

Supplementary Information

Identification and engineering of potent cyclic peptides with selective or promiscuous binding through biochemical profiling and bioinformatic data analysis

Thomas P. Smith^{a,†}, Bhaskar Bhushan^{b,†}, Daniele Granata^c, Christian S. Kaas^d, Brigitte Andersen^e, Klaas W. Decoene^d, Qiansheng Ren^f, Haimo Liu^f, Xinping Qu^f, Yang Yang^f, Jia Pan^f, Quijia Chen^f, Martin Münzel^{d,‡}, Akane Kawamura^{a,b,‡}

^a Chemistry - School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK

^b Chemistry Research Laboratory, Department of Chemistry, University of Oxford, OX1 3TA, UK

^c Digital Science and Innovation, Novo Nordisk A/S, Novo Nordisk Park, 2760 Måløv, Denmark

^d Global Research Technologies, Novo Nordisk A/S, Novo Nordisk Park, 2760 Måløv, Denmark

^e Global Drug Discovery, Novo Nordisk A/S, Novo Nordisk Park, 2760 Måløv, Denmark

^f Novo Nordisk Research Center China, Novo Nordisk A/S, Shengmingyuan West Ring Rd, Changping District, Beijing, China

†These authors contributed equally

‡Corresponding authors:

Akane Kawamura, E-mail: akane.kawamura@ncl.ac.uk

Martin Münzel, E-mail: myzm@novonordisk.com

Table of Contents

Protein Expression and Purification of hFGFR ECDs	1
<i>In vitro</i> selection of cyclic peptide libraries against FGF-R ECD targets.....	7
Next generation sequencing and clustering analysis	10
Chemical synthesis of selected macrocyclic peptides	10
BLI binding measurements.....	19
NNK single concentration BLI initial screen.....	21
NNK purified peptide BLI binding data	22
Trimer-18 single concentration BLI initial screen	25
Trimer-18 Purified peptide BLI binding data	27
Alphascreen binding assay	31
Statistical analysis of peptide fitness for affinity prediction	32
Supporting datasets	33
References.....	34

Protein Expression and Purification of hFGFR ECDs

Materials: EndoFree Plasmid Maxi Kit was purchased from QIAGEN. Optimum Growth™ Flasks were purchased from THOMSON. FreeStyle™ 293 Expression Medium, Geneticin™ Selective Antibiotic, Pluronic™ F-68 Non-ionic Surfactant (100 X), 293fectin™ Transfection Reagent, Opti-MEM™ Reduced Serum Medium were purchased from Invitrogen. Stericup Quick Release-HV Sterile Vacuum Filtration System was purchased from Merck. NuPAGE™ 4 to 12%, Bis-Tris, 1.0–1.5 mm, Mini Protein Gels was purchased from Invitrogen. Superdex® 75 was purchased from GE Healthcare. SEC- HPLC was performed on TSKgel G2000 SWxl column coupled with Agilent HPLC. LC-MS was performed on Agilent PLRP-S column coupled with Waters UPLC Xevo G2 Q-TOF. Kinetic Turbidimetric LAL Test kit was purchased from Charles River Laboratories.

DNA sequences of FGFRs cloned for expression:

FGF-R1C-ECD: AGGCCCTCCCCCACTGCCGAACAAGCTCAACCTGGGGAGCTCCCGTGGAGGTGGAAAGCTTTCTGGTGCACCCCGGCGATTTACTGCAGCTGAGATGTCGTCTGAGAGACGACGTCCAGAGCATCAACTGGCTCAGAGACGGCGTGAGCTGGCCGAATCCAATCGTACCAGAATCACCGGCGAGGAGGTCGAGGTCCAAGATAGCGTGCCCGCTGATTCCGGTTTATATGCTTGCCTGACCAGCTCCCCTAGCGGCAGCGATACCACCTACTTCAGCGTGAATGTCTCCGACGCTTTACCCAGCAGCGAGGATGACGACGACGATGACGACTCCTCCTCCGAGGAGAAGGAGACAGATAACACC AAGCCCAATCGTATGCCCGTGGCCCCCTACTGGACATCCCCGAGAAGATGGAGAAGAACTCCATGCCGTGCCCGCTGCTAAGACCGTGAAATTCAAGTGTCCAGCTCCGGCACCCCAATCCTACTTTAAGGTGGCTCAAGAATGGCAAGGAGTTCAAGCCCCGACCATCGTATCGGCGGCTATAAGGTGAGGTACGCCACTTGGAGCATCATCATGGACTCCGTCTGTGCCAGCGACAAGGGCAATTACACTTGTATCGTGGAGAACGAGTACGGAAGCATCAACCACACCTACCAGCTGGA TGTGGTGGAGAGGAGCCCCATCGTCCTATTCTGCAAGCTGGACTGCCCGCTAACAGACCGTCGCTTTAGGCAGC AACGTGGAGTTCATGTGCAAGGTCTACTCCGATCCCCAGCCCCACATTCAAGTGGCTGAAGCACATTGAGGTGAATG GCAGCAAGATTGGCCCCGACAATTTACCCTACGTGCAGATTTTAAAGACCGCCGGCGTCAACACCACCGACAAGGA GATGGAGGTGCTGCACCTCAGAAACGTGAGCTTCGAGGACGCCGGCGAATATACTTGTCTGGCCGGCAACAGCAT CGGACTGAGCCACCACAGCGCTTGGCTGACAGTTTTAGAGGCTCTGGAAGAGAGGCCCGCCGTGATGACCTCCCCT CTGTATTTAGAG

FGF-R3C-ECD: GAGTCTTTAGGCACCGAACAAAGGGTCTGGGCAGAGCTGCCGAAGTGCCCGGTCCCGAACCCG GACAGCAAGAACAGCTGGTGTTCGGAAGCGGCGACGCTGTGAGCTCAGCTGTCCCTCCTCCCGGTGGCGGACCTAT GGGCCCCACAGTGTGGGTGAAGGACGGCACCGGTTTAGTCCCTAGCGAGAGAGTGCTGGTCCGGCCCTCAGAGGCT GCAAGTTCTGAACGCCTCCCACGAGGATTCCGGCGCCTACTCTTGTAGACAGAGACTCACCCAGAGGGTGCTGTGC CACTTCAGCGTGAGGGTGACAGATGCCCCAGCTCCGGCGATGACGAGGATGGAGAGGATGAGGCCGAGGACAC TGGTGTGGATACCGGCGCTCCCTACTGGACAAGGCCCGAGAGGATGGATAAGAAATTATTAGCCGTGCCCGCCGCT AACACCGTGAGATTTGTTGTCCCGCTGCCGGAACCCACCCCTAGCATTAGCTGGCTGAAGAACGGTCTGTGAGT TTAGAGGCGAGCACAGAATCGGCGGCATCAAGCTGAGGCACCAGCAGTGGTCTTTAGTGATGGAGAGCGTGGTGC

CCAGCGACAGAGGAAACTACACTTGTGTCGTGGAGAACAAGTTCGGCAGCATTTCGTCAGACCTACACTTTAGACGT
TTTAGAGAGATCCCCCACAGACCCATTTTACAAGCTGGACTGCCCGCTAATCAGACAGCCGTGCTGGGCTCCGATG
TCGAGTTCCACTGCAAGGTGTACTCCGACGCCAGCCCCACATCCAGTGGCTGAAGCACGTGGAGGTGAATGGCAG
CAAAGTGGGACCCGATGGCACCCCTTACGTGACCGTGTGAAGACCGCCGGCGCTAACACCACCGATAAGGAGCT
GGAGGTGCTGTCTTTACACAACGTGACCTTCGAGGATGCCGGCGAGTACACTTGTCTCGCCGGCAATTCATCGGCT
TTTCCACCACAGCGCTTGGCTGGTGGTGTGCCCGCTGAGGAAGAAGTGGTGAAGCTGACGAGGCCGGAAGCG
TGTATGCCGGT

FGF-R4-ECD: CTGGAGGCCAGCGAGGAGGTGGAAGTGGAGCCTTGTTTAGCTCCTTCTTTAGAGCAGCAAGAACAA
GAACTGACAGTGGCCCTCGGACAGCCCGTGAGACTGTGCTGTGGTCGTGCTGAGAGGGGAGGCCACTGGTACAAG
GAGGGCTCCAGACTGGCCCCGCTGGAAGAGTGAGAGGCTGGAGAGGAAGGCTGGAGATCGCCAGCTTCTCCCC
GAAGACGCCGGTCGTTATTTATGTCTGGCTCGTGGCTCCATGATCGTGTGTCAGAAATTAACACTGATCACCGGCGA
CTCTTTAACAGCAGCAACGACGATGAAGATCCTAAGTCCCATAGAGACCCAGCAATCGTCACAGCTATCCTCAGC
AAGCTCCTTACTGGACCCACCCAGAGGATGGAGAAGAAGCTGCATGCCGTGCCCGCCGCAACACAGTGAAGTT
TCGTTGTCCCGTCCGGAACCCACCCACCCACCATTCGTTGGCTGAAGGACGGCCAAGCTTCCACGGCGAGAATC
GTATCGGCGGCATTCTGTGAGGCACCAGCACTGGTCTTTAGTGATGGAGAGCGTCGTGCCCTCCGATCGTGGCAC
CTATACATGTTTAGTGGAGAACGCCGTGGGAAGCATTGTTACAATTATTTACTCGACGTGCTGGAAAGGTCCCCC
ACAGACCTATTTACAAGCTGGACTGCCCGCTAACACAACCGCCGTGGTGGGCTCCGACGTGGAGCTGCTGTGTAA
GGTGTACAGCGATGCCAGCCTCACATCCAGTGGCTGAAGCACATCGTGATCAACGGCTCCAGCTTTGGCGCCGAC
GGCTTCCCCTATGTCCAAGTTCTGAAAACCGCCGACATTAACAGCAGCGAGGTGGAGGTGCTGTATCTGAGAAACG
TGTCCGCCGAGGATGCTGGCGAGTACACTTGTCTCGCCGGAACAGCATCGGCCTCTCTACCAGTCCGCTTGGCTG
ACCGTTTTACCCGAAGAAGATCCTACATGGACAGCCGCCGCTCCCGAGGCTAGATACACCGAT

Preparation of plasmids: Three different DNA constructs for the extracellular domain (ECD) of human FGF-R1C (uniprotKB P11362), FGF-R3C (uniprotKB P22607), and FGF-R4 (uniprotKB 22455) were generated and cloned into the expression vector pJSV002 using seamless assembly cloning technology (Taihe Biotechnology Co.,LTD)(Figure S 1). Sequences also contained C-terminal Avi and HPC4 sequences. The expression vectors were then transformed separately into DH5 α *Escherichia coli* (*E. coli*), plated onto a LB plate containing ampicillin (100 μ g/mL), and subsequently incubated overnight at 37 °C. The resulting culture medium of transformed cells was then used to obtain plasmid DNA, which was sequenced to confirm the proper reading frame of constructs. For mammalian cell transfection, plasmid preparation kits (EndoFree Plasmid Maxi Kit) were used to purify endotoxin-free plasmids from large-volume *E. coli* cultures.

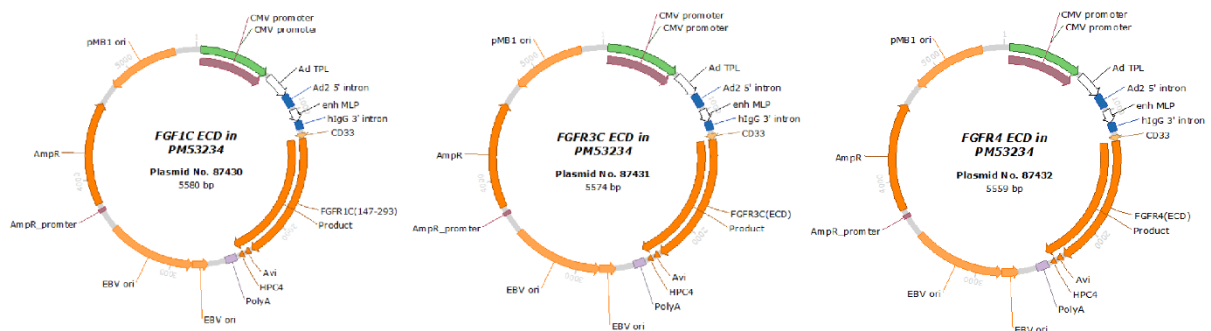


Figure S 1 Plasmids prepared for recombinant FGFR ECD protein expression

Expression of recombinant FGF-Rs: Plasmids of FGF-Rs were transiently transfected in suspension-growing mammalian cells using a FreeStyle 293 Expression Medium (Gibco). Briefly, HEK 293-6E cell suspension culture was maintained in FreeStyle 293 expression medium with 1% F-68 and 25 $\mu\text{g}/\text{mL}$ Geneticin (Gibco). A shake flask with a vented cap was used to produce the culture at 37 $^{\circ}\text{C}$ in a humidified atmosphere containing 5% CO_2 with 50 mm orbital shaking (135 rpm/min). For transient transfection of recombinant protein, the cells were split on at density of $0.7\sim 1 \times 10^6$ cells/mL the day before transfection and allowed to grow overnight. Cells were then diluted to a final volume of 560 mL with the density of 1.07×10^6 viable cells/mL. Transfection was carried out by: 1) 600 μg expression plasmid(s) and Opti-MEM were gently mixed in a final volume of 20 mL; 2) diluted DNA was then added to the diluted 293fectin (600 μL 293fectin in 20 mL Opti-MEM, incubated at room temperature for 5 min) and mixed followed by 20 minutes incubation at room temperature for 20 minutes; 3) DNA/293fectin mixture was added drop wise to 560 mL cells in a 1.6-L Optimum GrowthTM flask and incubated at 37 $^{\circ}\text{C}$ on an orbital shaker with a humidified atmosphere of 5% CO_2 . After 5 days, cell culture medium was harvested by centrifugation for 30 minutes at 6000 rpm and then filtered with a 0.45 μm membrane. Protein expression level was then determined by SDS-PAGE.

Purification of recombinant FGFRs: All the purification steps were performed at 4 $^{\circ}\text{C}$. The cell culture was loaded onto an anti-HPC4 column (prepared in house) and washed with washing buffer (20 mM Tris-HCl, 100 mM NaCl, 1 mM CaCl_2 , pH 7.5), and then eluted with elution buffer (20 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 7.5). Eluted fractions were biotinylated using BirA biotin ligase (in house produced). Finally, the biotinylated protein was purified using gel filtration with Superdex[®] 75 (PBS, pH 7.4). The eluted filtrate fraction was then concentrated using Amicon Ultra concentrator (Millipore) with a 10 kDa cut-off filter. Protein purity and integrity was analysed by SDS-PAGE and SEC-HPLC (TSKgel G2000 SWxl column on Agilent HPLC, 1x PBS pH 7.4) respectively (Supplementary Figure S2-10). LC-MS (Agilent PLRP-S column on Waters UPLC Xevo G2 Q-TOF) was used to confirm the protein mass, and the endotoxin was determined using Kinetic Turbidimetric LAL Test kit.

