This is a repository copy of *Experimental Validation of Plant Peroxisomal Targeting Prediction Algorithms by Systematic Comparison of In Vivo Import Efficiency and In Vitro PTS1 Binding Affinity*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/84173/

Version: Supplemental Material

**Experimental validation of plant peroxisomal targeting prediction algorithms by systematic comparison of *in vivo* import efficiency and *in vitro* PTS1 binding affinity**

Nicola S. Skoulding[#1], Gopal Chowdhary[#2,3], Mara J. Deus[2], Alison Baker[4], Sigrun Reumann[2,5], and Stuart L. Warriner[1]

[#]These authors contributed equally to this work.

[1] - School of Chemistry and the Astbury Centre, University of Leeds, Leeds, LS2 9JT, UK.

[2] - Centre for Organelle Research (CORE), Faculty of Science and Technology, University of Stavanger, Richard Johnsons gate 4, N-4021 Stavanger, Norway.

[3] - KIIT School of Biotechnology, Campus XI, KIIT University, I-751024 Bhubaneswar, India.

[4] – Centre for Plant Sciences, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK.

[5] - Department of Biology, Biocentre Klein Flottbek, University of Hamburg, D-22609 Hamburg, Germany.

Correspondence to Stuart Warriner, s.l.warriner@leeds.ac.uk, +44 113 343 6437.
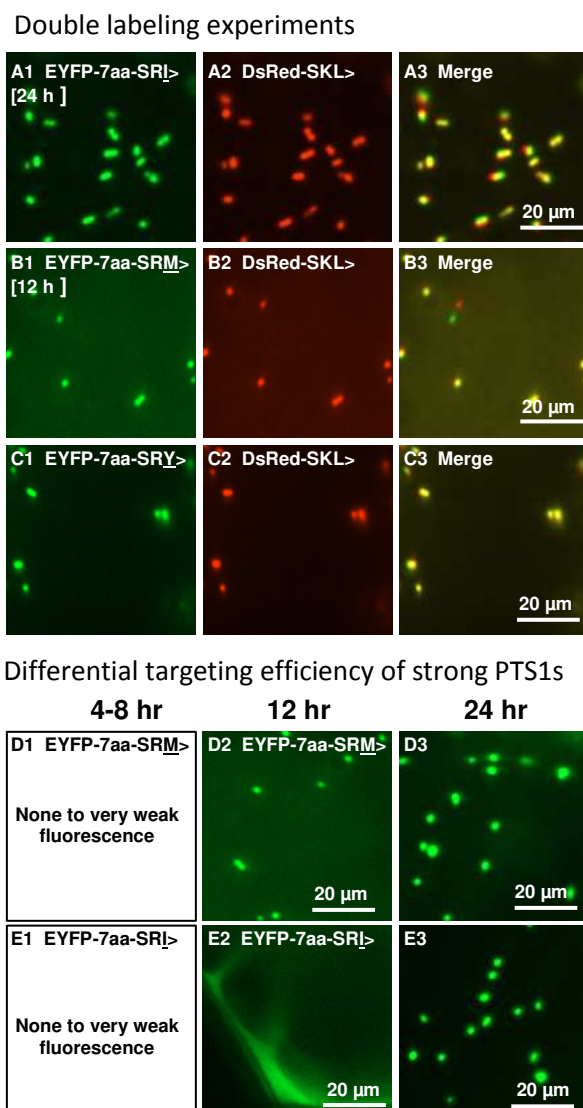
## Supporting Information

## Contents

# 1   Supplementary Figures

Double labeling experiments



Differential targeting efficiency of strong PTS1s



## Supplementary Figure 1

**Verification of peroxisome targeting for selected point mutated reporter fusions (a-c) and analysis of peroxisome targeting at very early time points p.t. for two canonical strong PTS1s, SRM> (d) and SRI> (e).** Onion epidermal cell were biolistically transformed with EYFP fusion constructs that were C-terminally extended by the decapeptide PTS1 domain of a model sequence, acyl-CoA oxidase isoform 4 from *Zinnia elegans* (ZeACX4, VAKTTRP-SRV>) or various mutant versions. Subcellular targeting was analyzed by fluorescence microscopy at different time intervals post transformation (p.t.). The mutated aa are underlined. Peroxisome targeting was verified by colocalization of EYFP fluorescence with the peroxisomal marker DsRed-SKL for the original model sequence (see Fig. 1b) and the PTS1 tripeptide point mutations from SRV> to SRI> (a), to SRM> (b) and to SRY> (c). To investigate putative peroxisome targeting efficiencies for the two canonical strong PTS1 domains terminating with SRM> and SRI> and document the efficiency of peroxisome targeting, peroxisome targeting was investigated at very early time points p.t. (4 h, 8 h, 12 h and 24 h). The EYFP images were not modified for brightness or contrast in single transformants (d and e).
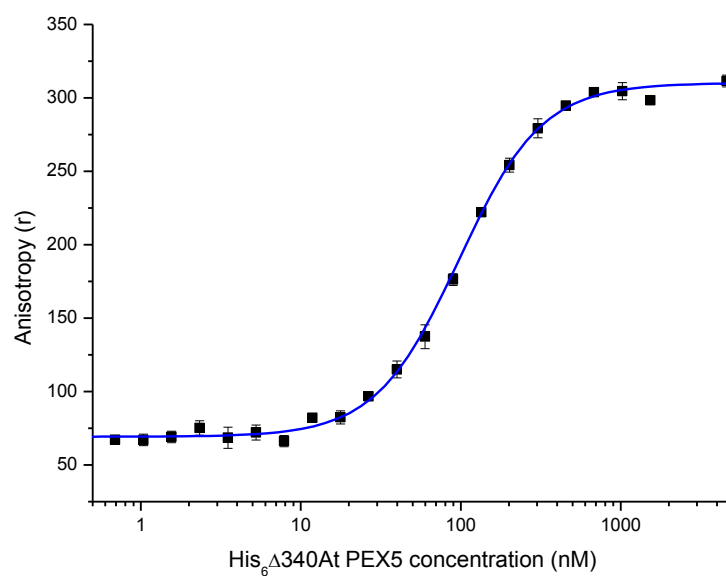
**Supplementary Figure 2.** Determination of the maximum and minimum anisotropy values by titration of protein (PEX5C) into a fixed concentration of fluorescent tracer (lissamine labelled YQSKL, 100 nM). The anisotropy was measured using the methods detailed in Section 3.3 and the curve is a fit to Equation 3.
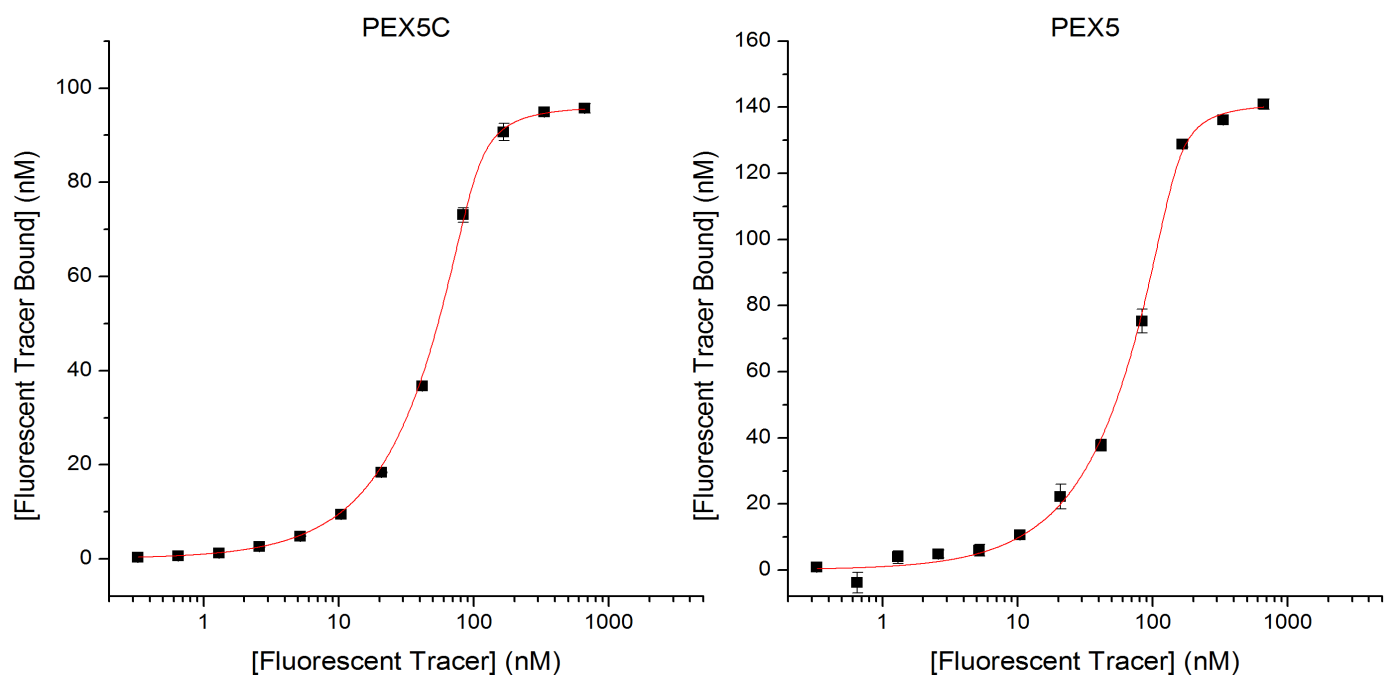
**Supplementary Figure 3.** Determination of $K_d$ of the Tracer Peptide for PEX5C (left) and PEX5 (right). Fluorescent tracer (lissamine labelled YQSKL) was titrated into a fixed concentration of protein (~100nM). The anisotropy was measured and converted using appropriate equations to give the concentration of bound tracer (Section 3.4). The curves are the result of non-linear least squares fitting for the data to Equation 5. PEX5C: $K_d$ = 3.98 ± 0.53 nM, $P_T$ = 96.2 ± 0.71 nM, PEX5: $K_d$ = 4.52 ± 1.2 nM, $P_T$ = 141.2 ± 2.0 nM.
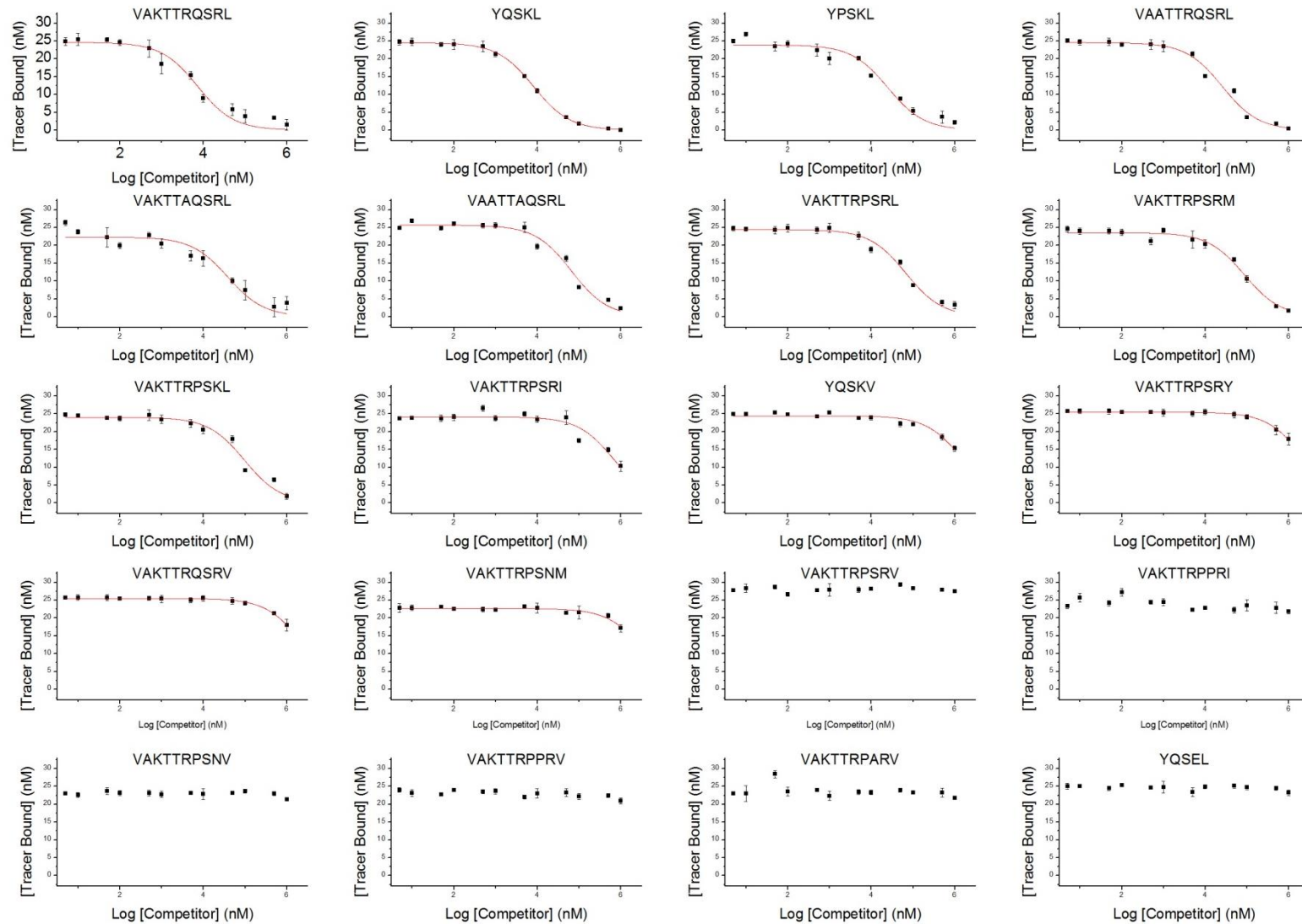
**Supplementary Figure 4.** Competition experiments to determine the affinity of peptides for PEX5C. The curves are fits to equation 6 and were used to extract $IC_{50}$ values which were then converted to $K_i$ (Section 3.5).

