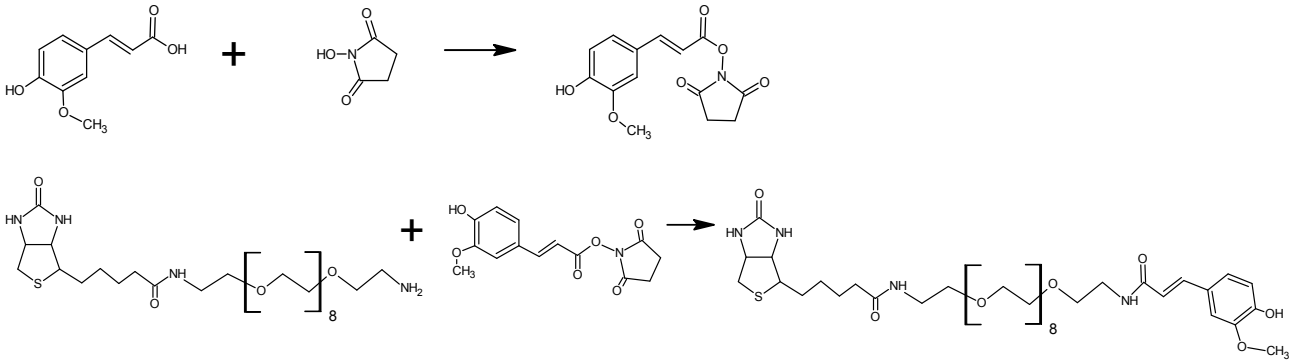
Electronic supplementing information (ESI)

Chemical synthesis of biotinylated haptens

The desired compounds were synthesized starting from the commercially available O-(2-aminoethyl)-O'-[2-(biotinylamino)ethyl]octaethylene glycol (Santa Cruz Biotechnology Inc.) (PEG8-B). Amide bond formation with cinnamic acids was done by the activated esters method. N-Hydroxysuccinimide esters (NHS) were used to activate carboxyl groups.



The esterification of p-coumaric, ferulic and caffeic acids is a well reported procedure which usually gives high yields. Preparation of N-hydroxysuccinimide-esters of cinnamic acids was published as a general method quite long ago [J. Stöckigt, M. Zenk. Chemical synthesis and properties of hydroxycinnamoyl-Coenzyme A derivatives. Z. Naturforsch. 30c, 352-358, 1975] and is actively cited. However, we had to optimize the reaction conditions for each particular acid used in the present study because of their different solubility. Thus, NHS-esters of p-coumaric and ferulic acids were prepared in the mixture of solvents tetrahydrofurane : acetonitrile 1:2 using EDC•HCl as a coupling agent. The NHS-ester of p-coumaric acid was crystallized from water yielding 65% of the pure substance. The crystallization of the NHS-ester of ferulic acid from water was not successful. Instead, the compound was dissolved in ethyl acetate, washed with water and precipitated with n-hexane, providing pure substance with the 52% yield. The NHS-ester of caffeic acid was prepared in dimethylformamide (DMF) using N,N'-dicyclohexylcarbodiimide (DCC) as a coupling agent with further purification of the product by column chromatography on Silica gel, yielding 33% of the target substance.

The formation of amide bond in all three cases was carried out under the same conditions. The reaction was run in DMF under argon in dark at room temperature. Isolation of p-coumaroyl-PEG8-B was performed by preparative TLC on aluminum oxide. For the isolation of feruloyl-PEG8-B and caffeoyl-PEG8-B we used preparative TLC on Silica gel. Silica is preferred for the feruloyl derivative because of the similar chromatographic mobility on Al_2O_3 of the product and remaining NHS-ester of ferulic acid; and for the caffeloyl derivative -–and because the reaction product is immobile on Al_2O_3. All in all, the compounds p-coumaroyl-PEG8-B, feruloyl- PEG8-B and caffeoyl- PEG8-B were synthesized with the yields 49%, 47.3% and 41.7% on the last step, respectively. According to NMR-spectra, all these products are *trans*-isomers (with the characteristic coupling constant 15.5 Hz for the double bond).