FGF-R1C analytical data:

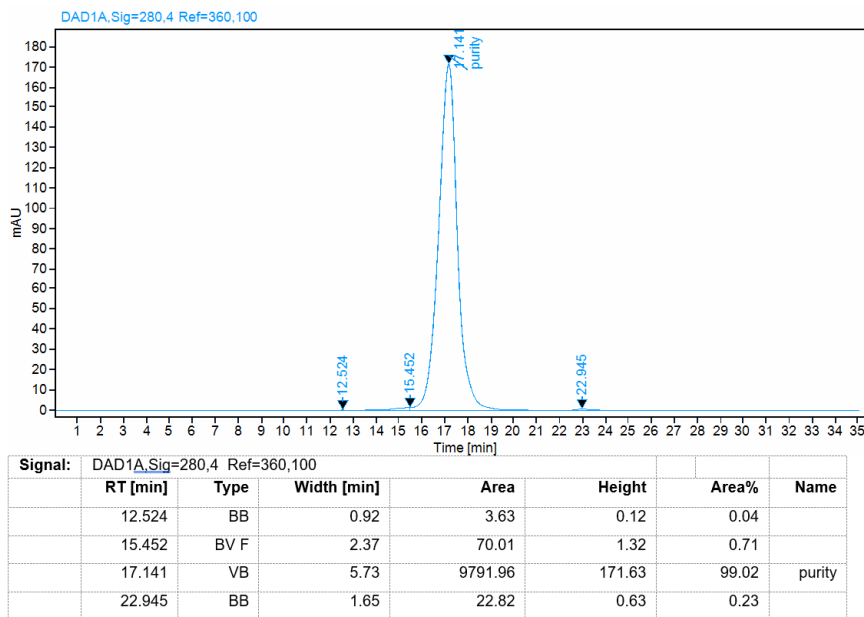


Figure S 2 SEC-HPLC chromatogram of purified FGF-R1C ECD. Column: TSKgel G2000 SWxl

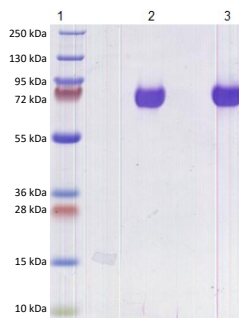


Figure S 3 SDS-PAGE of purified biotinylated FGF-R1C ECD. Lane 1: Molecular weight marker. Lane 2: FGF-R1C ECD (reducing conditions), Lane 3: FGF-R1C ECD (oxidising conditions)

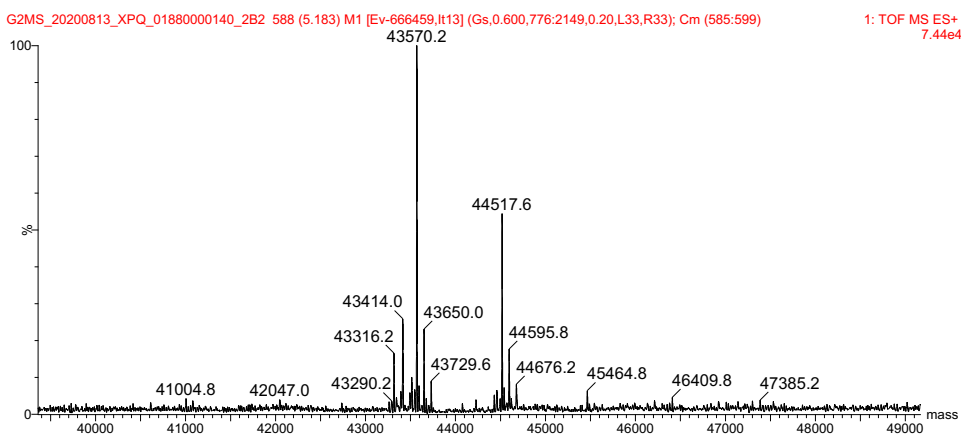
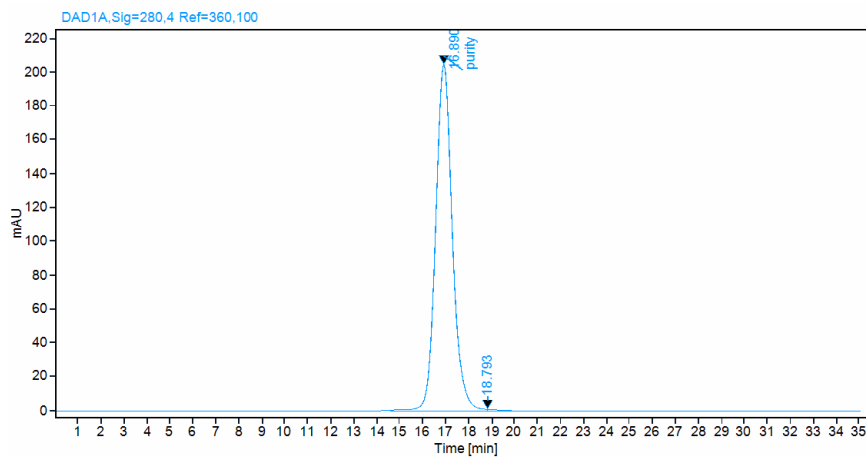


Figure S 4 LC-MS spectrum of purified biotinylated FGF-R1C ECD. Theoretical m/z biotinylated FGF-R1C ECD: 43570.3 Da, observed m/z: 43570.2. Theoretical m/z biotinylated FGF-R1C ECD + O-Glycan modification: 44517.3, observed m/z: 44517.6.

FGF-R3C analytical data:



Signal: DAD1A, Sig=280,4 Ref=360,100							
RT [min]	Type	Width [min]	Area	Height	Area%	Name	
16.89	VV B	4.24	10381.24	204.68	99.64	purity	
18.793	VB B	2.59	37.39	0.96	0.36		

Figure S 5 SEC-HPLC chromatogram of purified FGF-R3C ECD. Column: TSKgel G2000 SWxl.

G2MS_20200813_XPQ_01880000141b_2B2 571 (5.037) M1 [Ev0,It12] (Gs,0.600,5 1.72e5

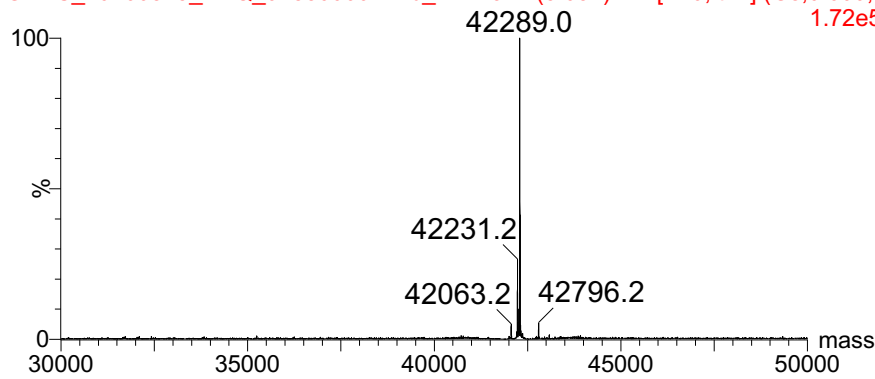


Figure S 6 LC-MS spectrum of purified biotinylated FGF-R3C ECD. Theoretical m/z biotinylated FGF-R3C ECD: 42290.1 Da, observed m/z: 42289.0.

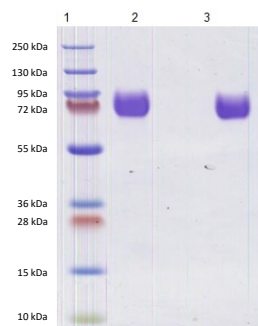


Figure S 7 SDS-PAGE of purified biotinylated FGF-R3C ECD. Lane 1: Molecular weight marker. Lane 2: FGF-R3C ECD (reducing conditions), Lane 3: FGF-R3C ECD (oxidising conditions)

FGF-R4 analytical data:

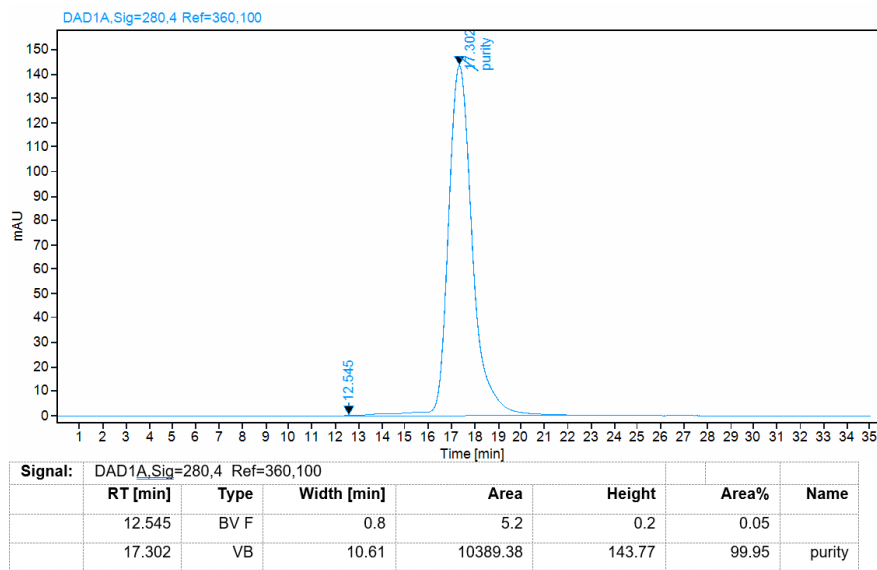


Figure S 8 SEC-HPLC chromatogram of purified FGF-R4 ECD. Column: TSKgel G2000 SWxl.

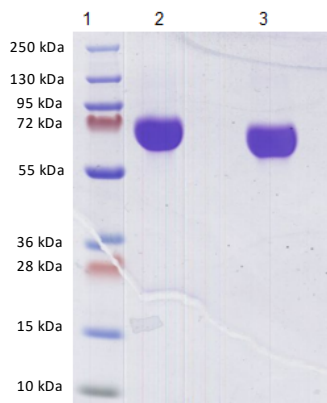


Figure S 9 SDS-PAGE of purified biotinylated FGF-R4 ECD. Lane 1: Molecular weight marker. Lane 2: FGF-R4 ECD (reducing conditions), Lane 3: FGF-R4 ECD (oxidising conditions)

G2MS_20200813_XPQ_0188000142b_2B2 600 (5.285) M1 [Ev0,It13] (Gs,0.600,50 4.57e4

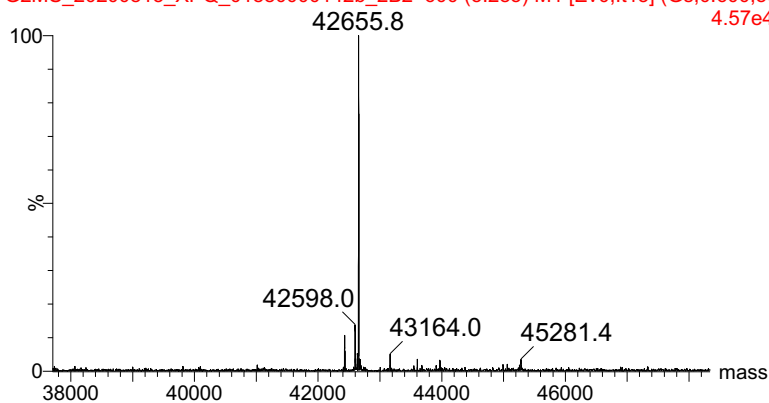


Figure S 10 LC-MS spectrum of purified biotinylated FGF-R4 ECD. Theoretical m/z biotinylated FGF-R4 ECD: 42657.5 Da, observed m/z: 42655.8.

In vitro selection of cyclic peptide libraries against FGF-R ECD targets

Preparation of puromycin-fused mRNA library

RNAs consisting of 4–12 repeated NNK random sequences (DNA: 5'-TAA TAC GAC TCA CTA TAG GGT TAA CTT TAA GAA GGA GAT ATA CAT **ATG TGC (NNK)_m TGC GGC AGC GGC AGC GGC AGC** TAG GAC GGG GGG CGG AAA-3', $m = 4-12$) were prepared by *in vitro* transcription according to the reported method.¹ The resulting RNAs were mixed in the following ratio: (NNK)₄:(NNK)₅:(NNK)₆:(NNK)₇:(NNK)₈:(NNK)₉:(NNK)₁₀:(NNK)₁₁:(NNK)₁₂=1:2:4:8:16:32:64:64:64. The mRNA library was ligated with a puromycin linker (5'-CTCCC GCCCC CCGTC C-(SPC18)₅-CC-puromycin-3') by T4 RNA ligase (Thermo Fisher). The ligated product was purified by phenol–chloroform extraction and ethanol precipitation.

Trimer-18 (T18) RNAs were constructed as above using 10–12 repeated random T18 codons purchased from Ella Biotech (DNA: 5'-TAA TAC GAC TCA CTA TAG GGT TAA CTT TAA GAA GGA GAT ATA CAT ATG TGC (XXX)_n TGC GGC AGC GGT TCT GGT AGT TAG GAC GGG GGG CGG AAA-3', $n = 10-12$). The RNAs were then mixed in the following ratio: (XXX)₁₀:(XXX)₁₁:(XXX)₁₂=1:1:1 or 1:18:324.

***In vitro* selection of cyclic peptides binding to FGF-R**

Translation of the first selection rounds was performed using 25 pmol mRNA-puromycin in 11.5 μ L of translation mixture (PURExpress *in vitro* protein synthesis kit, NEB) at 37 °C for 30 min in the presence of PURExpress disulfide-bond enhancers (NEB), followed by incubation at 60 °C for 10 min to dissociate ribosomes from the mRNA constructs.

This solution was then diluted up to 18.7 μ L with M-MLV reverse transcriptase (Promega, 5.3 U/ μ L), reverse primer P2 (5'-TGC GGC AGC GGC AGC GGC AGC TAG GAC GGG GGG CGG AAA-3', 1.8 μ M), M-MLV buffer (Promega), dNTPs (0.3 mM), and MilliQ water. The contents were reverse transcribed at 42 °C for 1 h. Following this, the contents were diluted to 40 μ L with selection buffer (PBS pH7.4 supplemented with 0.05% Tween-20 and 0.01% BSA). For FGF-R4, yeast total RNA (Roche diagnostics) was added to the diluted reverse-transcribed pool to a final concentration of 1 mg/mL.

Negative selections were performed by incubation of the diluted reverse-transcription solution for 30 min at 4 °C with a suspension of streptavidin-functionalised Dynabeads magnetic beads (M-280, Invitrogen) which were pre-loaded with biotinylated Avi-tag peptide for NNK selections or biocytin (Sigma Aldrich) for T18 selections. Following this, the beads were precipitated using a magnet, the supernatant collected, and transferred to a fresh tube for another round of negative selection. After three rounds of negative selection, the resulting supernatant was then incubated for 30 min at 4 °C with biotinylated FGF-R loaded

magnetic beads. For T18 selections, FGF-R4 was pre-incubated with the supernatant for 30 min at 4 °C before incubation for a further 10 min with magnetic beads.

All bead pellets were washed three times with milliQ water, and the bound cDNA sequences eluted by incubation at 95 °C for 5 min in 100 µL PCR buffer (1x NH₄ buffer, 2.5 mM MgCl₂, 0.25 mM dNTPs, and 0.5 µM each of reverse primer P2 (5'-TGC GGC AGC GGC AGC GGC AGC TAG GAC GGG GGG CGG AAA-3') and forward primer P1 (5'-TAA TAC GAC TCA CTA TAG GGT TAA CTT TAA GAA GGA GAT ATA CAT ATG TGC-3'). For T18 selections PCR buffer used reverse primer CPrevCys.R39 (5'-TTT CCG CCC CCC GTC CTA ACT ACC AGA ACC GCT GCC GCA-3') and forward primer CPfwdCys.F51 (5'-TAA TAC GAC TCA CTA TAG GGT TAA CTT TAA GAA GGA GAT ATA CAT ATG TGC-3'). Absolute quantities of recovered DNA after each selection step were determined by qPCR (Figure S 11).

After addition of Taq DNA polymerase to the positive selection eluate, the mixture was used for PCR amplification. The amplified DNA was purified by phenol/chloroform extraction followed by ethanol precipitation. The resulting DNA library was transcribed *in vitro*, the resultant RNA purified by phenol/chloroform extraction, and quantified using an Invitrogen™ Qubit™ 4 Fluorometer (Thermo Fisher). The RNA was then ligated with the puromycin linker and used for the next round of selection.

Subsequent selection rounds were performed as described but using 12.5 pmol puromycylated mRNA in 5.5 µL IVTT reactions instead. Dynabead (mass per pmol of puromycylated mRNA library) used for both negative and positive selections for each target was 8-12 µg/pmol. Protein (8-10 pmol) was loaded onto the Dynabeads to a final concentration of 420-500 nM. After the final rounds of the selection, the resulting cDNA libraries were sequenced by next-generation sequencing.

For mRNA display cross screens mRNA display procedure was performed as above. Round 5 (R5) input puromycin-ligated RNA from all three NNK selections were subjected to two R5 selections in parallel against the two receptors not selected for in the previous 4 rounds, i.e. R5 input ligated RNA from FGF-R1C selection was selected against both FGF-R3C and FGF-R4 in parallel.

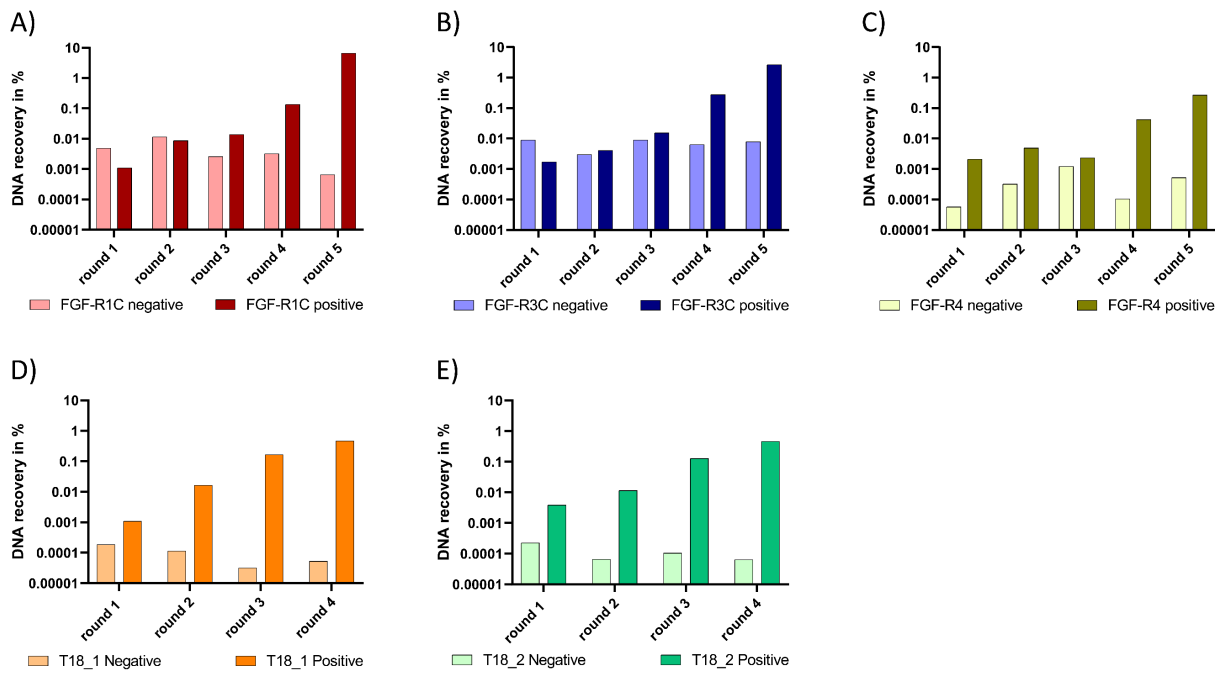


Figure S 11 Enrichment data from individual selections against each FGF-R target using NNK input libraries (1:2:4:8:16:32:64:64:64 4-12mer variable codons) (A: FGF-R1C, B: FGF-R3C, C: FGF-R4) and against FGF-R4 using two different Trimer-18 input libraries. T18_1: 1:1:1 10-12mer variable codons, T18_2: 1:18:324 10-12mer variable codons libraries (D-E). % DNA recovery as determined by qPCR shown for each round of selection against biotinylated FGF-R ECD. Negative data obtained after 3 selections against beads immobilised with biotinylated avitag peptide (A-C) or biocytin (D-E), positive data obtained after single selection against beads immobilised with target receptor.

