromen-4-yl)butanoic acid (mono-Boc) and (S)-2-((tert-Butoxycarbonyl)amino)-4-(7-((tert-butoxycarbonyl)oxy)-2-oxo-2H-chromen-4-yl)butanoic acid (di-Boc)
A solution of the coumaryl amino acid (87 mg, 0.23 mmol) in 5% aqueous $NaHCO_3$/dioxane (1:1 v/v) (20 mL) was cooled in an ice bath and Boc anhydride (0.542 g, 2.4 mmol) was added. The reaction was stirred on ice for 1 h then allowed to warm to rt over 16 h. The reaction mixture was acidified to a pH of 3 with 10% (w/v) aqueous citric acid (8 mL) and the solution was extracted with EtOAc (3 x 20 mL), washed with water (3 x 50 mL), brine (3 x 50 mL), dried over Na_2SO_4 and concentrated *in vacuo*. The resultant oil was dissolved in a 30% aqueous acetonitrile and lyophilised to give a mixture of the *N*α-Boc protected coumaryl amino acid (76 mg, 90%) and the di-Boc protected coumaryl amino acid (8 mg, 8%) as an off-white solid in a ratio of 10:1. Mono-Boc from the mixture: ^{1}H NMR (600 MHz, DMSO-*d6*, δ) ppm: 10.58 (br s, O*H*, 1H), 7.67 (d, J = 8.6 Hz, Ar*H*, 1H), 6.84 – 6.75 (m, Ar*H*, 1H), 6.75 – 6.68 (m, Ar*H*, 1H), 6.07 (s, Ar*H*, 1H), 4.06 – 3.92 (m, $NHCHCH_2CH_2$, 1H), 2.83 – 2.70 (m, CH_2, 1H), 2.07 – 1.78 (m, CH_2, 2H), 1.47 (s, $(CH_3)_3$, 9H). Di-Boc: ^{1}H NMR (600 MHz, DMSO-d6, δ) ppm: 10.58 (br s, O*H*, 1H), 7.91 (d, J = 8.9 Hz, Ar*H*, 1H), 7.42 – 7.33 (m, Ar*H*, 1H), 7.00 - 6.84 (m, Ar*H*, 1H), 6.33 (s, Ar*H*, 1H), 3.92 – 3.80 (m, $NHCHCH_2CH_2$, 1H),), 2.93 – 2.81 (m, CH_2, 2H), 2.07 – 1.78 (m, CH_2, 2H), 1.51 (s, $(CH_3)_3$, 9H), 1.40 (s, $(CH_3)_3$, 9H). Mixture (mono-Boc and di-Boc): ^{13}C NMR (75 MHz, DMSO-*d6*, δ) ppm: 173.80, 161.19, 160.39, 158.49, 156.27, 155.70, 155.20, 126.41, 112.99, 111.07, 109.75, 106.23, 102.48, 78.23, 53.15, 29.91, 28.26, 27.96, 27.90, 27.26, 26.92.
Spectral data matches literature:[4, 5]

3.2 Peptide Synthesis

Procedure for Fmoc-based peptide synthesis (Fmoc-SPPS)
Automated peptide synthesis was carried out on a Liberty Microwave Peptide Synthesiser (CEM Corporation) using the Fmoc/*t*Bu strategy [6, 7]. Deprotection: The *N*α-Fmoc group was deprotected during each cycle by treatment of the resin with 20% v/v piperidine in DMF containing HOBt (0.10 M) for 30 s, followed by a second deprotection for 3 min. A maximum microwave power of 60 W and the maximum temperature was set to 50 °C for both deprotections. Amino acid couplings: The *N*α-Fmoc protected amino acid in DMF (5 equiv, 0.20 M), HATU in DMF (4.5 equiv, 0.45 M) and DIPEA in NMP (10 equiv, 2.0 M) were added to the resin and subjected to 25 W microwave irradiation for 6 min. Each coupling was performed twice with a maximum temperature set to 50 °C for both couplings. The exception was Fmoc-Arg(Pbf)-OH, which was coupled for 25 min at rt, followed by coupling for 5 min at 25 W microwave irradiation and the maximum temperature was set to 50 °C for both couplings.

On completion of the synthesis, a solution comprised of TFA: TIPS: DODT: water (92.5: 2.5: 2.5: 2.5 v/v/v/v) was added to the resin-bound peptide (10 mL/250 mg resin) and the mixture was shaken at rt for 3 h. The resin beads were filtered and the filtrate concentrated to approximately 2 mL under a stream of nitrogen. Ice cooled diethyl ether (40 mL) was added, and the precipitate isolated by centrifugation. The solid material was

washed with ice cooled diethyl ether (3 x 40 mL), dissolved in a 30% aqueous acetonitrile containing 0.1 % TFA and lyophilised, followed by purification by RP-HPLC as described above.

Procedure for manual peptide synthesis:-*in situ* neutralisation Boc SPPS[8]

Manual peptide synthesis was performed using the *in situ* neutralisation Boc SPPS methodology.[8] Deprotection: The *N*α-Boc group was removed by treatment with TFA (neat) (2 x 1 min) followed by flow washing of the resin with DMF for 30 s. Amino acid coupling: The *N*α-Boc protected amino acid (4 equiv) was pre-activated with a solution of HBTU in DMF (3.8 equiv, 0.38 M) and DIPEA (neat) (20 equiv) for 2 min at rt. The activated solution was added to the resin and the mixture was shaken for 10 min at rt. The solution was drained and the resin flow washed with DMF for 30 s. with the exception of the *N*α-Boc protected coumaryl amino acid (Boc-GluCou-OH) (1.2 equiv), which was pre-activated with a solution of HBTU in DMF (1.1 equiv, 0.4 M) and DIPEA (neat) (5 equiv) for 2 min at rt. The activated solution was added to the resin and the mixture was shaken at rt for 18 h. All coupling reactions were monitored *via* the Kaiser test[9] for primary amines, and where applicable the acetaldehyde/chloranil test[10, 11] for secondary amines was used.

Immediately prior to resin cleavage, the *N*α-Boc group was removed by two successive treatments with TFA (neat) for 1 min. The resin was flow washed with DMF (30 s), DCM (30 s), diethyl ether (30 s) and dried under vacuum for 30 min. The resin-bound peptide was cleaved using anhydrous HF and *p*-cresol (9:1 v/v) with stirring at 0 °C for 1 h. The HF was evaporated, ice cooled diethyl ether (40 mL) was added, and the precipitate isolated by centrifugation. The solid material was washed with ice cooled diethyl ether (3 x 40 mL), dissolved in a 30% aqueous acetonitrile solution containing 0.1 % TFA and lyophilised, followed by purification by RP-HPLC, as described above.

Confirmation of coupling and deprotection reactions:

Kaiser resin test[9] - A qualitative confirmation test was carried out after each coupling or deprotection reaction using the Kaiser test for free primary amino groups. The test involved sampling approximately twenty resin beads and treating these with 1:1:1 v/v/v of 5% (w/v) solution of ninhydrin in ethanol, phenol (80 g) in ethanol (20 mL); and 0.001 M aqueous KCN (2 mL) in pyridine (98 ml)). The mixture was heated at 100 °C for 5 min. A positive test was denoted by an intense blue coloration of the beads, indicating the presence of a free amino group. Where the test was positive, the *N*α-protected amino acid was coupled again under the same conditions.

Acetaldehyde/chloranil resin test [10, 11] - A qualitative confirmation test was carried out after each coupling or deprotection reaction using the chloranil test for free secondary amino groups The test involved sampling approximately twenty resin beads and treating these with a 2% (v/v) solution of acetaldehyde in DMF (25 µL) followed by the addition of of a 2% (w/v) solution of chloranil in DMF (25 µL). A positive test was denoted by a an dark blue/green coloration of the beads, indicating the presence of a free amino group Where the test was positive, the *N*α-protected amino acid was coupled again under the same conditions. Solutions were stored in the fridge and were kept for a maximum of 1 week.

Procedure for solid phase extraction (SPE)

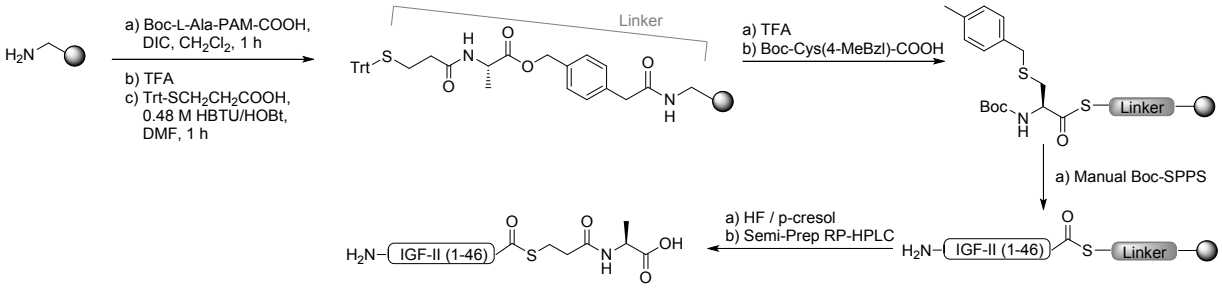
The crude peptide mixture was isolated using a solid phase extraction (SPE) cartridge (Alltech, C18LP, 600 mg). The cartridge was prepared by washing with MeOH (5.0 mL) and equilibrated with 5% aqueous acetonitrile containing 0.1% TFA (10.0 mL). The ligation mixture was diluted with water and acidified with TFA if necessary to give a pH of 3, and loaded onto the prepared SPE cartridge at a rate of approximately 2 mL/min. Salts and non-binding contaminants were eluted with 5% aqueous acetonitrile containing 0.1% TFA (10.0 mL) and the crude peptide was eluted on washing the cartridge with 50% aqueous acetonitrile containing 0.1% TFA (10.0 mL). Fractions containing the desired peptide were combined and lyophilised.

4 Procedure for oxidative folding/disulphide bond formation

Oxidative folding of the IGF-II peptides was carried out as described by Delaine *et al.*[12]. Specifically, the purified peptide (4.0 mg/mL) was dissolved in a buffer consisting of 8.0 M urea, 0.10 M Tris and 40 mM glycine containing 20 mM dithiothreitol (DTT). The solution was incubated at rt for 1 h and reduction of any disulphide adducts was monitored by analytical RP-HPLC. After 1 h the reduced solution was rapidly diluted with refolding buffer (1.0 M Tris-HCl, 2.0 mM EDTA, pH 9.1 and sterilised through 1.0 μm filter) to give a solution with a final concentration of 2.5 M urea, 0.7 M Tris, 12.5 mM glycine, 2.0 mM EDTA, 0.50 mM DTT, 1.25 mM 2-hydroxyethyldisulfide (2-HED) and a final protein concentration of 0.10 mg/mL This reaction mixture was slowly stirred at rt and formation of the folded IGF-II protein was monitored by analytical RP-HPLC. The reaction was determined complete when no more of the reduced peptide remained, typically after 120 min. The reaction mixture was quenched with conc. HCl to give a pH of 3. The acidified solution was filtered through a 0.45 μM syringe filter and the filtrate purified by semi-preparative RP-HPLC on an HP 1100 series HPLC.

5 Synthesis of peptide fragments

5.1 Synthesis of N-terminal IGF-II (1-46) thioester



Aminomethyl resin (100 mg, 0.10 mmol)[1, 2] was washed with DCM (3 x 5.0 mL), swollen in a mixture of DCM and DMF (1:1 v/v) for 15 min and washed with further DCM (3 x 5.0 mL). A solution of Boc-L-Ala-PAM-COOH (67.8 mg, 0.2 mmol) and DIC (33 μL, 0.2 mmol) in DCM (3.0 mL) was added to the resin and the mixture was shaken for 1 h at rt. The solution was drained and the resin flow washed with DCM (30 s) and DMF (30 s). The Boc group was removed from the resin bound PAM linker on treatment with TFA (neat) (5.0 mL) (2 x 1 min), followed by flow washing with DMF (30 s). A solution of S-Trityl-β-mercaptopropionic acid (191 mg, 0.55 mmol) activated with HBTU (0.4 M in DMF, 1.3 mL) was added to the resin followed by DIPEA (198 μL, 1.1 mmol). The resin was shaken for 20 min and then flow washed with DMF (30 s). The S-trityl group was removed by treatment of the resin twice with a solution of TFA, TIPS and water (95:2.5:2.5, v/v/v) (5.0 mL) (2 x 1 min), followed by flow washing with DMF (30 s). A solution of Boc-Cys(4-MeBzl)-OH (180 mg, 0.55 mmol) in HBTU in DMF (0.4 M in DMF, 1.3 mL) and DIPEA (198 μL, 1.1 mmol) was added to the resin and the mixture was shaken for 1 h and finally flow washed with DMF (30 s) to give the cysteine alkyl thioester linked resin.

Cysteine alkyl thioester linked resin was elongated using the Boc-SPPS protocol described in Section 3.2 to give the resin-bound IGF-II (1-46) thioester. The resin was cleaved with HF and the resulting crude peptide (Figure S1A) was dissolved in a 5% aqueous acetonitrile containing 0.1% TFA and purified by semi-preparative RP-HPLC to afford the IGF-II (1-46) thioester (12.5 mg, 7%) as a white solid (Figure S1B). LCMS (ESI+) calculated mass for $C_{228}H_{359}N_{69}O_{71}S_4$: 5330.9859Da (average isotopes); observed m/z: 1777.72 ($[M+3H]^{+3}$), 1333.47 ($[M+4H]^{+4}$), 1067.00 ($[M+5H]^{+5}$), 889.36 ($[M+6H]^{+6}$), 762.41 ($[M+7H]^{+7}$), 667.25 ($[M+8H]^{+8}$) and 593.21 ($[M+9H]^{+9}$) (Figure S1C)

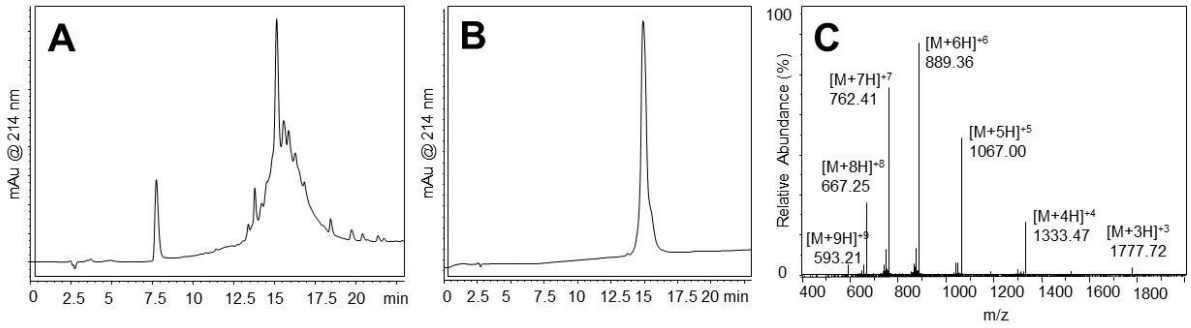
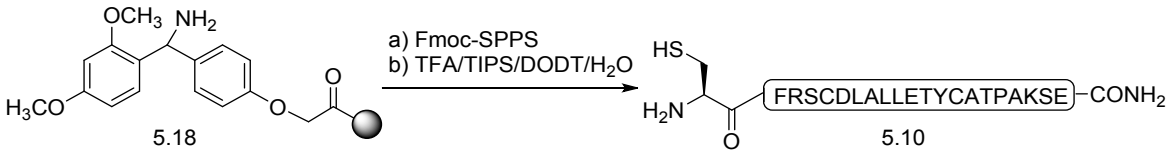


Figure S1: Analysis of the N-terminal IGF-II (1-46) thioester. a) RP-HPLC analysis of the crude IGF-II (1-46) thioester; b) RP-HPLC analysis of the purified IGF-II (1-46) thioester; c) ESI-MS of the purified IGF-II (1-46) thioester. Calculated mass for $C_{228}H_{359}N_{69}O_{71}S_4$: 5330.9859Da (average isotopes); observed m/z: 1777.72 ($[M+3H]^{+3}$), 1333.47 ($[M+4H]^{+4}$), 1067.00 ($[M+5H]^{+5}$), 889.36 ($[M+6H]^{+6}$), 762.41 ($[M+7H]^{+7}$), 667.25 ($[M+8H]^{+8}$) and 593.21 ($[M+9H]^{+9}$).

5.2 Synthesis of C-terminal IGF-II (47-67) fragment



IGF-II (47-67) resin-bound peptide was synthesised on a 0.10 mmol scale from Tentgel® S RAM resin using the Fmoc-based SPPS approach described in Section 3.2. The crude peptide (Figure S2A) was dissolved in a 5% aqueous acetonitrile containing 0.1% TFA and purified by semi-preparative RP-HPLC to give the IGF-II (47-67) fragment (30 mg, 13%) as a white solid (Figure S2B). LCMS (ESI+) calculated mass for $C_{99}H_{158}N_{26}O_{32}S_3$: 2320.6720Da (average isotopes); observed m/z: 1160.84 ($[M+2H]^{+2}$), 774.25 ($[M+3H]^{+3}$) and 581.02 ($[M+4H]^{+4}$) (Figure S2C).