# PEX5
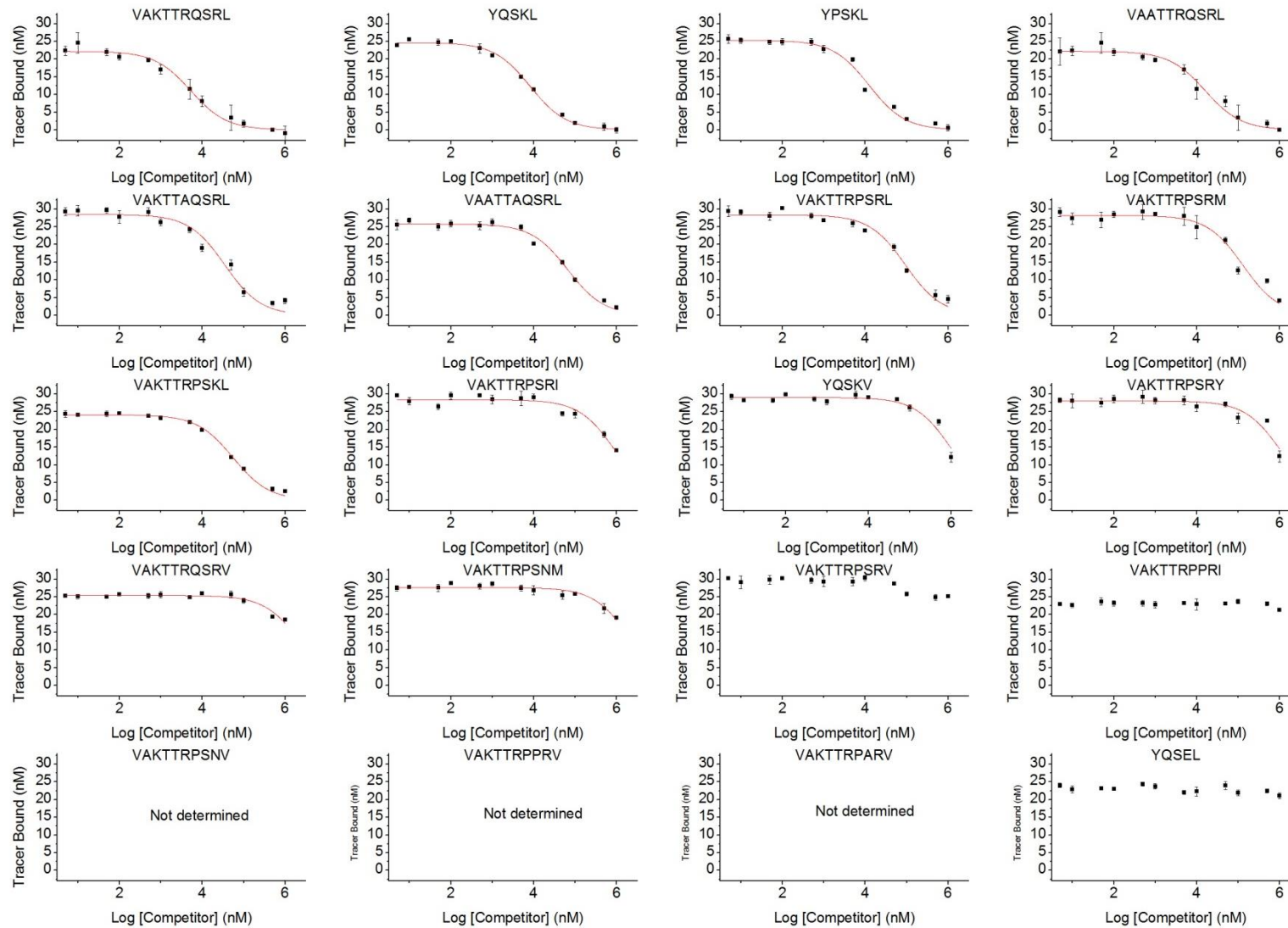


**Supplementary Figure 5** Competition experiments to determine the affinity of peptides for PEX5. The curves are fits to equation 6 and were used to extract $IC_{50}$ values which were then converted to $K_i$ (Section 3.5)

# 2  Supplementary Tables

## Supplementary Table I :PWM Scoring Matrix

In the PWM-based PTS1 protein prediction model, each of the 20 possible amino acid (aa) residues of the C-terminal 14-aa positions is assigned a position-specific score, which indicates whether a specific residue at a particular sequence position is predicted to enhance (more positive score) or reduce peroxisome targeting (more negative score) and to what extent.

| Res. | **Position** -14 | -13 | -12 | -11 | -10 | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | Res. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | -0,06 | -0,06 | -0,07 | 0,01 | -0,01 | -0,10 | -0,06 | -0,10 | -0,12 | -0,04 | -0,01 | 0,34 | -0,17 | -0,22 | A |
| R | -0,08 | -0,09 | 0,01 | -0,06 | -0,10 | -0,10 | -0,06 | -0,07 | 0,03 | -0,06 | 0,03 | -0,16 | 0,46 | -0,23 | R |
| N | -0,09 | -0,12 | -0,08 | -0,11 | -0,01 | -0,05 | -0,06 | -0,07 | -0,08 | -0,13 | -0,09 | -0,17 | 0,01 | -0,16 | N |
| D | -0,09 | -0,05 | -0,08 | 0,02 | -0,04 | -0,01 | -0,10 | -0,07 | -0,08 | -0,15 | -0,17 | -0,23 | -0,15 | -0,13 | D |
| C | -0,12 | -0,06 | -0,16 | -0,13 | -0,16 | -0,19 | -0,09 | -0,03 | -0,17 | -0,08 | -0,18 | 0,12 | -0,10 | -0,22 | C |
| Q | -0,06 | -0,09 | -0,07 | -0,02 | -0,10 | -0,06 | 0,00 | -0,04 | -0,02 | -0,07 | -0,04 | -0,14 | -0,06 | -0,18 | Q |
| E | -0,06 | -0,08 | -0,11 | -0,07 | -0,07 | -0,02 | -0,17 | 0,00 | -0,12 | -0,07 | -0,15 | -0,17 | -0,12 | -0,13 | E |
| G | -0,06 | -0,08 | -0,08 | -0,04 | -0,11 | -0,09 | -0,03 | -0,08 | -0,08 | -0,07 | -0,16 | -0,20 | -0,18 | -0,13 | G |
| H | -0,10 | -0,01 | -0,06 | -0,05 | -0,03 | 0,01 | -0,04 | -0,03 | -0,07 | -0,08 | 0,03 | -0,09 | -0,07 | -0,18 | H |
| I | -0,04 | -0,11 | -0,06 | -0,12 | -0,08 | -0,10 | -0,02 | -0,06 | -0,04 | 0,00 | -0,10 | -0,19 | -0,13 | 0,33 | I |
| L | -0,12 | -0,09 | -0,07 | -0,10 | -0,06 | -0,10 | -0,08 | -0,10 | -0,03 | -0,01 | -0,08 | -0,19 | -0,12 | 0,66 | L |
| K | -0,12 | -0,08 | -0,10 | -0,02 | -0,12 | 0,00 | -0,05 | -0,01 | -0,03 | -0,11 | 0,00 | -0,12 | 0,44 | -0,21 | K |
| M | -0,10 | -0,03 | 0,00 | -0,11 | -0,08 | -0,10 | -0,05 | -0,07 | 0,07 | -0,02 | -0,11 | -0,13 | -0,12 | 0,64 | M |
| F | -0,07 | -0,18 | -0,13 | -0,08 | -0,14 | -0,04 | -0,05 | -0,15 | -0,10 | -0,04 | -0,05 | -0,02 | -0,19 | -0,09 | F |
| P | -0,07 | -0,12 | -0,05 | -0,10 | 0,03 | -0,08 | -0,07 | 0,02 | -0,05 | 0,03 | 0,00 | 0,13 | -0,18 | -0,19 | P |
| S | -0,05 | -0,08 | -0,05 | -0,04 | -0,09 | -0,08 | 0,00 | -0,10 | -0,04 | -0,11 | -0,06 | 0,48 | -0,06 | -0,19 | S |
| T | -0,14 | -0,11 | -0,02 | -0,02 | 0,00 | -0,08 | -0,06 | -0,05 | -0,04 | -0,05 | -0,06 | -0,14 | -0,16 | -0,24 | T |
| W | 0,15 | 0,15 | 0,00 | 0,01 | -0,09 | -0,12 | -0,19 | -0,14 | -0,33 | -0,17 | -0,10 | -0,26 | -0,15 | -0,21 | W |
| Y | 0,01 | -0,03 | -0,14 | -0,28 | -0,11 | 0,02 | -0,07 | -0,12 | 0,00 | -0,09 | 0,01 | -0,13 | -0,12 | -0,16 | Y |
| V | -0,09 | -0,04 | -0,03 | -0,08 | 0,00 | -0,07 | -0,11 | -0,08 | -0,06 | -0,05 | -0,07 | -0,11 | -0,20 | -0,12 | V |
|  | -14 | -13 | -12 | -11 | -10 | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 |  |

## Supplementary Table II: Correlation analysis between PTS1 protein prediction scores and in vivo peroxisome targeting efficiency

Correlation analysis between PTS1 protein prediction scores and *in vivo* peroxisome targeting efficiency of reporter protein fusions with mutated PTS1 domains.

| Reporter protein fusion | PWM pred. score | Post. prob. (%) | Bal. post. prob. (%) | Perox. targ. predict.[1] | Perox. targ. efficiency | Reference |
|---|---|---|---|---|---|---|
| EYFP-VAKTTRP**SRM** | 0.969 | 100 | 100 | P (P) | P (very strong) | this study |
| EYFP-VAKTTRP**SRI** | 0.664 | 99.6 | 99.5 | P (P) | P (strong) | this study |
| EYFP-AKTTRP**SNM** | 0.523 | 93.3 | 97.8 | P (P) | P (mod.) | this study |
| EYFP-VAKTTRP**SRV** | 0.216 | 0.4 | 57.6 | C (P) | P (weak) | Lingner et al. (2011), this study |
| EYFP-VAKTTRP**SRY** | 0.173 | 0.1 | 45.5 | C (C) | P (mod.) | this study |
| EYFP-VAKTTRP**SRK** | 0.119 | 0.0 | 31.1 | C (C) | C | this study |
| EYFP-AGTTGG**SRV** | 0.073 | 0.0 | 21.2 | C (C) | P (weak) | this study |
| EYFP-VAKTTRD**SRV** | 0.045 | 0.0 | 16.4 | C (C) | C | this study |
| EYFP-VAETTDP**SRV** | 0.011 | 0.0 | 11.7 | C (C) | C | this study |
| EYFP-VAKTTRP**PRV** | -0.135 | 0.0 | 2.5 | C (C) | C | this study |
| EYFP-VAKTTRP**SNV** | -0.229 | 0.0 | 0.9 | C (C) | P (weak) | this study |
| EYFP-VAKTTRP**SNY** | -0.272 | 0.0 | 0.50 | C (C) | C | this study |
| EYFP-VAKTTRP**STV** | -0.401 | 0.0 | 0.1 | C (C) | C | this study |

Mutations introduced into the model decapeptide of ZeACX4 (VAKTTRP**SRV**, shaded in grey) are underlined, and the C-terminal tripeptides are printed bold. The constructs are sorted by PWM prediction score; C, cytosol; p, peroxisome; n.d., not determined. For additional information see Tables 1 & 2 in the manuscript.
[1] Peroxisome targeting prediction (column 5) is presented for the standard posterior probability and (in parenthesis) for the balanced posterior probability.