Next generation sequencing and clustering analysis

For Illumina MiSeq sequencing, the DNA output of each selection round was amplified in two PCR steps to append the required barcodes and constant regions. The PCR amplification was performed using Roche KAPA HiFi HotStart Library Amp Kit following the manufacturer's protocol. The PCR products were then gel purified, quantified using Qubit Broad range dsDNA assay kit (Thermo Fisher) and pooled into libraries of 20 pM. The libraries also contained 10% PhiX DNA following manufacturers recommendations. Finally, the NGS library was sequenced using MiSeq® Reagent Kit v3 (150 cycles). See already published procedures for further details and clustering procedure (Bhushan, B. et al, Chem Sci 2022).²

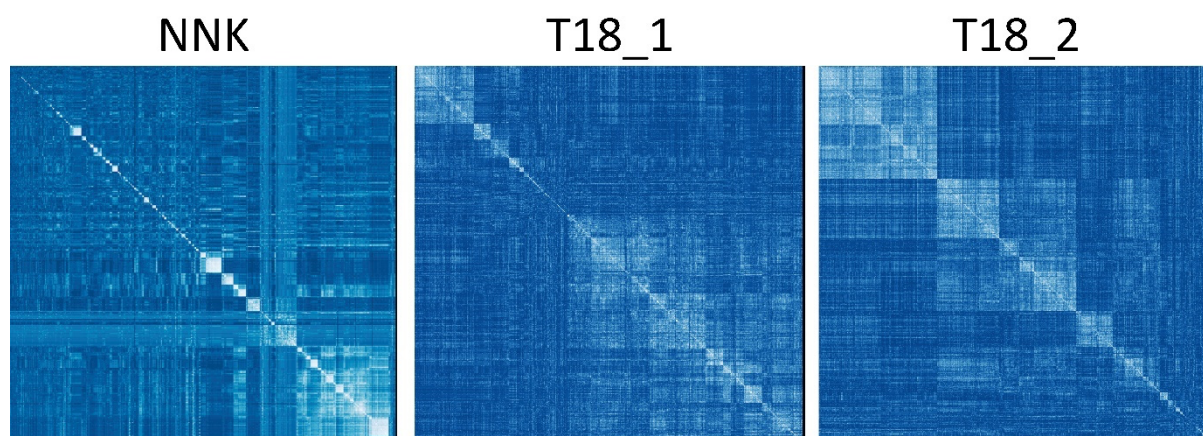


Figure S 12 Comparison of clustering of three different mRNA display libraries against FGF-R4 target. NNK library comprised 4-12mer variable NNK codons (1:2:4:8:16:32:64:64:64), T18_1 comprised 10-12mer variable Trimer-18 codons (1:1:1), T18_2 comprised 10-12mer variable trimer-18 codons (1:18:324). NGS hierarchical clustering diagrams of the top enriched sequences in 5th NNK selection round or 4th T18 selection round. Peptide similarity for each enriched hits are compared individually for each target with low (dark) to high (light) similarity scores shaded as a heat-map: NNK FGF-R4 selection (left), T18_1 FGF-R4 selection (middle), and T18_2 FGFR4 selection (right).

Chemical synthesis of selected macrocyclic peptides

General: All chemicals were of analytical grade or higher. Acetonitrile (MeCN) (LiChrosolve), trifluoroacetic acid S3 (TFA), and diethyl ether were purchased from Merck KGaA (Darmstadt, Germany). Triisopropylsilane (TIPS), N,N'-diisopropylcarbodiimide (DIC), dithiothreitol (DTT), collidine, dimethyl sulfoxide (DMSO) were from Sigma–Aldrich, Chemie GmbH (Steinheim, Germany). Water was obtained from a MilliQ equipment (Advantage A10) from Millipore (Molsheim, France). N-Methyl pyrrolidone (NMP) dimethylformamide (DMF) and piperidine were from Biosolve (Dieuze, France). Standard Fmoc amino acids, resins and coupling reagents, Oxyma Pure were from Novabiochem (Darmstadt, Germany) or Protein Technologies (Tucson, USA). PS resin, and Fmoc-Rink amide PS resin, were all purchased from Merck Millipore (Novabiochem).

Array synthesis of peptides selected for primary screen/ full amino acid scans: Peptides selected for initial high throughput screening, as well as those for mutational scans, were synthesised in a 96-well plate format using Intavis Multiprep RSI instruments. The synthesis was performed on a 5 µmol scale using 0.3 mM of

Fmoc-PAL AM resin (Novabiochem) pre-loaded with the sequence: GSGSDYKDDDDK-NH₂. This corresponds to the spacer sequence present in our mRNA display libraries as well as a Flag-tag added for solubility and potential detection in bioassays. Synthesis was performed using standard Fmoc chemistry. Double deprotection (5 + 8 min) was achieved at room temperature using 20% (v/v) piperidine with 0.1 M Oxyma in DMF (200 μL). Double couplings (60 + 120 min) were performed at room temperature with 3 M DIC (9 μL), 3 M collidine (9 μL), and 0.1 M amino acid solution containing 0.1 M Oxyma (83 μL), all in DMF. A capping step was added after each amino acid coupling for some peptides using 0.3 M AcOH in DMF/ 0.1 M Oxyma with coupling reagents. Peptides shown in supplementary table S3/4/5.

After completion of the synthesis, peptides were cleaved from the resin using a mixture of trifluoroacetic acid (TFA)/ water/ triisopropylsilane (TIPS)/ dithiothreitol (DTT) (94:2:2:2). Approximately 1.2 mL of the cleavage mixture was added stepwise each well over two hours. After incubation, addition of diethylether was used to precipitate the peptides and the subsequent mixture of cleavage solution/diethylether was separated from the precipitated peptide using gravity filtration. The peptides were then washed an additional 3 times with 500 μL diethylether and dried using vacuum filtration. Each of the dried peptides were re-dissolved in DMSO (1 mL) and transferred to a master plate (96 deep-well). Concentration in the master plate was mostly in the range 1 – 2 mM. Finally, peptides were diluted to 50 μM in folding buffer (100 mM HEPES pH 7.4, 20% DMSO) to promote disulfide bond formation and peptide macrocyclization. 50 mM ammonium bicarbonate was also used for some 96 well plates instead of 100 mM HEPES.

Otherwise, peptides were purchased from Apigenex, Prague, Czech Republic, or GL-Biochem, Shanghai, China.

Purification of selected peptides: Purification was performed on Gilson LC systems, equipped with a 322 pump module, 155 UV/vis detector and GX-271 automated sample collector. A typical gradient of 20% - 60% solvent B over 30 min (flow rate of 25 mL/min, Axia Gemini 5μM NX-C18 110 Å columns, 250x30 mm) was employed (Solvent A: 0.1% TFA in water, Solvent B: 0.1% TFA in Acetonitrile). Finally, each peptide was isolated by lyophilisation.

Peptide characterisation: 96 well plate peptides were characterised by LC-MS (Waters Acquity system with QDA detector, KDA3351, and TUV detector, F08UPT321M) using Waters Acquity UPLC BEH C18 column (130 Å, 1.7 μm, 2.1 mm x 50 mm). A 3 min linear gradient of 10% to 90% solvent B was run; solvent A: 0.05% TFA in water, solvent B: 0.05% TFA in MeCN. NNK Purified peptide identity was determined either by method 1: Agilent 1220 system with a Phenomenex BioZen Peptide PS-C18 column (3 μm, 150 x 4.6 mm) using a 20 min linear gradient of 2% to 98% solvent B in solvent A (solvent A 0.1% TFA in water, solvent B 0.085% TFA in MeCN); method 2: Waters Acquity system with a PDA detector (MIOUPD33A), and SQ detector (BA699) using Waters Acquity UPLC BEH C18 column (130 Å, 1.7 μm, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 80% solvent D in solvent C (solvent C: 10 mM Tris 15 mM(NH₄)₂SO₄, solvent D:

80% MeCN in water); method 3: Waters Acquity UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 60% solvent D in solvent A (solvent A: 0.1% TFA in water, solvent D: 80% MeCN in water). Purified peptide identity shown in Table S 1 and Table S 2. Analytical HPLC traces shown in Figure S 13-17.

Peptides were quantified using either NanoDrop (Thermo Scientific One Microvolume UV-Vis Spectrophotometer, for peptides 4.9-4.25 only), or UHPLC-CAD (Thermo Fisher Vanquish system) with a Waters Acquity UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 mm x 50 mm) coupled to a charged aerosol detector H. Gradient: 6 min, 5-95% solvent F in solvent E, flow rate 0.45 mL/min.; solvent E: 0.05% TFA in water, solvent F: 0.05% TFA in MeCN.

Table S 1 Sequences and UPLC-MS identity of purified peptides from NNK selections and initial screening. 1C.9 and 1C.10 LC-MS data show HPLC fraction data prior to fraction pooling.

Name	Sequence	N-terminus	C-terminus	HPLC method	HPLC retention time [min]	HPLC Purity [%]	Exact mass found	Exact mass calculated [M+xH]/x	charge state MS x =
1C.1	MCIGKWYTYQYWTRC	H	GSGSDYKDDDDK-NH ₂	2	7.4	81.6	1095.1	1094.5	3
1C.2	MCLLWTRHFDRYWVC	H	GSGSDYKDDDDK-NH ₂	2	7.3	61.4	1104.1	1103.5	3
1C.3	MCSNMWKLNRWWAC	H	GSGSDYKDDDDK-NH ₂	2	7.6	65.3	1060.9	1060.4	3
1C.4	MCVPCFSHTFLCC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.1	84.3	937.9	938.0	3
1C.5	MCVPCFSHTFLCC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.8	81.2	1406.2	1406.5	2
1C.6	MCYRLSDFWVVC	Acetyl	GSGSDYKDDDDK-NH ₂	3	5.9	55.0	1479.3	1478.6	2
1C.7	MCYRLSDFWVEC	Acetyl	GSGSDYKDDDDK-NH ₂	1	3.8	93.1	1494.0	1493.6	2
1C.8	MCTKFPWPPIFWAC	H	GSGSDYKDDDDK-NH ₂	2	2.0	49.5	1052.4	1051.8	3
1C.9	MCIRWPARQPLFWLC	H	GSGSDYKDDDDK-NH ₂	2	8.4	75.7	1067.6	1067.2	3
1C.10	MCFRWRLFHHLTPFC	H	GSGSDYKDDDDK-NH ₂	2	7.8	53.0	1092.4	1091.8	3
3C.1	MCIFDLFLPFIFC	Acetyl	GSGSDYKDDDDK-NH ₂	2	6.5	91.3	1466.4	1465.6	2
3C.2	MCIFDIHMPFIPC	H	GSGSDYKDDDDK-NH ₂	3	6.0	95.7	950.0	949.4	3
3C.3	MCLFDSFVPFIFC	H	GSGSDYKDDDDK-NH ₂	3	6.7	83.0	1425.5	1424.6	2
3C.4	MCQQWNLYTRHWHSC	Acetyl	GSGSDYKDDDDK-NH ₂	1	3.3	97.0	1105.9	1105.4	3
3C.5	MCHPTRLHLWCCNC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.2	62.1	1051.7	1051.1	3
3C.6	MCHYCFNKLFLCC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.3	91.9	1473.2	1473.6	2
4.1	MCLVSIYKIMRTC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.2	100.0	962.1	961.4	3
4.2	MCHVILWRSWRYC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.4	69.6	1538.8	1537.7	2
4.3	MCHLAYLAFRSC	Acetyl	GSGSDYKDDDDK-NH ₂	1	4.5	88.8	1450.6	1450.1	2
4.4	MCVTWYKVKTLRC	Acetyl	GSGSDYKDDDDK-NH ₂	1	3.7	89.4	985.5	984.8	3
4.5	MCFKHPWFFAWLC	Acetyl	GSGSDYKDDDDK-NH ₂	1	5.7	79.2	1521.1	1519.1	2
4.6	MCIKMFLAHRFC	Acetyl	GSGSDYKDDDDK-NH ₂	3	5.0	92.3	984.4	983.8	3
4.7	MCPERLFYWLLVC	Acetyl	GSGSDYKDDDDK-NH ₂	1	5.4	51.9	1000.0	998.8	3
4.8	MCEPWITIKVLKCC	Acetyl	GSGSDYKDDDDK-NH ₂	3	5.3	79.9	1060.1	1059.5	3

Table S 2 Sequences and UPLC-MS identity of purified peptides from trimer-18 selections and initial screening. Peptides purchased from GL-biochem (shanghai). GSGSEE sequence included for peptides with calculated isoelectric point (pI) of approximately 7

Name	Sequence	N-terminus	C-terminus	HPLC retention time [min]	HPLC Purity [%]	Exact mass found	calculated [M+xH]/x	state MS x =
4.9	MCTTLYYYKLWRRRC	Acetyl	GSGSEEGSGSDYKDDDDK-NH ₂	12.40	85.19	1257.2	1256.5	3
4.10	MCKHLYYYLLWKKC	Acetyl	GSGSEEGSGSDYKDDDDK-NH ₂	13.36	88.75	1254.6	1253.9	3
4.11	MCNVLLYKVLRLHC	Acetyl	GSGSDYKDDDDK-NH ₂	10.41	86.88	1458.1	1457.2	2
4.12	MCESDAYRWKLWWVC	Acetyl	GSGSDYKDDDDK-NH ₂	10.64	86.50	1100.1	1099.8	3
4.13	MCYTRFYNWEKWQLC	Acetyl	GSGSDYKDDDDK-NH ₂	11.62	87.56	1132.2	1131.5	3
4.14	MCYHRPPNSKIWRKC	Acetyl	GSGSDYKDDDDK-NH ₂	12.61	88.36	1621.6	1620.7	2
4.15	MCEREVLYWLLVTC	Acetyl	GSGSDYKDDDDK-NH ₂	10.94	85.07	1541.1	1540.2	2
4.16	MCYLDDFYSRYWAC	Acetyl	GSGSDYKDDDDK-NH ₂	12.20	88.29	1661.5	1660.6	2
4.17	MCNVLLYKDLRHC	Acetyl	GSGSDYKDDDDK-NH ₂	9.26	86.47	1465.7	1465.1	2
4.18	MCNDFYLLWRKAC	Acetyl	GSGSDYKDDDDK-NH ₂	8.38	86.11	1088.1	1087.5	3
4.19	MCYNKFITLYKVLVC	Acetyl	GSGSDYKDDDDK-NH ₂	10.66	85.53	1075.7	1075.1	3
4.20	MCENKVLFFYLLWC	Acetyl	GSGSDYKDDDDK-NH ₂	9.14	85.09	1092.7	1092.2	3
4.21	MCDKYWVKYWIAWRC	Acetyl	GSGSDYKDDDDK-NH ₂	12.12	88.62	1125.4	1124.8	3
4.22	MCFWKKLHVFWVKDC	Acetyl	GSGSDYKDDDDK-NH ₂	9.56	95.71	1098.4	1097.8	3
4.23	MCEHLYYYLLSHRRC	Acetyl	GSGSDYKDDDDK-NH ₂	9.92	85.17	1104.4	1103.5	3
4.24	MCYTKAYWHEPWRPC	Acetyl	GSGSDYKDDDDK-NH ₂	10.08	92.50	1098.6	1098.1	3
4.25	MCNKHLYYYLLFKQC	Acetyl	GSGSDYKDDDDK-NH ₂	12.46	92.86	823.4	822.9	4