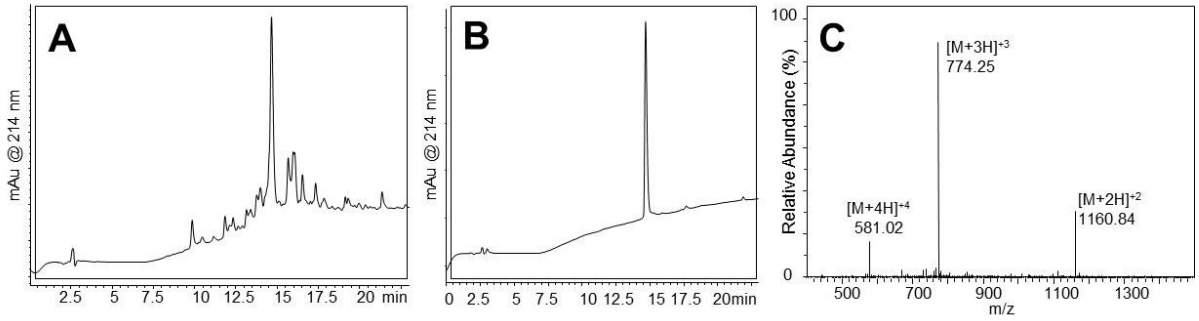
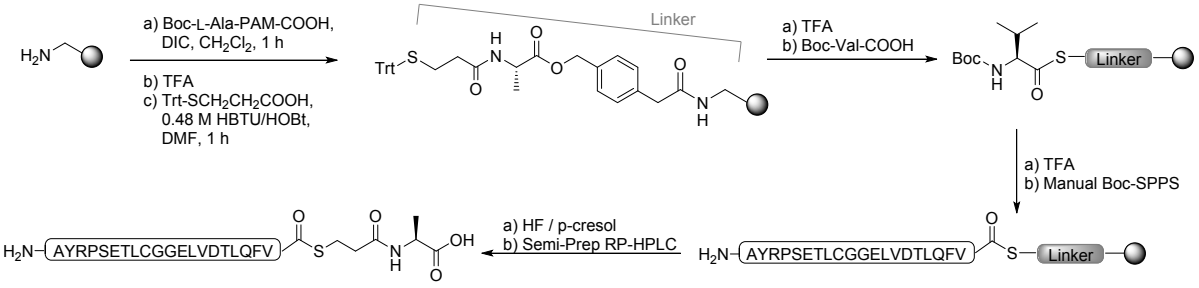


Figure S2: Analysis of the C-terminal IGF-II (47-67) fragment. a) RP-HPLC of the crude IGF-II (47-67) fragment; b) RP-HPLC of the purified IGF-II (47-67) fragment; c) ESI-MS of the purified IGF-II (47-67) fragment. Calculated mass for $C_{99}H_{158}N_{26}O_{32}S_3$: 2320.6720Da (average isotopes); observed m/z: 1160.84 ($[M+2H]^{+2}$), 774.25 ($[M+3H]^{+3}$) and 581.02 ($[M+4H]^{+4}$).

5.3 Synthesis of IGF-II (1-20) thioester



Aminomethyl resin (100 mg, 0.10 mmol)[1, 2] was washed with DCM (3 x 5.0 mL), swollen in a mixture of DCM and DMF mixture (1:1 v/v) for 15 min and washed with further DCM (3 x 5.0 mL). A solution of Boc-L-Ala-PAM-COOH (71.6 mg, 0.21 mmol) and DIC (33 µL, 0.21 mmol) in DCM (3.0 mL) was added to the resin and the mixture was shaken for 1 h at rt. The solution was drained and the resin flow washed with DCM (30 s) and DMF (30 s). The Boc group was removed on treatment of the resin with TFA (neat) (5.0 mL) twice for 1 min. The resin was flow washed with DMF (30 s). A solution of S-trityl-β-mercaptopropionic acid (191 mg, 0.55 mmol) in HBTU (0.4 M in DMF, 1.3 mL) was added to the resin followed by DIPEA (198 µL, 1.1 mmol). The resin was shaken for 20 min and then flow washed with DMF (30 s). The S-trityl group was removed on treatment with a solution of TFA, TIPS and water (95:2.5:2.5, v/v/v) (5.0 mL) (2 x 1 min) and then flow washed with DMF (30 s). Thioesterification was achieved by adding a solution of Boc-Val-OH (119 mg, 0.55 mmol) in HBTU (0.4 M in DMF, 1.3 mL) and DIPEA (198 µL, 1.1 mmol) to the resin. The mixture was shaken for 1 h, the solution drained and the resin flow washed with DMF (30 s) to give the valine alkyl thioester linked resin.

Valine alkyl thioester linked resin was elongated using the Boc-SPPS protocol described in Section 3.2 to give the resin-bound IGF-II (1-20) thioester. The resin was cleaved with HF and the resulting crude peptide (Figure S3A) was dissolved in a 5% aqueous acetonitrile containing 0.1% TFA and purification by semi-preparative RP-HPLC gave the IGF-II (1-20) thioester (41.3 mg, 17%) as a white solid (Figure S3B). LCMS (ESI+) calculated mass for $C_{103}H_{161}N_{25}O_{34}S_2$: 2357.6648Da (average isotopes); observed m/z: 1179.53 ($[M+2H]^{+2}$), 786.70 ($[M+3H]^{+3}$) and 590.26 ($[M+4H]^{+4}$) (Figure S3C).

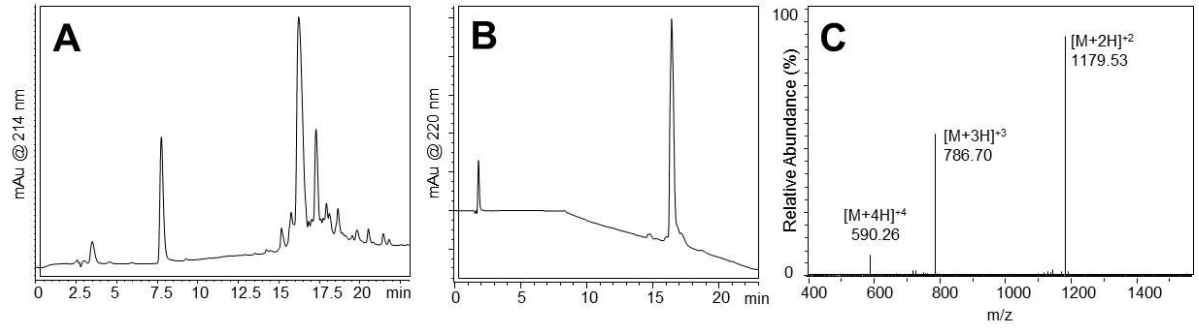
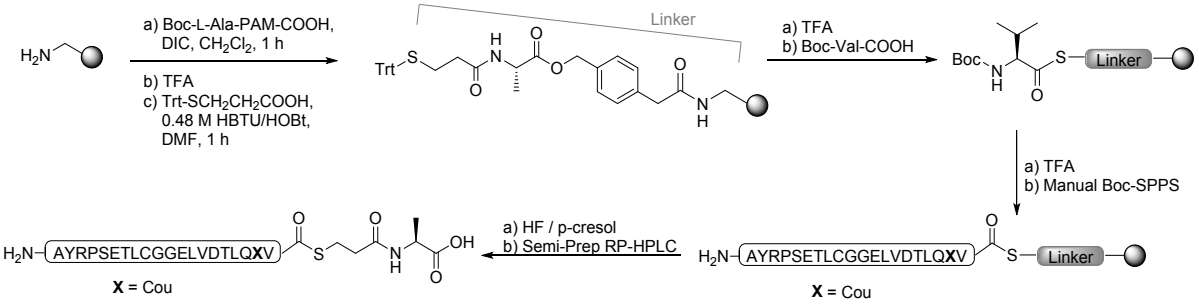


Figure S3 Analysis of the N-terminal IGF-II (1-20) thioester. a) RP-HPLC analysis of the crude IGF-II (1-20) thioester; b) RP-HPLC analysis of the purified IGF-II (1-20) thioester; c) ESI-MS of the purified IGF-II (1-20) thioester. Calculated mass for $C_{103}H_{161}N_{25}O_{34}S_2$: 2357.6648Da (average isotopes); observed m/z: 1179.53 ($[M+2H]^{+2}$), 786.70 ($[M+3H]^{+3}$) and 590.26 ($[M+4H]^{+4}$).

5.4 Synthesis of the F19Cou IGF-II (1-20) thioester



Aminomethyl resin (100 mg, 0.10 mmol)[1, 2] was washed with DCM (3 x 5.0 mL), swollen in a mixture of DCM and DMF mixture (1:1 v/v) for 15 min and washed with further DCM (3 x 5.0 mL). A solution of Boc-L-Ala-PAM-COOH (71.6 mg, 0.21 mmol) and DIC (33 µL, 0.21 mmol) in DCM (3.0 mL) was added to the resin and the mixture was shaken for 1 h at rt. The solution was drained and the resin flow washed with DCM (30 s) and DMF (30 s). The Boc group was removed on treatment of the resin with TFA (neat) (5.0 mL) twice for 1 min. The resin was flow washed with DMF (30 s). A solution of S-trityl-β-mercaptopropionic acid (191 mg, 0.55 mmol) in HBTU (0.4 M in DMF, 1.3 mL) was added to the resin followed by DIPEA (198 µL, 1.1 mmol). The resin was shaken for 20 min and then flow washed with DMF (30 s). The S-trityl group was removed on treatment with a solution of TFA, TIPS and water (95:2.5:2.5, v/v/v) (5.0 mL) (2 x 1 min) and then flow washed with DMF (30 s). Thioesterification was achieved by adding a solution of Boc-Val-OH (119 mg, 0.55 mmol) in HBTU (0.4 M in DMF, 1.3 mL) and DIPEA (198 µL, 1.1 mmol) to the resin. The mixture was shaken for 1 h, the solution drained and the resin flow washed with DMF (30 s) to give the valine alkyl thioester linked resin.

Valine alkyl thioester linked resin was elongated using the Boc-SPPS protocol described in Section 3.2 to give the resin-bound F19Cou IGF-II (1-20) thioester. The resin was cleaved with HF and the resulting crude peptide (Figure S4A) was dissolved in a 5% aqueous acetonitrile containing 0.1% TFA and purification by semi-preparative RP-HPLC gave the F19Cou IGF-II (1-20) thioester (3.7 mg, 2%) as a white solid (Figure S4B). LCMS (ESI+) Calculated mass for C107H163N25O37S2: 2455.7218Da (average isotopes); observed m/z: 1228.48 ([M+2H]+2), 819.36 ([M+3H]+3) and 614.82 ([M+4H]+4) (Figure S4C).

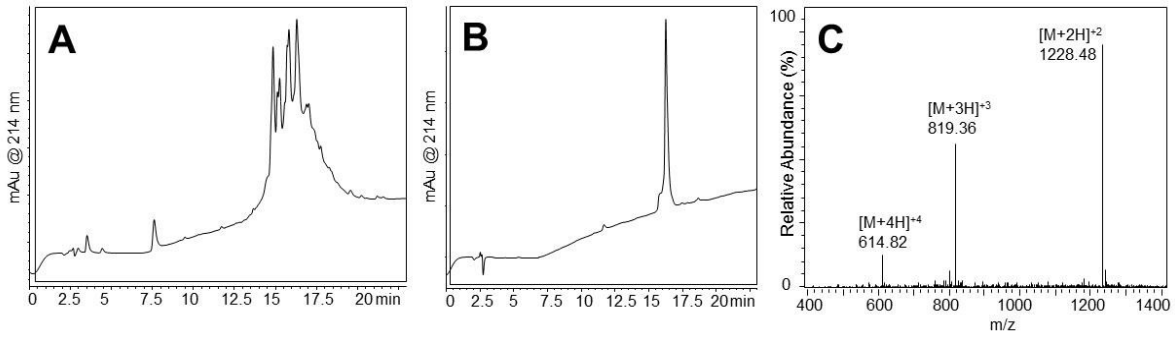
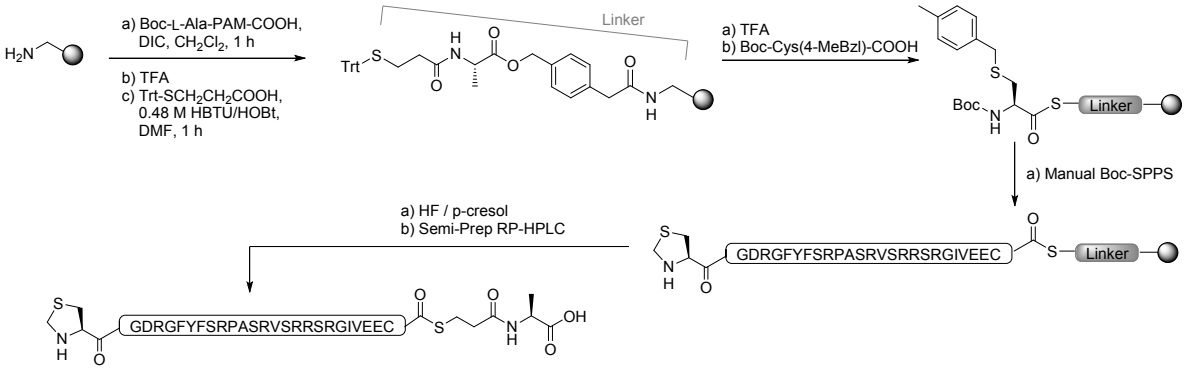


Figure S4: Analysis of the N-terminal F19Cou IGF-II (1-20) thioester. a) RP-HPLC analysis of the crude F19Cou IGF-II (1-20) thioester; b) RP-HPLC analysis of the purified F19Cou IGF-II (1-20) thioester; c) ESI-MS of the purified F19Cou IGF-II (1-20) thioester. Calculated mass for $C_{107}H_{163}N_{25}O_{37}S_2$: 2455.7218Da (average isotopes); observed m/z: 1228.48 ($[M+2H]^{+2}$), 819.36 ($[M+3H]^{+3}$) and 614.82 ($[M+4H]^{+4}$).

5.5 Synthesis of IGF-II (Thz-46) thioester



Aminomethyl resin (100 mg, 0.10 mmol)[1, 2] was washed with DCM (3 x 5.0 mL), swollen in a mixture of DCM and DMF mixture (1:1 v/v) for 15 min and washed with further DCM (3 x 5.0 mL). A solution of Boc-L-Ala-PAM-COOH (71.6 mg, 0.21 mmol) and DIC (33 µL, 0.21 mmol) in DCM (3.0 mL) was added to the resin and the mixture was shaken for 1 h at rt. The solution was drained and the resin flow washed with DCM (30 s) and DMF (30 s). The Boc group was removed on treatment of the resin with TFA (neat) (5.0 mL) twice for 1 min. The resin was flow washed with DMF (30 s). A solution of S-trityl-β-mercaptopropionic acid (191 mg, 0.55 mmol) in HBTU (0.4 M in DMF, 1.3 mL) was added to the resin followed by DIPEA (198 µL, 1.1 mmol). The resin was shaken for 20 min and then flow washed with DMF (30 s). The S-trityl group was removed on treatment with a solution of TFA, TIPS and water (95:2.5:2.5, v/v/v) (5.0 mL) (2 x 1 min) and then flow washed with DMF (30 s). Thioesterification was achieved by adding a solution of Boc-Cys(4-MeBzl)-OH (180 mg, 0.55 mmol) in HBTU in DMF (0.4 M in DMF, 1.3 mL) and DIPEA (198 µL, 1.1 mmol) to the resin.The mixture was shaken for 1 h, the solution drained and the resin flow washed with DMF (30 s) to give the cysteine alkyl thioester linked resin.

Cysteine alkyl thioester linked resin was elongated using the Boc-SPPS protocol described in Section 3.2 to give the resin-bound IGF-II (Thz-46) thioester. The resin was cleaved with HF and the resulting crude peptide (Figure S5A) was dissolved in a 5% aqueous acetonitrile containing 0.1% TFA and purified by semi-preparative RP-HPLC to give the IGF-II (Thz-46) thioester (12.3 mg, 11%) as a white solid (Figure S5B). LCMS (ESI+) Calculated mass for $C_{136}H_{209}N_{45}O_{40}S_3$: 3155.5546Da (average isotopes); observed m/z: 1054.93 ($[M+3H]^{+3}$), 791.52 ($[M+4H]^{+4}$), 633.40 ($[M+5H]^{+5}$), 528.00 ($[M+6H]^{+6}$) and 452.75 ($[M+7H]^{+7}$) (Figure S5C).

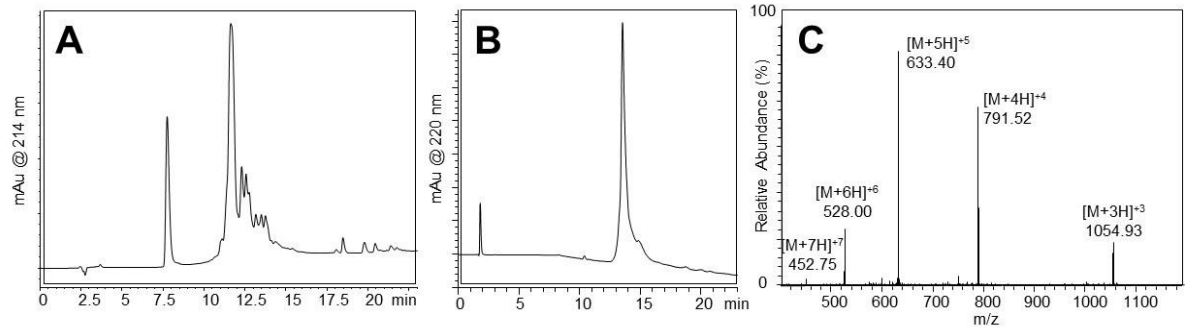
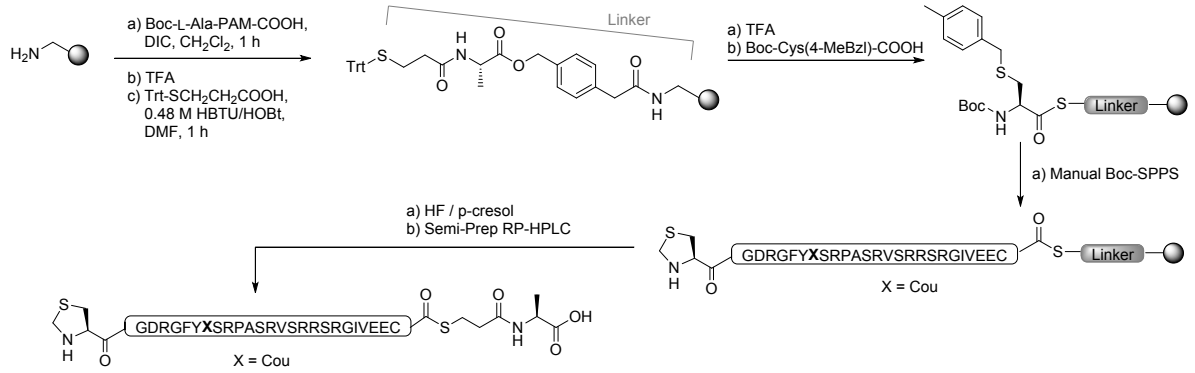


Figure S5: Analysis of the IGF-II (Thz-46) thioester. a) RP-HPLC analysis of the crude IGF-II (Thz-46) thioester; b) RP-HPLC analysis of purified IGF-II (Thz-46) thioester; c) ESI-MS of purified IGF-II (Thz-46) thioester. Calculated mass for $C_{136}H_{209}N_{45}O_{40}S_3$: 3162.5546Da (average isotopes); observed m/z: 1054.93 ($[M+3H]^{+3}$), 791.52 ($[M+4H]^{+4}$), 633.40 ($[M+5H]^{+5}$), 528.00 ($[M+6H]^{+6}$) and 452.75 ($[M+7H]^{+7}$).