# Supplementary Table III: Statistical Analysis of the Prediction Models

P-Values for statistical analysis of the correlation between prediction of peroxisomal protein targeting and observed results from in vivo targeting experiments. (a) PWM Score (b) Posterior Probability (c) Balanced Posterior Probability.  In each case ANOVA analysis showed that the means of at least one of the predictive parameters were different between classes.  Post-Hoc two-tailed Student T-Test between the categories highlighted where the parameters were significantly different (highlighted).  The threshold was the 95% confidence band with the P-value for significance scaled using the Bonferroni correction.

(a) PWM Score

|            | P(strong) | P(mod) | P(weak) | C      |
|------------|-----------|--------|---------|--------|
| P(strong)  |           | 0.1810 | 0.0298  | 0.0015 |
| P(mod)     | 0.1810    |        | 0.2225  | 0.0381 |
| P(weak)    | 0.0298    | 0.2225 |         | 0.4239 |
| C          | 0.0015    | 0.0381 | 0.4239  |        |

(b) Posterior Probability

|            | P(strong)               | P(mod) | P(weak)                 | C                       |
|------------|-------------------------|--------|-------------------------|-------------------------|
| P(strong)  |                         | 0.3726 | $2.63\times10^{-08}$    | $4.80\times10^{-17}$    |
| P(mod)     | 0.3726                  |        | 0.2725                  | 0.0777                  |
| P(weak)    | $2.63\times10^{-08}$    | 0.2725 |                         | 0.1705                  |
| C          | $4.80\times10^{-17}$    | 0.0777 | 0.1705                  |                         |

(c) Balanced Posterior Probability

|            | P(strong) | P(mod) | P(weak) | C      |
|------------|-----------|--------|---------|--------|
| P(strong)  |           | 0.3950 | 0.0419  | 0.0001 |
| P(mod)     | 0.3950    |        | 0.2173  | 0.0070 |
| P(weak)    | 0.0419    | 0.2173 |         | 0.2548 |
| C          | 0.0001    | 0.0070 | 0.2548  |        |

## Supplementary Table IV: Oligonucleotide primers used for reporter gene cloning

One forward primer was used for EYFP amplification and introduced a 5'-NcoI site into the PCR products (5'-AAGT<u>CCATGG</u>TGAGCAAGGGCGAGGA-3'). In the reverse primers the XbaI sites are underlined. Mutations introduced into the model decapeptide of ZeACX4 (VAKTTRP**SRV**) are underlined, and the C-terminal tripeptides are printed bold.

| Construct | Primer acronym | Reverse primer sequence (5' to 3') | Reference |
|---|---|---|---|
| EYFP-VAKTTRP**SRV** | | TATA<u>TCTAGA</u>GTCAcacgcggctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | Lingner et al. (2011) |
| EYFP-VAKTTRP**SRI** | **SR580r** | TATG<u>TCTAGA</u>GTCAgatgcggctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**SRM** | **SR581r** | TATG<u>TCTAGA</u>GTCAcatgcggctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**SRY** | **GC60r** | TATG<u>TCTAGA</u>GTCAATAgcggctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**SRK** | **GC111r** | TATG<u>TCTAGA</u>GTCActtgcggctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**SNV** | **GC62r** | TATG<u>TCTAGA</u>GTCAcac<u>GTT</u>gctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**STV** | **GC112r** | TATG<u>TCTAGA</u>GTCAcacggtgctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**SNM** | **GC61r** | TATG<u>TCTAGA</u>GTCAcat<u>GTT</u>gctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**SNY** | **GC115r** | TATG<u>TCTAGA</u>GTCAATA<u>GTT</u>gctggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTTRP**PRV** | **GC113r** | TATG<u>TCTAGA</u>GTCAcacgcgtggggggcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VA<u>GTTGG</u>**SRV** | **GC64r** | TATG<u>TCTAGA</u>GTCAcacgcggcttcctccggtcgttcctgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VA<u>ETTD</u>P**SRV** | **GC116r** | TATG<u>TCTAGA</u>GTCAcacgcggctggggtcggtcgtttctgcaacCTTGTACAGCTCGTCCATGCC | This study |
| EYFP-VAKTT<u>RD</u>**SRV** | **GC118r** | TATG<u>TCTAGA</u>GTCAcacgcggctgtcgcgggtcgtctttgcaacCTTGTACAGCTCGTCCATGCC | This study |

# 3 Determination of Binding Constants using Fluorescence Anisotropy Measurements.

## 3.1 Measurement of Anisotropy.

Anisotropy is a fundamental molecular property which reflects the interaction of a fluorescent molecule with plane polarised light. Low values of anisotropy indicate that the molecule is tumbling rapidly relative to the fluorescence lifetime of the fluorophore whereas higher values of anisotropy indicate that the tumbling times are slow compared to this value. The fluorescence lifetimes of the fluorophores used in this study are typically 1-10 ns and hence low values of anisotropy are shown when the fluorophore is part of a small peptide molecule. However, when the fluorophore labelled peptide binds to the larger protein the tumbling rate is related to the overall complex size and consequently slows significantly with a resulting increase in anisotropy. Anisotropy is an intrinsic molecular property and should be used for the determination of binding constants (rather than the related polarization values). A more detailed discussion of the principles of fluorescence anisotropy are found in reference [1].

Fluorescence anisotropy measurements were carried out on an Envision™ 2103 multilabel plate reader (PerkinElmer) using the following optics:

|  | Manufacturer's description | Wavelength (nm) | Bandwidth (nm) |
|---|---|---|---|
| Dichroic Mirror | BODIPY TMR FP dichroic mirror | 555 | - |
| Excitation Filter | BODIPY TMR FP 531 | 531 | 25 |
| Emission Filter 1: | BODIPY TMR FP P-pol 595 | 595 | 60 |
| Emission Filter 2: | BODIPY TMR FP S-pol 595 | 595 | 60 |

Assays were performed using black 384 well plates (Optiplate, PerkinElmer) and the plates were incubated at 25 °C for 20 minutes with agitation (within the plate reader) prior to reading to ensure equilibration. The reading protocol was optimised (instrument gain, read height) using the instrument software to ensure that the detector was not saturated. 30 flashes were used per measurement.

The anisotropy (r) was calculated using Equation 1.

$$r = 1000 \times \frac{s - gp}{s + g2p} \qquad \textit{Equation 1}$$

The total fluorescence intensity (I) is defined by Equation 2.

$$I = s + g2p \qquad \textit{Equation 2}$$

Where s is the intensity of the light emitted with the same polarisation as the excitation light, p is the intensity of the light emitted with perpendicular polarisation and g is a grating factor which accounts for the different responsiveness of an instrument to light with differing polarisations. In these assays this was set to 1 as part of the optimisation process.

Blank corrections were performed by subtracting the average of the appropriate blank intensities in the s and p channels <u>prior</u> to calculation of the anisotropy. [2]

## 3.2   Preparation of Assay Plates

All assays were performed using an FA buffer (HEPES (20 mM), NaCl (150 mM), pH 7.5 with NaOH).

In order to minimise non-specific interactions the plates were pre-treated with a gelatine solution:[3]  All wells to be used were filled with 80 µl of FA buffer containing 0.32 mg/ml of gelatine solution using a Multidrop combi (ThermoElectron), sealed using self adhesive plate seals (StarLab) and stored at 4 $^{o}$C the evening prior to use.

The lissamine rhodamine-YQSKL peptide was stored as a stock solution at 500 µM in H$_2$O, stored at either 4 $^{o}$C as a working stock with dilutions made fresh for each plate or at -80 $^{o}$C for longer term storage. All stock solutions (protein, lissamine rhodamine-YQSKL and inhibitor) were stored at 4 $^{o}$C or kept on ice at all times.

## 3.3   Determination of limits of anisotropy

60 µl of gelatine buffer solution was removed from plate rows A-F, wells 1-24. 40 µl of protein stock solution (3x desired concentration) was added to well A1 and the contents of the well agitated with a pipette. 40 µl of solution was removed from A1 and placed in A2, then agitated and placed into A3. Following addition and agitation of column A24, 40 µl of solution was removed and discarded. The above process repeated across the entire series and for Rows B-F.  20 µl of 200 nM of lissamine rhodamine-YQSKL solution in FA Buffer (final concentration 100 nM) was added to all wells in the first three (A-C) rows. 20 µl of FA buffer was then added to the wells of the other three rows (D-F) to act as the blank. The anisotropy of each well was measured and the average intensities measured in wells containing no tracer peptide were used to provide a blank correction for each protein concentration (see section 3.1).  The anisotropy was plotted against protein concentration with the standard deviation of the replicates used as the error.  The data were fitted to a logistic function (equation 3) to find the lower and upper limits of the anisotropy r$_{max}$, r$_{min}$ which reflect the anisotropy of the tracer when in the unbound and bound state respectively (Supplementary Figure 2).

$$r = \frac{r_{min} - r_{max}}{1 + \left(x/x_0\right)^p} + r_{max}$$

*Equation 3*
Where $x$ is the protein concentration
$x_0$ is the midpoint of the transition
and p is the slope

Across a number of repeats, average values for the limits of anisotropy were found to be r$_{min}$ = 69  r$_{max}$ = 310 for both full length PEX5 and PEX5C and these values were used throughout subsequent assays. These data were also used to check for any change in quantum yield between the bound and free states of the assay by checking for any variation in total intensity (Equation 2) over the course of the titration.  No change was observed (data not shown).

## 3.4   Determination of the dissociation constant of the fluorescent tracer.

60 µl of gelatine buffer solution was removed from wells 1-3 of rows A-P. A serial dilution of lissamine rhodamine labelled YQSKL in FA buffer was made starting with 2000 nM (3 x final concentration) and diluted by ½ through a series. 20 µl of each was added to columns 1 - 14. 20 µl of FA buffer was added to A15-C15 to act as a blank. 20 µl of a ~300 nM solution of protein in FA buffer was then added to all wells being used (3 x final concentration). The anisotropy was then measured (Section 3.1) and calculated by subtracting the average intensities from the s and p channels in the blank wells from the wells of interest.  The anisotropy in each well was converted into the fraction of tracer bound using equation 4 and finally to the amount of tracer bound by multiplying by the total concentration of tracer in the well of interest.

$$\frac{L_B}{L_T} = \left[\frac{\lambda(r_{max} - r)}{(r - r_{min})} + 1\right]^{-1}$$

*Equation 4*
Where $L_B$ is the concentration of fluorescent tracer bound to PEX5, $L_T$ is the total tracer concentration. $\lambda$ reflects the difference in quantum yields of the bound and free states which was determined to be 1 (Section 3.3)

The data were plotted and fitted to equation 5 using non-linear least squares fitting algorithm in OriginPro to determine the $K_d$ of the tracer peptide for the protein. The error in $K_d$ is obtained from the fitting error within the procedure (Supplementary Figure 3).

$$L_B = \frac{(L_T + P_T + K_d) - \sqrt{(L_T + P_T + K_d)^2 - 4L_TP_T}}{2}$$

*Equation 5*
Where $L_B$ is the concentration of fluorescent tracer bound to PEX5, $L_T$ is the total tracer concentration. $P_T$ is the total concentration of PEX5

## 3.5 Competition experiments to determine the dissociation constant of unlabelled peptides.

70 µl gelatin/buffer solution was removed from rows A-F, wells 1-14. A concentration series of unlabelled peptide from 4 mM to 0 mM was made into FA buffer. 10 µl of each concentration was added to the columns 1 – 14. 10 µl of a 120 nM lissamine rhodamine-YQSKL solution added to all wells on rows A-C. For blank samples, 10 µl of FA buffer was then added to all wells on rows D-F. Finally 10 µl of an 800 nM solution of protein then added to all wells being used. The anisotropy was then measured (Section 3.1) and calculated by subtracting the average intensities from the s and p channels in the blank wells from the wells of interest. The anisotropy in each well was converted into the fraction of tracer bound using equation 4 and finally to the amount of tracer bound by multiplying by the final tracer concentration (30 nM). The bound amount was plotted against the $Log_{10}$ of the competitor concentration and the data fitted to a single site competition model (Equation 6) in Origin pro. The lower asymptote (Min) was fixed at 0 in all the fittings. The error in $IC_{50}$ was taken from the fitting error of the data. The data for competition of unlabelled peptides for the binding site of PEX5C and PEX 5 are shown in Supplementary Figures 4 and 5.

$$L_B = Min + \frac{(Max - Min)}{1 + 10^{(x - LogIC_{50})}}$$

*Equation 6.*
Where x is the $Log_{10}$ of the unlabelled peptide concentration. Max and Min are the upper and lower asymptotes of the curve. Min was fixed at zero in all fittings for consistency.