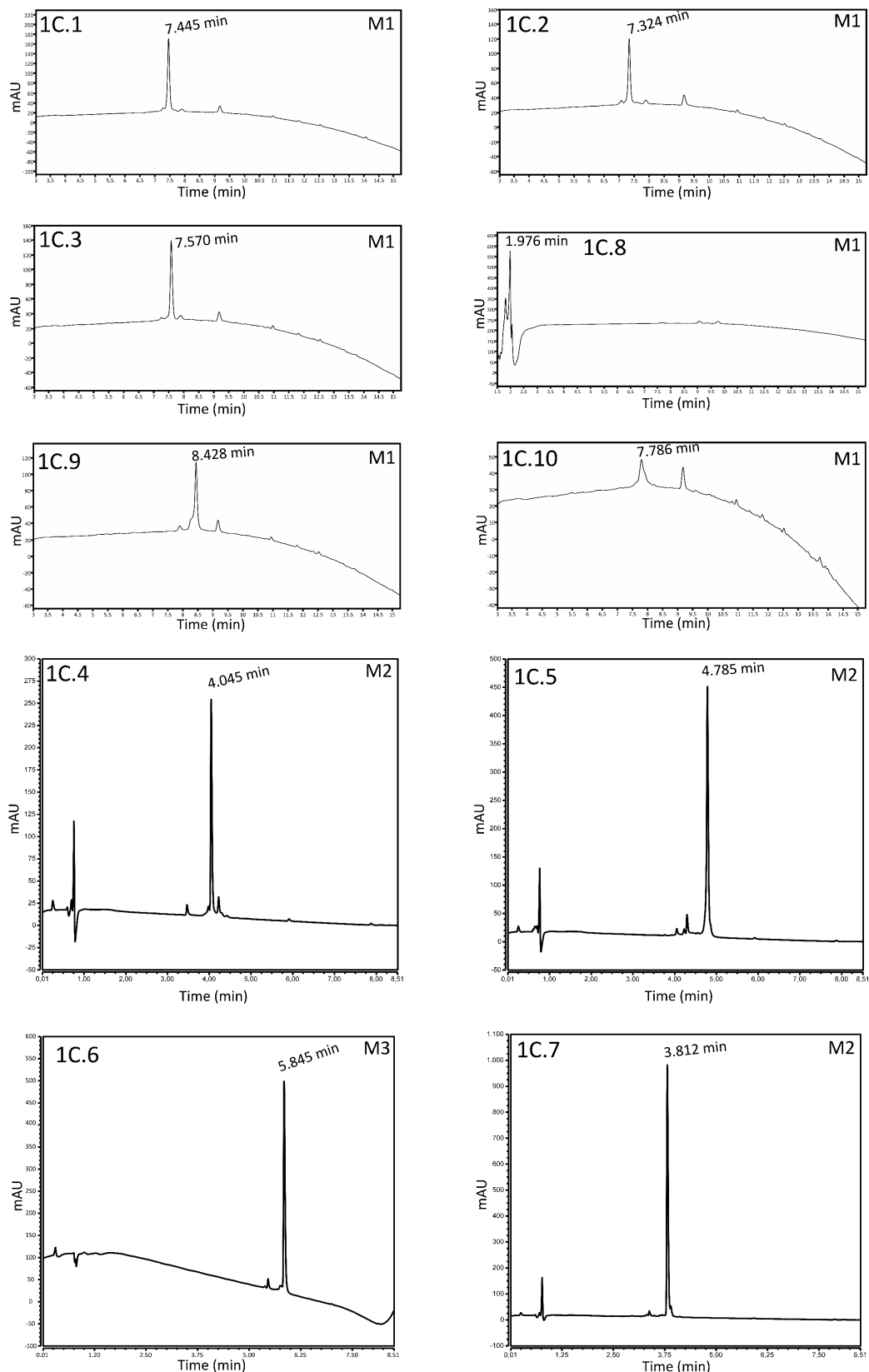


Figure S 13 HPLC traces of purified peptides from FGF-R1C NNK selection and initial screening. Method 1 (M1): Agilent 1220 system with a Phenomenex BioZen Peptide PS-C18 column (3 μ m, 150 x 4.6 mm) using a 20 min linear gradient of 2% to 98% solvent B in solvent A (solvent A 0.1% TFA in water, solvent B 0.085% TFA in MeCN); method 2 (M2): Waters Acquity system with a PDA detector (MIOUPD33A), and SQ detector (BA699) using Waters Acquity UPLC BEH C18 column (130 \AA , 1.7 μ m, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 80% solvent D in solvent C (solvent C: 10 mM Tris 15 mM(NH₄)₂SO₄, solvent D: 80% MeCN in water); method 3 (M3): Waters Acquity UPLC BEH C18 column (130 \AA , 1.7 μ m, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 60% solvent D in solvent A (solvent A: 0.1% TFA in water, solvent D: 80% MeCN in water).

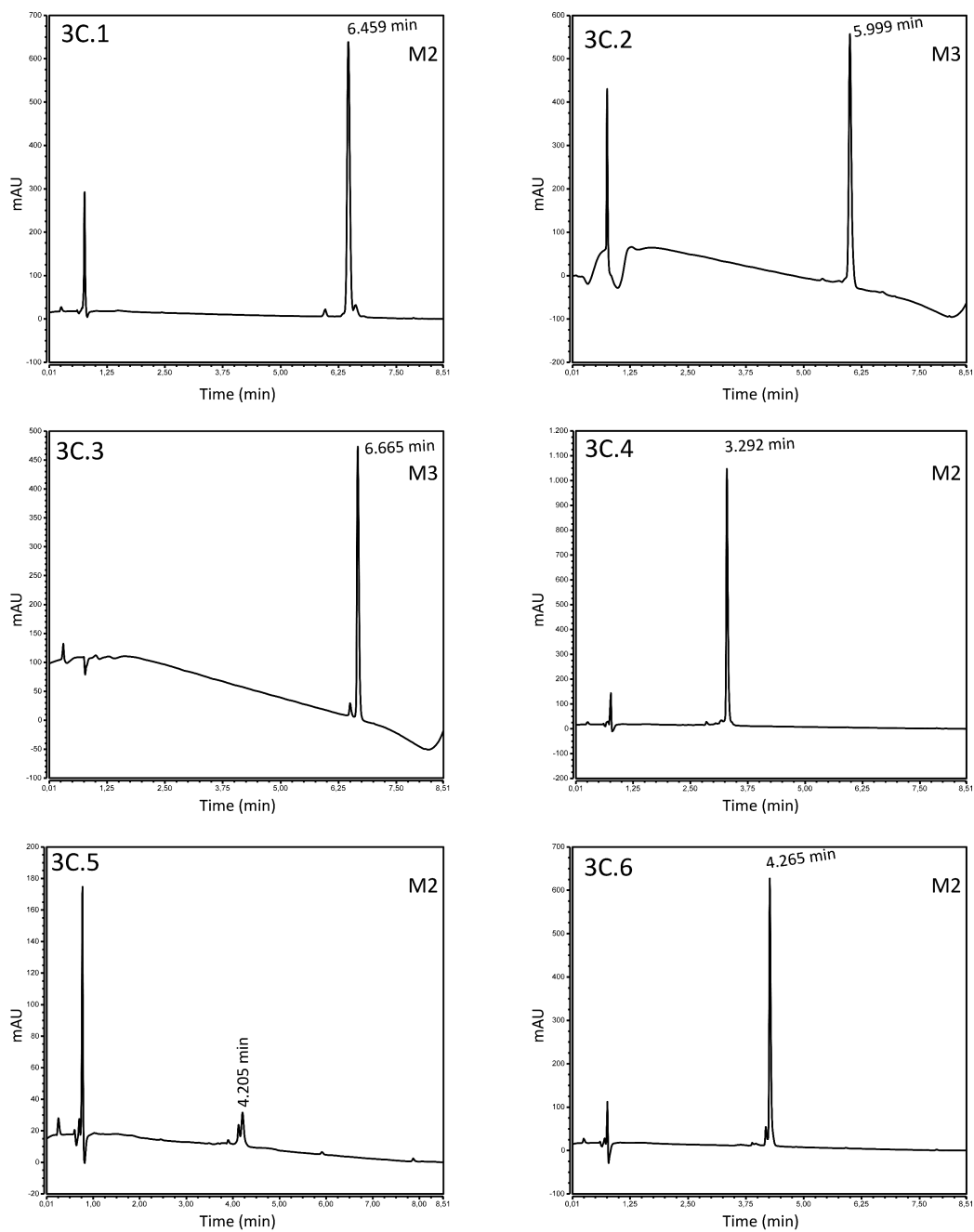


Figure S 14 HPLC traces of purified peptides from FGF-R3C NNK selection and initial screening. Method 2 (M2): Waters Acquity system with a PDA detector (MIOUPD33A), and SQ detector (BA699) using Waters Acquity UPLC BEH C18 column (130 Å, 1.7 μm, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 80% solvent D in solvent C (solvent C: 10 mM Tris 15 mM(NH₄)₂SO₄, solvent D: 80% MeCN in water); method 3 (M3): Waters Acquity UPLC BEH C18 column (130 Å, 1.7 μm, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 60% solvent D in solvent A (solvent A: 0.1% TFA in water, solvent D: 80% MeCN in water).

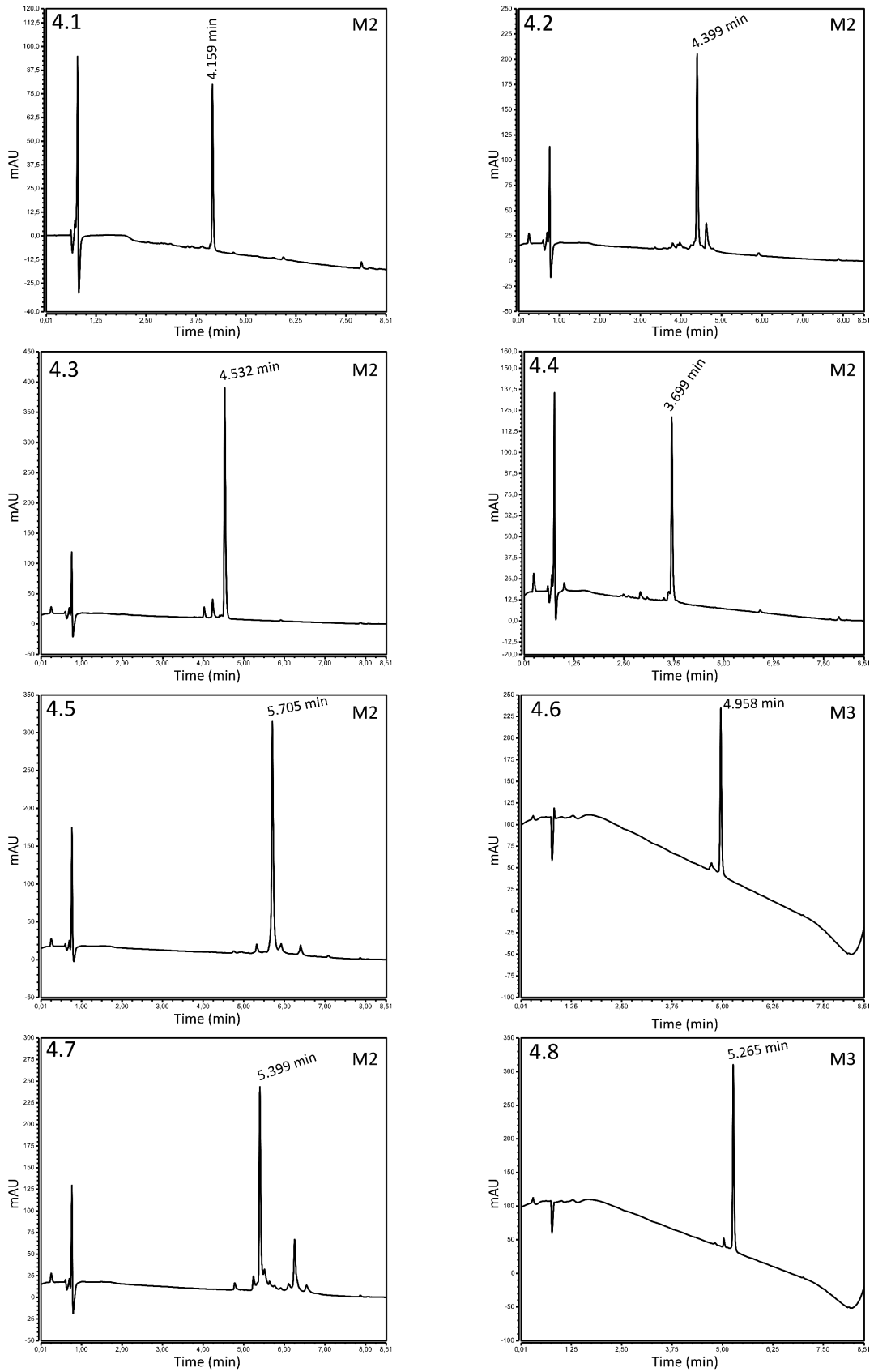


Figure S 15 HPLC traces of purified peptides from FGF-R4 NNK selection and initial screening. Method 2 (M2): Waters Acquity system with a PDA detector (MIQUPD33A), and SQ detector (BA699) using Waters Acquity UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 80% solvent D in solvent C (solvent C: 10 mM Tris 15 mM(NH₄)₂SO₄, solvent D: 80% MeCN in water); method 3 (M3): Waters Acquity UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 mm x 50 mm) on a 6 min linear gradient of 0% to 60% solvent D in solvent A (solvent A: 0.1% TFA in water, solvent D: 80% MeCN in water).

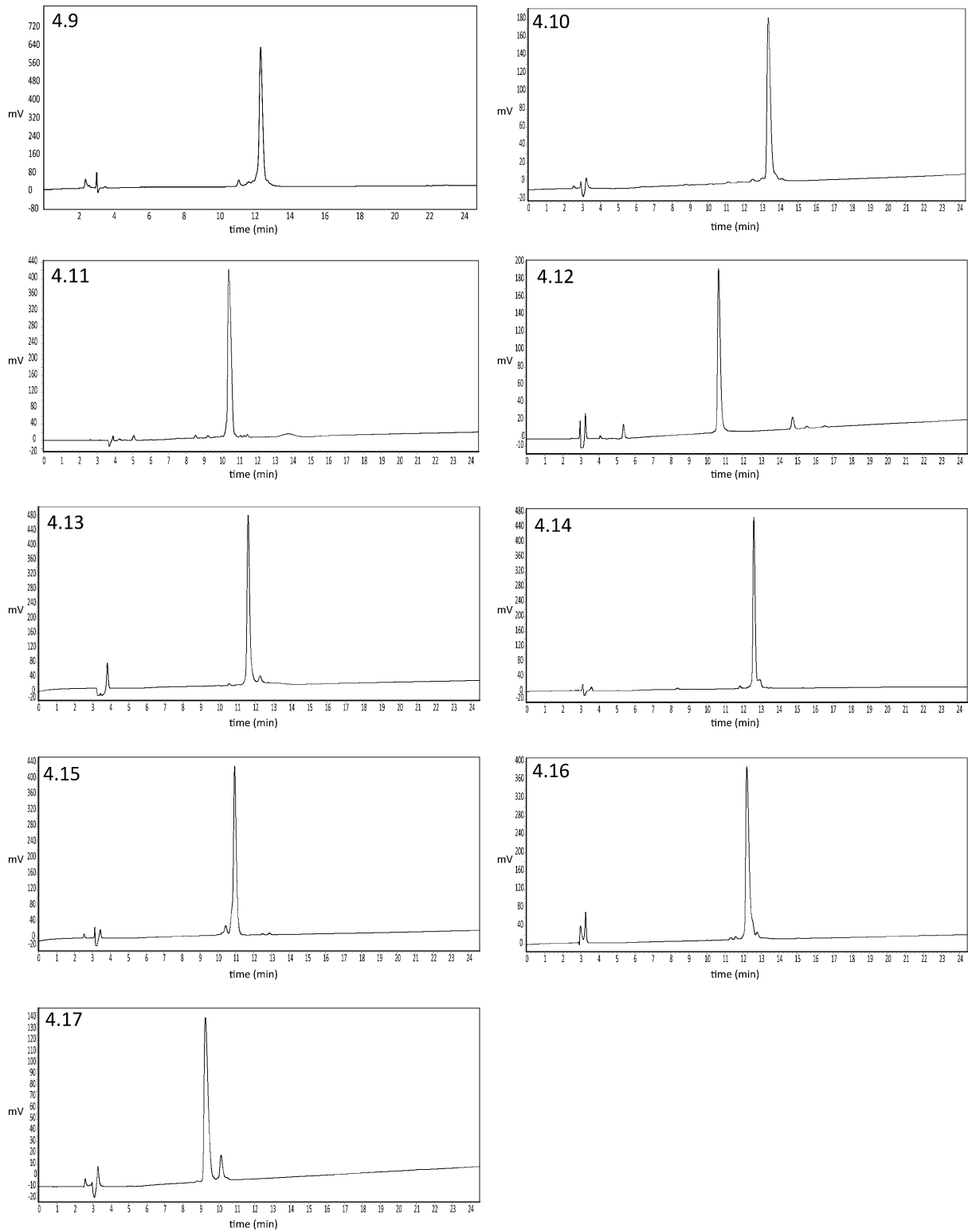


Figure S 16 HPLC traces of purified peptides from FGF-R4 T18_1 selection and initial screening. Peptides purchased from GL-biochem (shanghai). 250 x 4.6 mm, Kromasil-C18-5 μ m column used on a 20 minute linear gradient of 25% to 50% solvent G in solvent A (solvent A: 0.1% TFA in water, solvent G: 0.1% TFA in MeCN).