5.6 F28Cou IGF-II (Thz-46) thioester



Aminomethyl resin (100 mg, 0.10 mmol)[1, 2] was washed with DCM (3 x 5.0 mL), swollen in a mixture of DCM and DMF mixture (1:1 v/v) for 15 min and washed with further DCM (3 x 5.0 mL). A solution of Boc-L-Ala-PAM-COOH (71.6 mg, 0.21 mmol) and DIC (33 µL, 0.21 mmol) in DCM (3.0 mL) was added to the resin and the mixture was shaken for 1 h at rt. The solution was drained and the resin flow washed with DCM (30 s) and DMF (30 s). The Boc group was removed on treatment of the resin with TFA (neat) (5.0 mL) twice for 1 min. The resin was flow washed with DMF (30 s). A solution of S-trityl-β-mercaptopropionic acid (191 mg, 0.55 mmol) in HBTU (0.4 M in DMF, 1.3 mL) was added to the resin followed by DIPEA (198 µL, 1.1 mmol). The resin was shaken for 20 min and then flow washed with DMF (30 s). The S-trityl group was removed on treatment with a solution of TFA, TIPS and water (95:2.5:2.5, v/v/v) (5.0 mL) (2 x 1 min) and then flow washed with DMF (30 s). Thioesterification was achieved by adding a solution of Boc-Cys(4-MeBzl)-OH (180 mg, 0.55 mmol) in HBTU in DMF (0.4 M in DMF, 1.3 mL) and DIPEA (198 µL, 1.1 mmol) to the resin.The mixture was shaken for 1 h, the solution drained and the resin flow washed with DMF (30 s) to give the cysteine alkyl thioester linked resin.

Cysteine alkyl thioester linked resin was elongated according to the Boc-SPPS protocol described in Section 3.2 to give the resin-bound F28Cou IGF-II (Thz-46) thioester. The resin was cleaved with HF and the resulting crude peptide (Figure S6A) was dissolved in a 5% aqueous acetonitrile containing 0.1% TFA and purified by semi-preparative RP-HPLC to afford the F28Cou IGF-II (Thz-46) thioester (11.8 mg, 3%) as a white solid (Figure S6B). LCMS (ESI+) Calculated mass for $C_{13}6H_{211}N_{45}O_{43}S_3$ for: 3260.6117Da (average isotopes); observed m/z: 1630.93 ($[M+2H]^{+2}$), 1087.58 ($[M+3H]^{+3}$), 815.99 ($[M+4H]^{+4}$), 653.00 ($[M+5H]^{+5}$), 544.33 ($[M+6H]^{+6}$) and 466.73 ($[M+7H]^{+7}$) (Figure S6C).

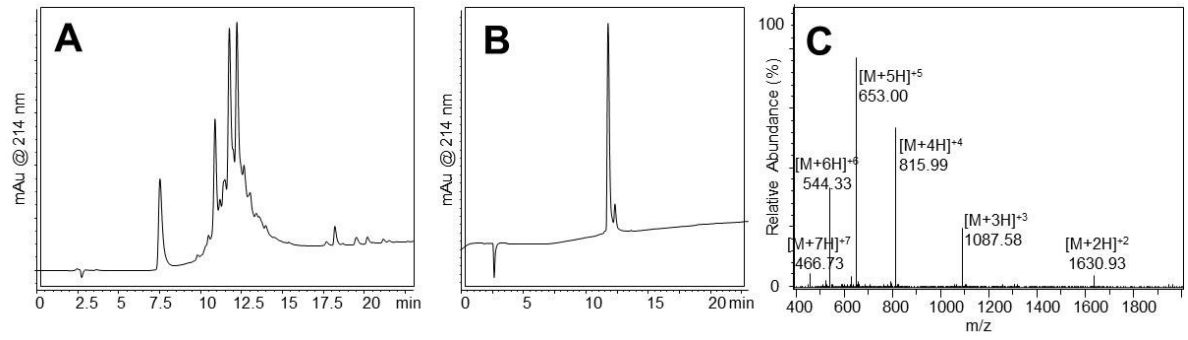
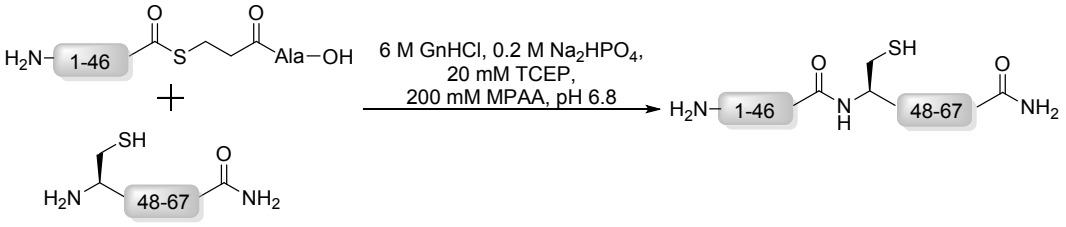


Figure S6: Analysis of the F28Cou IGF-II (Thz-46) thioester. a) RP-HPLC analysis of the crude F28Cou IGF-II (Thz-46) thioester; b) RP-HPLC analysis of the purified F28Cou IGF-II (Thz-46) thioester; c) ESI-MS of the purified F28Cou IGF-II (Thz-46) thioester. Calculated mass for. $C_{136}H_{211}N_{45}O_{43}S_3$ for: 3260.6117Da (average isotopes); observed m/z: 1630.93 ($[M+2H]^{+2}$), 1087.58 ($[M+3H]^{+3}$), 815.99 ($[M+4H]^{+4}$), 653.00 ($[M+5H]^{+5}$), 544.33 ($[M+6H]^{+6}$) and 466.73 ($[M+7H]^{+7}$).

5.7 Two fragment synthesis of the native IGF-II protein (IGF-II (3 Fragment)) by native chemical ligation



N-terminal IGF-II (1-46) thioester (6.8 mg, 1.5 µmol, 3.0 mM), C-terminal IGF-II (47-67) fragment (3.4 mg, 1.3 µmol, 3.5 mM), TCEP (20.0 mM) and MPAA (200 mM) were dissolved in degassed native chemical ligation buffer (6.0 M GnHCl and 0.20 M Na_2HPO_4, 423 µL). The pH of the solution was adjusted to 6.9 using aqueous NaOH (10 M and 2 M), the reaction mixture was sparged with Ar (10 s) and shaken at rt for 1 h. The products were isolated using SPE according to the procedure described in Section 3.2 and purified by semi-preparative RP-HPLC, using a linear gradient of 0-40% acetonitrile over 40 min to give the native IGF-II peptide (0.4 mg, 4%) as a white solid. The native IGF-II peptide (0.06 mg, 8 nmol) was subsequently folded as detailed in Section 4, to give the native IGF-II protein (2.7 µg, 5%) as a white solid (Figure S8A). HRMS (ESI+) Calculated mass for $C_{321}H_{500}N_{94}O_{100}S_6$: 7468.3875Da (average isotopes); observed m/z: 1494.74 ($[M+5H]^{+5}$), 1245.79 ($[M+6H]^{+6}$) and 1067.93 ($[M+7H]^{+7}$) (Figure S8B). Analytical data for the ligation of the N-terminal IGF-II (1-46) thioester to the C-terminal IGF-II (47-67) fragment in shown in Figure S7.

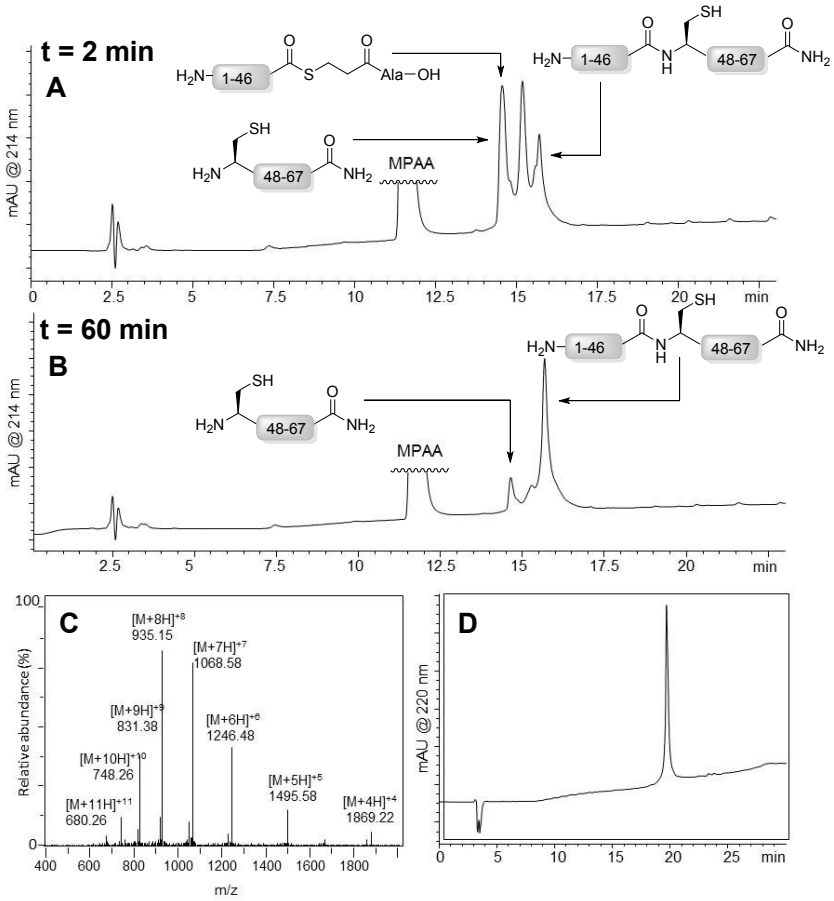


Figure S7: LCMS analysis of the cysteine-based ligation between the IGF-II (1-46) thioester and C-terminal IGF-II (47-67) fragment after a) 2 min and b) 60 min; c) ESI-MS of IGF-II (1-67) peptide; d) RP-HPLC of the purified IGF II (1-67) peptide.

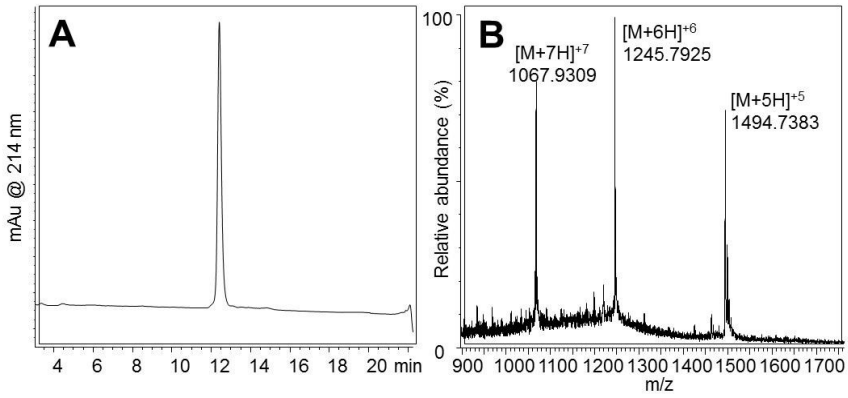
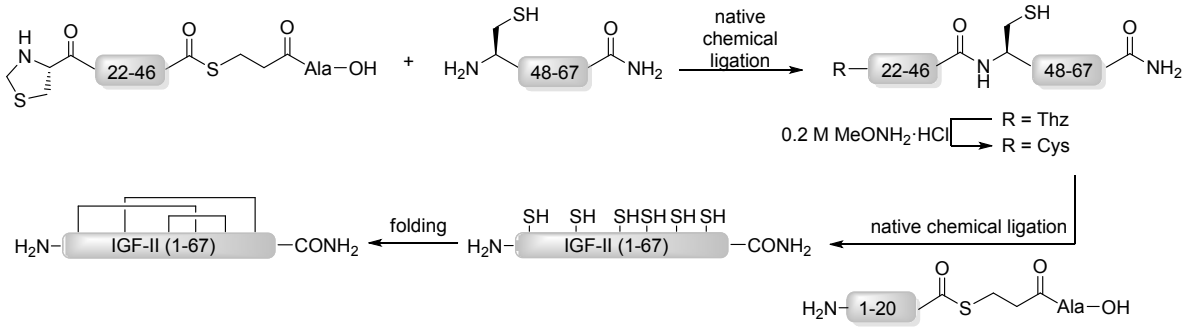


Figure S8: Characterisation of the synthetic native IGF-II protein synthesised using the two fragment approach. a) RP-HPLC trace of the purified native IGF-II protein; b) HRMS of the purified native IGF-II protein. Calculated mass. for $C_{321}H_{500}N_{94}O_{100}S_6$: 7468.3875Da (average isotopes); observed m/z: 1494.7383 ([M+5H]$^{+5}$), 1245.7925 ([M+6H]$^{+6}$) and 1067.9309 ([M+7H]$^{+7}$).

5.8 Three fragment synthesis of the native IGF-II protein (IGF-II (3 Fragment))



IGF-II (Thz-46) thioester (4.6 mg, 1.4 µmol, 3.00 mM), C-terminal IGF-II (47-67) fragment (3.8 mg, 1.6 µmol, 3.35 mM), TCEP (20.0 mM) and MPAA (200 mM) were dissolved in degassed native chemical ligation buffer (6.0 M GnHCl and 0.20 M Na_2HPO_4, 487 µL), pH of the solution was adjusted to 6.9 using aqueous NaOH (10.0 M and 2.0 M) and shaken at rt for 1 h. $MeONH_2$•HCl (8.1 mg, 97 µmol, 0.20 M) was then added, the pH of the solution was adjusted to 3.9 with aqueous HCl (5.0 M), the reaction mixture was shaken at rt for 6 h. The pH of the solution was adjusted to 6.8 with aqueous NaOH (10.0 M and 2.0 M) and the N-terminal IGF-II (1-20) thioester (3.9 mg, 1.6 µmol, 3.37 mM) was added. The reaction mixture was shaken at rt for 47 h. The products were isolated using SPE according to the procedure described in Section 3.2 and purified by semi-preparative RP-HPLC to give the native IGF-II peptide (0.44 mg, 4%) as a white solid. The native IGF-II peptide (0.44 mg, 0.058 µmol) was subsequently folded as detailed in Section 4 to give the native IGF-II protein (25 µg, 6%) as a white solid (Figure S10A). HRMS (ESI+) Calculated mass for $C_{321}H_{500}N_{94}O_{100}S_6$: 7468.3875Da (average isotopes); observed m/z: 1494.7144Da ([M+5H]$^{+5}$), 1245.7516 ([M+6H]$^{+6}$) and 1067.9352 ([M+7H]$^{+7}$) (Figure S10B). Analytical data for the one-pot three fragment synthesis of the native IGF-II protein is summarised in Figure S9.

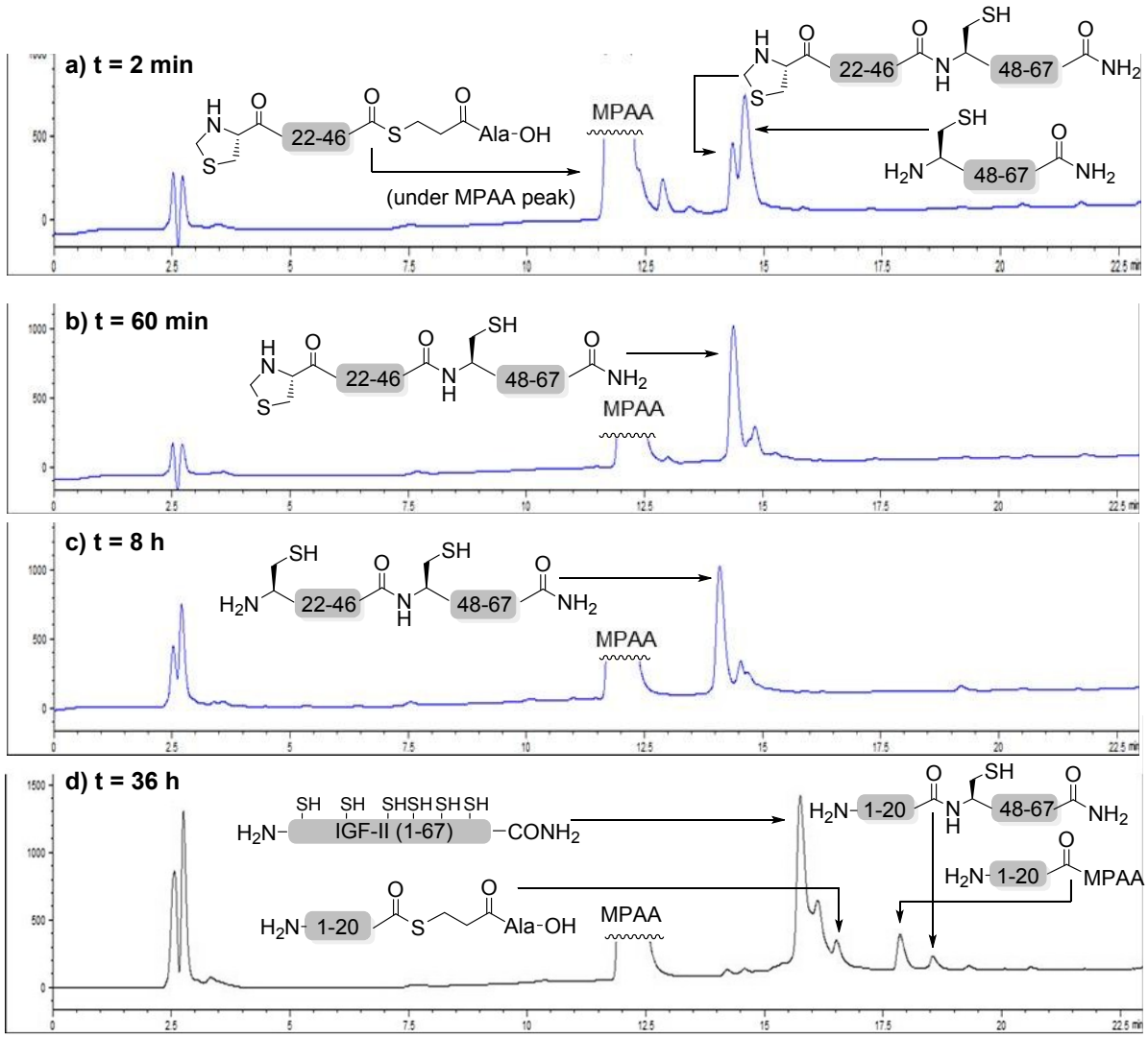


Figure S9: LCMS analysis of the one-pot three fragment synthesis of the native IGF-II peptide. Analysis of the cysteine-based ligation between the IGF-II (Thz-46) thioester and C-terminal IGF-II (47-67) fragment after a) 2 min and b) 60 min; c) Analysis of the thiazolidone deprotection of the IGF-II (Thz-67) fragment after 8 h; d) LCMS analysis of the valine-based ligation between the N-terminal IGF-II (1-20) thioester and IGF-II (21-67) fragment after 36 h.