Experimental

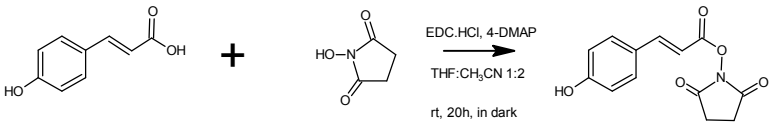
All reagents and solvents (HPLC grade) were purchased from Sigma-Aldrich Co. or Tokyo Chemical Industry Co. O-(2-Aminoethyl)-O'-[2-(biotinylamino)ethyl]octaethylene glycol was purchased from Santa Cruz Biotechnology Inc.

The monitoring of reactions was carried out by thin layer chromatography (TLC) using aluminum sheets precoated with silica gel 60 F_{254} or aluminum oxide 60 F_{254} neutral (Merck).

NMR spectra were measured using a Varian Mercury 300 MHz spectrometer (Varian Inc.). All chemical shifts are given in ppm relative to tetramethylsilane (TMS) as an internal standard.

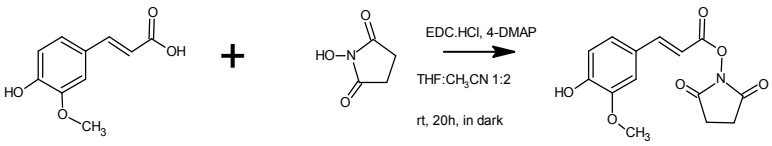
Mass-spectra were recorded using a high resolution ESI-TOF LCT Premier XE mass-spectrometer (Waters Corp.). The sample of an analyte was dissolved in appropriate solvent at a concentration of ca. 0.01 mg/mL and infused at a 15 μL/min rate. The reference solution of Leucine Enkephaline (50 pg/mL) was infused simultaneously. Original spectra were centered, and lock mass TOF correction was applied.

N-Hydroxysuccinimide ester of trans-p-coumaric acid (1).



p-Coumaric acid (200 mg, 1.22 mmol), N-hydroxysuccinimide (154 mg, 1.34 mmol) and 1-(3-dimethylaminopropyl)-3-ethyl-carbodiimide hydrochloride (EDC) (281 mg, 1.5 mmol) were dissolved in the mixture of tetrahydrofuran : acetonitrile 1:2 v/v (7.5 mL) under argon. Then 4-dimethylaminopyridine (DMAP) (5 mg, 0.04 mmol) was added and the reaction mixture was stirred under argon in dark for 24 hours. The solvents were removed in vacuum, and the yellow oil residue was washed with water (3 x 15 mL) until solidification. The white solid was washed with water (80 mL) and dried in a vacuum exsiccator over sodium hydroxide. Yield: 208 mg (65.4%), R_f 0.37 (Silica gel, chloroform:methanol 95:5). ^{1}H NMR (300 MHz; $CDCl_3$; Me_4Si): δ_H, ppm 7.80 (d, J=16 Hz, 1H,-C**H**=CH-), 7.42 (d, J=8.5 Hz, 2H, -C_6H_4-), 6.76 (d, J=8,79 Hz, 2H, -C_6H_4-), 6.35 (d, J=16 Hz, 1H, -CH=C**H**-), 2.89 (m, 4H, -CO-C**H$_2$**-C**H$_2$**-CO-); MS (ESI-TOF, $CHCl_3$/MeOH 10:15 v/v): *m/z*=260.0607 $[M-H]^-$ (calcd for $C_{13}H_{11}NO_5$ 260.0559).

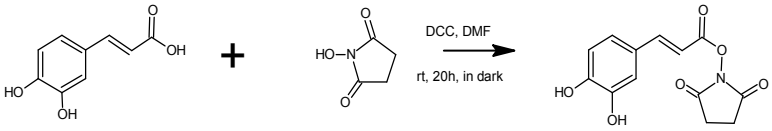
N-Hydroxysuccinimide ester of trans-ferulic acid (2).