The $IC_{50}$ values were then converted to the $K_i$ of the competitor peptide by combination with $K_d$ of the protein using the method of Nicolovska-Coleska [4] using Equations 7-10.

$$K_i = \frac{I_{50}}{\frac{L_{50}}{K_d} + \frac{P_0}{K_d} + 1}$$

*Equation 7*
Where $K_i$ is the dissociation constant of unlabelled peptide and $K_d$ is the dissociation constant of fluorescent tracer. The remaining parameters are computed using Equations 8-10

$P_0$, the unbound concentration of protein at zero inhibitor concentration is calculated from:

$$P_0 = \frac{-(K_d + L_T - P_T) + \sqrt{(K_d + L_T - P_T)^2 - 4K_d P_T}}{2}$$

*Equation 8*
Where $K_d$ is the dissociation constant of fluorescent tracer, $L_T$ is the total tracer concentration. $P_T$ is the total concentration of PEX5

$L_{50}$ the concentration of unbound fluorescent tracer at the $IC_{50}$ point is calculated from:

$$L_{50} = L_T - \frac{(P_T - P_0)}{2}$$

*Equation 9*
Where $L_T$ is the total tracer concentration. $P_T$ is the total concentration of PEX5 and $P_0$ is the unbound concentration of protein at zero inhibitor concentration calculated from Equation 8

$I_{50}$, the unbound inhibitor concentration at the $IC_{50}$ point is calculated from:

$$I_{50} = IC_{50} - P_T + \left(\frac{P_T - P_0}{2}\right) \cdot \left(1 + \frac{K_d}{L_{50}}\right)$$

*Equation 10*
Where $IC_{50}$ is determined from the curve fit of inhibitor titration. $P_T$ is the total concentration of PEX5 and $K_d$ is the dissociation constant of fluorescent tracer. $P_0$ is calculated from Equation 8 and $L_{50}$ From Equation 9

As the error in the final $K_i$ involves components from the error in the measurement of the $K_d$ of the tracer and the $IC_{50}$ of the inhibitor the error was assessed using simulation. A series of simulations were run with a simulated value of $K_d$ and $IC_{50}$ generated around the observed mean, the probability of the value occurring was based upon a normal distribution with standard deviation equal to the fitting error of each value. The simulated value was then propagated to give a simulated $K_i$. 2000 simulations were performed with the final error estimated from the standard deviation of the simulated results. The calculated $K_i$ and errors are shown in the Table II of the main manuscript.

# 4 Peptide Synthesis and Characterisation

## 4.1 General Information

N-terminal Fmoc protected amino acids (Merck Millipore) were used with the following side chain protection: Arg (Pbf), Asn and Gln (Trt), Glu (OtBu) Lys, (N-Boc), Ser, Thr and Tyr (tBu). Peptides were synthesised on 2-Chlorotrityl resin which was purchased pre-loaded with the appropriate C-terminal amino acid. Peptide elongations were performed manually using standard Fmoc solid phase peptide synthesis as outlined by Chan and White (Chan and White, 1999) using SPE tubes fitted with a polyethylene frit (Grace Davidson, Baltimore, USA) and a vacuum tank attached to a water aspirator. DCM, Diethyl ether (Fisher Scientific, Hampton, NH, USA), DIPEA (Sigma-Aldrich), DMF (Fluka, Buchs, Switzerland) HCTU (Merck-Millipore), Piperidine (Fluka) and triisopropylsilane (Sigma-Aldrich) were used without further purification. TFA (Sigma-Aldrich) was freshly distilled before use. Lissamine$^{TM}$ rhodamine B sulfonyl chloride was purchased from Invitrogen (Invitrogen Molecular Probes). Mass spectrometry and LC-MS was performed on a Bruker HCT Ultra spectrometer using an electrospray source. High Resolution accurate mass spectra were obtained on a Bruker MicrOTOF using electrospray ionisation. Analytical HPLC was performed on an Agilent 1290 Infinity LC system and preparative HPLC on a Gilson preparative system. Reverse phase solvents were acetonitrile and Ultra pure water both containing 0.1 % TFA.

## 4.2 General Procedure for Peptide Synthesis

### 4.2.1 Peptide Assembly

Peptides were prepared on 2 Cl-Trt resin preloaded with the C-terminal amino acid (0.085 mmol, 0.85 mmol/g). The resin was swollen in DCM for 20 mins then washed 2 x 5 mins with DMF. Fmoc amino acid (5 mmol) and HCTU (5 mmol) were dissolved in the minimum volume of DMF. DIPEA (10 mmol) was added and transferred to the tube containing the resin (1 mmol). The mixture was gently agitated for 1 hr on a laboratory rotator and the resin was isolated by filtration and washed with DMF (3 × 5 ml × 2 mins). A negative Kaiser colour test showed a complete reaction. The resin was treated with a solution of 20 % piperidine in DMF (5 × 5 ml × 2 mins) and then washed with DMF (5 × 5 ml × 2 mins). Couplings were repeated until the desired sequence had been assembled. The functionalised resin was washed with DMF (3 × 3 ml, 2 mins), Isopropanol (3 × 3 ml × 5 mins) and hexane (4 × 3 ml × 2 mins), sucked dry on a manifold tank (10 min) then dried *in-vacuo* over KOH for 16 hrs.

### 4.2.2 Cleavage of Peptide from resin

The dried functionalised resin was treated with a cleavage cocktail of TFA/H$_2$O/TIPS (95:2.5:2.5, 1 ml/25 mg resin) with the exception of Met containing peptides in which case a cocktail of TFA/H$_2$O/EDT/TIPS (95:2.5:2.5:1, 1 ml/25 mg resin) was used and all solvents degassed with nitrogen prior to use. The mixture was gently agitated for 1 hr. The resin beads were removed by filtration and rinsed with fresh TFA (2 × 2 ml). The combined filtrates were concentrated under reduced pressure to ~500 μl and the peptide then isolated by precipitation in chilled Et$_2$O (5 ml). The peptide was then dissolved in H$_2$O until the Et$_2$O contained no further precipitate and the aqueous solution

then washed with further chilled Et$_2$O (2 × 5 ml) then lyophilised. The resulting peptides were purified by preparative reverse phase HPLC, lyophilised and the purity of the peptides confirmed by electrospray MS and analytical HPLC. Recovered yields for the purified peptides are calculated based on the manufactures quoted loading of the resin.

### 4.2.3  Synthesis of the fluorescent tracer peptide

Fmoc cleaved, side-chain protected resin bearing the sequence YQSKL, (20 mg, 11.2 µmol) was suspended in anhydrous DMF (2 ml) and DIPEA (19.4 µL, 112 µmol, 10 eq) then Lissamine$^{TM}$ rhodamine B sulfonyl chloride (20 mg, 3 eq) was added to the reaction mixture.  The reaction was stirred in the dark on ice for 18 hrs under nitrogen. The resin washed with DMF (2 × 3 ml × 2 mins) and DCM (2 × 3 ml × 2 min), then cleaved and purified as detailed in the general method then purified *via* preparative HPLC to yield a dark pink solid

## 4.3  Characterisation data for peptides

### 4.3.1  *Lissamine Rhodamine-YQSKL-CO$_2$H (Fluorescent tracer peptide)*
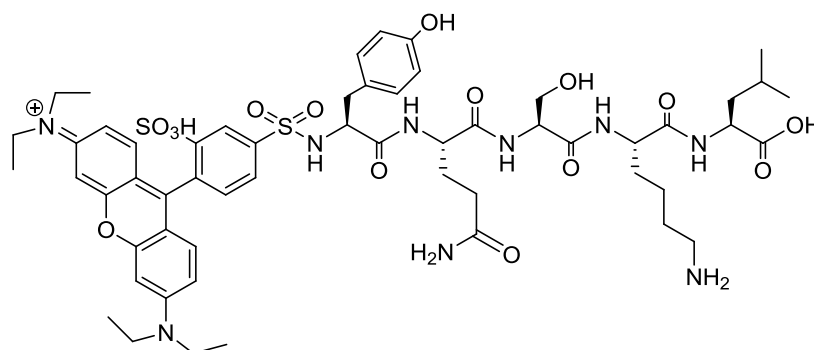


Yield = 2.4 mg, 20 %; RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.7 min; *m/z* (ES) found [M+H]$^+$ 1179.4966; C$_{56}$H$_{77}$N$_9$O$_{15}$S$_2$ requires 1179.4970.
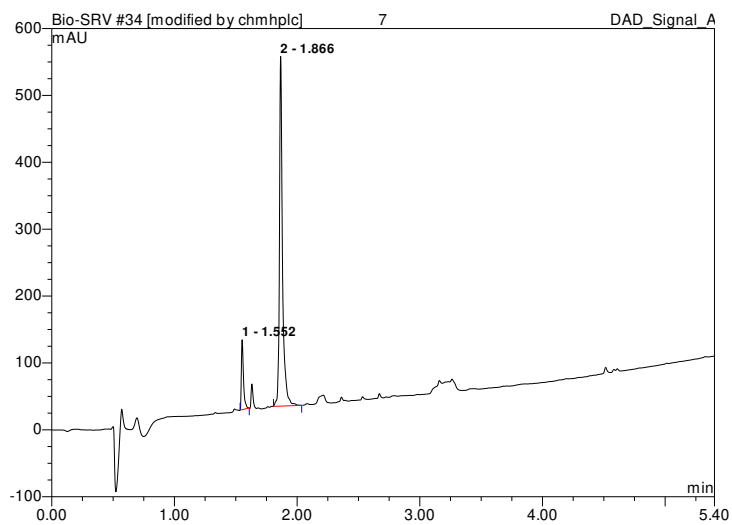
*4.3.2   H₂N-YQSKL-CO₂H*
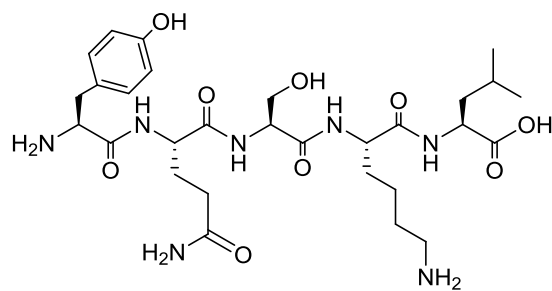
Yield = 33 mg; 62%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:$H_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 0:100, t = 5.4 min 98:2; Rt 1.7 min, *m/z* (ES) found [M+H]$^+$ 638.3507 ; $C_{29}H_{48}N_7O_9$ requires 638.3508.
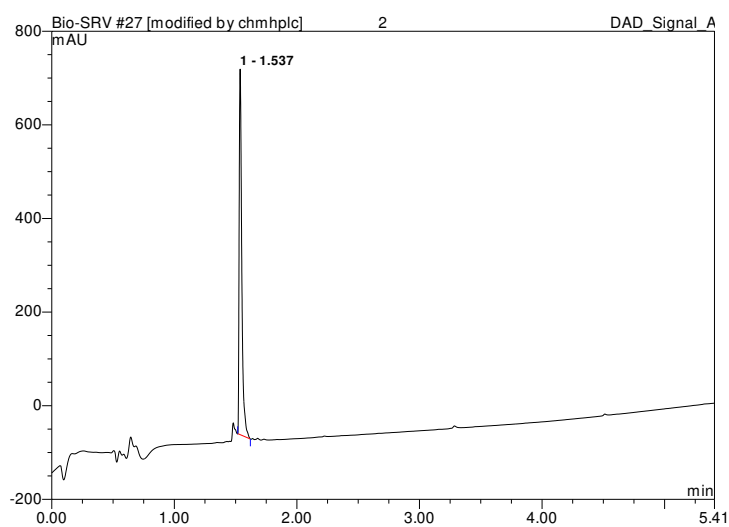
### 4.3.3   H₂N-YPSKL-CO₂H
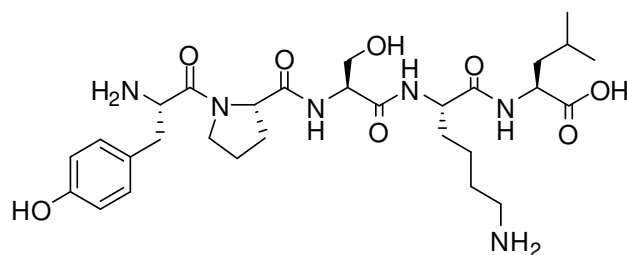