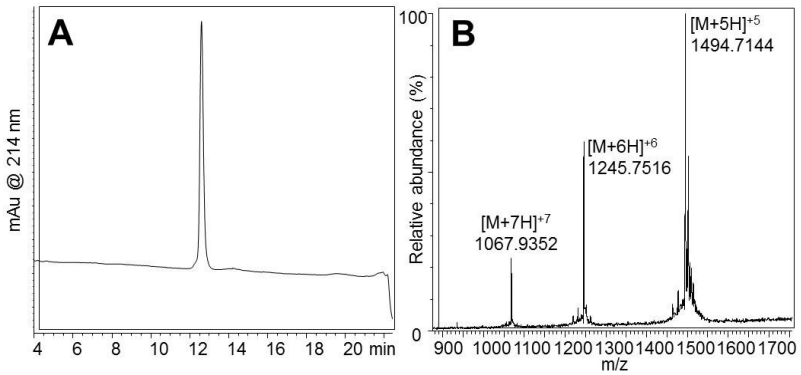
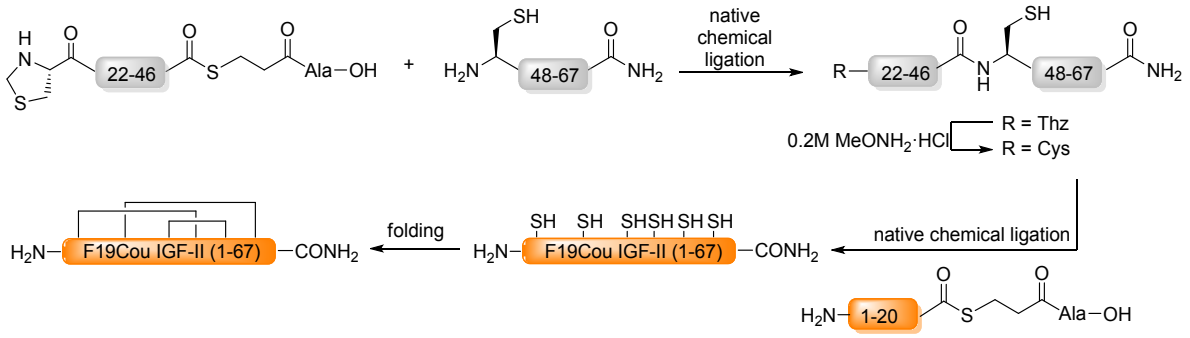


Figure S10: Characterisation of the synthetic native IGF-II protein synthesised using the three fragment approach. a) RP-HPLC trace of the native IGF-II (3 Fragment); b) HRMS of the native IGF-II (3 Fragment).

Calculated mass for. for $C_{321}H_{500}N_{94}O_{100}S_6$: 7468.3875Da (average isotopes); observed m/z: 1494.7144 ($[M+5H]^{+5}$), 1245.7516 ($[M+6H]^{+6}$) and 1067.9352 ($[M+7H]^{+7}$).

5.9 Three fragment synthesis of the F19Cou IGF-II protein



C-terminal IGF-II (47-67) fragment (3.15 mg, 1.36 µmol, 3.25 mM), IGF-II (Thz-46) thioester (3.96 mg, 1.25 µmol, 3.00 mM), TCEP (20.0 mM) and MPAA (200 mM) were dissolved in degassed native chemical ligation buffer (6.0 M GnHCl and 0.20 M Na_2HPO_4, 418 µL), pH of the solution was adjusted to 6.9 using aqueous NaOH (10.0 M and 2.0 M) and shaken at rt for 1 h. $MeONH_2$•HCl (7.0 mg, 84 µmol, 0.20 M) was then added, the pH of the solution was adjusted to 4.0 with aqueous HCl (5.0 M), the reaction mixture was shaken at rt 16 h. The pH of the solution was adjusted to 7.0 using aqueous NaOH (10.0 M and 2.0 M) and the N-terminal F19Cou IGF-II (1-20) thioester (3.47 mg, 1.40 µmol, 3.38 mM) was added. The reaction mixture was shaken at rt for 27 h. The products were isolated using SPE according to Section 3.2 and purified by semi-preparative RP-HPLC to give the F19Cou IGF-II peptide (0.84 mg, 9%) as a white solid. F19Cou IGF-II peptide 4.23 (0.84 mg, 0.11 µmol) was subsequently folded according to Section 4 to give the F19Cou IGF-II protein (89 µg, 11%) as a white solid (Figure S13A). HRMS (ESI+) Calculated mass for $C_{325}H_{500}N_{94}O_{103}S_6$: 7566.4445Da (average isotopes); observed m/z: 1514.1628 ($[M+5H]^{+5}$), 1261.9639 ($[M+6H]^{+6}$), 1081.8225 ($[M+7H]^{+7}$) and 946.7245 ($[M+8H]^{+8}$) (Figure S13B). Analytical data for the one-pot three fragment synthesis of the F19Cou IGF-II protein is summarised in Figure S11 and Figure S12.

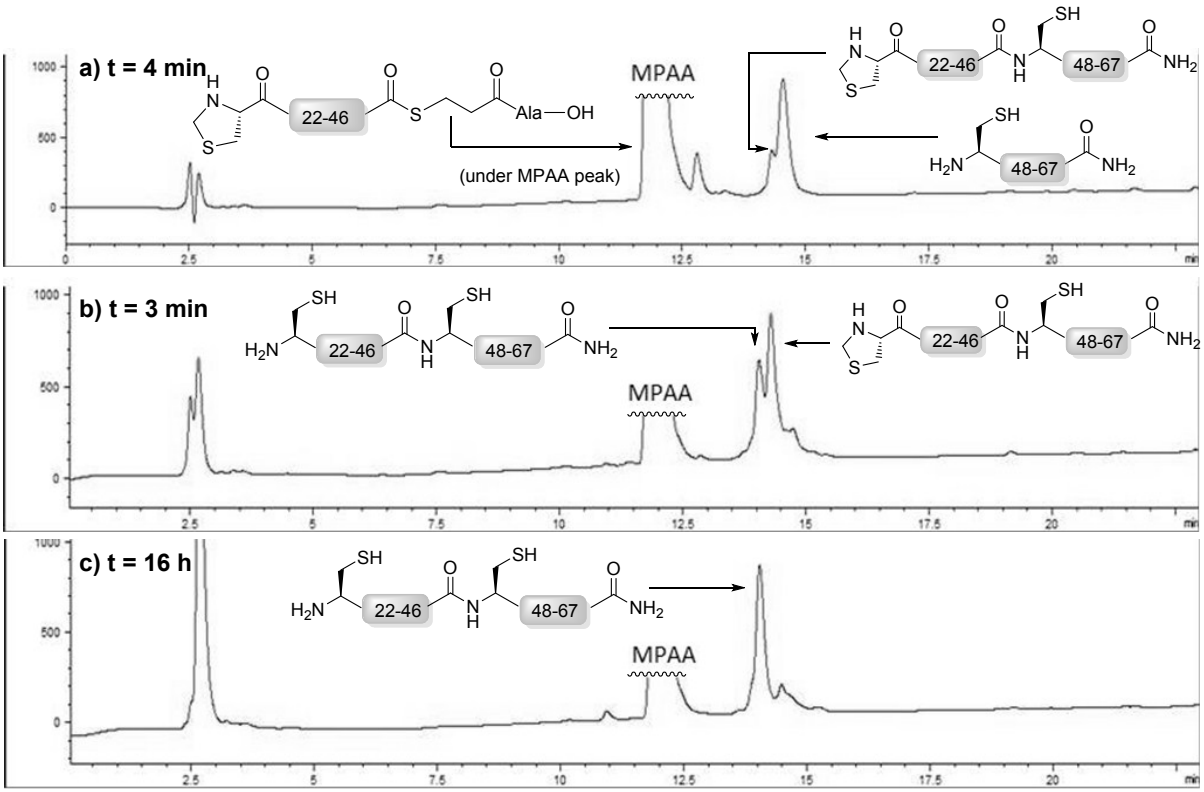


Figure S11: LCMS analysis of a) the cysteine-based ligation between the IGF-II (Thz-46) thioester and C-terminal IGF-II (47-67) fragment after 4 min; LCMS analysis of the thiazolidone deprotection of the IGF-II (Thz-67) fragment after b) 3 min and c) 16 h.

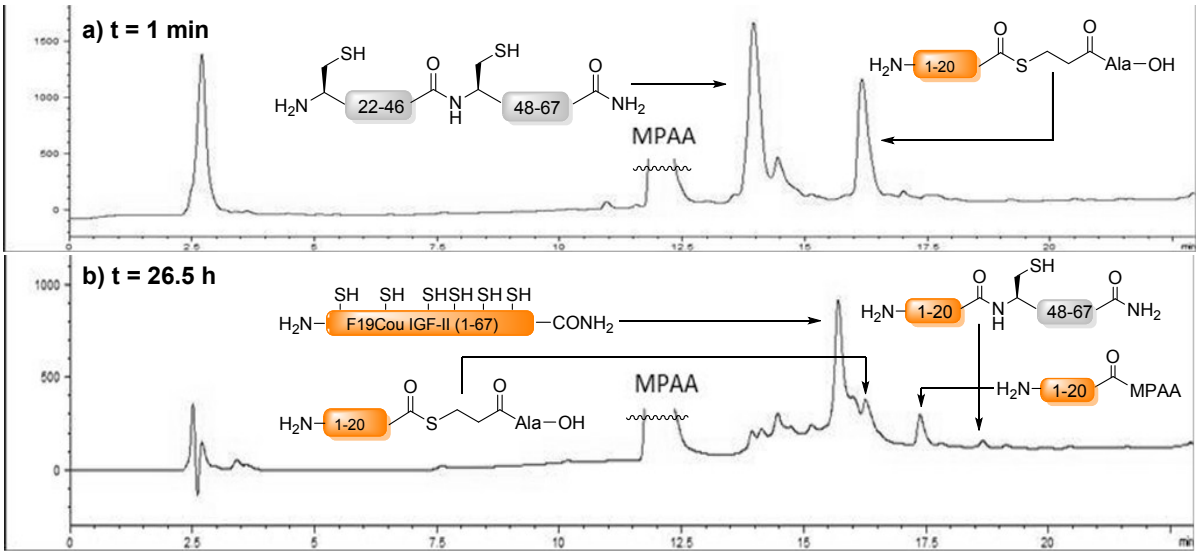


Figure S12: LCMS analysis of the valine-based ligation between the IGF-II (21-67) fragment and N-terminal F19Cou IGF-II (1-20) thioester after a) 1 min and b) 26.5 h.

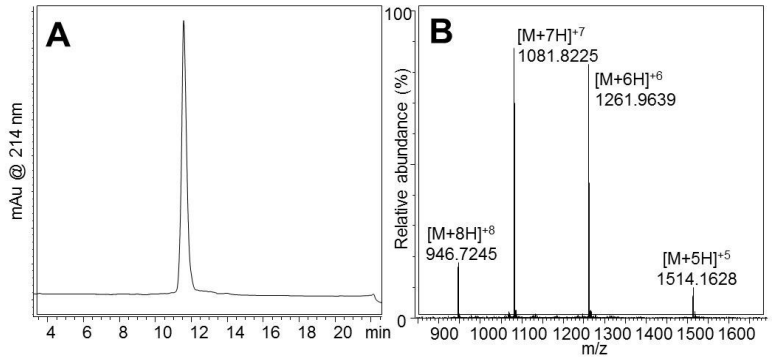
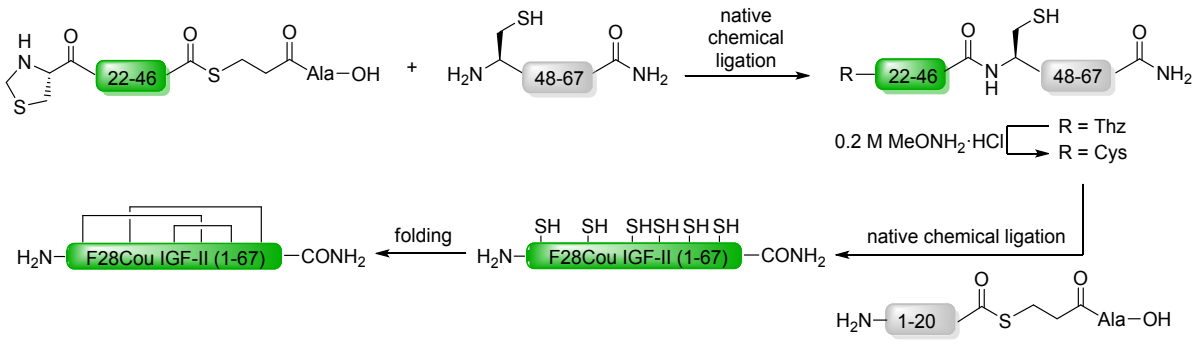


Figure S13: Characterisation of the synthetic F19Cou IGF-II protein. a) RP-HPLC of the synthetic F19Cou IGF-II protein; b) HRMS of the synthetic F19Cou IGF-II protein. Calculated mass for $C_{325}H_{500}N_{94}O_{103}S_6$: 7566.4445Da (average isotopes); observed m/z: 1514.1628 ($[M+5H]^{+5}$), 1261.9639 ($[M+6H]^{+6}$), 1081.8225 ($[M+7H]^{+7}$) and 946.7245 ($[M+8H]^{+8}$).

5.10 Three fragment synthesis of the F28Cou IGF-II protein using native chemical ligation



C-terminal IGF-II (47-67) fragment (7.09 mg, 3.00 µmol, 3.12 mM), F28Cou IGF-II (Thz-46) thioester (9.57 mg, 2.9 µmol, 3.00 mM), TCEP (20.0 mM) and MPAA (200 mM) were dissolved in degassed native chemical ligation

buffer (6.0 M GnHCl and 0.20 M Na_2HPO_4, 976 µL), pH of the solution was adjusted to 6.9 using aqueous NaOH (10.0 M and 2.0 M) and shaken at rt for 1 h. $MeONH_2$•HCl (16.4 mg, 0.19 mmol, 0.20 M) was then added, the pH of the solution was adjusted to 4.0 with aqueous HCl (5.0 M) and the reaction mixture was shaken at rt 16 h. The pH of the solution was adjusted to 7.0 using aqueous NaOH (10.0 M and 2.0 M) and the N-terminal IGF-II (1-20) thioester (7.54 mg, 3.20 µmol, 3.28 mM) was added. The reaction mixture was shaken at rt for 47 h. The products were isolated using SPE according to Section 3.2 and purified by semi-preparative RP-HPLC to give the F28Cou IGF-II peptide (0.70 mg, 3%) as a white solid. F28Cou IGF-II peptide (0.70 mg, 0.09 µmol) was subsequently folded according to Section 4 to give the F28Cou IGF-II protein (29 µg, 4%) as a white solid (Figure S16A). HRMS (ESI+) Calculated mass for $C_{325}H_{500}N_{94}O_{103}S_6$: 7566.4445Da (average isotopes); observed m/z: 1514.2609 ($[M+5H]^{+5}$), 1261.8750 ($[M+6H]^{+6}$), 1081.7289 ($[M+7H]^{+7}$) and 946.7548 ($[M+8H]^{+8}$) (Figure S16B). Analytical data for the one-pot three fragment synthesis of the F19Cou IGF-II protein is summarised in Figure S14 and Figure S15.

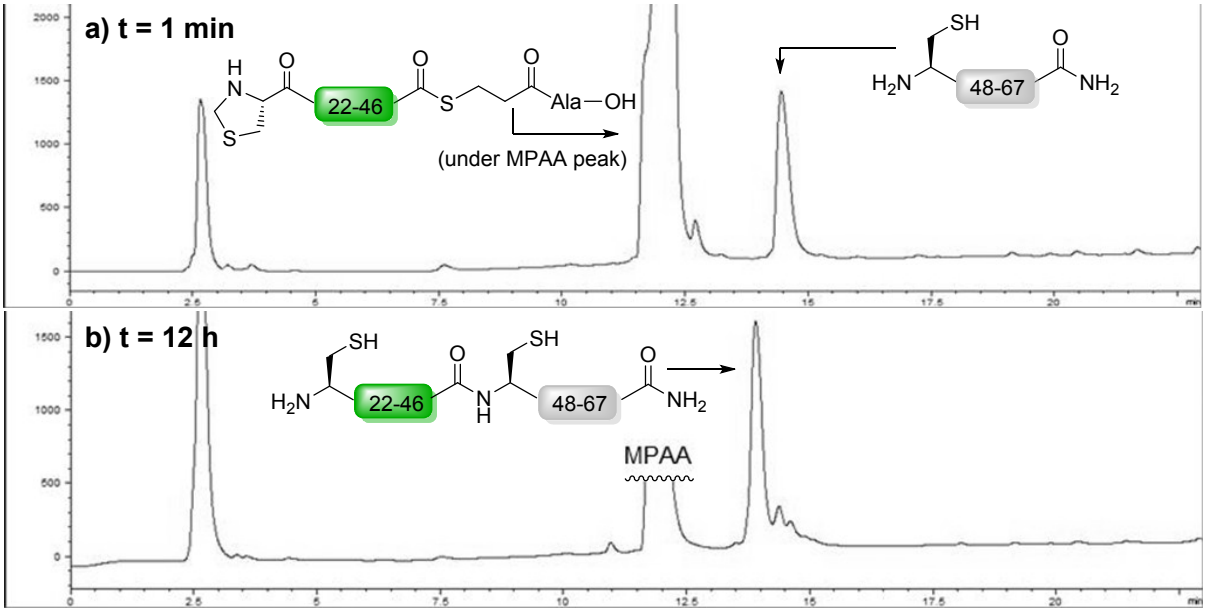


Figure S14: LCMS analysis of a) the cysteine-based ligation between the F28Cou IGF-II (Thz-46) thioester and C-terminal IGF-II (47-67) fragment after 1 min; b) the thiazolidone deprotection of the F28Cou IGF-II (Thz-67) fragment after 12 h.