Ferulic acid (237 mg, 1.22 mmol), N-hydroxysuccinimide (154 mg, 1.34 mmol) and 1-(3-dimethylaminopropyl)-3-ethyl-carbodiimide hydrochloride (EDC) (281 mg, 1.5 mmol) were dissolved in the mixture of tetrahydrofuran: acetonitrile 1: 2 v/v (7.5 mL) under argon. Then 4-dimethylaminopyridine (DMAP) (5 mg, 0.04 mmol) was added and the reaction mixture was stirred under argon in dark for 24 hours. The solvents were removed in vacuum, and the bright yellow oil residue was washed with water (3 x 15 mL), then dissolved in ethyl acetate (20 mL) and washed again with water (2 x 20mL). The organic phase was dried over sodium sulfate and concentrated in vacuum. The oily residue was crystallized from hexane (40 mL) and dried in a vacuum exsiccator over sodium hydroxide. Yield: 186 mg (52.4%), R_f 0.56 (Silica gel, chloroform: methanol 95:5). ^{1}H NMR (300 MHz; $CDCl_3$; Me_4Si): δ_H, ppm 7.84 (d, J=15.8 Hz, 1H,-C**H**=CH-), 7.14 (m, 2H, -C_6H_3-), 6.95 (m, 1H, -C_6H_3-), 6.43 (d, J=15.8 Hz, 1H, -CH=C**H**-), 3.94 (s, 3H, -

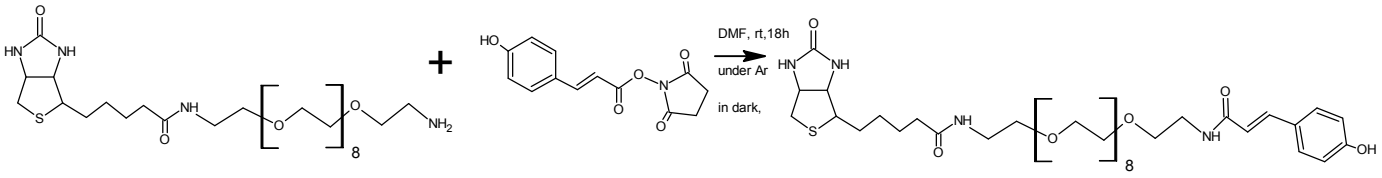
OC**H$_3$**), 2.91 (m, 4H, -CO-C**H$_2$**-C**H$_2$**-CO-); MS (ESI-TOF, $CHCl_3$/MeOH 10:15 v/v): *m/z*=290.0610 $[M-H]^-$ (calcd for $C_{14}H_{13}NO_6$ –H 290.0665).

N-Hydroxysuccinimide ester of trans-caffeic acid (3).