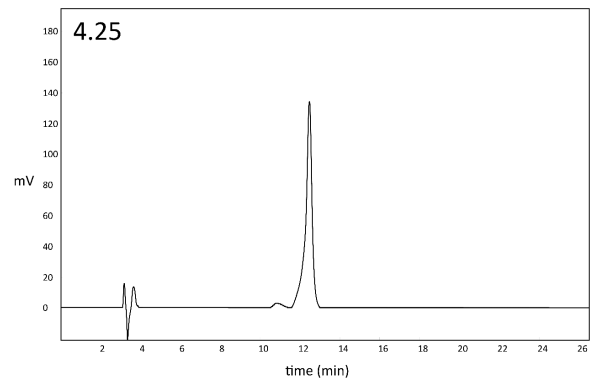
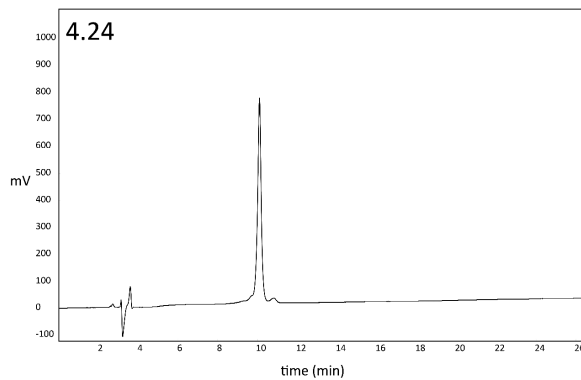
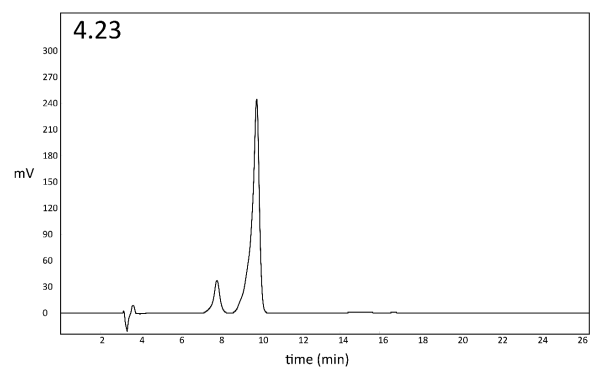
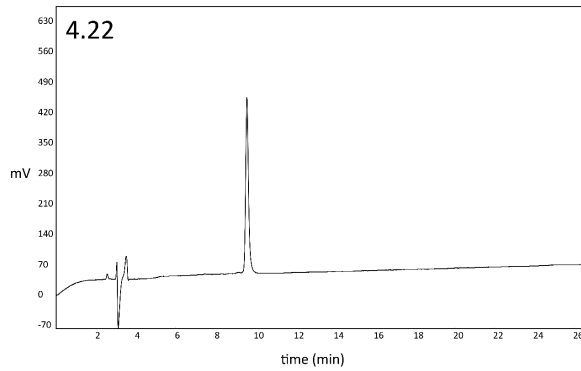
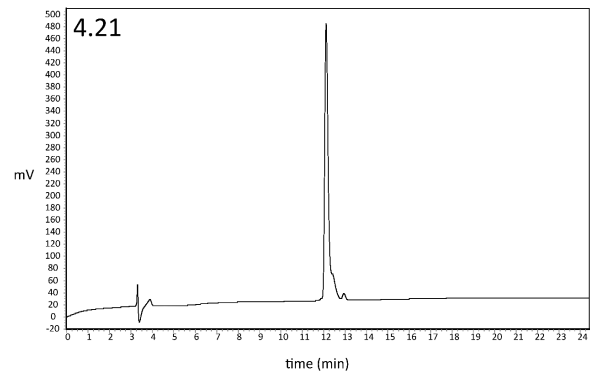
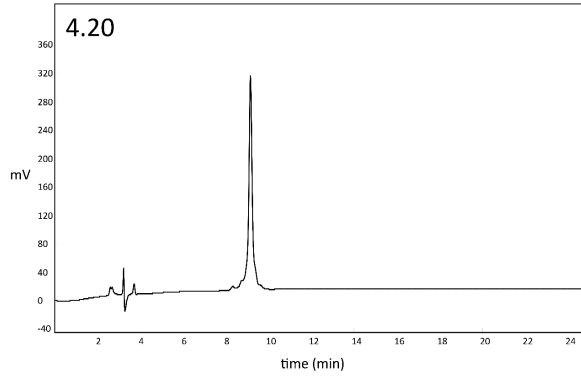
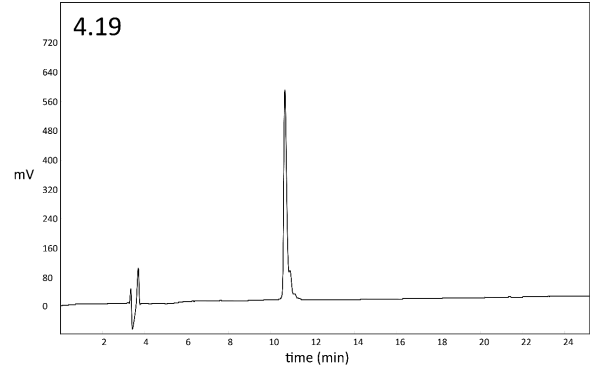
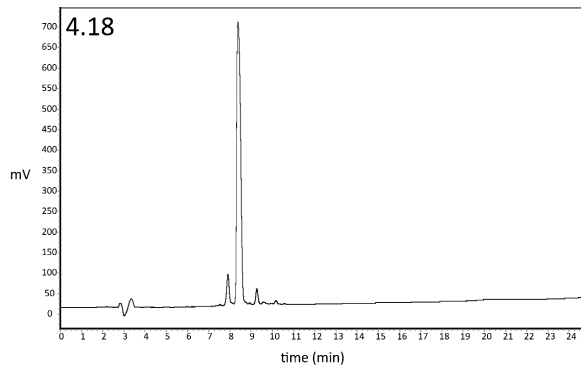


Figure S 17 HPLC traces of purified peptides from FGF-R4 T18_2 selection and initial screening. Peptides purchased from GL-biochem (shanghai). 250 x 4.6 mm, Kromasil-C18-5 μ m column used on a 20 minute linear gradient of 25% to 50% solvent G in solvent A (solvent A: 0.1% TFA in water, solvent G: 0.1% TFA in MeCN).

BLI binding measurements

K_D affinity binding measurements were obtained by biolayer interferometry (BLI) instruments (ForteBio Octet HTX, or Sartorius Octet RH16 using either 8 or 16 channel mode).

For single concentration binding experiments, 50 μM peptide stocks in folding buffer were diluted 1:50 in mRNA display selection buffer (PBS pH7.4 supplemented with 0.05% Tween-20 and 0.01% BSA) in either a 96 well black plate, or a 384 well tilted bottom microplate (ForteBio). Biotinylated FGF-R protein was loaded onto Streptavidin-functionalised biosensors (ForteBio) and immersed into $\sim 1 \mu\text{M}$ peptide to obtain association kinetics curves. Dissociation kinetics curves were obtained by subsequent immersion of ligand-associated loaded sensors into a 1:50 dilution of folding buffer in mRNA display selection buffer. Fitted K_D data was adjusted for concentration as determined by UPLC-CAD.

For accurate multi-concentration K_D determination, dilution series of each peptide were made, and association and dissociation kinetics curves were obtained as above for each concentration to fit a multi-point K_D binding model. Data collection and analysis was performed using the Octet Instrument Control and Data Analysis software packages (ForteBio, Ver 9.0.0.10) or Prism (GraphPad Software, Ver. 7.05). Reference sensor was subtracted for all assays and where non-specific binding was observed, double referencing was used. Exemplary BLI kinetics curves of peptides showing different levels of selectivity to FGF-Rs is shown in Figure S 18. Single concentration BLI data shown in supplementary table S3-5 and multi-concentration assay kinetics curves shown in supplementary Figure S 19-24.

Peptide concentration (nM)

— 1000 — 500 — 250 — 125 — 62.5 — 31.25 — 15.625

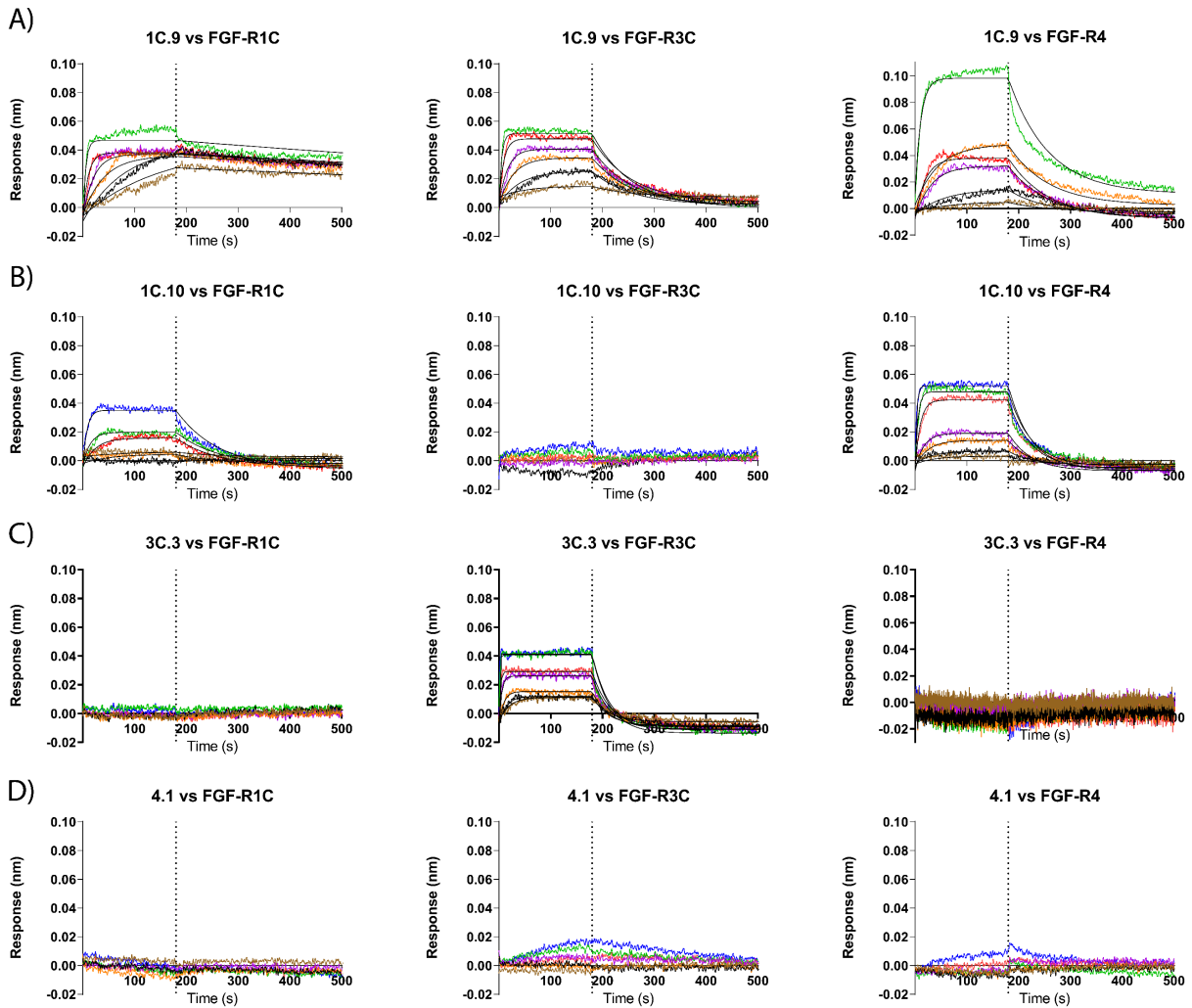


Figure S 18 Exemplary BLI kinetics curves from peptides displaying differing FGF-R selectivity. (A) pan binding 1C.9; (B) FGF-R1C and FGF-R3C binding 1C.10; (C) FGF-R3C selective 3C.3; (D) non-binding 4.1. Data from multi-concentration association and dissociation kinetics curves binding of FGF-R ECD to purified NNK CP. Data generated by ForteBio Octet HTX.

NNK purified peptide BLI binding data

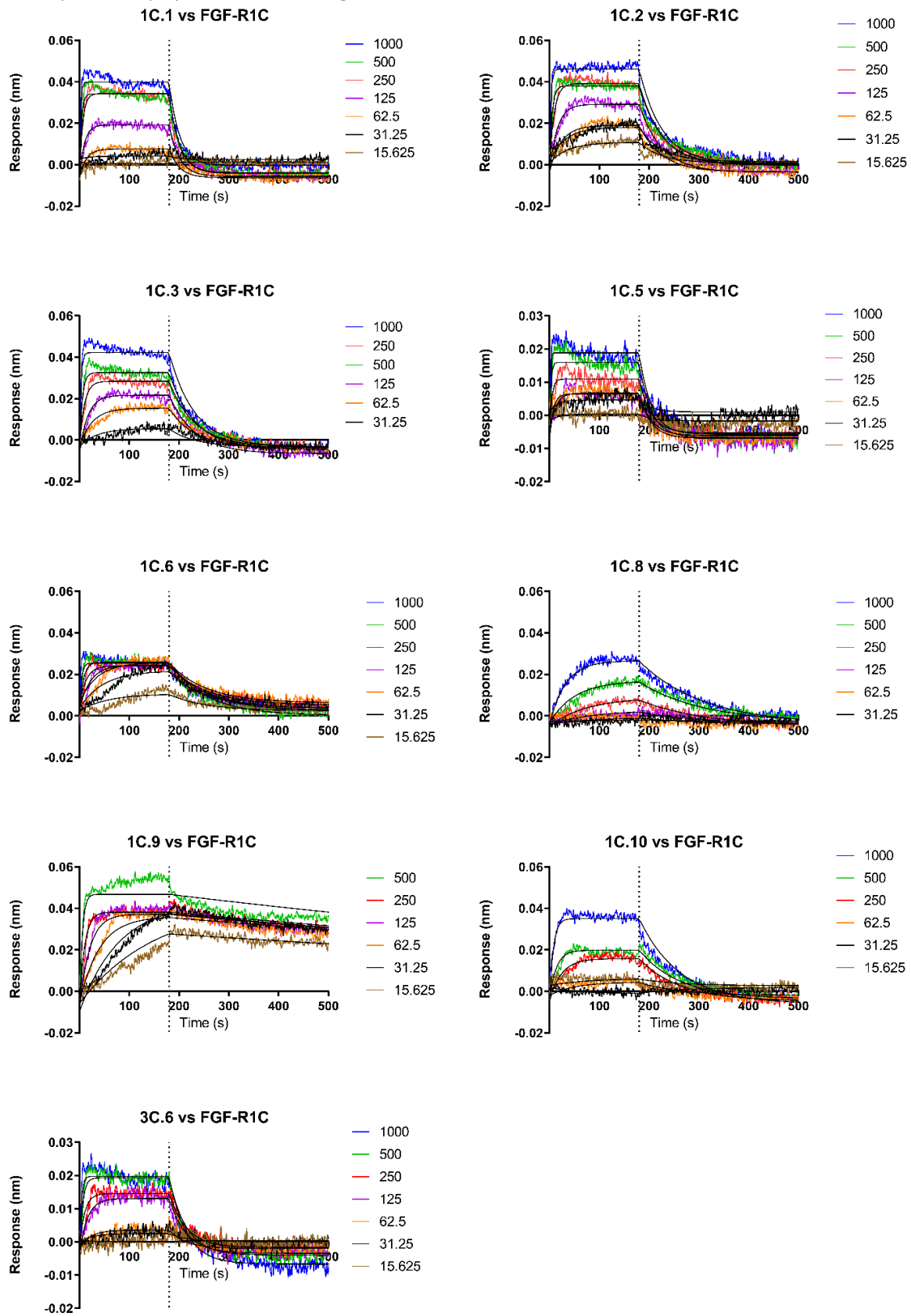


Figure S 19 Purified NNK CP multi-concentration association and dissociation kinetics curves binding to FGF-R1C ECD. Data generated by ForteBio Octet HTX. Concentrations shown in nM

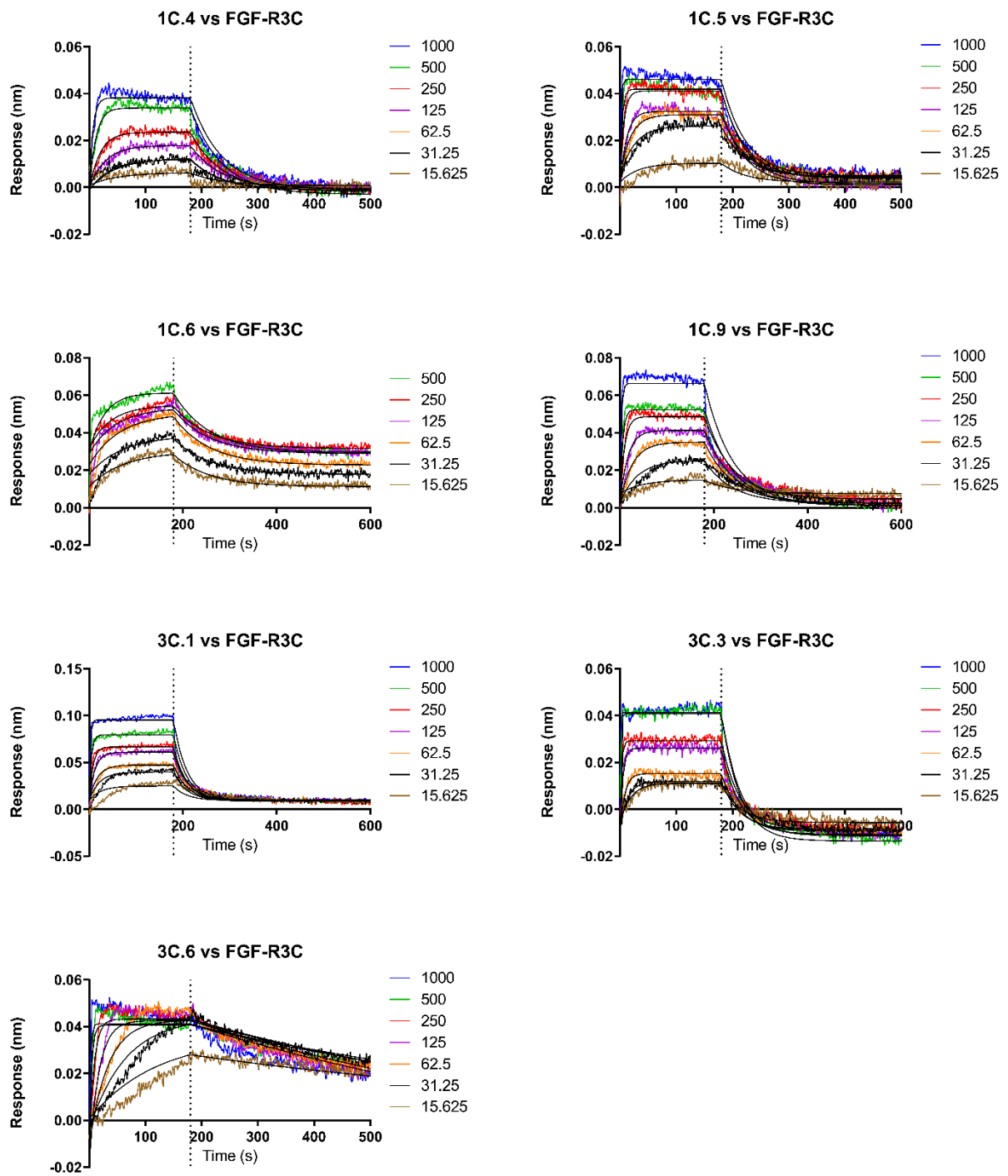


Figure S 20 Purified NNK CP multi-concentration association and dissociation kinetics curves binding to FGF-R3C ECD. Data generated by Fortebio Octet HTX. Concentrations shown in nM.