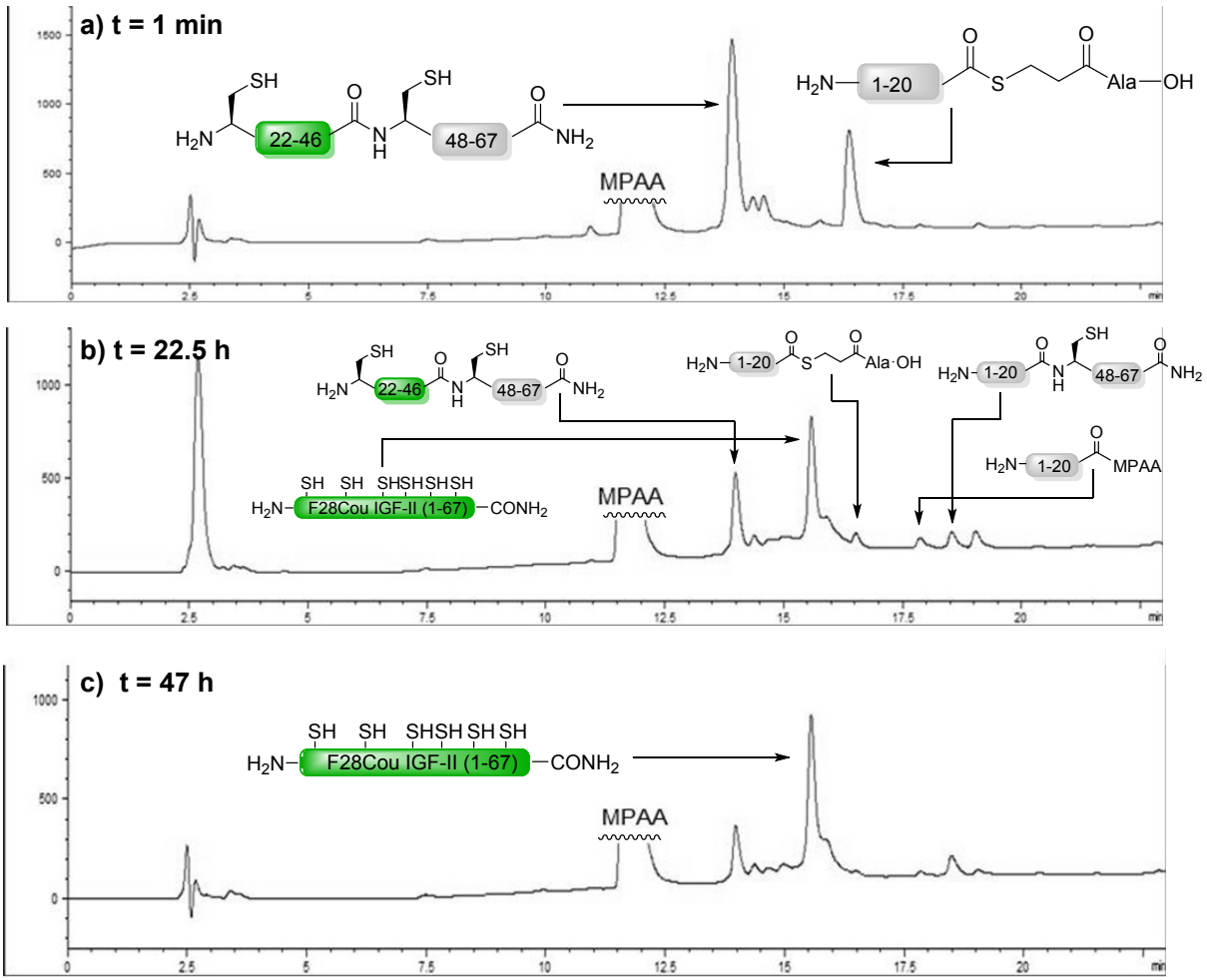


Figure S15: LCMS analysis of the valine-based ligation between the F28Cou IGF-II (21-67) fragment and N-terminal IGF-II (1-20) thioester after a) 1 min; b) 22.5 h and c) 47 h.

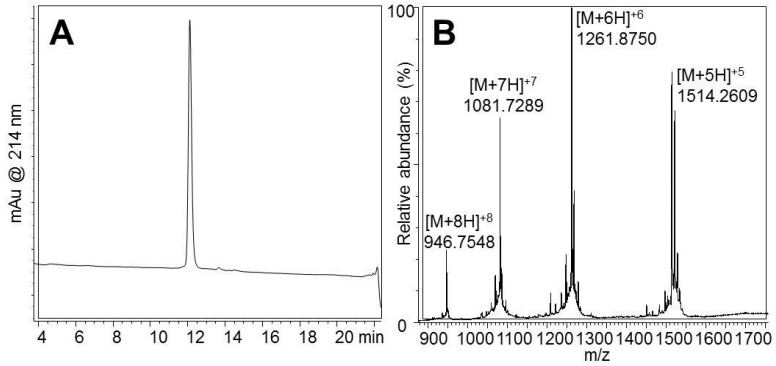


Figure S16: Characterisation of the synthetic F28Cou IGF-II protein. a) RP-HPLC of the purified F28Cou IGF-II protein; b) HRMS of the purified F28Cou IGF-II protein. Calculated mass for. for $C_{325}H_{500}N_{94}O_{103}S_6$: 7566.4445Da (average isotopes); observed m/z: 1514.2609 ($[M+5H]^{+5}$), 1261.8750 ($[M+6H]^{+6}$), 1081.7289 ($[M+7H]^{+7}$) and 946.7548 ($[M+8H]^{+8}$).

6 Competition binding assays

6.1 Materials

Long^{TM}Arg^{3}IGF-I and human IGF-II were purchased from Novozymes *GroPep* Pty Ltd. (Adelaide, South Australia). Greiner Lumitrac 600 96-well plates were purchased from Omega scientific (Tarzana, USA). The DELFIA® europium-labeling kit and DELFIA® enhancement solution were purchased from PerkinElmer Life Sciences. EuIGF-II was produced as described by Denley *et al.*[3] according to the manufacturer's instructions. The anti-IGF-1R antibody 24-31 was a kind gift from Prof. K. Siddle (Cambridge, UK). P6 IGF-1R cells (BALB/c3T3 cells overexpressing the human IGF-1R)[13] were a kind gift from Prof. R. Baserga (Philadelphia, PA).

Competition binding assay buffers

Lysis Buffer	20.0 mM HEPES, 150.0 mM NaCl, 1.50 mM $MgCl_2$, 10% (v/v) glycerol, 1% (v/v) Triton X-100, 1.0 mM EGTA, and 1.0 mM phenylmethylsulfonyl fluoride, pH 7.5
TBST Buffer	20.0 mM Tris, 150.0 mM NaCl, and 0.1% (v/v) Tween 20

6.2 Procedure for competition binding assays

The binding of IGF-II analogues to the IGF-1R were measured essentially as described by Denley *et al.*[3]. Briefly, P6 IGF-1R cells were serum-starved for 4 h and then lysed in lysis buffer (see recipe above) for 1 h at 4 °C. Lysates were centrifuged for 10 min at 3,500 rpm, then an aliquot (100 µL) was added per well to a white Greiner Lumitrac 600 96-well plate previously coated with anti-IGF-1R antibody 24-31[14]. Europium labelled IGF-II (EuIGF-II) (100 µL) with a fluorescent count of 50,000 was added to each well along with increasing concentrations of unlabelled competitor and then the plate was incubated for 16 h at 4 °C. The next day, the wells were washed four times with TBST buffer (see recipe above), then twice with water, and then DELFIA enhancement solution (100 µL/well) was added and the plate was incubated at rt for 10 min. Time-resolved fluorescence of the EuIGF-II was measured using 340 nm excitation and 612 nm emission filters with a Perkin Elmer VICTOR X5 Multilabel Plate Reader (Waltham, MA, U.S.A). IC_{50} values were calculated, using GraphPad Prism 6.01, by curve-fitting with a one-site competition model. The baseline used to calculate all IC_{50} values was set at the % bound/total value of the highest competing IGF-II concentration. Assays were performed in triplicate at least three times, unless otherwise stated.

7 Experimental details for Fluorescence Resonance Energy Transfer (FRET)

7.1 Experimental materials

FRET experiments were performed on a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, CA, USA), using Spectrosil quartz micro cuvettes (100 µL) (Starna) with a 10 mm pathlength at 25 °C, at a pH of 7.0. The excitation wavelength was set to 280 nm and emission spectra were collected between 290 to 600 nm with a scan speed of 600 nm/min. Excitation and emission slit widths were set to 5 nm. Each single spectrum was averaged from three consecutive scans. All solvents, buffers and reagents used were of spectroscopic grade or higher. If spectroscopic grade was not available then solutions and solvents were filtered through a 0.22 µm filter before use.

7.2 Procedure for FRET experiments (native IGF-II, F19Cou IGF-II, F28Cou IGF-II in *presence* of sIGF-1R)

Solublised immunocaptured receptor (sIGF-1R) was prepared according to the procedure described by Surinya *et al.*[15] and used as a 0.20 µM solution in 0.10 M sodium phosphate buffer (pH 7.2). The IGF-II proteins (synthetic native IGF-II (3 fragment), F19Cou IGF-II and F28Cou IGF-II) were dissolved in 10 mM HCl. FRET experiments were performed by titrating the synthetic native IGF-II (3 Fragment) (0.20 µM; 1.7 µL) or the F19Cou IGF-II protein (0.20 µM; 1.7 µL) against the sIGF-1R (0.20 µM; 100 µL). FRET experiments for the F28Cou IGF-II protein were performed by titration of the F28Cou IGF-II protein (0.19 µM; 1.6 µL) against the sIGF-1R (0.19 µM; 91 µL). After each addition the solution was manually mixed, incubated at rt for 30 min and then spectra after excitation at 280 and 320 nm were recorded. Each IGF-II analogue was added until the protein and sIGF-1R were present in an equimolar ratio.

7.3 Procedure for control experiments (native IGF-II, F19Cou IGF-II, F28Cou IGF-II in *absence* of sIGF-1R)

The IGF-II proteins (synthetic native IGF-II (3 fragment), F19Cou IGF-II and F28Cou IGF-II) were dissolved in 10 mM HCl. Control experiments were carried out under the same conditions as the FRET experiments by titrating the synthetic native IGF-II (3 Fragment) (0.20 µM; 1.7 µL, F19Cou IGF-II protein (0.20 µM; 1.7 µL) or F28Cou IGF-II protein (0.19 µM; 1.6 µL) into buffer (in 0.10 M sodium phosphate buffer (pH 7.2). After each addition the solution was manually mixed, incubated at rt for 30 min and then spectra after excitation at 280 and 320 nm were recorded.

7.4 Summary of results from FRET experiments

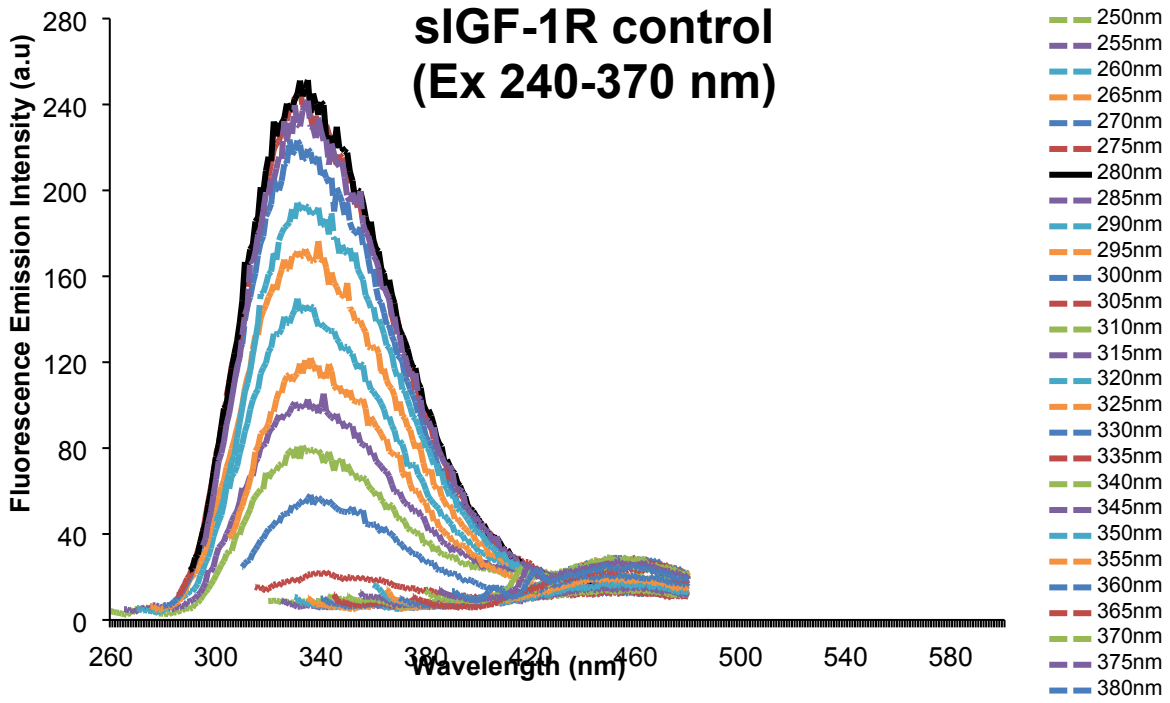


Figure S17: Fluorescence emission profiles for the sIGF-IR after excitation at varying wavelengths (250 - 380 nm). Fluorescence emission profiles for the sIGF-IR are shown as *coloured dotted lines* and the emission maximum after irradiation at 280 nm is shown as a *black solid line* Spectra are derived from a single experiment and are not corrected for background fluorescence.

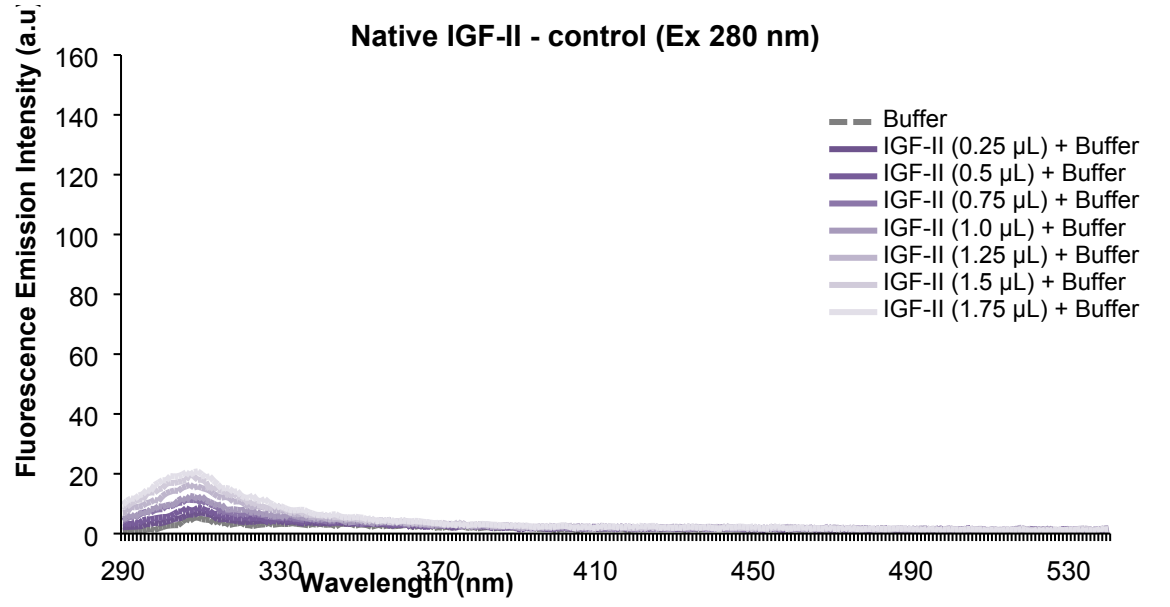


Figure S18: Fluorescence emission profiles for the synthetic native IGF-II protein (3 Fragment) at varying concentrations (0-0.20 µM) in buffer (0.10 M sodium phosphate buffer (pH 7.2). All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment and are not corrected for background fluorescence.

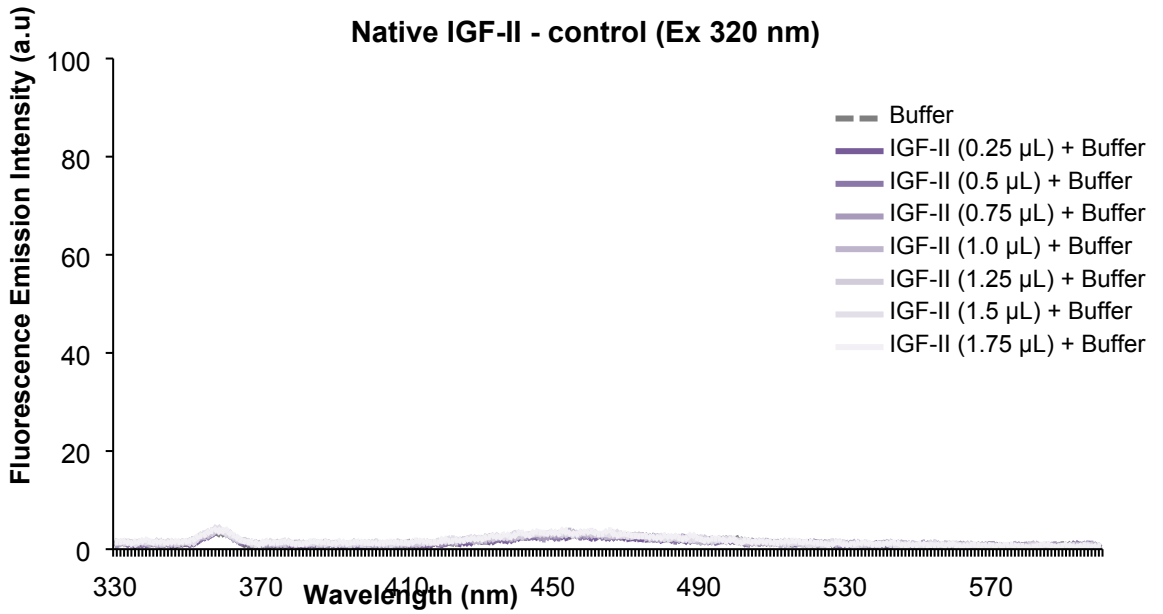


Figure S19: Fluorescence emission profiles for the synthetic native IGF-II protein (3 Fragment) at varying concentrations (0-0.20 µM) in buffer (0.10 M sodium phosphate buffer (pH 7.2) after excitation at 320 nm. All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment and are not corrected for background fluorescence.