Yield = 33 mg; 62%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.7 min, *m/z* (ES) found [M+H]$^+$ 607.3455; C$_{29}$H$_{47}$N$_6$O$_8$ requires 607.3450.
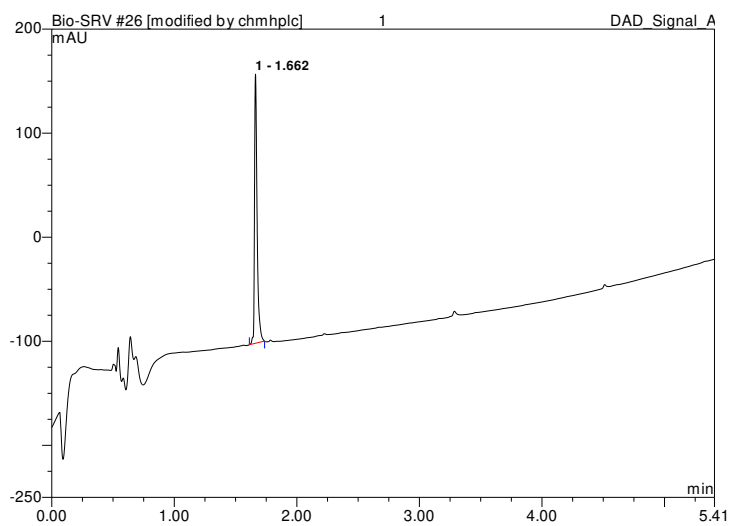
### 4.3.4  H₂N-YQSKV-CO₂H
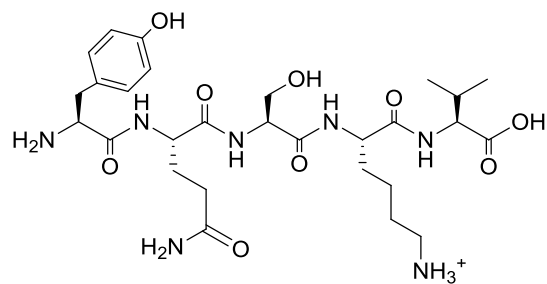


Yield = 30.2 mg; 59%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.3 min, *m/z* (ES) found [M+H]$^+$ 624.3354; C$_{28}$H$_{46}$N$_7$O$_9$ requires 624.3352.
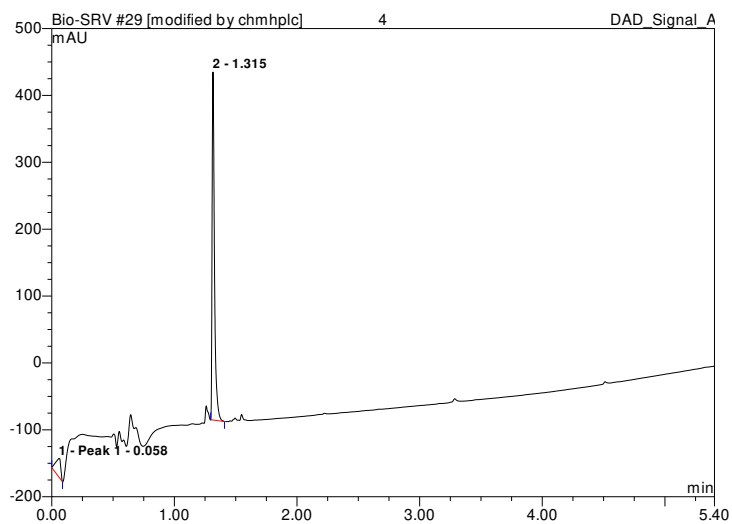
*4.3.5   H₂N-VAKTTRPSRL-CO₂H*
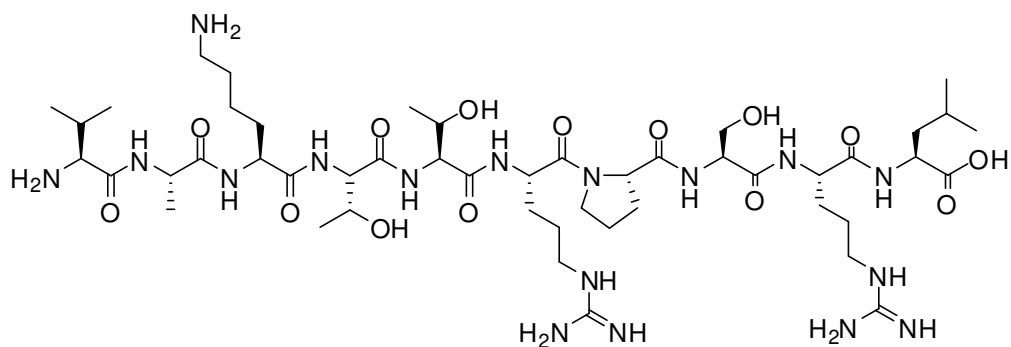


Yield = 47.9 mg, 50%. RP-HPLC Jupiter 4 µ Proteo 90 Å (C12 10 cm X 2 mm); Flow rate 1 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t =30 min, 25:75, t = 32 min 98:2; Rt 8.3 min, *m/z* (ES) found [M+H]$^{+}$ 1128.6852; C$_{48}$H$_{90}$N$_{17}$O$_{14}$ requires 1128.6848.
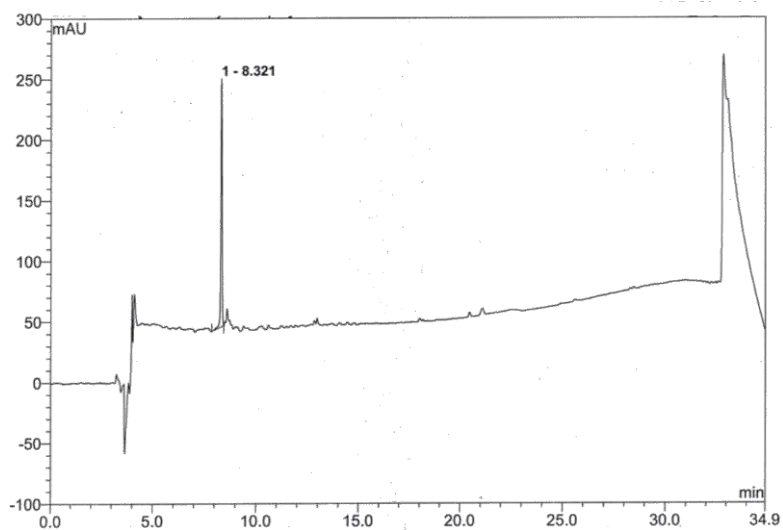
*4.3.6   H<sub>2</sub>N-VAKTTRPSRM-CO<sub>2</sub>H*
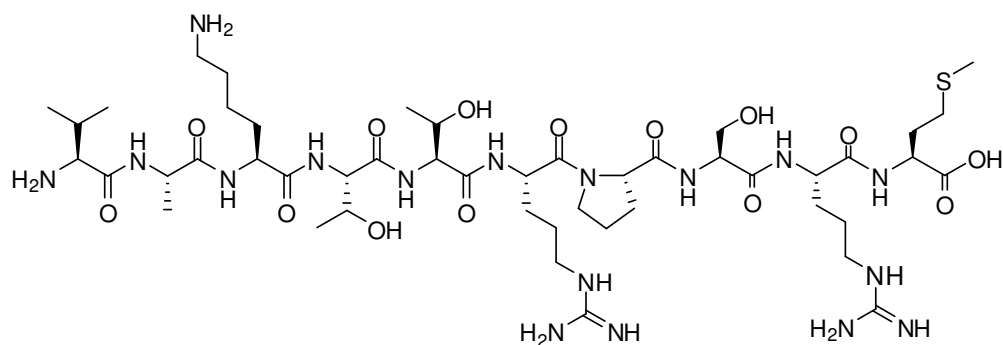


Yield = 50.2 mg, 52%. RP-HPLC Jupiter 4 μ Proteo 90 Å (C12 10 cm X 2 mm); Flow rate 1 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 0:1;  t =30 min, 15:85, t = 32 min 98:2; Rt 7.6 min; *m/z* (ES) found [M+H]$^+$ 1146.6447; C$_{47}$H$_{88}$N$_{17}$O$_{14}$S requires 1146.6442.
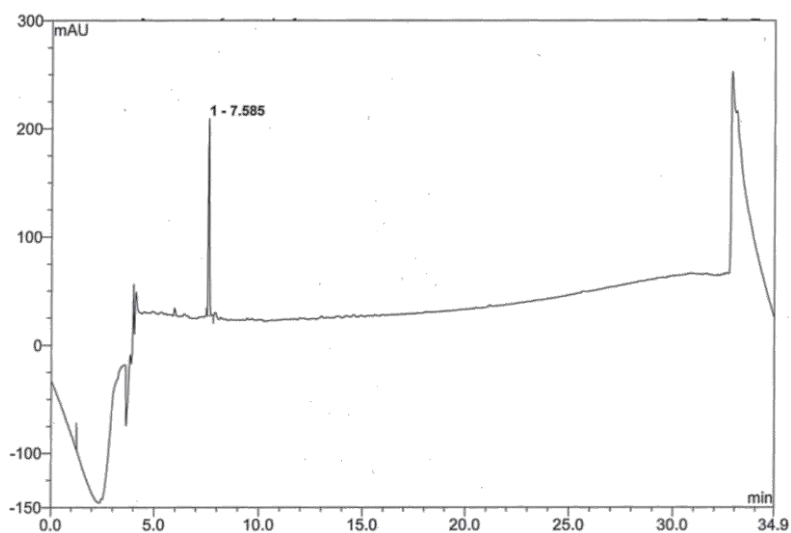
## 4.3.7 H₂N-VAKTTRPSRI-CO₂H
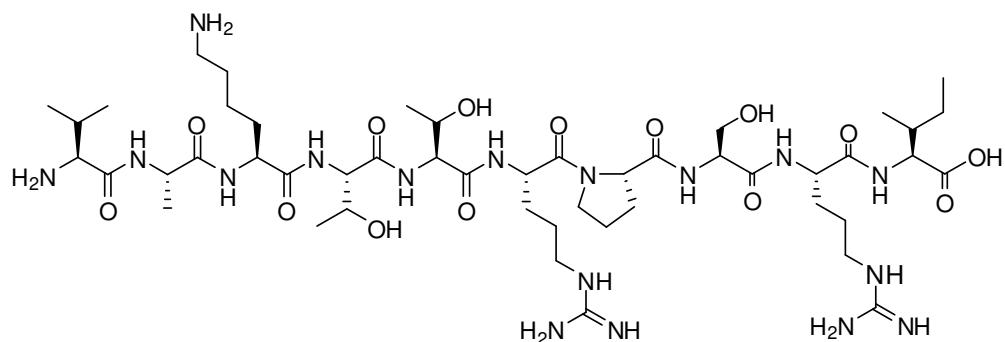