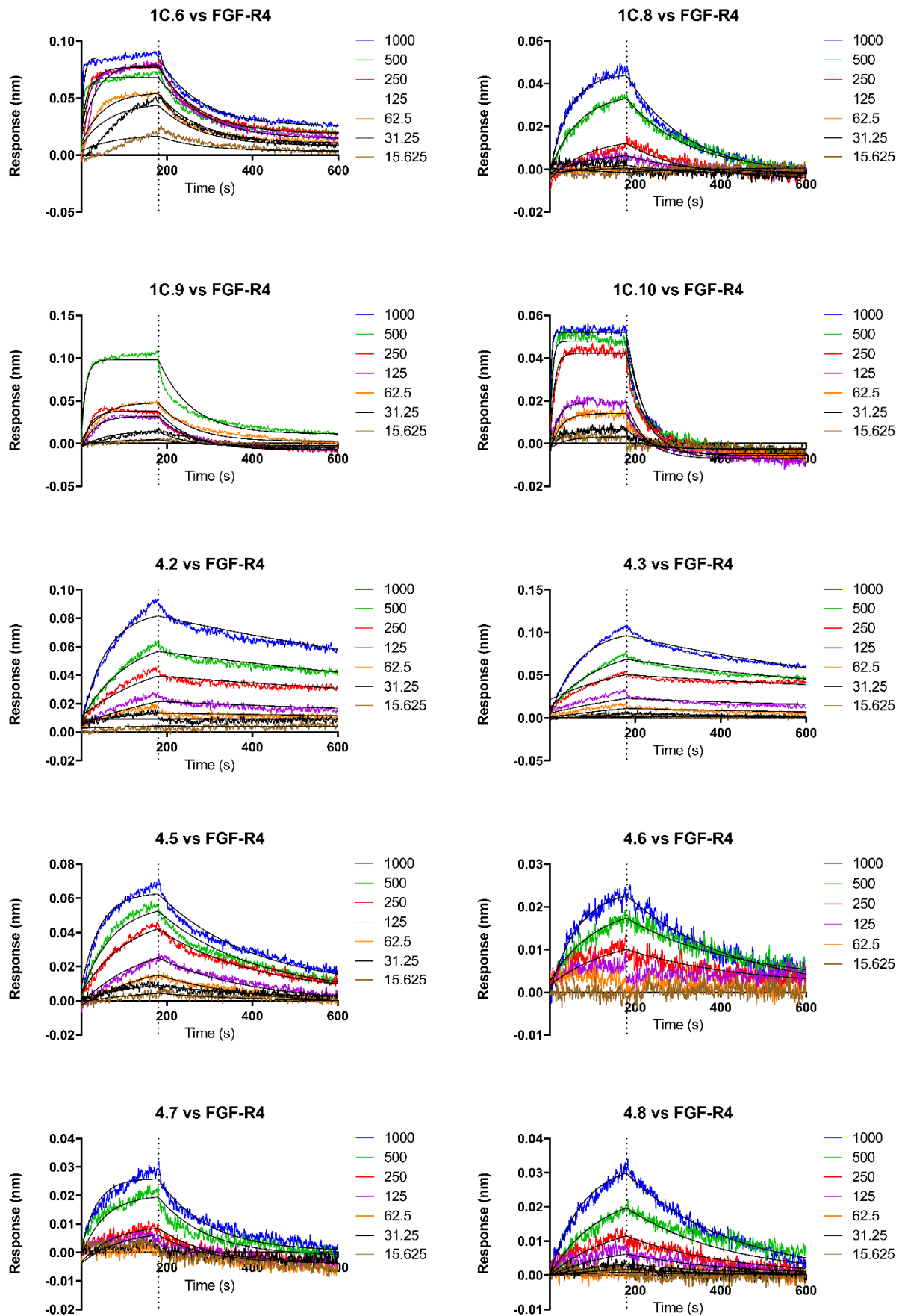


Figure S 21 Purified NNK CP multi-concentration association and dissociation kinetics curves binding to FGF-R4 ECD. Data generated by Fortebio Octet HTX. Concentrations shown in nM

Trimer-18 single concentration BLI initial screen

Table S 4 Single concentration K_D values of crude synthesised peptides against different FGF-R ECDs and corresponding FGF-R selected against during mRNA display from T18 selection1. Peptides shown were identified from trimer-18 selection 1 using a 1:1:1 library of 10-12mer variable codons. Affinity determined using single concentration BLI assays. All peptides synthesised with a C-terminal (GS)₂ spacer-flag-amide (GSGSDYKDDDDK-NH₂), N-terminal Acetyl, and cyclised via disulfide bond. GSGSEE sequence included for peptides with calculated isoelectric point (pI) of approximately 7. Purity determined by UPLC.

Name	Purity (in %)	Peptide sequence	K_D (1C) (nM)	K_D (3C) (nM)	K_D (4) (nM)	selected against
4.11	78.00	M C N V L L L Y K V L R H C	>10000	>10000	11	FGF-R4
4.17	65.14	M C N V L L L Y K G L R H C	>10000	>10000	>10000	FGF-R4
	54.80	M C H L L L Y Y K L F H S C	63	>10000	114	FGF-R4
4.9	77.24	M C T T L Y Y Y K L W R R C	699	2247	1	FGF-R4
	30.34	M C T Y V Y Y Y K L F H R C	>10000	>10000	3	FGF-R4
	19.95	M C R E I R Y W Q L Y R A F C	>10000	>10000	A.D.	FGF-R4
	42.58	M C R H L F Y W K L W H Q R C	>10000	>10000	2	FGF-R4
	47.25	M C R W V K Y Y Q L V H R K C	>10000	>10000	1	FGF-R4
	40.28	M C R E I R Y W Q L W Q C	>10000	>10000	27	FGF-R4
	44.06	M C Y N I Y W Y K Y V R R E C	387	>10000	1	FGF-R4
	59.54	M C H I R Y W Q L V R K H C	>10000	>10000	10	FGF-R4
	33.86	M C E V K Y F L I W K R W C	650	>10000	15	FGF-R4
	50.47	M C Y T V L W Y K Y Q K K C	>10000	>10000	1	FGF-R4
	49.15	M C P Y A I Y W Y K V R R K T C	>10000	>10000	11	FGF-R4
	33.38	M C E H L F Y Y L L V K R Q C	>10000	>10000	1	FGF-R4
	47.25	M C E D L F Y Y I J N R Y G C	>10000	>10000	18	FGF-R4
	41.72	M C S I L W Y K I T K P L C	>10000	>10000	13	FGF-R4
	38.42	M C E V I W Y K I Q W R K C	>10000	>10000	12	FGF-R4
4.10	47.42	M C K H L Y Y Y L L W K K C	1029	928	1	FGF-R4
	19.38	M C E D L F Y Y L L F H H C	4537	>10000	81	FGF-R4
	89.29	M C Y K H I L W Y K I Q K L C	33	>10000	1	FGF-R4
	42.02	M C Y E I I W Y K L Q R V C	>10000	>10000	A.D.	FGF-R4
	27.66	M C L W Y Y L Q F H R R E C	21	>10000	5	FGF-R4
	22.18	M C A V Y Y K L W L K H S R C	>10000	A.D.	7	FGF-R4
	48.94	M C K E Y E I L W Y K L Q Y C	>10000	>10000	17	FGF-R4
	26.58	M C D L Y T V L W Y K I Q K E C	21	>10000	19	FGF-R4
	36.14	M C A K Y F L I W R A Y Q Q C	>10000	>10000	10	FGF-R4
	12.76	M C A K A F L L W K G K F C	>10000	>10000	>10000	FGF-R4
	47.51	M C P I H F F K I V Y R H C	10	>10000	10	FGF-R4
	57.28	M C L W H Y Y K I F R R Q K C	96	>10000	8	FGF-R4
	38.86	M C E W F Y Y H K S W Y Y C	191	>10000	>10000	FGF-R4
	55.61	M C H R Y I Y L D Y W H C	>10000	>10000	44	FGF-R4
	63.49	M C L R H L Y E P Y W H C	>10000	>10000	>10000	FGF-R4
	70.25	M C Y K N I S T G I F L Y C	>10000	>10000	A.D.	FGF-R4
	65.58	M C Y K N I D N G V L V F C	>10000	>10000	14	FGF-R4
	36.61	M C T Y D F V L G H Y L I C	>10000	>10000	>10000	FGF-R4
	58.48	M C T Y E F L F E E L Y L H C	>10000	>10000	>10000	FGF-R4
	62.65	M C Y Y N F L N K V L V Y C	>10000	>10000	8	FGF-R4
	49.84	M C F L D H K T K V F V Y C	>10000	>10000	130	FGF-R4
	55.20	M C D S L R A Q Y A R W W J C	>10000	>10000	183	FGF-R4
4.12	59.42	M C E S D A Y R W K L W W V C	1567	>10000	10	FGF-R4
	46.75	M C Y D L V F D Y I T V Y H C	>10000	>10000	21	FGF-R4
	39.17	M C V W Q Y S H H Y R N W H C	>10000	>10000	>10000	FGF-R4
	29.62	M C V W K Y R L D R N W Y C	78	>10000	2043	FGF-R4
	34.21	M C W W F K Q H E F W F K V C	>10000	>10000	2	FGF-R4
	58.39	M C W Y D Q H S F Y W E R L C	>10000	>10000	>10000	FGF-R4
	34.92	M C K Y A T L Q W Y K W F Y C	114	7341	63	FGF-R4
	53.42	M C P Y D Y Y D W H D W R D C	>10000	>10000	1585	FGF-R4
	11.27	M C W S K L I I K Y Y L S Q C	>10000	>10000	16	FGF-R4
	100.00	M C N L S I J H G Y V A Y J C	>10000	>10000	1	FGF-R4
4.16	23.91	M C Y L D D F Y S Y R Y W A C	54	>10000	>10000	FGF-R4
	40.00	M C H I W Y F N N Q F H Y V C	>10000	A.D.	6	FGF-R4
	43.18	M C K I W Y V D H K F W W V C	>10000	>10000	2	FGF-R4
	31.04	M C V W W S H P S I W I C	65	>10000	>10000	FGF-R4
	38.56	M C H R H L Y Y W R L F H S C	>10000	>10000	5	FGF-R4
4.15	54.53	M C E R E V L Y W L L V T C	48	>10000	128	FGF-R4
	49.40	M C K R H I R Y W Q L Y V C	265	838	A.D.	FGF-R4
	21.16	M C K R E I K Y W L V F C	>10000	>10000	8	FGF-R4
	45.38	M C A K E H L F Y Y L L W C	>10000	>10000	41	FGF-R4
	68.57	M C K R H L I Y Y K I A R C	>10000	1540	1	FGF-R4
	40.68	M C I R Y L F Y Y L I S H C	>10000	>10000	2	FGF-R4
	36.54	M C P Y Y V Q Y W R V F C	>10000	>10000	9	FGF-R4
	52.99	M C T G R H L Y Y W K L F H C	>10000	>10000	10	FGF-R4
	47.51	M C D A Y Y V R Y W Q V L H C	>10000	>10000	36	FGF-R4
	34.45	M C K Y S K D L W L Y R L W C	>10000	>10000	2	FGF-R4
	38.33	M C Y K Y S K D L W L Y R L W C	502	>10000	22	FGF-R4
	59.90	M C I K D L H L F Y Y L L C	30	>10000	14	FGF-R4
	46.62	M C D K F V K W Y I V F R H C	>10000	>10000	1	FGF-R4
	54.15	M C V K Y I K Y W S V F H C	>10000	>10000	17	FGF-R4
	34.72	M C D W Y V T I Y K V L K K C	>10000	>10000	8	FGF-R4
	28.19	M C Y E D E I W L Y K V L Y C	>10000	>10000	9	FGF-R4
	50.13	M C Y H R H I F Q N H W Y K C	A.D.	A.D.	A.D.	FGF-R4
4.14	71.65	M C Y H R P P N S K I W R K C	>10000	>10000	214	FGF-R4
	69.02	M C E Q R S R W S N Y W Y P C	>10000	>10000	>10000	FGF-R4
	29.58	M C Y T R F W T E W W Q P C	>10000	>10000	>10000	FGF-R4
	64.78	M C Y T R H P Y E A L W H Q C	>10000	>10000	>10000	FGF-R4
	73.30	M C T H R H P Y E T L W H Q C	>10000	>10000	>10000	FGF-R4
4.13	62.39	M C Y T R F Y N W E K W Q L C	17	>10000	>10000	FGF-R4
	61.82	M C I T R K N K W E E W R P C	>10000	>10000	>10000	FGF-R4
	58.43	M C E H R H P Y G H W E Q C	>10000	>10000	>10000	FGF-R4

Table S 5 Single concentration K_D values of crude synthesised peptides against different FGF-R ECDs and corresponding FGF-R selected against during mRNA display from T18 selection 2. Peptides shown were identified from trimer-18 selection 2 using a 1:18:324 library of 10-12mer variable codons. Affinity determined using single concentration BLI assays. All peptides synthesised with a C-terminal (GS)₂ spacer-flag-amide (GSGSDYKDDDDK-NH₂), N-terminal Acetyl, and cyclised via disulfide bond. GSGEE sequence included for peptides with calculated isoelectric point (pI) of approximately 7. Purity determined by UPLC.