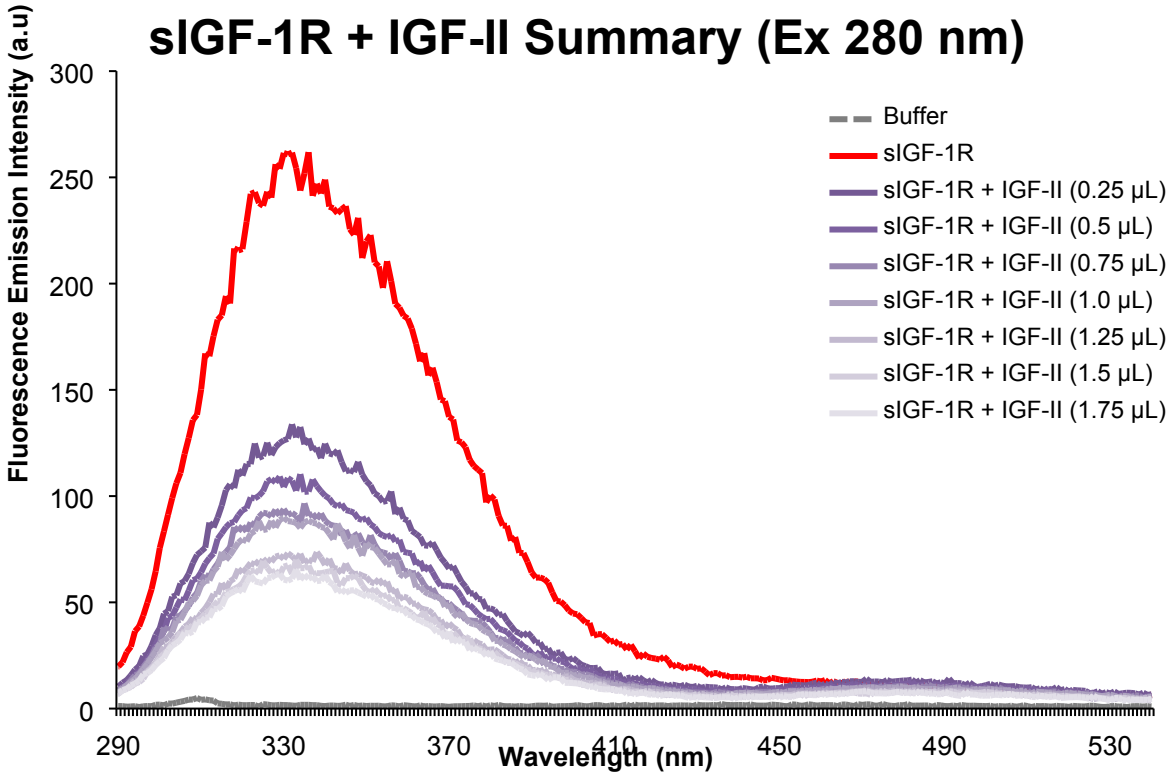


Figure S20: Changes in fluorescence emission spectra during the titration of the synthetic native IGF-II protein (3 Fragment) against the sIGF-1R. The sIGF-1R (alone) is shown as a *red solid line,* the sIGF-1R in the presence of IGF-II are shown as *purple solid lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment, and are not corrected for background fluorescence

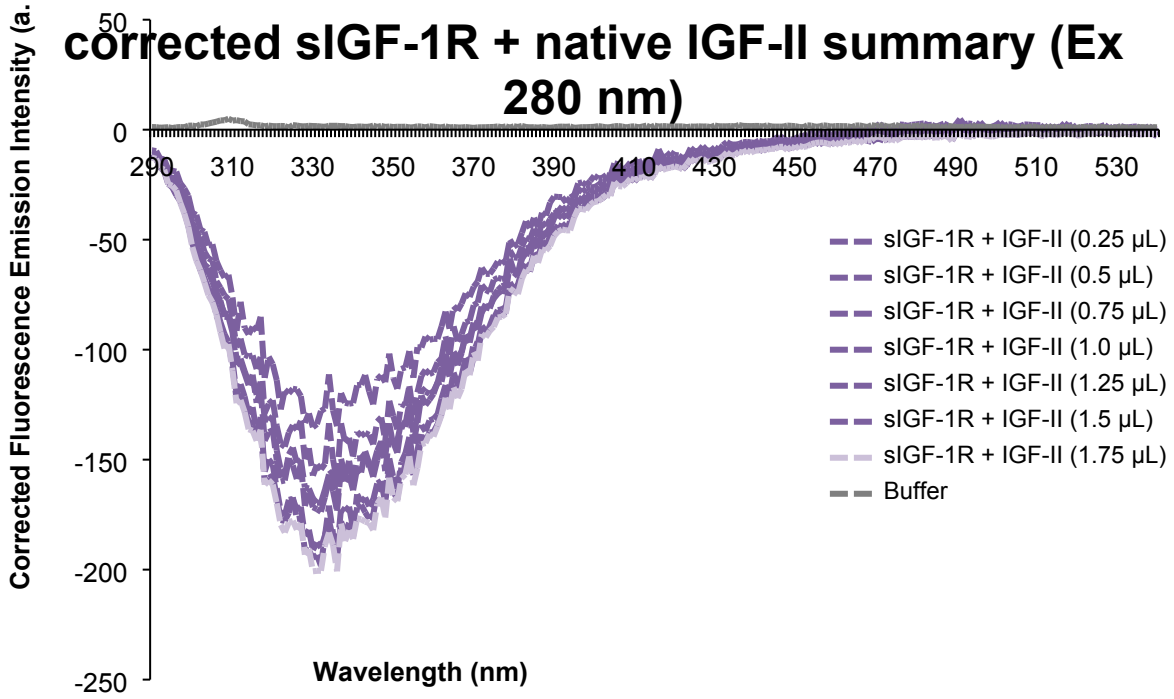


Figure S21: Extracted acceptor emission for the native IGF-II protein at varying concentrations after excitation at 280 nm. Spectra were obtained by subtracting the spectrum for the sIGF-1R (donor only) from the spectrum of the sIGF-1R in complex with the native IGF-II protein (donor and control). Where the corrected spectra for the sIGF-1R in the presence of the native IGF-II are shown as *purple dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment.

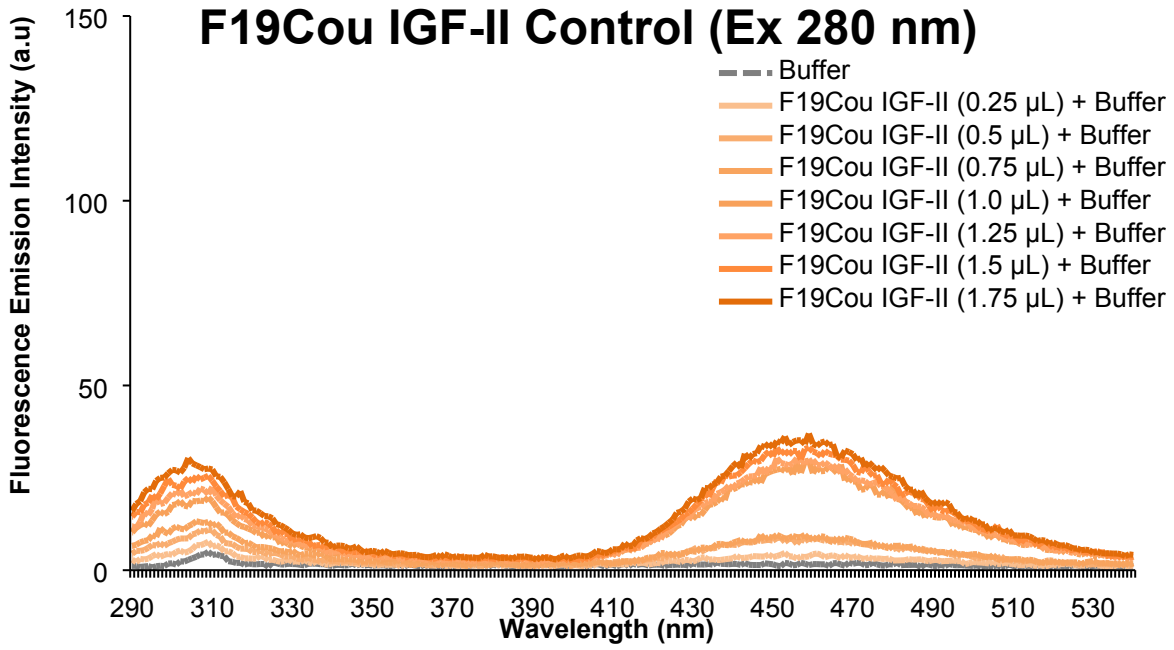


Figure S22: Fluorescence emission profiles for the synthetic F19Cou IGF-II protein (3 Fragment) at varying concentrations (0-0.20 µM) in buffer (0.10 M sodium phosphate buffer (pH 7.2). All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment and are not corrected for background fluorescence.

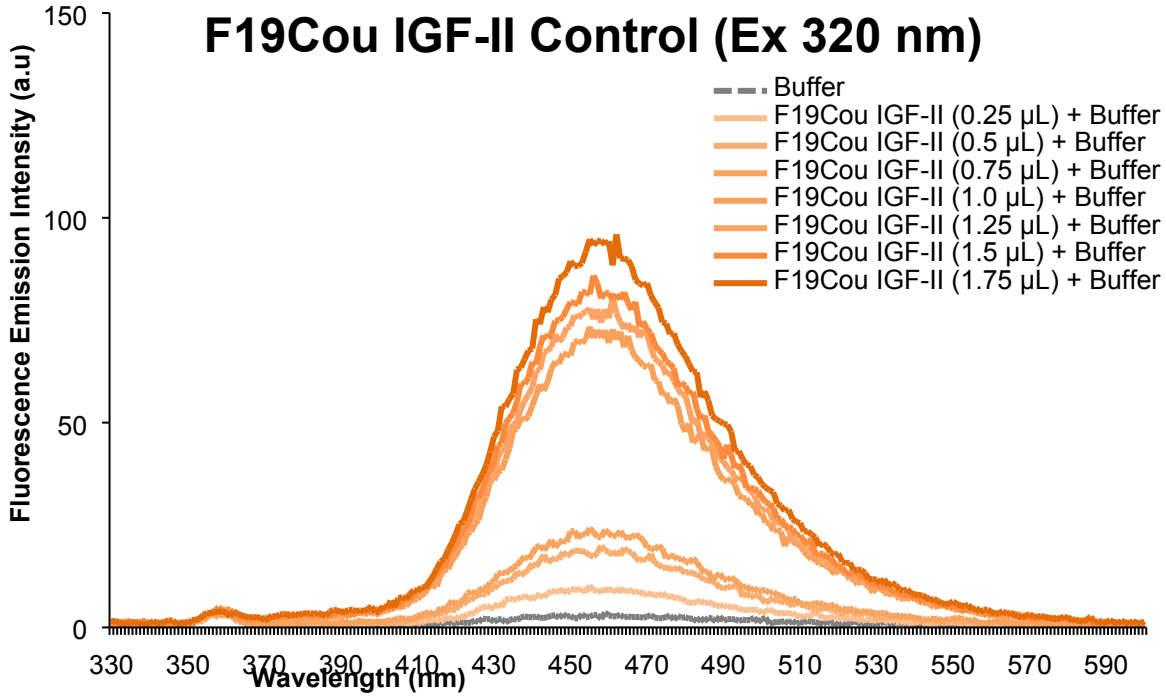


Figure S23: Fluorescence emission profiles for the synthetic F19Cou IGF-II protein (3 Fragment) at varying concentrations (0-0.20 µM) in buffer (0.10 M sodium phosphate buffer (pH 7.2). All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment and are not corrected for background fluorescence.

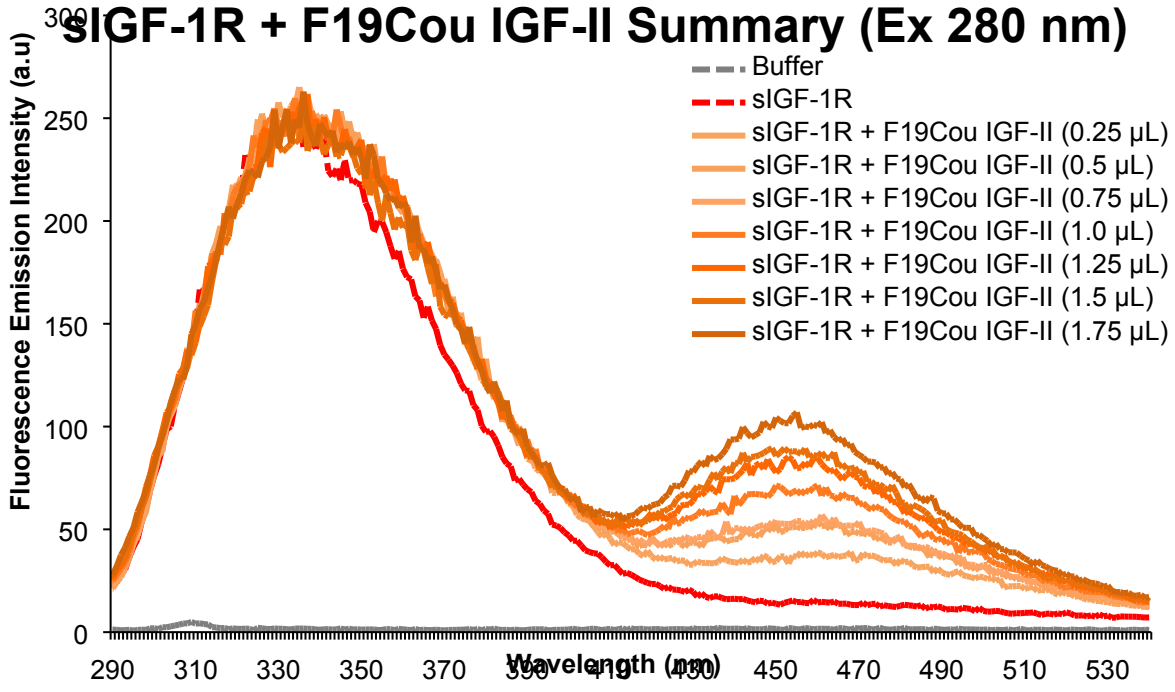


Figure S24: Changes in fluorescence emission spectra during the titration of the F19Cou IGF-II protein (3 Fragment) against the sIGF-1R. Where the sIGF-1R (alone) is shown as a *red dotted line,* the sIGF-1R in the presence of IGF-II are shown as *orange solid lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment, and are not corrected for background fluorescence

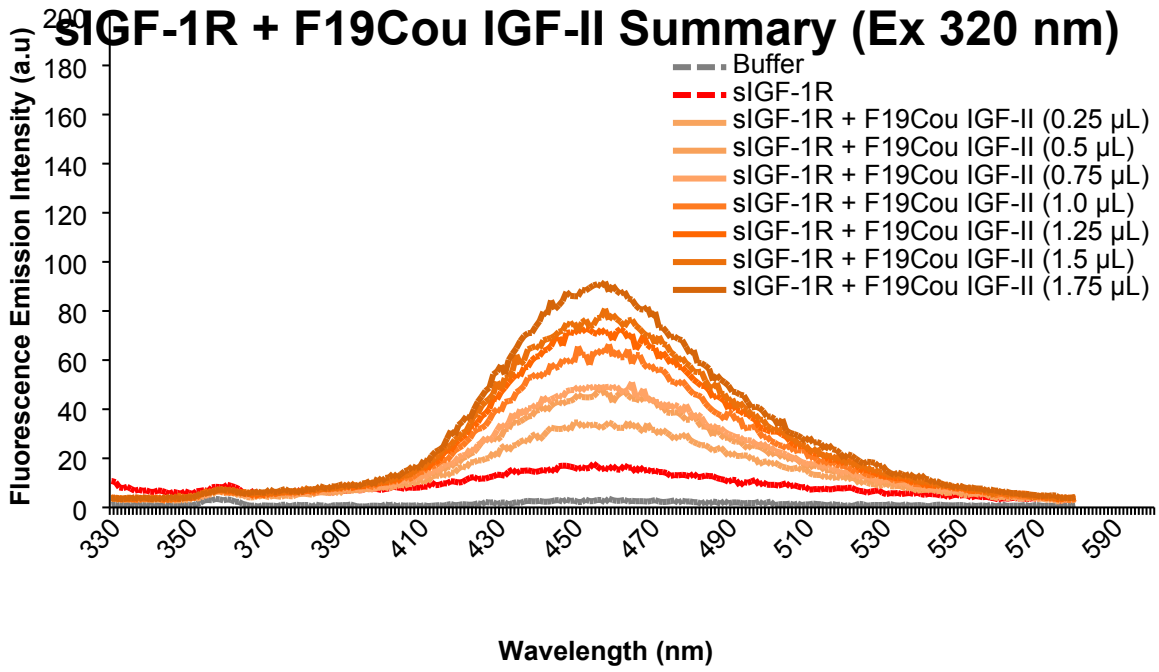


Figure S25: Changes in fluorescence emission spectra during the titration of the F19Cou IGF-II protein (3 Fragment) against the sIGF-1R. Where the sIGF-1R (alone) is shown as a *red dotted line,* the sIGF-1R in the presence of F19Cou IGF-II are shown as *orange solid lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment, and are not corrected for background fluorescence.