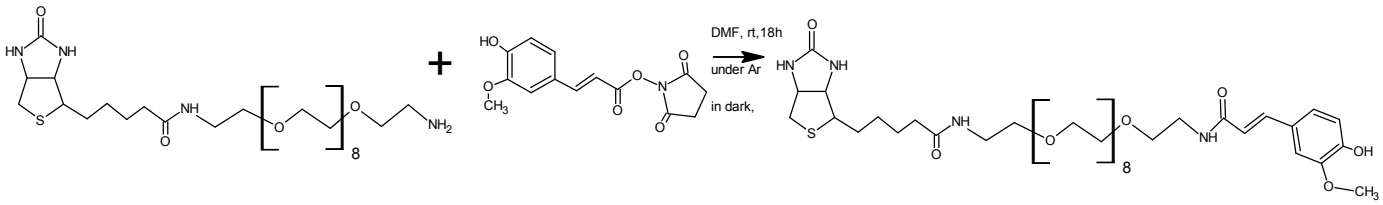
Solution of caffeic acid (220 mg, 1.22 mmol) and N-hydroxysuccinimide (140 mg, 1.22 mmol) in dimethylformamide (DMF) (2 mL) was stirred under argon until complete dissolution. Then a solution of N,N'-dicyclohexylcarbodiimide (DCC) (277 mg, 1.34 mmol) in DMF (2 mL) was added. The reaction mixture was stirred under argon in dark for 24 hours. The precipitate of dicyclohexylurea (DCU) was removed by filtration, the filtered liquid was concentrated in vacuum and the residue was chromatographed on Silica gel 60 (column 3 x 13 cm, elution with chloroform: methanol 95:5) and dried in vacuum exsiccator over sodium hydroxide to yield 112 mg of the target product (33.2%). Rf 0.5 (Silica gel, chloroform:methanol 9:1). ^{1}H NMR (300 MHz; DMSO-D_6 ; Me_4Si): δ_H, ppm 7.76 (d, J=16 Hz, 1H, -C**H**=CH-), 7.19 (m, 2H, -C_6H_3-), 6.81 (d, J=7.92 Hz, 1H, -C_6H_3-), 6.57 (d, J=16 Hz, 1H, -CH=C**H**-), 2.83 (m, 4H, -CO-C**H$_2$**-C**H$_2$**-CO-); MS (ESI-TOF, $CHCl_3$/MeOH 10:15 v/v): *m/z*=276.0408 $[M-H]^-$ (calcd for $C_{13}H_{11}NO_6$ 276.0508).

O-[2-(p-Coumaroylamino)ethyl]-O'-[2-(biotinylamino)ethyl]octaethylene glycol (4).



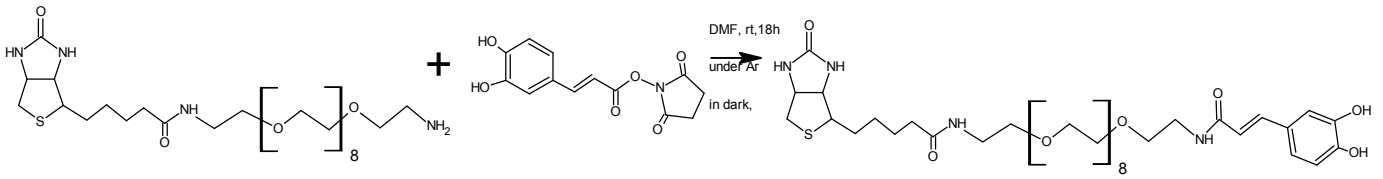
Solution of the compound **1** (N-hydroxysuccinimide ester of trans-p-coumaric acid) (8.34 mg, 0.032 mmol) in DMF (1 mL) was added to a solution of O-(2-Aminoethyl)-O'-[2-(biotinylamino)ethyl]octaethylene glycol (19 mg, 0.028 mmol) in DMF (0.6 mL). The reaction mixture was stirred under argon in dark for 18 hours. The solvent was removed in vacuum, and the residue was purified by preparative TLC on aluminum oxide (4 sheets 20×10 cm) in chloroform: methanol 7:1. Yield: 11.3 mg (49%); R_f 0.6 (Al_2O_3, chloroform: methanol 7:1). ^{1}H NMR (300 MHz; $CDCl_3$; Me_4Si): δ_H, ppm 7.52 (d, J=15.5 Hz, 1H, -C**H**=CH-), 7.37 (d, J=8.8 Hz, 2H, -C_6H_4-), 7.19 (m, 1H, PEG-N**H**-CO-), 7.05 (m, 1H, PEG-N**H**-CO-), 6.85 (d, J=8.5 Hz, 2H, -C_6H_4-), 6.39 (d, J=15.5 Hz, 1H, -CH=C**H**-), 6.29 (br.s, 1H, -N**H**-CO-NH-in biotin), 5.53 (m, 1H, -NH-CO-N**H**-in biotin), 4.46 (m, 1H, >C**H**-CH< in biotin), 4.26 (m, 1H, >CH-C**H**< in biotin), 3.59 (m, 40H, -C**H$_2$**- in PEG), 3.08 (m, 1H, -S-C**H**< in biotin), 2.80 (m, 2H, -S-C**H$_2$**-CH< in biotin), 2.2 (t, 2H, -C**H$_2$**-CO-NH-PEG-), 1.35-1.75 (m, 6H, -C**H$_2$**-C**H$_2$**-C**H$_2$**-CH_2-CO- in biotin); ^{13}C NMR (75Hz, $CDCl_3$): 173.94, 167.08, 164.16, 158.89, 140.57, 129.71, 126.91, 118.40, 116.25, 77.68, 77.46, 77.26, 76.84, 70.63, 70.32, 70.11, 62.03, 60.43, 55.73, 51.05, 40.83, 39.69, 39.46, 36.02, 28.40, 28.25, 25.82, 25.43; MS (ESI-TOF, $CHCl_3$/MeOH 10:15 v/v): *m/z*=851.4140 $[M+Na]^+$ (calcd for $C_{39}H_{64}N_4O_{13}SNa$ 851.4088).