Name	Purity (in %)	Peptide sequence	K_D (1C) (nM)	K_D (3C) (nM)	K_D (4) (nM)	selected against
28.05	M	C F H I K A N Y G F V R V C	>10000	>10000	>10000	FGF-R4
31.99	M	C L H I K A N Y G F V R V C	>10000	>10000	>10000	FGF-R4
24.56	M	C T Y T Y V F N Q T Y V H C	>10000	A.D.	>10000	FGF-R4
29.92	M	C V Y E Y V F D R L T V T C	>10000	>10000	>10000	FGF-R4
28.18	M	C E Y T Y L F G Y H Y I K C	>10000	>10000	>10000	FGF-R4
15.76	M	C A Y E Y V F D R L T V T C	A.D.	>10000	28	FGF-R4
30.02	M	C H N K Y L Y Y Y L L Y S C	>10000	>10000	36	FGF-R4
35.72	M	C Y K R D L Y Y Y L V F K C	>10000	A.D.	8	FGF-R4
26.24	M	C D K R H L Y Y Y L L V C C	>10000	>10000	32	FGF-R4
22.3	M	C D K Y W V K Y W I A W R C	A.D.	>10000	1	FGF-R4
28.38	M	C E R E F L Y Y Y L Y Y K C	>10000	A.D.	>10000	FGF-R4
17.18	M	C W E E F T K F G W T I C	>10000	>10000	1	FGF-R4
17.17	M	C F K Q H L F Y Y L L H C	>10000	>10000	22	FGF-R4
27.72	M	C Y N K F I T L Y K V L Y C	505	79	13	FGF-R4
25.37	M	C H D K F I T L Y K V L Y C	>10000	A.D.	48	FGF-R4
27.34	M	C Y Y E V L W Y Q Q Q R H C	>10000	>10000	12	FGF-R4
30.15	M	C R D F D J I W Y K L Q F C	5364	340	20	FGF-R4
42.08	M	C K Y S V Y W Y K W Q R K C	>10000	A.D.	A.D.	FGF-R4
16.57	M	C E F E I R W Y Q Y Q R K C	>10000	>10000	16	FGF-R4
18.94	M	C W E E I T W Y K L Q H K C	A.D.	>10000	7	FGF-R4
18.42	M	C D N R R W V R Y Y Q I F C	>10000	>10000	11	FGF-R4
27.69	M	C E N K K V L F Y Y L L W C	>10000	>10000	<1	FGF-R4
19.7	M	C D L K R E L F Y Y L L V C	>10000	A.D.	27	FGF-R4
32.31	M	C D H K R H L Y Y Y L I W C	>10000	>10000	5	FGF-R4
30.31	M	C K F D L D L L F Y K I W C	>10000	>10000	9	FGF-R4
30.35	M	C H F K R D Y Y W Y L W Q C	>10000	>10000	73	FGF-R4
27.43	M	C K Y H R D L L Y Y F L H C	>10000	>10000	23	FGF-R4
7.25	M	C V H K K W I R Y W Q I F G S G S E E	A.D.	>10000	6	FGF-R4
21.06	M	C F W H K R S D F W L Q F C	A.D.	>10000	367	FGF-R4
36.05	M	C W W F R N H K F W Q Q C	>10000	>10000	9	FGF-R4
24.34	M	C F W K K L H V F W V K D C	23	>10000	A.D.	FGF-R4
26.53	M	C W Y V Q H S F Y W E R L C	>10000	A.D.	33	FGF-R4
41.92	M	C W W Q H P L F W S K I C	A.D.	>10000	22	FGF-R4
33.09	M	C F R H K R S D F W L Q F C	>10000	>10000	>10000	FGF-R4
38.29	M	C E H R H P Y S N F W I L C	>10000	>10000	>10000	FGF-R4
48.58	M	C E H R H P Y S N F W I L C	>10000	>10000	>10000	FGF-R4
53.91	M	C E H R H P Y E A F W H E C	>10000	>10000	>10000	FGF-R4
63.29	M	C V H R H R F E P E W H V C	>10000	>10000	>10000	FGF-R4
45.07	M	C A Y R H I Y E P Y W H S C	>10000	>10000	>10000	FGF-R4
42.83	M	C E H K H P F E R V W R S C	>10000	>10000	>10000	FGF-R4
44.26	M	C E Y R H K F E R W H V C	>10000	>10000	>10000	FGF-R4
54.81	M	C E H R H P Y S N F W I P C	>10000	>10000	>10000	FGF-R4
33.02	M	C Y T R H P F Q K R W H Y C	>10000	>10000	>10000	FGF-R4
49.75	M	C Y T K A Y W H E P W R R P C	>10000	3148	>10000	FGF-R4
29.41	M	C Y T R K P F E F W R R L C	>10000	A.D.	168	FGF-R4
37.7	M	C Y T R H K F D H R W R A C	>10000	>10000	>10000	FGF-R4
34.3	M	C Y T R D R Y E L H W R L C	>10000	>10000	>10000	FGF-R4
52.97	M	C Y T R H K F D Q Y W Q P C	>10000	>10000	>10000	FGF-R4
21.14	M	C I T R H R F E Q F W H L C	>10000	>10000	>10000	FGF-R4
27.67	M	C Y T K A T V F S Q W V L C	>10000	>10000	>10000	FGF-R4
54.73	M	C Y T R R F K W Q D W H A C	>10000	A.D.	>10000	FGF-R4
33.84	M	C Y T K Y F Y F E P W R I C	>10000	>10000	>10000	FGF-R4
40.25	M	C T T R K N K W E E W R P C	>10000	>10000	>10000	FGF-R4
38.6	M	C Y T K S F Y W P I W R S C	>10000	A.D.	>10000	FGF-R4
41.33	M	C Y T R L N K H A K W Q K C	>10000	A.D.	149	FGF-R4
27.09	M	C T D L Y Y Y L W Q R Y C	>10000	>10000	10	FGF-R4
19.04	M	C N D L F Y Y L L W R K A C	A.D.	>10000	1	FGF-R4
52.09	M	C K H L F Y Y K L I R K A C	>10000	>10000	7	FGF-R4
22.55	M	C V D L F Y Y K L R K L C	>10000	>10000	9	FGF-R4
42.02	M	C P Y T K Y W I V W R H S C	>10000	>10000	3	FGF-R4
42.61	M	C E H L Y Y Y L S H R R C	172	>10000	7	FGF-R4
24.68	M	C R E L F Y Y R F W Q K T C	>10000	>10000	14	FGF-R4
45.99	M	C R H L A Y W K L W H R K C	>10000	A.D.	8	FGF-R4
12.32	M	C D I V I Y K I L R S L C	>10000	3	4	FGF-R4
21.21	M	C D L F Y Y L Q H K K V C	>10000	>10000	8	FGF-R4
18.41	M	C Y V Q Y Y R V W R S I C	>10000	>10000	4	FGF-R4
25.46	M	C E V T W Y K I Q Y R L C	>10000	A.D.	21	FGF-R4
17.33	M	C D L F Y Y L F I K R R K C	>10000	A.D.	10	FGF-R4
19.76	M	C H L W Y Y L V H R R R D C	>10000	A.D.	6	FGF-R4
11.65	M	C V I V I Y K I L R S L C	>10000	A.D.	>10000	FGF-R4
60.49	M	C Y N V H W Y R W I R R R C	>10000	>10000	>10000	FGF-R4
19.23	M	C Y E V L W Y K L Q R K H C	A.D.	>10000	3	FGF-R4
13.76	M	C Y T E L W Y K Y I R R H C	1120	>10000	6	FGF-R4
15.39	M	C Y D V L W Y R Y R Y K Q C	>10000	>10000	22	FGF-R4
24.43	M	C K I I W L Y K A A R R F C	>10000	A.D.	7	FGF-R4
15.96	M	C Y E V L W Y Q Y W R K V C	>10000	>10000	22	FGF-R4
24.32	M	C F N V H W Y R W I R R R C	>10000	A.D.	3	FGF-R4
26.09	M	C N K H L Y Y L L F K H C	>10000	A.D.	9	FGF-R4
25.95	M	C P R Y V L Y Y K L F H H C	>10000	>10000	8	FGF-R4
13.54	M	C K R H I T Y W K L W H F G S G S E E	A.D.	A.D.	10	FGF-R4
16.46	M	C H D Y I V F Y K V L R H C	>10000	>10000	17	FGF-R4
39.1	M	C R R E L Y Y L L Y S K C	>10000	>10000	5	FGF-R4
32.24	M	C D R L I L Y Y R I V H K C	>10000	>10000	22	FGF-R4
28.11	M	C P Y E I K Y W I V W Q Q C	>10000	>10000	14	FGF-R4
22.94	M	C N K H L Y Y Y L L F K Q C	>10000	>10000	<1	FGF-R4

Trimer-18 Purified peptide BLI binding data

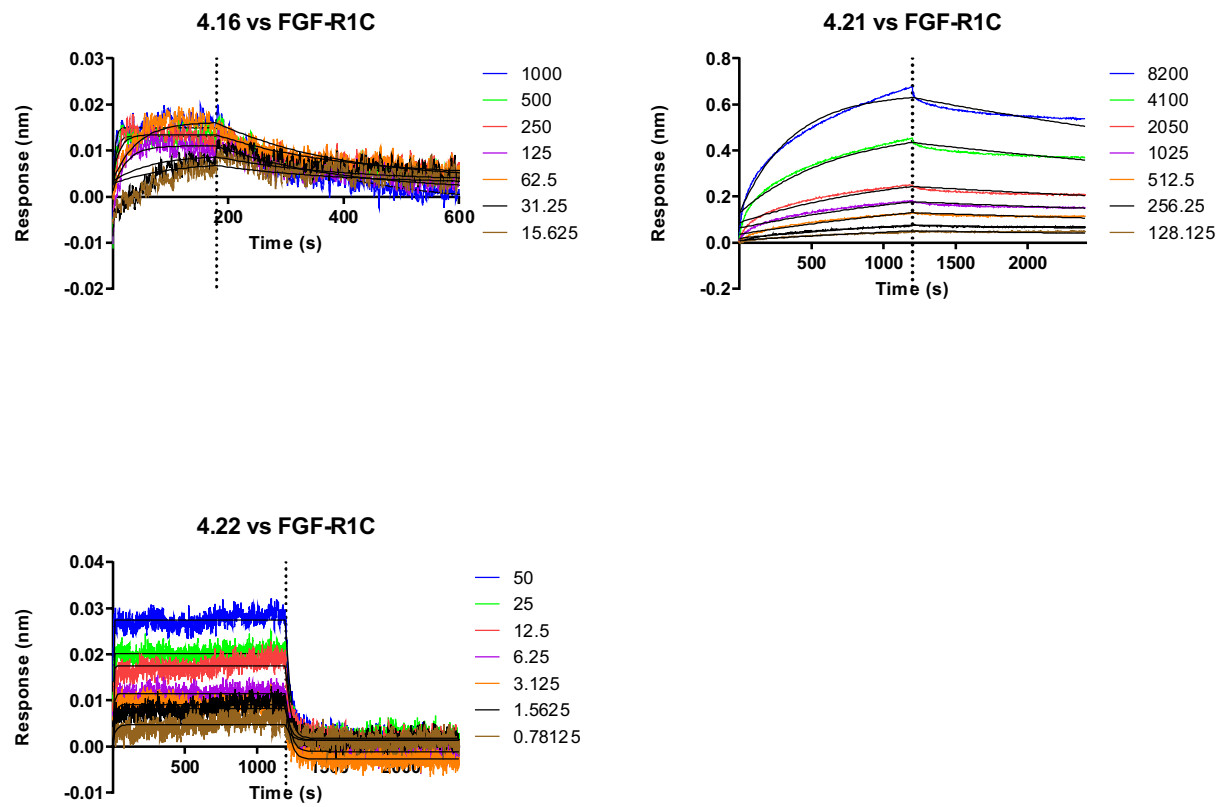


Figure S 22 Purified trimer-18 CP multi-concentration association and dissociation kinetics curves binding to FGF-R1C ECD. Data generated by Sartorius Octet RH16. Concentrations shown in nM.

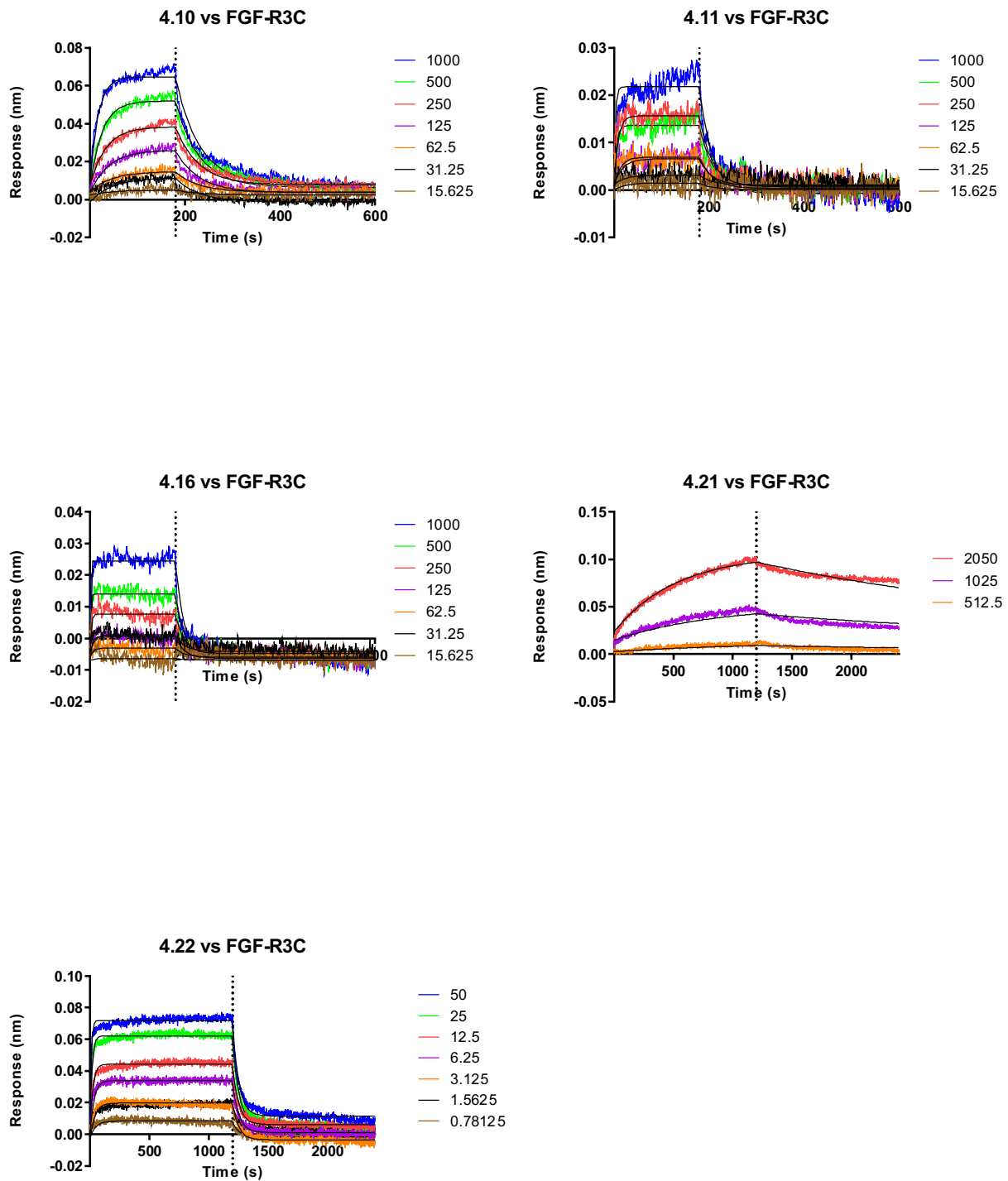
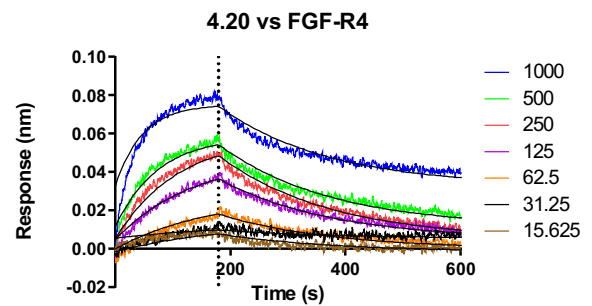
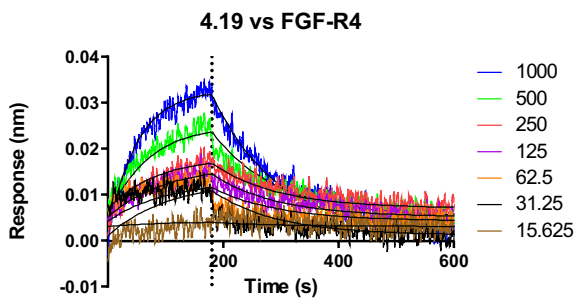
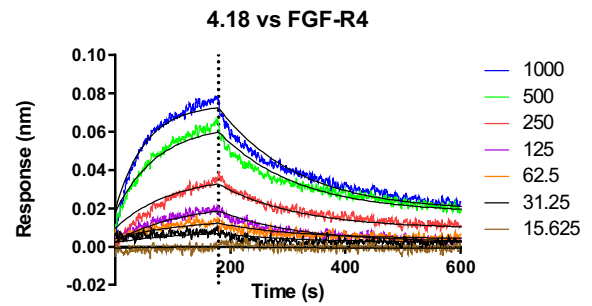
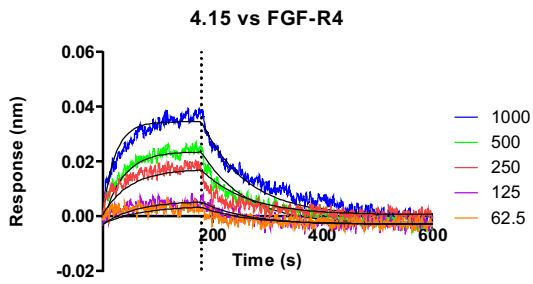
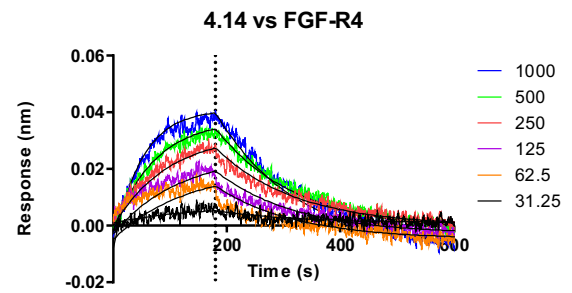
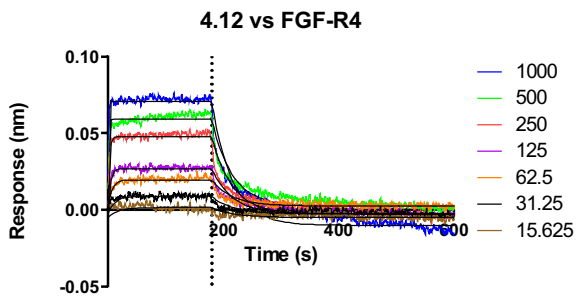
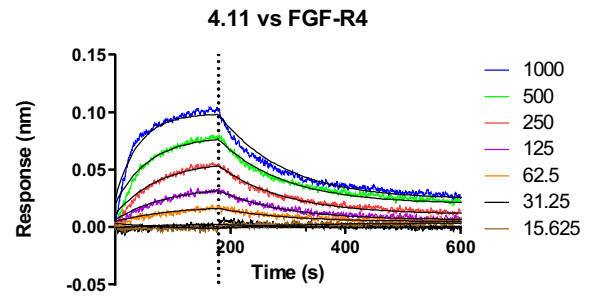
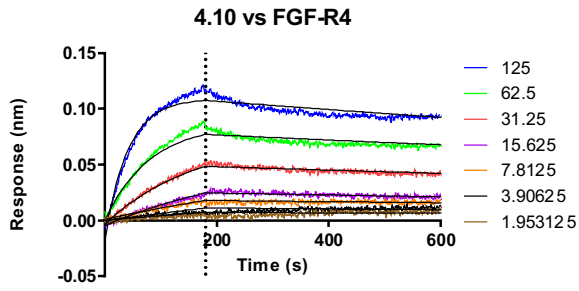