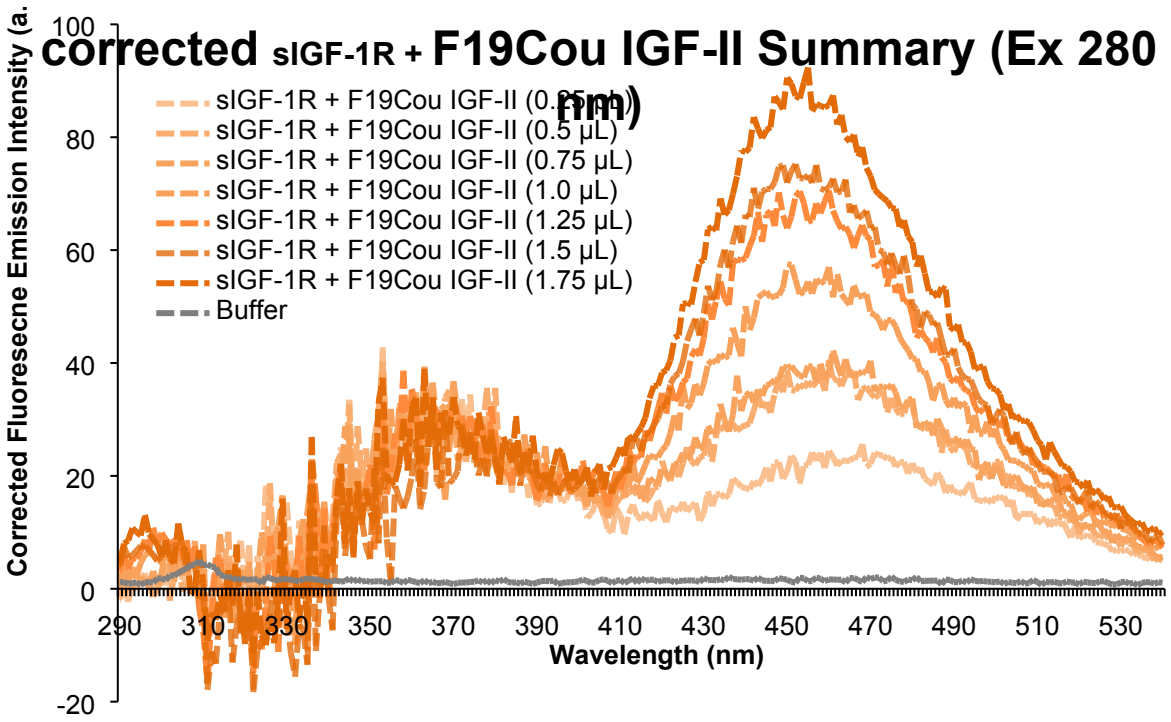


Figure S26: Extracted acceptor emission for F19Cou IGF-II at varying concentrations after excitation at 280 nm. Spectra were obtained by subtracting the spectrum for the sIGF-1R (donor only) from the spectrum of the sIGF-1R in complex with the F19Cou IGF-II protein (donor and acceptor). Where the corrected spectra for the sIGF-1R in the presence of F19Cou IGF-II are shown as *orange dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment.

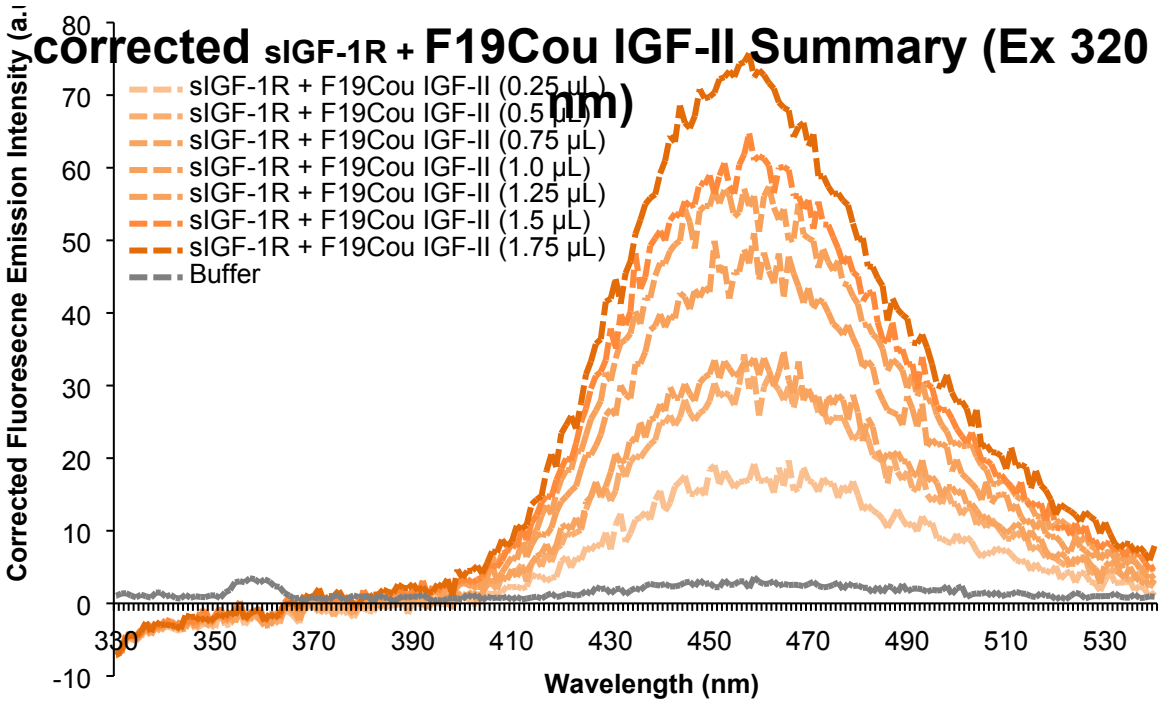


Figure S27: Extracted acceptor emission for F19Cou IGF-II at varying concentrations after excitation at 320 nm. Spectra were corrected by subtracting the spectrum for the sIGF-1R (donor only) from the spectrum of the sIGF-1R in complex with the F19Cou IGF-II protein (donor and acceptor). Where the corrected sIGF-1R in the presence of F19Cou IGF-II are shown as *orange dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment.

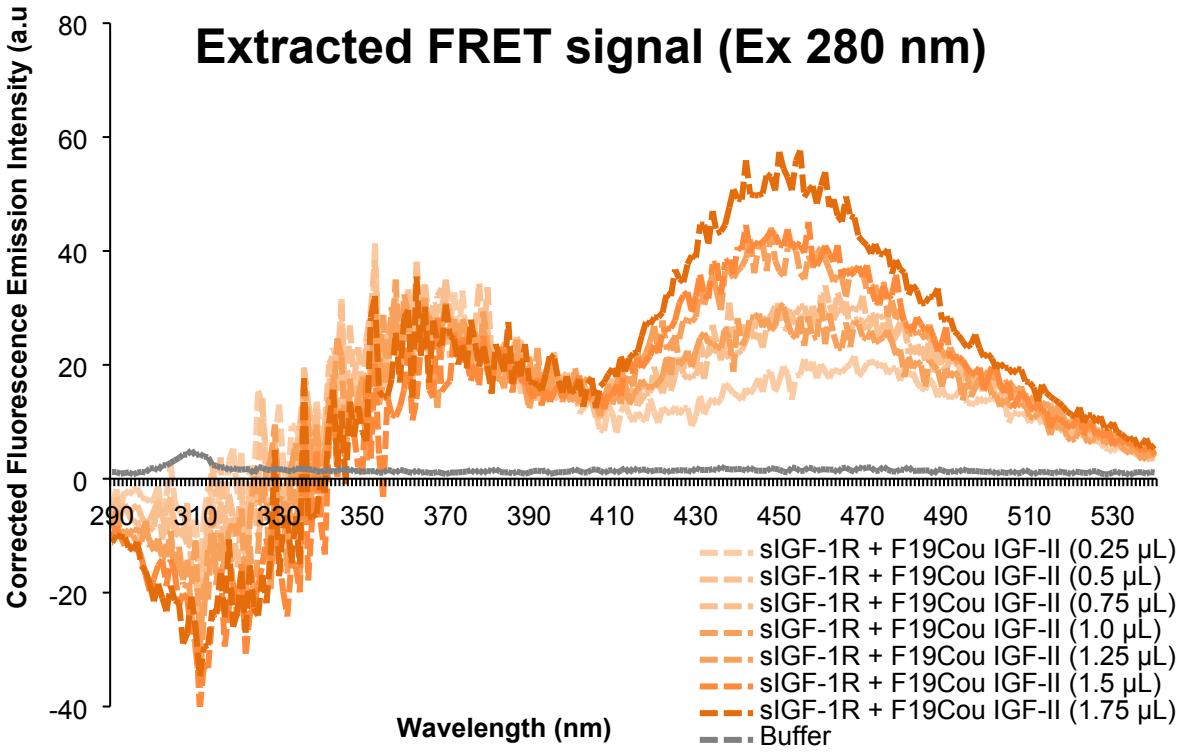


Figure S28: Extracted FRET emission from the sIGF-1R in the presence of the F19Cou IGF-II. Spectra were obtained by subtracting the spectrum for the sIGF-1R (donor only) and for the F19Cou IGF-II (acceptor) from the spectrum of the sIGF-1R in complex with the F19Cou IGF-II protein (donor and acceptor). Where the extracted FRET emissions are shown as *orange dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment.

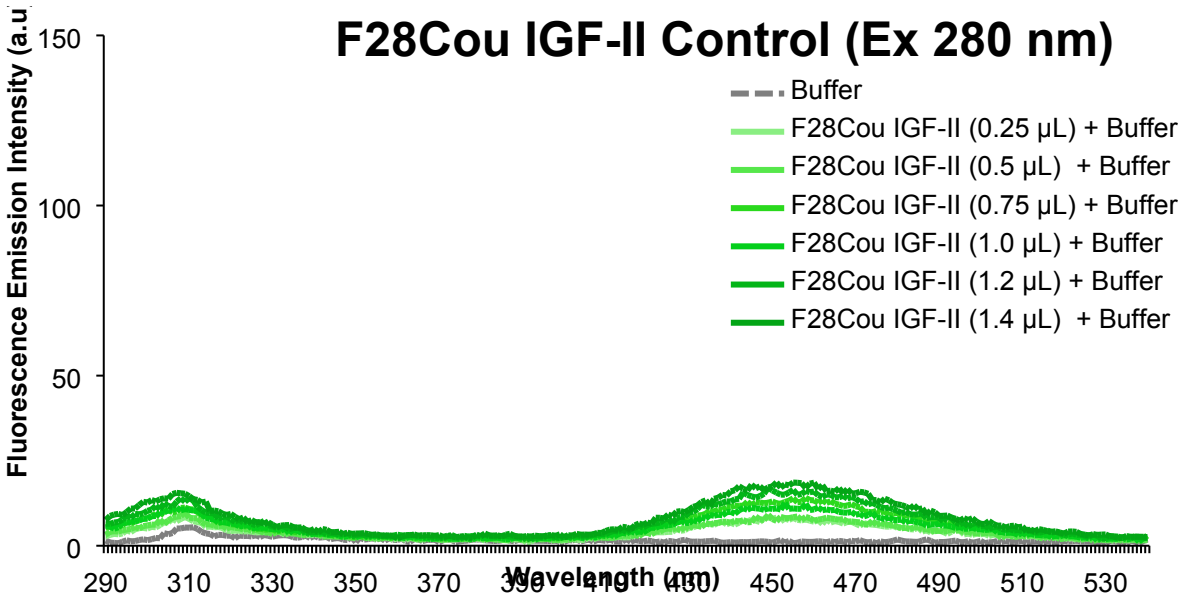


Figure S29: Fluorescence emission profiles for the synthetic F28Cou IGF-II protein (3 Fragment) at varying concentrations (0-0.19 µM) in buffer (0.10 M sodium phosphate buffer (pH 7.2). All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment and are not corrected for background fluorescence.

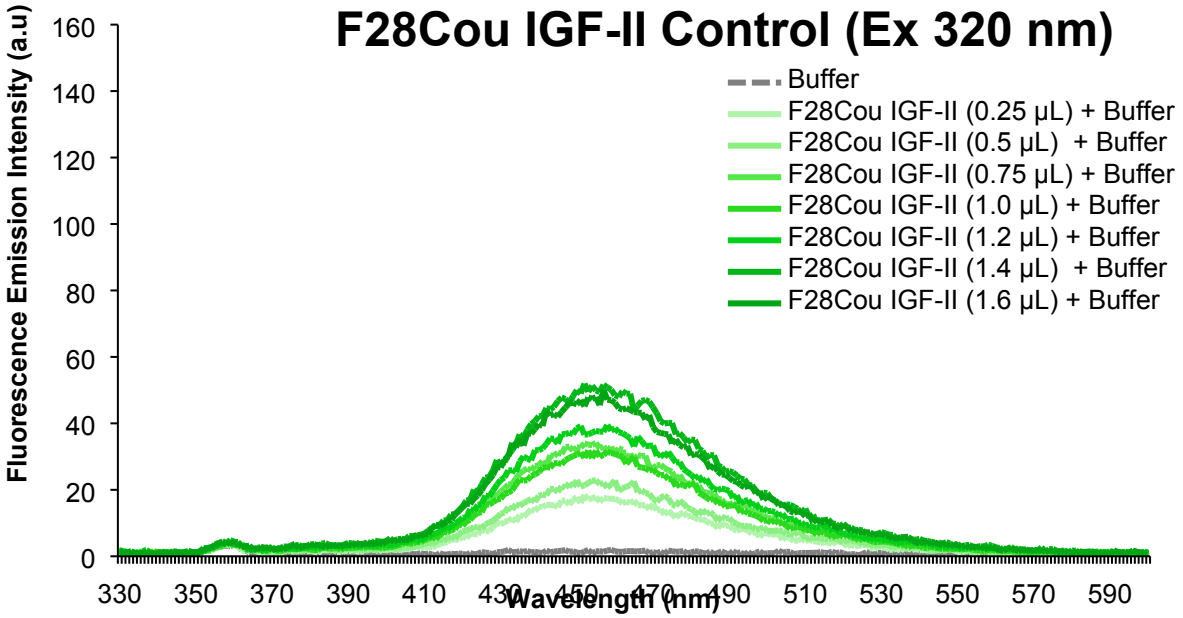


Figure S30: Fluorescence emission profiles for the synthetic F28Cou IGF-II protein (3 Fragment) at varying concentrations (0-0.19 µM) in buffer (0.10 M sodium phosphate buffer (pH 7.2). All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment and are not corrected for background fluorescence.

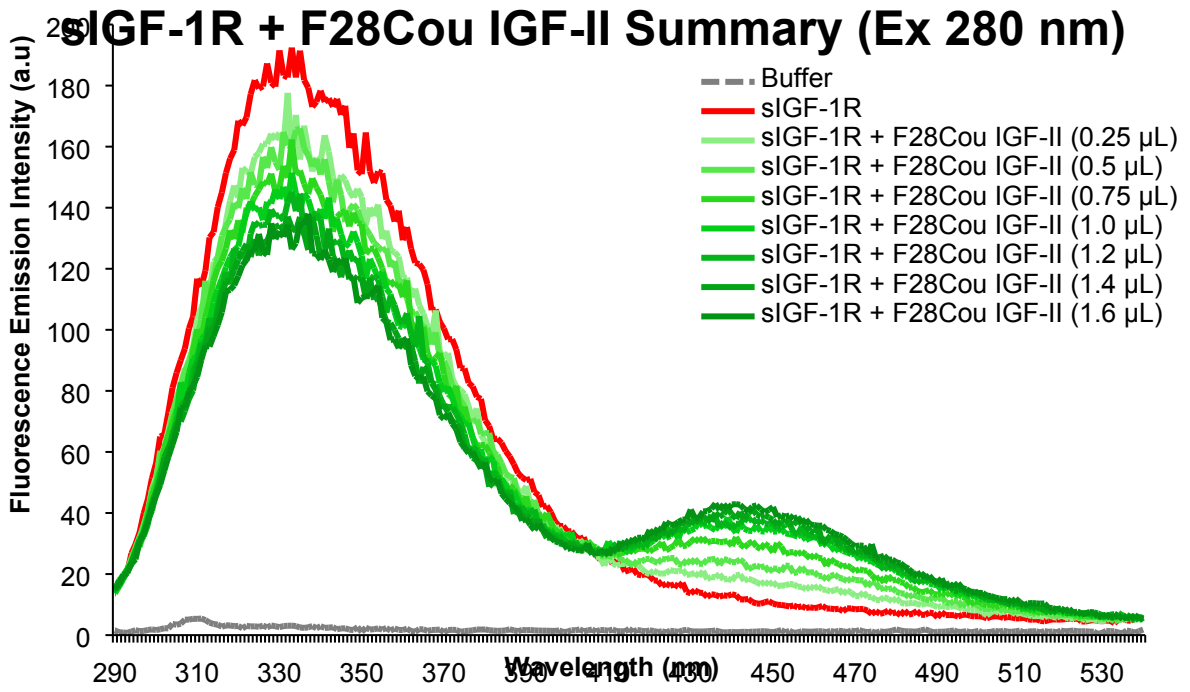


Figure S31: Changes in fluorescence emission spectra during the titration of the F28Cou IGF-II protein (3 Fragment) against the sIGF-1R. Where the sIGF-1R (alone) is shown as a *red solid line,* the sIGF-1R in the presence of F28Cou IGF-II are shown as *green solid lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment, and are not corrected for background fluorescence

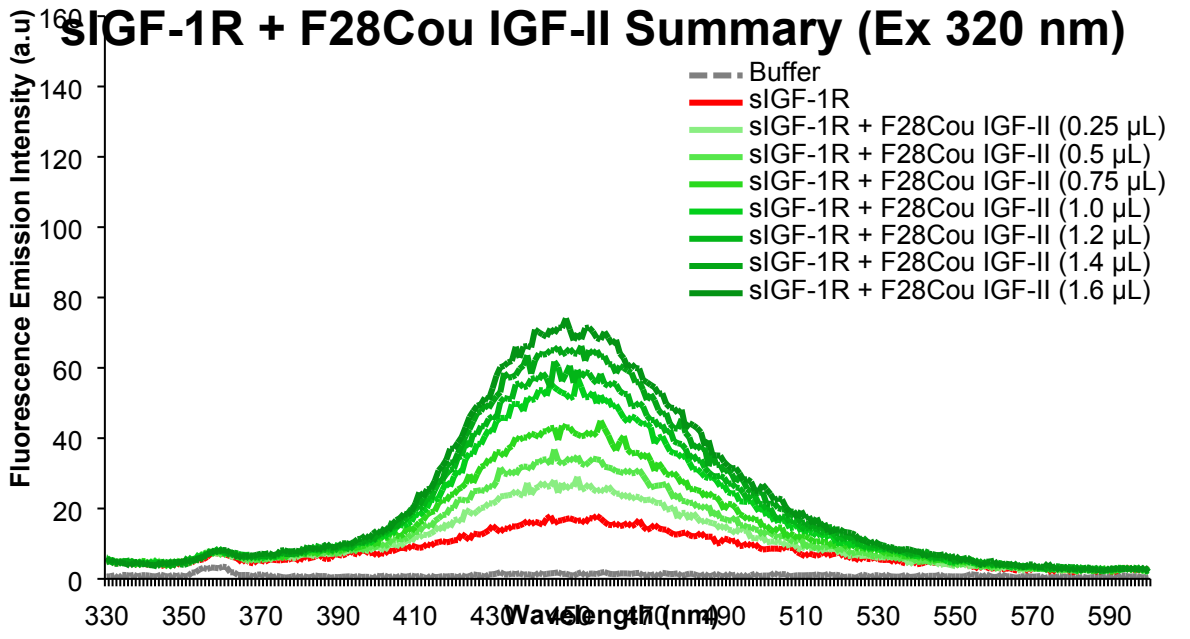


Figure S32: Changes in fluorescence emission spectra during the titration of the F28Cou IGF-II protein (3 Fragment) against the sIGF-1R. Where the sIGF-1R (alone) is shown as a *red solid line,* the sIGF-1R in the presence of F28Cou IGF-II are shown as *green solid lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment, and are not corrected for background fluorescence