Yield = 48.1 mg, 51%. RP-HPLC Jupiter 4 μ Proteo 90 Å (C12 10 cm X 2 mm); Flow rate 1 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t =30 min, 25:75, t = 32 min 98:2; Rt 8.0 min, *m/z* (ES) found [M+H]$^+$ 1128.6850; C$_{48}$H$_{90}$N$_{17}$O$_{14}$ requires 1128.6848
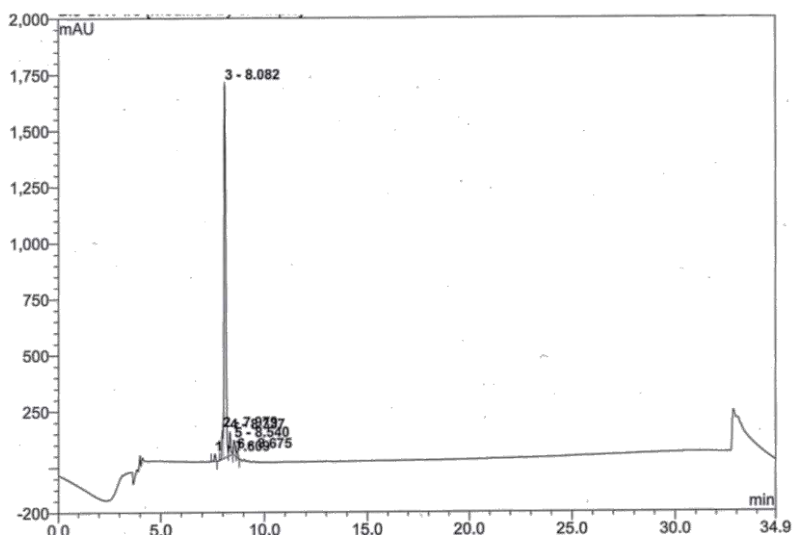
### 4.3.8   H₂N-VAKTTRPSRY-CO₂H
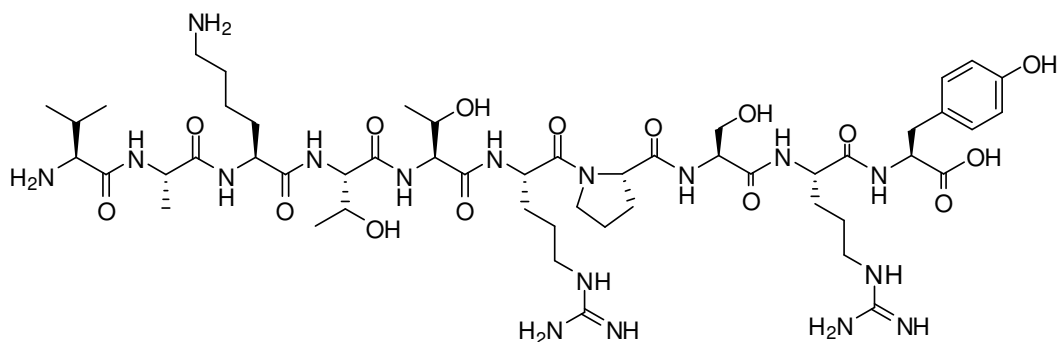


Yield = 48.1 mg, 51%. RP-HPLC Jupiter 4 μ Proteo 90 Å (C12 10 cm X 2 mm); Flow rate 1 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t =30 min, 25:75, t = 32 min 98:2; Rt 7.6 min, *m/z* (ES) found [M+H]$^+$ 1178.6638; C$_{51}$H$_{88}$N$_{17}$O$_{15}$ requires 1178.6640.
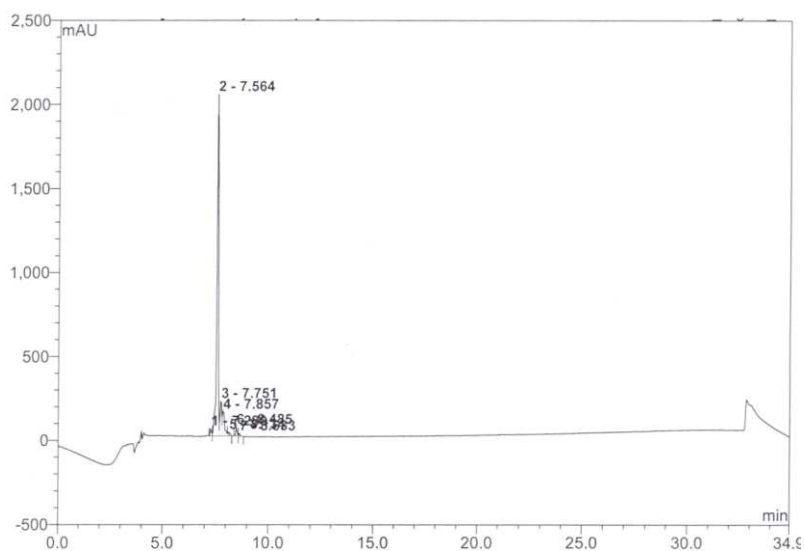
### 4.3.9  H<sub>2</sub>N-VAKTTRPSRV-CO<sub>2</sub>H
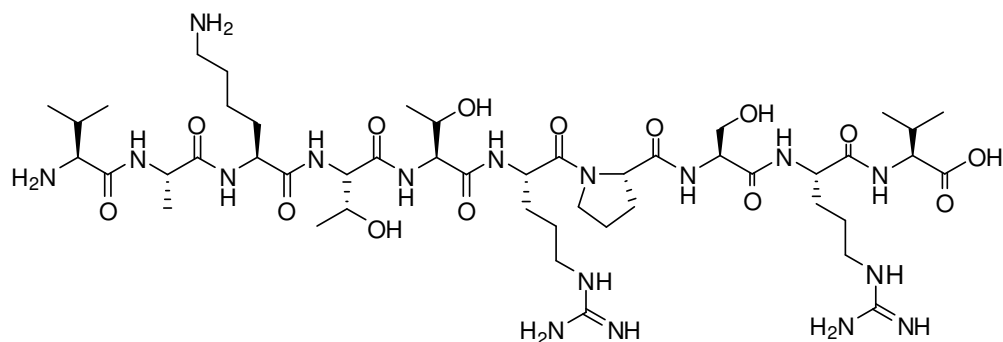


Yield = 47.4 mg, 52%. RP-HPLC Jupiter 4 μ Proteo 90 Å; Flow rate 1 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 0:1;  t =30 min, 15:85, t = 32 min 98:2; Rt 7.3 min; *m/z* (ES) found [M+H]$^+$ 1114.6696; C$_{48}$H$_{89}$N$_{16}$O$_{14}$ requires 1114.6691.
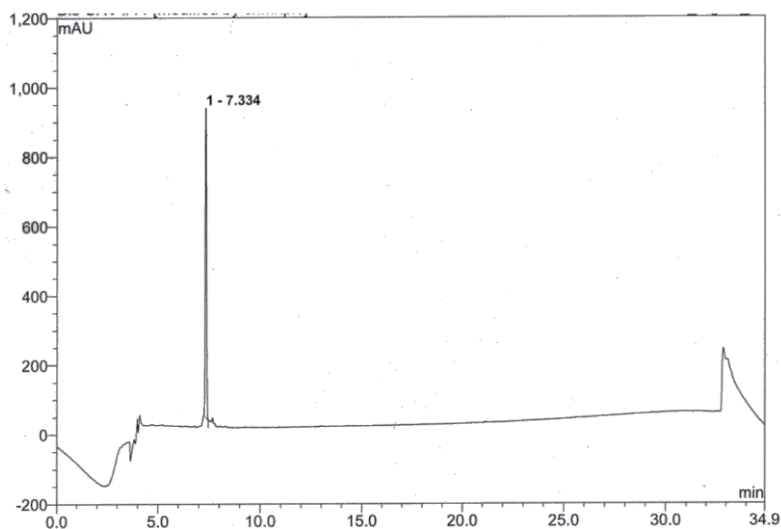
## 4.3.10 H₂N-VAKTTRPSKL-CO₂H
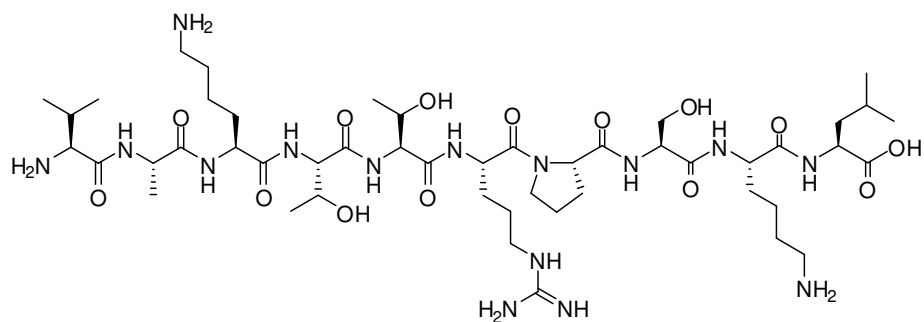


Yield = 46 mg, 50%. Ascentis Peptide; Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.5 *m/z* (ES) found [M+H]$^+$ 1100.6787; C$_{48}$H$_{90}$N$_{15}$O$_{14}$ requires 1100.6786
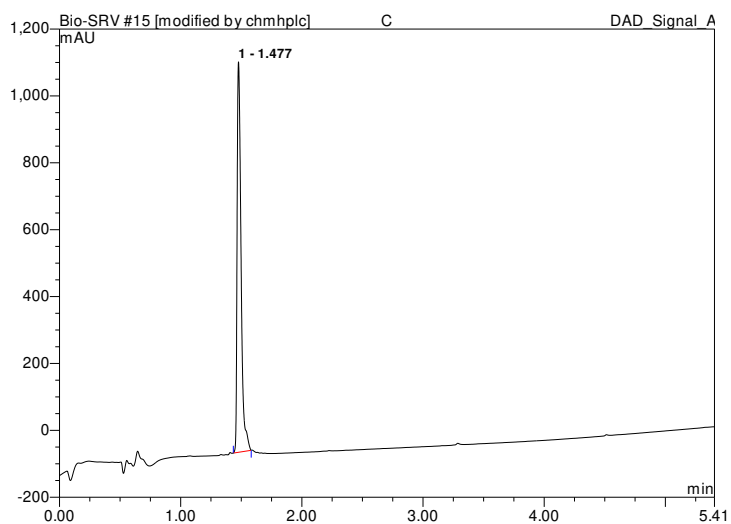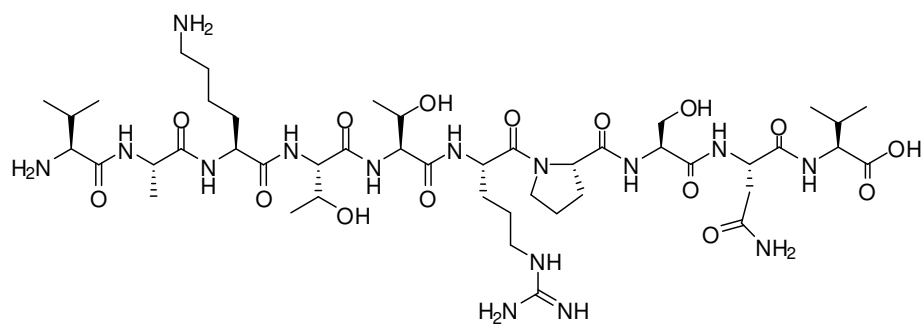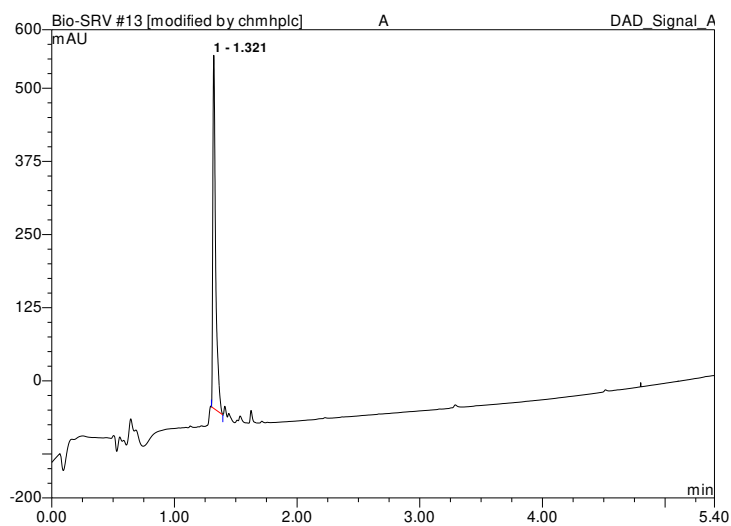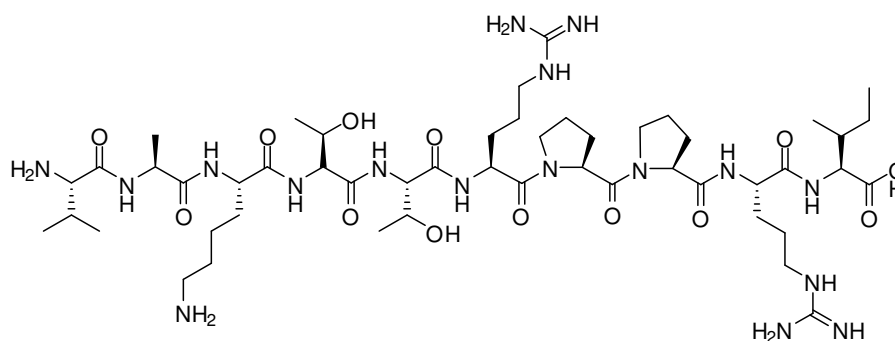
## 4.3.11 H₂N-VAKTTRPSNM-CO₂H
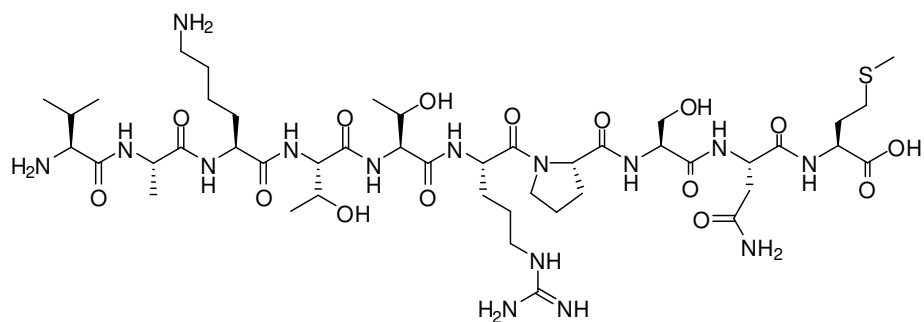