Figure S 23 Purified trimer-18 CP multi-concentration association and dissociation kinetics curves binding to FGF-R3C ECD. Data generated by Sartorius Octet RH16. Concentrations shown in nM



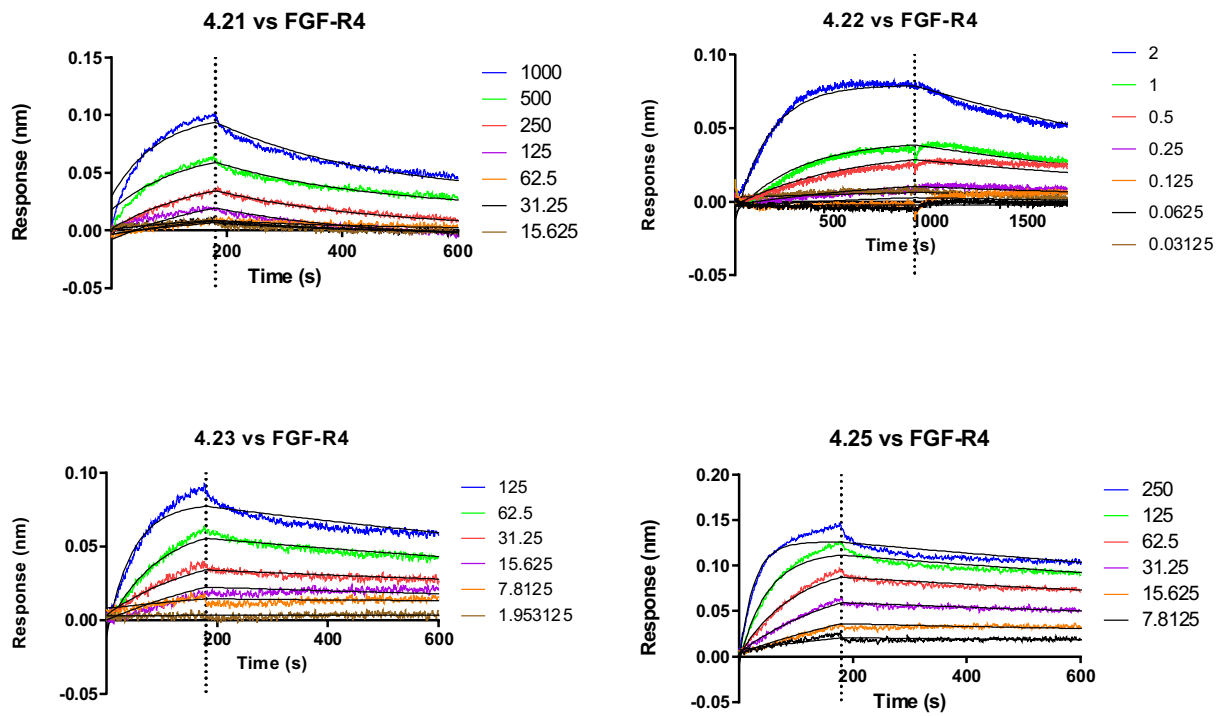


Figure S 24 Purified trimer-18 CP multi-concentration association and dissociation kinetics curves binding to FGF-R4 ECD. Data generated by Sartorius Octet RH16. Concentrations shown in nM

Alphascreen binding assay

The competitive binding assay was carried out according to the methods detailed in Hansen. et al using AlphaScreen.³ All reagents and samples were prepared in assay buffer: 40mM Hepes, 140mM NaCl, 1mM CaCl₂, 0,05% Tween20, pH=7.5, 0.1% BSA. In brief, recombinant human beta-klotho (R&D Systems) was preincubated for 60 min with the ectodomain, either Fc-fused human FGF-R1alpha(IIIc), FGF-R3alpha(IIIc) or FGF-R4 (R&D Systems), and the mixture was added to the test peptides in OptiPlate™ (PerkinElmer). Biotinylated FGF21 and heparin were further added to each well. AlphaScreen bead mix (1:1 mix of Protein A acceptor beads/ streptavidin-coated donor beads, PerkinElmer) were added to each well incubated overnight at RT in the dark, and measured on a plate reader (Envision). Assays were carried out in triplicates with concentrations ranging from 3, 10, 30, 100, 300 ,1000 and 3000 nM of peptide. No peptide control (media only) was used as a negative control, and unlabeled FGF21 and FGF1 as positive control. Data analysis was performed in Prism (GraphPad Software, Ver. 7.05, Figure S 25).

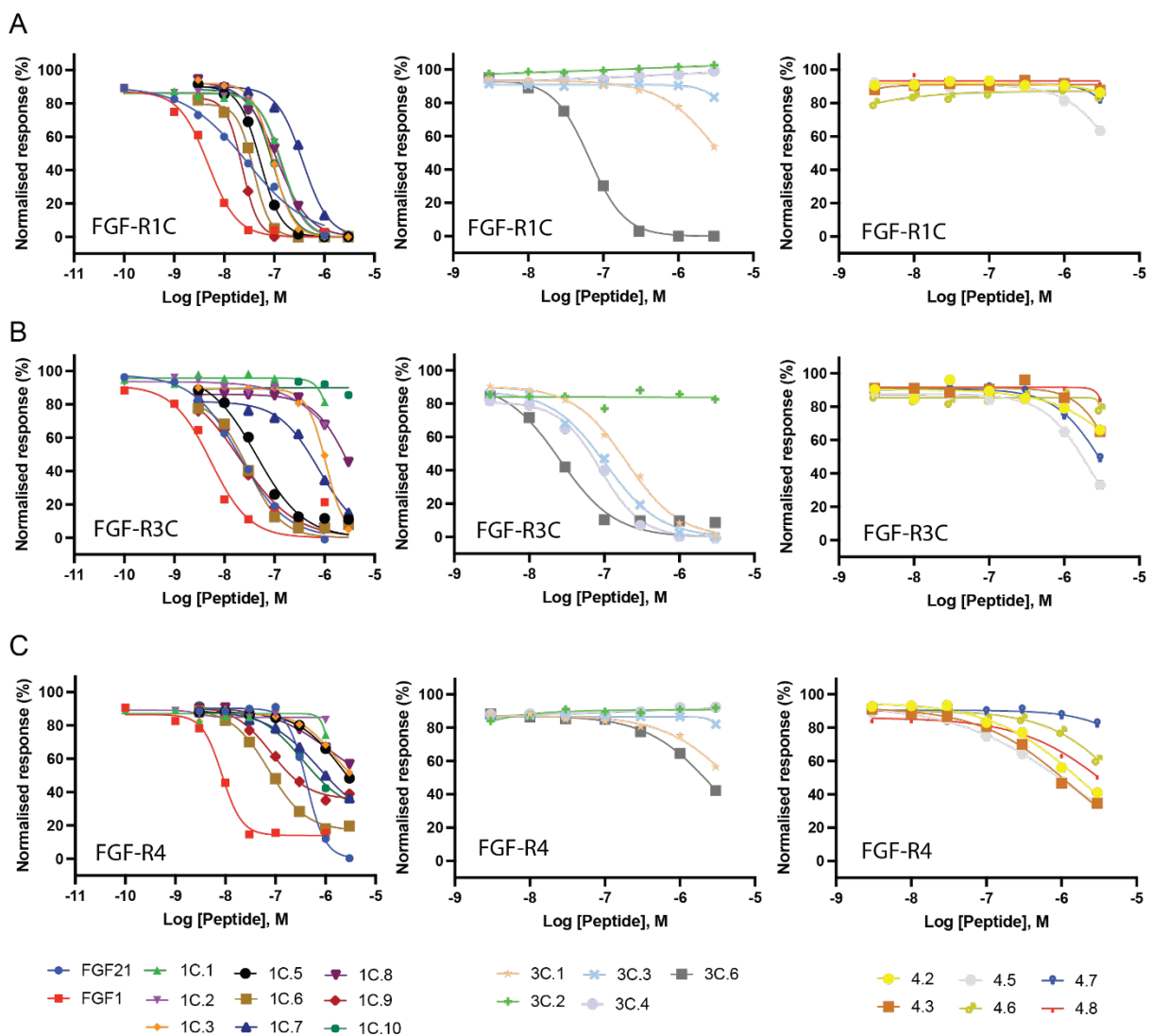


Figure S 25 Alpha screen displacement curves with purified peptides against recombinant human FGF-R Fc chimera proteins competing with biotinylated natural FGF21 ligand. A) Data against FGF-R alpha (IIIc) Fc chimera FGF-R1C, B) data against FGF-R3 (IIIc) Fc chimera FGF-R3C, C) data against FGF-R4 FGF-R4 Fc chimera proteins. IC₅₀ values for controls: unlabelled FGF21 (IC₅₀ = 23nM (FGF-R1C), 20nM (FGF-R3), 674 nM (FGFR-4), FGF1 (IC₅₀ = 4 nM (FGF-R1C), 3 nM (FGF-R3C), 7 nM (FGF-R4).

Statistical analysis of peptide fitness for affinity prediction

Biochemical selectivity prediction: From the NGS data of mRNA display cross screens (detailed above), fitness values were calculated using an in-house python script to assess the relative enrichment of the peptides at the last round, based on the log-ratio of normalised reads counts of the cross screen round five selection as a fraction of normalised reads of the previous round selection against the original FGF-R target:

$$F_i = \log \frac{c_i^R N^{R-1}}{N^R c_i^{R-1}}$$

Where F_i represents the fitness of peptides i , c_i^R refers to its reads counts and N^R the total number of peptides reads at round R (or $R-1$). The fitness values have been used as proxy for affinity prediction of the peptides.

Low fitness values (<0.1) show de-enrichment of the sequence against the new FGF-R target in the cross screen so was used to predict low affinity (>1000 nM K_D). High fitness values (>0.5) show high sequence enrichment against the new FGF-R target in the cross screen so was used to predict high affinity (<1000 nM K_D). Fitness values lying between 0.1 and 0.5 were assumed to be ambiguous and were not used to predict biochemical results. Fitness values were calculated for the purified NNK peptides for which both K_D (multi-concentration BLI) and IC_{50} (Alpha screen) data were present (24 peptides total). Fitness values were also calculated for original selections against the same target throughout all 5 rounds. These fitness values were then compared with the biochemical (K_D) data to evaluate success rate of fitness binding prediction. Success was obtained where fitness prediction matched biochemical data for each peptide (Fit <0.1 and K_D >1000 nM, or Fit >0.5 and K_D <1000 nM) and failure where fitness prediction did not match biochemical data for each peptide (Fit <0.1 and K_D <1000 nM, or Fit > 0.5 and K_D >1000 nM). Individual fitness value comparison for NNK purified peptides shown in heatmap (Figure S 26) made using Prism (GraphPad Software, Ver. 7.05). Overall results summarised as pie chart and violin plot using Prism (GraphPad Software, Ver. 7.05).

3C.3 SAR binding prediction: To predict SAR for peptide 3C.3 the relative abundance of each amino acid at each residue position was calculated using an in-house python script for all sequences present in the same cluster as 3C.3 (selection against FGF-R3C, total 3284 sequences). At each residue position, the occurrence of each amino acid was calculated relative to the fitness value (as above) of the sequence it is found in. A high relative occurrence (>200) was used to predict higher affinity binding to FGF-R3C, since this amino acid at this position was enriched more within the cluster, whereas a low relative occurrence (<200) was used to predict a lower affinity due to lower enrichment. A heatmap was then constructed of all values using Prism (GraphPad Software, Ver. 7.05).

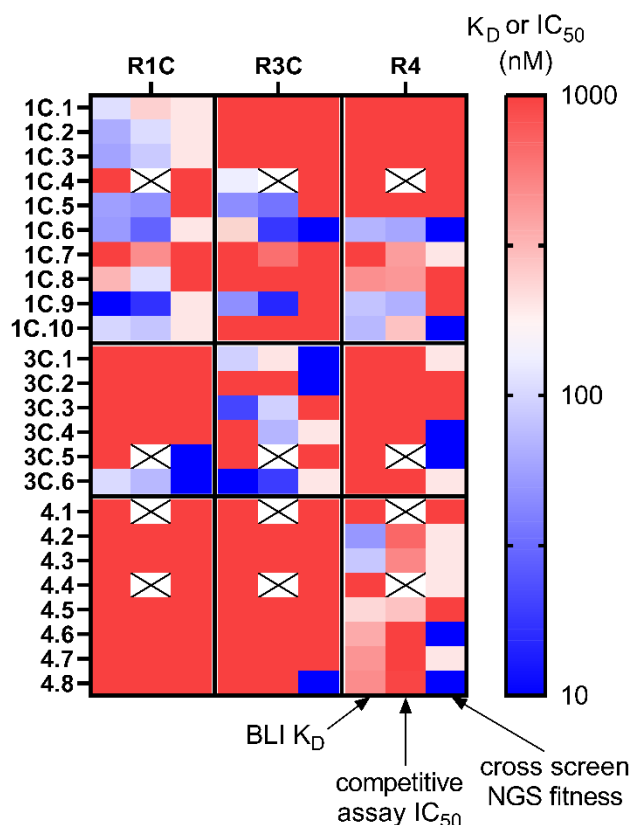


Figure S 26 heatmap showing cross screen fitness value prediction against individual peptide sequences. Columns are divided into three groups depending on whether data is for FGF-R1C, FGF-R3C, or FGF-R4. Rows depict individual sequences numbered according to the FGF-R target they were selected against. For each receptor, there are 3 sub-columns depicting multi-concentration BLI K_D (left), alpha screen IC_{50} (middle), and cross screen fitness binding prediction (right). K_D and IC_{50} affinity obtained using purified peptides and displayed on a log scale of high affinity (blue) to low affinity (red). Crosses denote where no IC_{50} data was determined. Cross screen NGS fitness values calculated as normalised read counts in cross screen round 5 as a fraction of normalised read counts in previous round. Fitness value >0.5 (Blue) was taken to predict binding to the receptor, fitness value <0.1 (red) was taken to predict poor binding to the receptor, and fitness between 0.1 and 0.5 was considered an ambiguous result and is shown in white. Where the colour of K_D/IC_{50} matches with the fitness prediction, then the prediction was considered successful.

Supporting datasets

Next generation sequencing: a complete list of peptide hits from each mRNA display campaign performed, including the number of reads for each round. NNK selections, trimer-18 selections and cross screens are included.

High throughput peptide synthesis compound identity: a complete list of the exact masses calculated and found for all peptides synthesised in 96 well plate format for high throughput BLI screening. All data obtained via LC-MS (Waters Acquity system with QDA detector, KDA3351, and TUV detector, F08UPT321M) using Waters Acquity UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 mm x 50 mm). A 3 min linear gradient of 10% to 90% solvent B was run; solvent A: 0.05% TFA in water, solvent B: 0.05% TFA in MeCN.

Peptide K_D : this dataset shows the K_D value of every peptide against each of the three receptors (FGF-R1C, FGF-R3C, and FGF-R4) including the purity state, origin of the peptide, target originally selected against, and type of library used during selection.

Statistical 3C.3 SAR prediction input dataset: contains all peptide sequences, along with NGS read count for each round in the NNK selection against FGF-R3C ECD belonging to the same cluster as peptide 3C.3. This dataset was used as the basis for SAR prediction heatmaps.

References

1. A. Kawamura, M. Munzel, T. Kojima, C. Yapp, B. Bhushan, Y. Goto, A. Tumber, T. Katoh, O. N. King, T. Passioura, L. J. Walport, S. B. Hatch, S. Madden, S. Muller, P. E. Brennan, R. Chowdhury, R. J. Hopkinson, H. Suga and C. J. Schofield, *Nat Commun*, 2017, **8**, 14773.
2. B. Bhushan, D. Granata, C. S. Kaas, M. A. Kasimova, Q. Ren, C. N. Cramer, M. D. White, A. M. K. Hansen, C. Fledelius, G. Invernizzi, K. Deibler, O. D. Coleman, X. Zhao, X. Qu, H. Liu, S. S. Zurmuhl, J. T. Kodra, A. Kawamura and M. Munzel, *Chem Sci*, 2022, **13**, 3256-3262.
3. A. M. K. Hansen, S. G. Vienberg, K. Lykkegaard, X. Zhao, G. Tingqing, D. Han, X. Zhang, H. Thogersen, K. Sass-Orum, T. Tagmose, K. Raun and B. Andersen, *Biochem J*, 2018, **475**, 2985-2996.