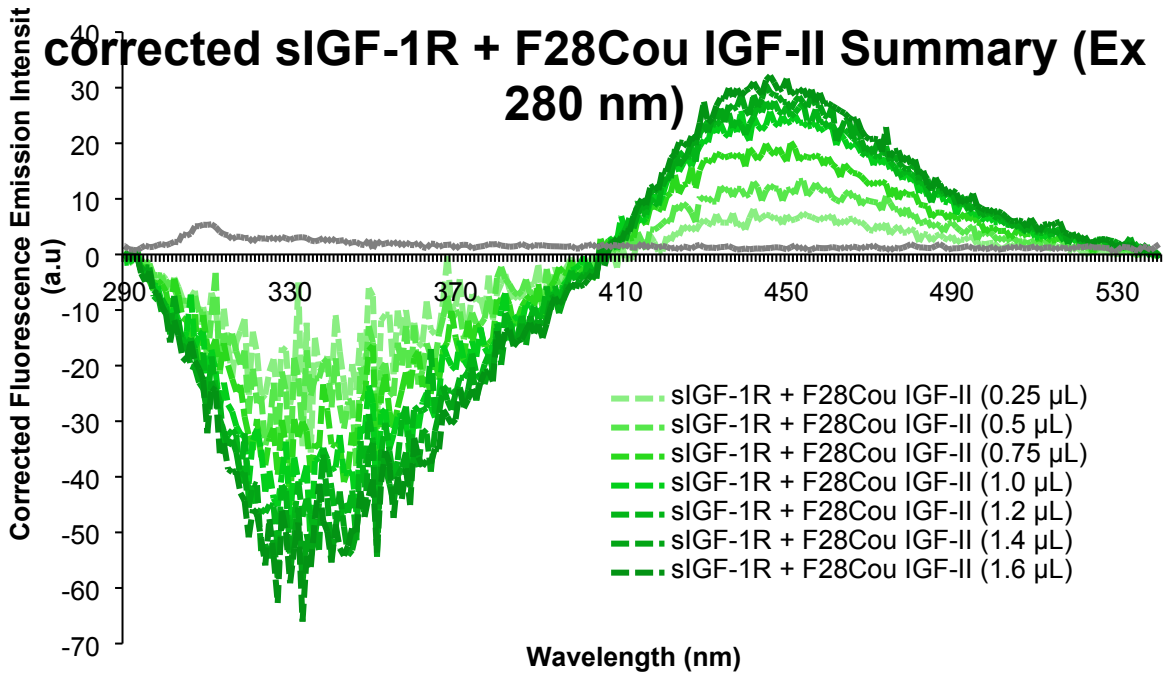


Figure S33: Extracted acceptor fluorescence emission spectra for the F28Cou IGF-II (acceptor), at varying concentrations and after excitation at 280 nm. Spectra were obtained by subtracting the spectrum for the sIGF-1R (donor only) from the spectrum of the sIGF-1R in complex with the F28Cou IGF-II protein (donor and acceptor). Where the corrected spectra for the sIGF-1R in the presence of F28Cou IGF-II are shown as *green dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment.

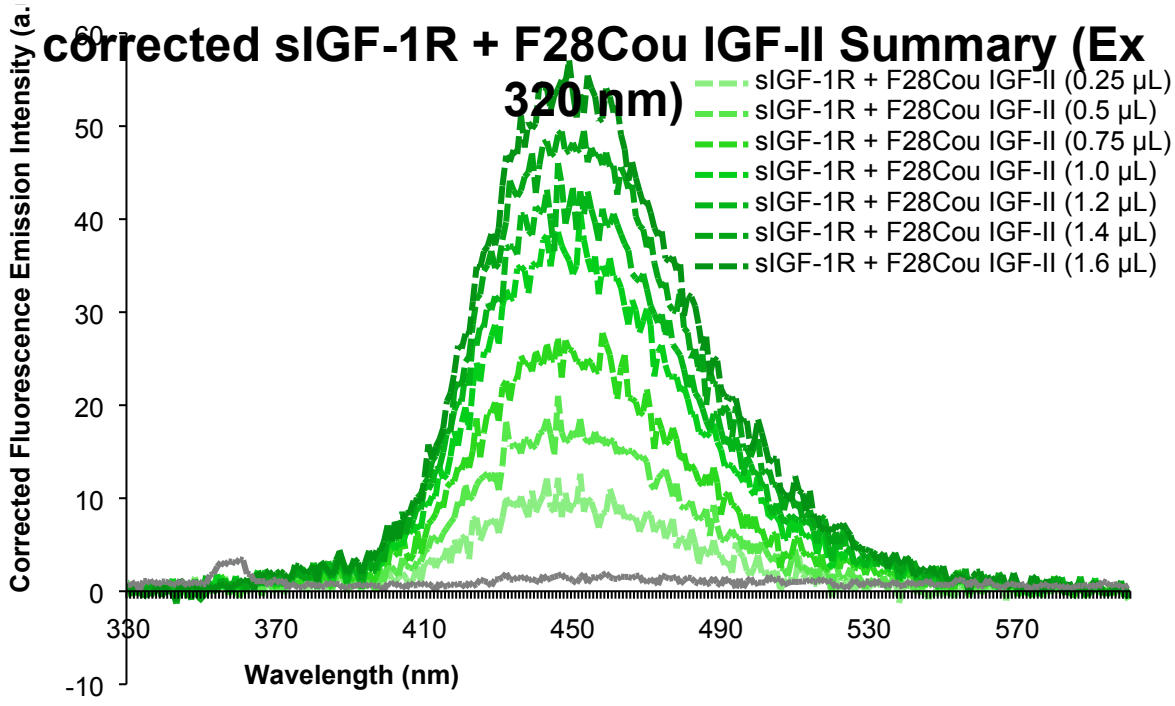


Figure S34: Extracted acceptor fluorescence emission spectra for the F28Cou IGF-II (acceptor), at varying concentrations after excitation at 320 nm. Spectra were obtained by subtracting the spectrum for the sIGF-1R (donor only) from the spectrum of the sIGF-1R in complex with the F28Cou IGF-II protein (donor and acceptor). Where the corrected spectra for the sIGF-1R in the presence of F28Cou IGF-II are shown as *green dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 320 nm. Spectra are derived from a single experiment.

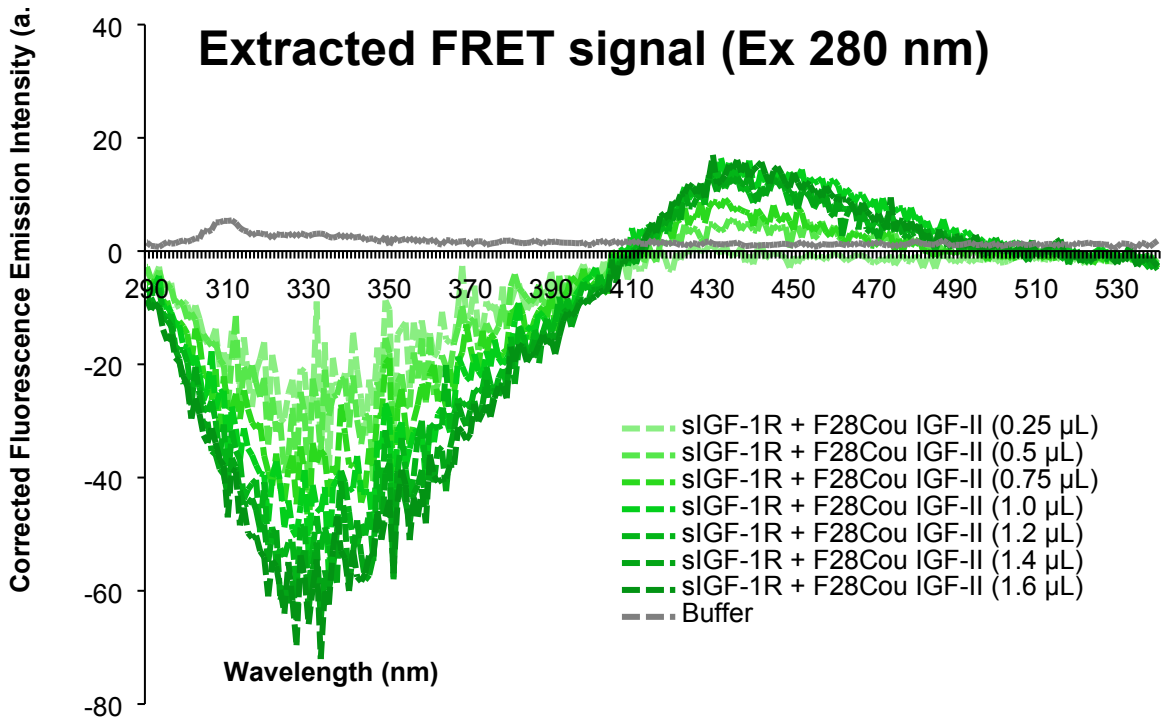


Figure S35: Extracted FRET emission for the F28Cou IGF-II (acceptor). Spectra were obtained by subtracting the spectrum for the sIGF-1R (donor only) and for the F28Cou IGF-II (acceptor) from the spectrum of the sIGF-1R in complex with the F28Cou IGF-II protein (donor and acceptor). Where the extracted FRET emissions are shown as *green dotted lines* and buffer is shown as a *grey dotted line*. All spectra were collected after excitation at 280 nm. Spectra are derived from a single experiment.

8 Schematic of naturally occurring tryptophan residues located adjacent to the putative IGF-II binding sites

A model for the binding of IGF-II to the IGF-1R is displayed in Figure S29 and shows potential Trp donors.[16-18] Both the receptor and ligand used in this model are in the unbound/ free state and thus distances in this model are approximate and are likely to be changed and likely reduced in the bound state. The site 1 analogue, the F28Cou IGF-II protein is proposed to interact with the L1 and the αCT segment of the IGF-1R. Both these binding determinants are represented in the model shown in Figure S29. Trp residues (Trp^{79}, Trp^{127}, and Trp^{176}) displayed in Figure S29 were hypothesised to be the residues involved in the FRET interaction with the F28Cou IGF-II protein, with particular emphasis on Trp^{79}.

The site 2 analogue, the F19Cou IGF-II protein is proposed to primarily interact with the FnIII (2-3) domains of the IGF-1R. Unfortunately, this model lacks these main binding determinants for the IGF-1R, thus it is likely the Trp residues which are could be potentially contributing to a FRET interaction, will not be visible in the current model.

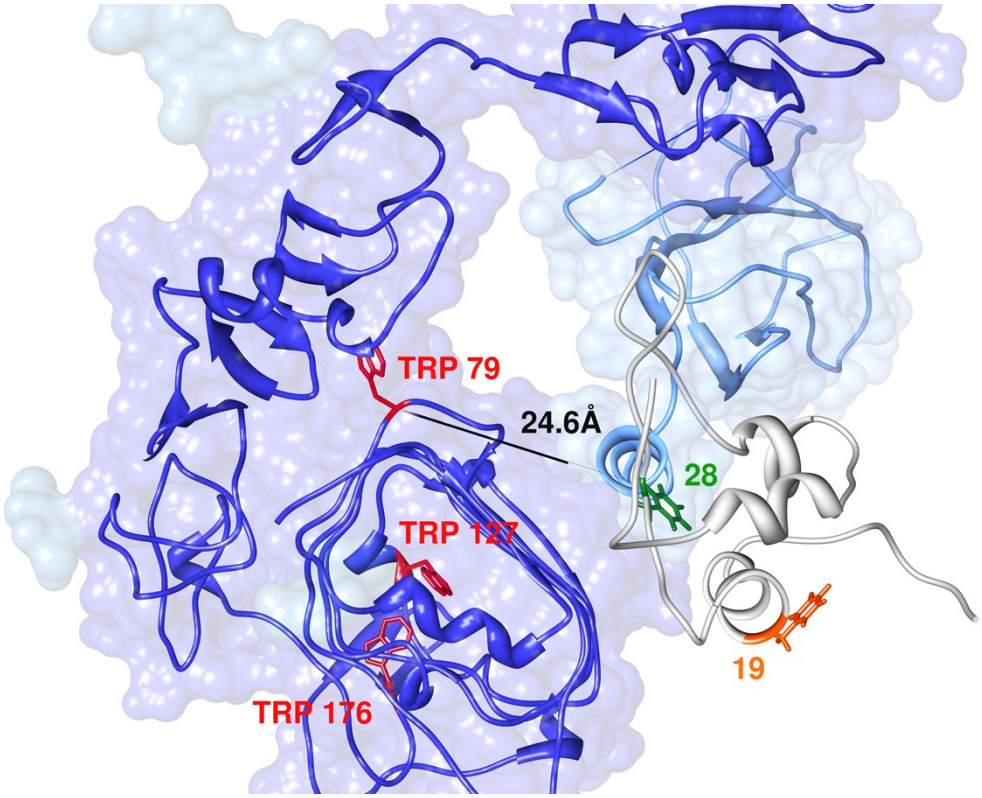


Figure S29: Model for the binding of the IGF-II protein to the IGF-1R. Model based on the crystal structure reported by Menting et al.[17] for insulin binding to the IR-A (IR593.αCT construct) (PDB: 3W14). In this model the crystal structure for the L1-CR-L2 domains reported by Garrett et al.[16] is used for the IGF-1R (PDB: 1IGR) and the NMR solution structure reported by Torres et al.[18] is used for IGF-II (PDB: 1IGL). The IGF-1R is superimposed and replacing the IR-A (L1-CR-L2), IGF-II is superimposed and replacing insulin and the (FnIII-1)-α-CT (704-719) segment is unchanged. Both the receptor and ligand used in this model are in the unbound/ free state. Where the IGF-1R is shown in medium blue; IGF-II is shown in grey; the (FnIII-1)-α-CT (704-719) segment is shown in light blue; the position of the coumarin-based fluorophore is shown in orange (F19Cou IGF-II) and green (28) (F28Cou IGF-II); Trp residues < 50 Å from residue 28 are shown in red; the distances (in Å) between Trp residues and position 28 (green) are designated by dashed black lines.

9 References

1. P. W. R. Harris, S. H. Yang and M. A. Brimble, *Tetrahedron Letters*, 2011, **52**, 6024-6026.
2. P. W. R. Harris and M. A. Brimble, *Synthesis*, 2009, **2009**, 3460,3466.
3. A. Denley, E. R. Bonython, G. W. Booker, L. J. Cosgrove, B. E. Forbes, C. W. Ward and J. C. Wallace, *Mol. Endocrinol.*, 2004, **18**, 2502-2512.
4. M.-P. Brun, L. Bischoff and C. Garbay, *Angewandte Chemie International Edition*, 2004, **43**, 3432-3436.
5. J. Wang, J. Xie and P. G. Schultz, *Journal of the American Chemical Society*, 2006, **128**, 8738-8739.
6. E. Atherton, H. Fox, D. Harkiss, C. J. Logan, R. C. Sheppard and B. J. Williams, *Journal of the Chemical Society, Chemical Communications*, 1978, 537-539.
7. R. Sheppard, *Journal of Peptide Science*, 2003, **9**, 545-552.
8. M. Schnölzer, P. Alewood, A. Jones, D. Alewood and S. Kent, *International Journal of Peptide Research and Therapeutics*, 2007, **13**, 31-44.
9. E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, *Analytical Biochemistry*, 1970, **34**, 595-598.
10. T. Christensen, *Acta Chemica Scandinavica Series B-Organic Chemistry and Biochemistry*, 1979, **33**, 763-766.
11. T. Vojkovsky, *Peptide research*, 1995, **8**, 236-237.
12. C. Delaine, C. L. Alvino, K. A. McNeil, T. D. Mulhern, L. Gauguin, P. De Meyts, E. Y. Jones, J. Brown, J. C. Wallace and B. E. Forbes, *J. Biol. Chem.*, 2007, **282**, 18886-18894.
13. C. Sell, G. Dumenil, C. Deveaud, M. Miura, D. Coppola, T. DeAngelis, R. Rubin, A. Efstratiadis and R. Baserga, *Mol. Cell. Biol.*, 1994, **14**, 3604-3612.
14. M. A. Soos, C. E. Field, R. Lammers, A. Ullrich, B. Zhang, R. A. Roth, A. S. Andersen, T. Kjeldsen and K. Siddle, *J. Biol. Chem.*, 1992, **267**, 12955-12963.
15. K. H. Surinya, B. E. Forbes, F. Occhiodoro, G. W. Booker, G. L. Francis, K. Siddle, J. C. Wallace and L. J. Cosgrove, *J. Biol. Chem.*, 2008, **283**, 5355-5363.
16. T. P. J. Garrett, N. M. McKern, M. Lou, M. J. Frenkel, J. D. Bentley, G. O. Lovrecz, T. C. Elleman, L. J. Cosgrove and C. W. Ward, *Nature*, 1998, **394**, 395-399.
17. J. G. Menting, J. Whittaker, M. B. Margetts, L. J. Whittaker, G. K. W. Kong, B. J. Smith, C. J. Watson, L. Zakova, E. Kletvikova, J. Jiracek, S. J. Chan, D. F. Steiner, G. G. Dodson, A. M. Brzozowski, M. A. Weiss, C. W. Ward and M. C. Lawrence, *Nature*, 2013, **493**, 241-245.
18. A. M. Torres, B. E. Forbes, S. E. Aplin, J. C. Wallace, G. L. Francis and R. S. Norton, *J. Mol. Biol.*, 1995, **248**, 385-401.