Yield = 45.2 mg, 48%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.4 min, *m/z* (ES) found [M+H]$^+$ 1104.5759; C$_{45}$H$_{81}$N$_{15}$O$_{15}$S requires 1104.5757
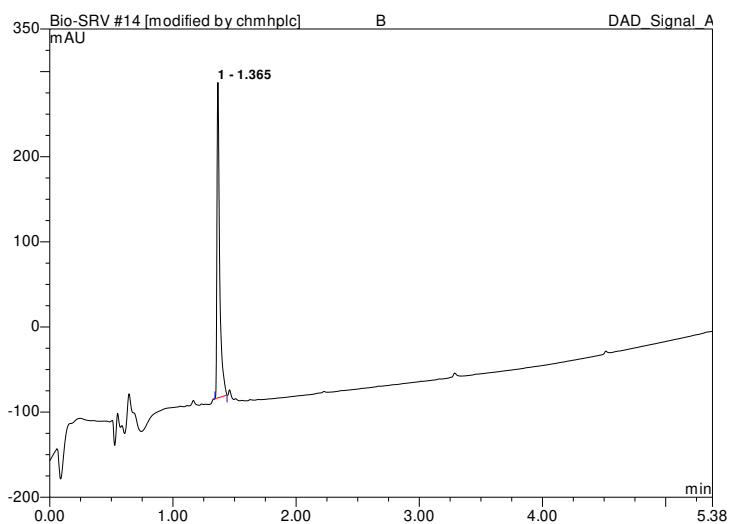
### 4.3.12  H₂N-VAKTTRPARV-CO₂H
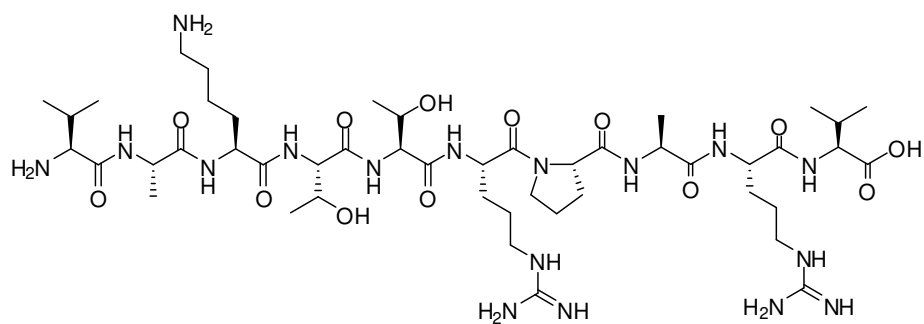
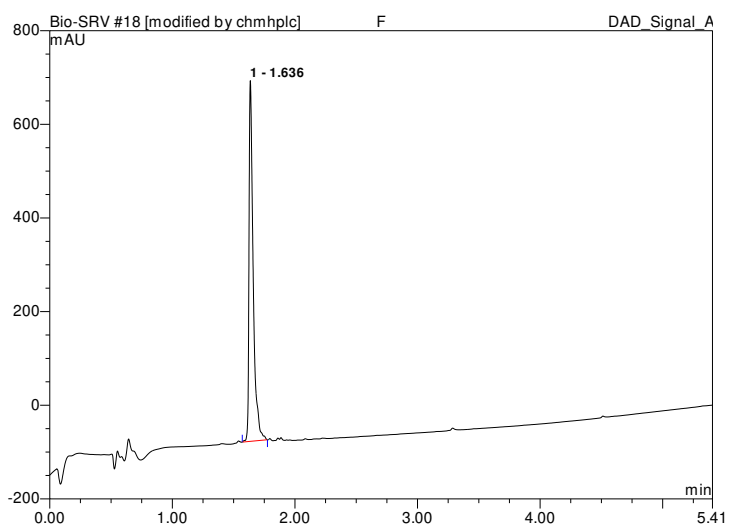

Yield = 47.1 mg; 51%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.4 min, *m/z* (ES) found [M+H]$^+$ 1098.6746; C$_{47}$H$_{88}$N$_{17}$O$_{13}$ requires 1098.6742.
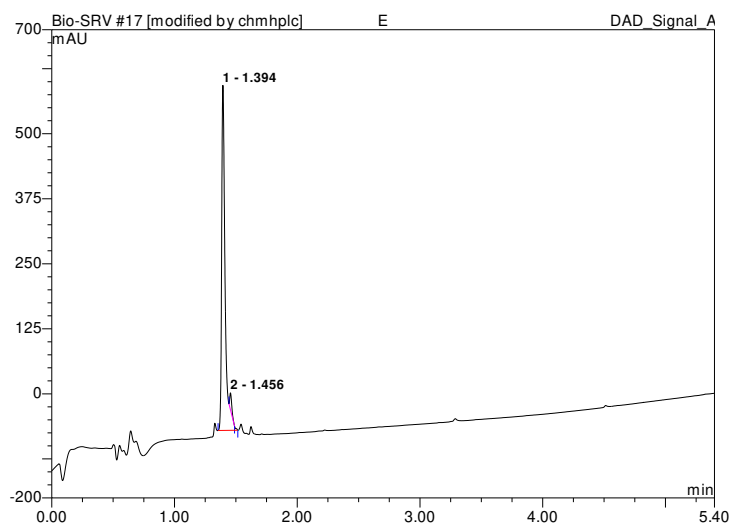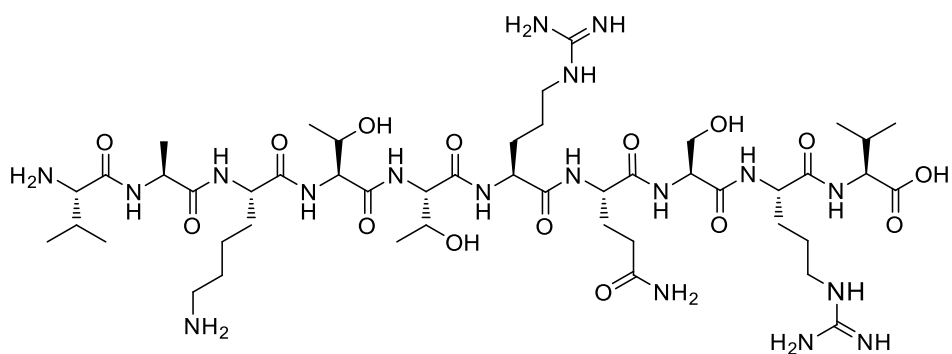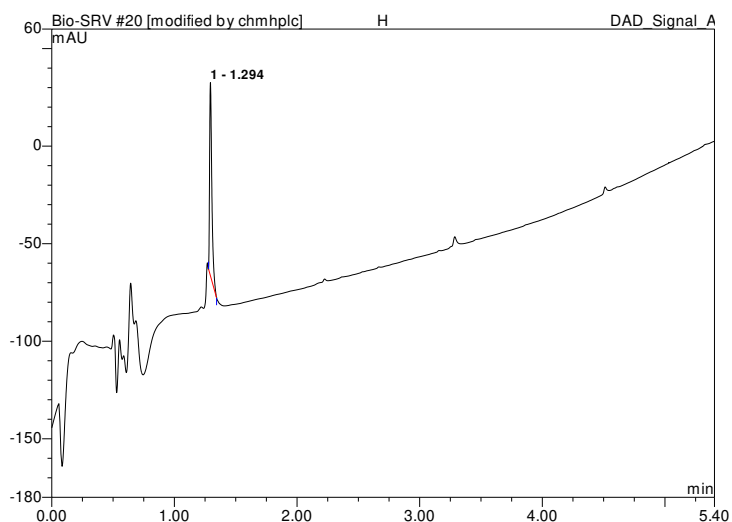
### 4.3.13 *H₂N-VAKTTRPPRV-CO₂H*
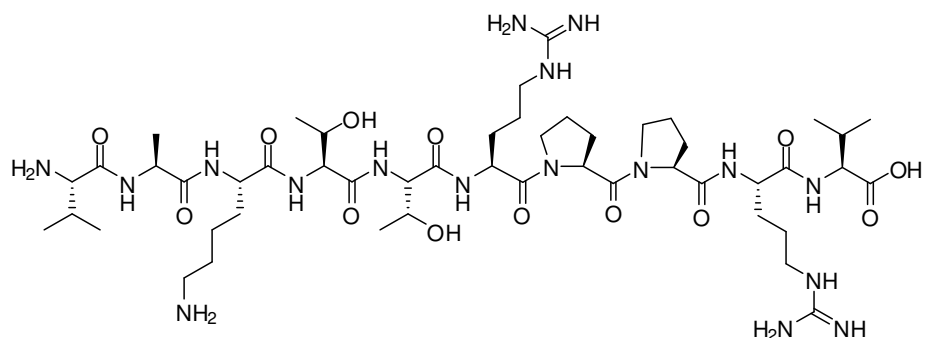


Yield = 46.9 mg; 50%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:$H_2O$ with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.5 min, *m/z* (ES) found [M+H]$^+$ 1124.6901; $C_{49}H_{90}N_{17}O_{13}$ requires 1124.6899.
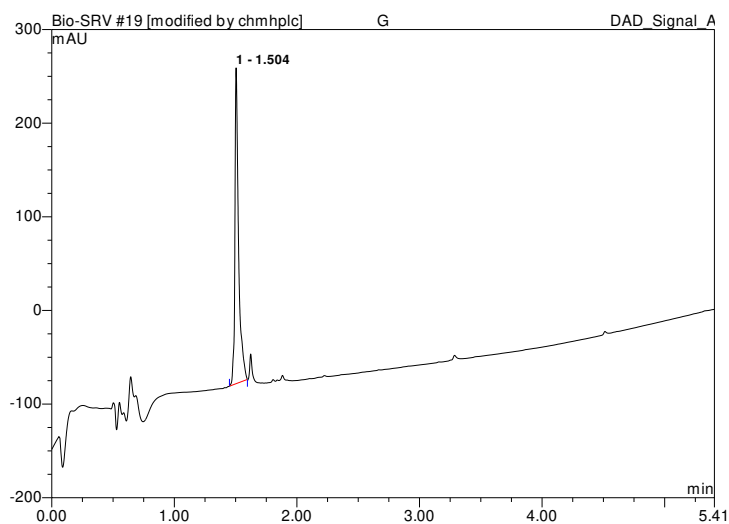
## 4.3.14  H₂N-VAKTTRPSNV-CO₂H



Yield = 39.6 mg; 45%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.3 min, *m/z* (ES) found [M+H]$^+$ 1072.6104; C$_{45}$H$_{82}$N$_{15}$O$_{15}$ requires 1072.6109.
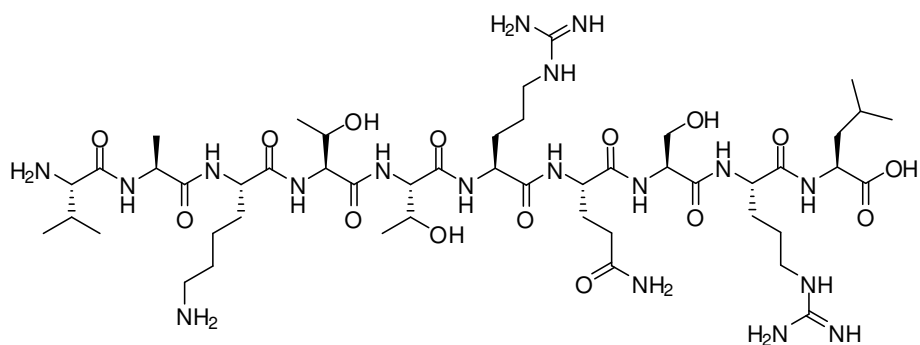
## 4.3.15 H₂N-VAKTTRPPRI-CO₂H


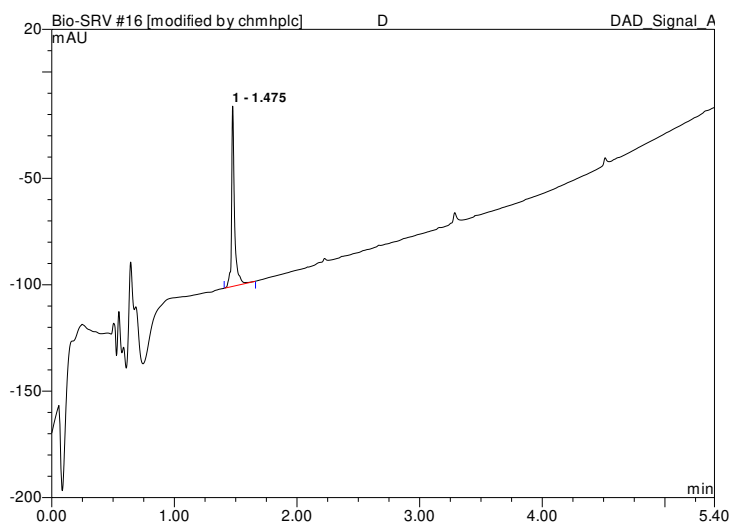
Yield = 39.6 mg; 48%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.6 min, *m/z* (ES) found [M+H]$^{+}$ 1138.7058; C$_{50}$H$_{92}$N$_{17}$O$_{13}$ requires 1138.7055.
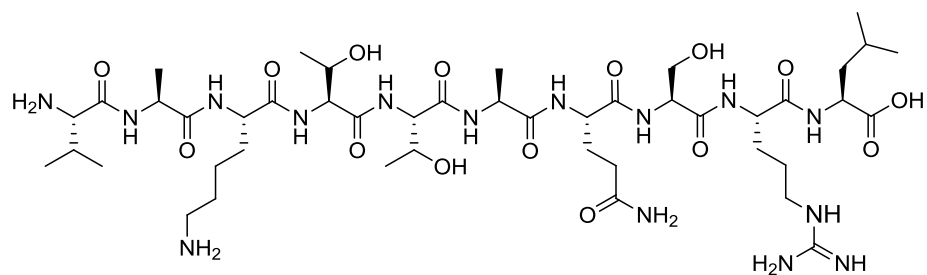
H₂N-VAKTTRQSRV-CO₂H

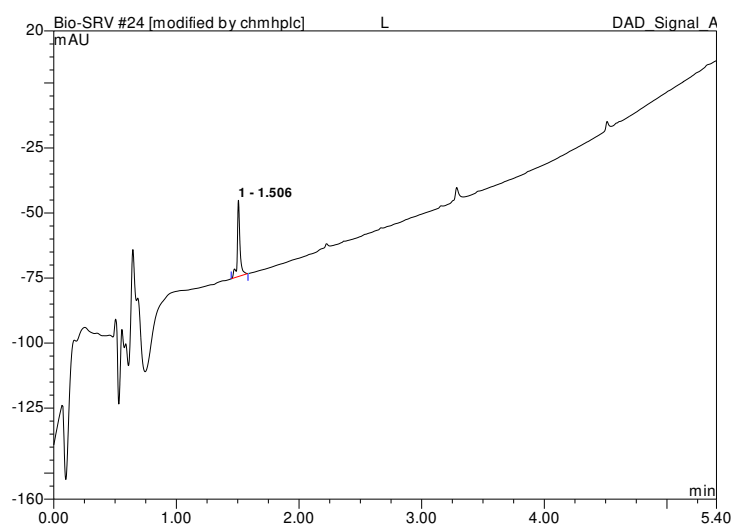

Yield = 39.6 mg; 58%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 95:5, t = 5.4 min 98:2; Rt 1.3 min, *m/z* (ES) found [M+3H]$^{3+}$ 382.5629; C$_{47}$H$_{91}$N$_{18}$O$_{15}$ requires 382.5632.

## 4.3.16  H<sub>2</sub>N-VAKTTRQSRL-CO<sub>2</sub>H



Yield = 54.1 mg; 55%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 0:100, t = 5.4 min 98:2; Rt 1.5 min, *m/z* (ES) found [M+2H]$^{2+}$ 580.3495; C$_{48}$H$_{92}$N$_{18}$O$_{15}$ requires 580.3490.
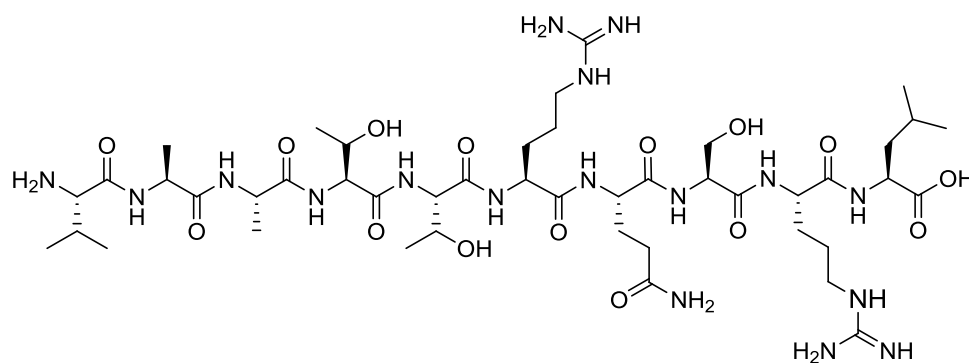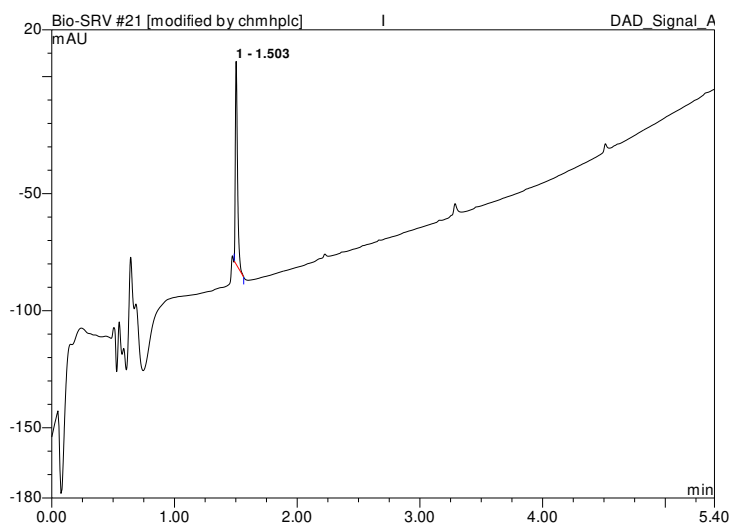
## 4.3.17 H$_2$N-VAKTTAQSRL-CO$_2$H



Yield = 36.9 mg; 39%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 0:100, t = 5.4 min 98:2; Rt 1.5 min, *m/z* (ES) found [M+H]$^+$ 1074.6261; C$_{46}$H$_{85}$N$_{14}$O$_{15}$ requires 1074.6266.
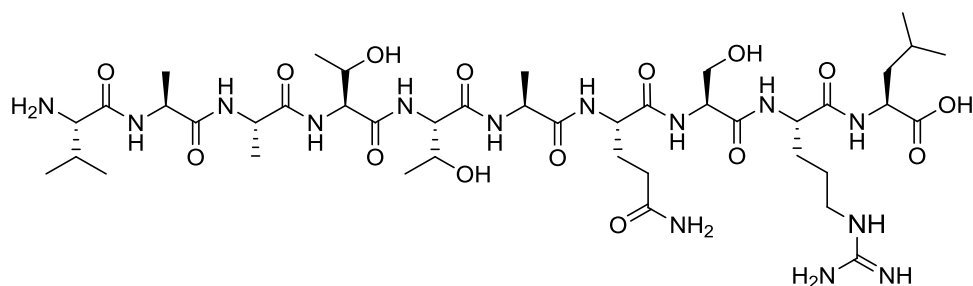
## 4.3.18 H_2N-VAATTRQSRL-CO_2H



Yield = 35.5 mg; 38%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 0:100, t = 5.4 min 98:2; Rt 1.5 min, *m/z* (ES) found [M+H]$^+$ 1102.6324; C$_{46}$H$_{85}$N$_{16}$O$_{15}$ requires 1102.6327

*4.3.19 H₂N-VAATTAQSRL-CO₂H*


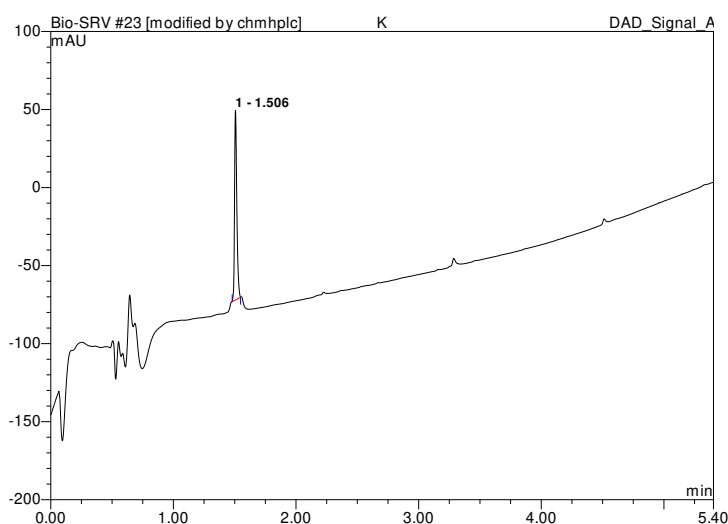
Yield = 25.9 mg; 30%. RP-HPLC Ascentis Express Peptide (ES - C18, 10 cm x 2.1 mm); Flow rate 0.5 ml min$^{-1}$; Gradient elution t = 0 min, MeCN:H$_2$O with 0.1 % TFA, 5:95; t = 5.0 min, 0:100, t = 5.4 min 98:2; Rt 1.5 min, *m/z* (ES) found [M+H]$^{+}$ 1016.5662; C$_{43}$H$_{78}$N$_{13}$O$_{15}$ requires 1016.5662.



# 5 References

[1] Lakowicz, J. R., Principles of Fluorescence Spectroscopy, **2006**, Springer.

[2] Moerke, N. J., "Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding", Current Protocols in Chemical Biology, **2009**, 1,1.

[3] Gatto, G. J, Maynard, E. J. , Guerrerio, A. L., Geisbrecht, B. V., Gould, S. J. , Berg, J. M., Correlating Structure and Affinity for PEX5:PTS1 Complexes", Biochemistry, **2003**, 1660.

[4] Nikolovska-Coleska, Z., Wang, R.,  Fang X., Pan, H.,Tomita, Y., Li, P., Roller,P. P., Krajewski, K., Saito,N. G.,  Stuckey, J. A., Wang, S.,  "Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization", Analytical Biochemistry, **2004**, 261-273.