O-[2-(Feruloylamino)ethyl]-O'-[2-(biotinylamino)ethyl]octaethylene glycol (5).



Solution of the compound **2** (N-hydroxysuccinimide ester of trans-ferulic acid) (10.2 mg, 0.035 mmol) in DMF (1 mL) was added to a solution of O-(2-Aminoethyl)-O'-[2-(biotinylamino)ethyl]octaethylene glycol (19 mg, 0.028 mmol) in DMF (0.6 mL). The reaction mixture was stirred under argon in dark for 18 hours. The solvent was removed in vacuum, and a residue was purified by preparative TLC on Silica gel (6 sheets) in chloroform – methanol 7:1. Yield: 11.3 mg (47.3%); R_f 0.18 (silica gel, chloroform : methanol 7:1). ^{1}H NMR (300 MHz; $CDCl_3$; Me_4Si): δ_H, ppm 7.53 (d, J=15.5 Hz, 1H, -C**H**=CH-), 6.87-7.1 (m, 3H, -C_6H_3-), 6.7-6.85 (m, 2H, -CO-N**H**-PEG-N**H**-CO-), 6.4 (d, J=15.5 Hz, 1H, -CH=C**H**-), 6.07 (s, 1H, -N**H**-CO-NH-in biotin), 5.4 (s, 1H, -NH-CO-N**H**-in biotin), 4.51 (m, 1H, >C**H**-CH< in biotin), 4.3 (m, 1H, >CH-C**H**< in biotin), 3.91 (s, 3H, -C_6H_3-OC**H$_3$**), 3.4-3.7 (m, 40H, -C**H$_2$**- in PEG), 3.12 (m, 1H, -S-C**H**< in biotin), 2.7-2.93 (m, 2H, -S-C**H$_2$**-CH< in biotin), 2.2 (t, 2H, -C**H$_2$**-CO-NH-PEG-), 1.37-1.75 (m, 6H, -C**H$_2$**-C**H$_2$**-C**H$_2$**-CH_2-CO- in biotin); MS (ESI-TOF, $CHCl_3$/MeOH 10:15 v/v): *m/z*=881.4238 $[M+Na]^+$ (calcd for $C_{40}H_{66}N_4O_{14}SNa$ 881.4194).

O-[2-(Caffeoylamino)ethyl]-O'-[2-(biotinylamino)ethyl]octaethylene glycol (6).



A solution of the compound **3** (N-hydroxysuccinimide ester of trans-caffeic acid) (10 mg, 0.036 mmol) in DMF (1 mL) was added to a solution of O-(2-Aminoethyl)-O'-[2-(biotinylamino)ethyl]octaethylene glycol (19 mg, 0.028 mmol) in DMF (0.6 mL). The reaction mixture was stirred under argon in dark for 18 hours. The solvent was removed in vacuum, and a residue was purified by preparative TLC on Silica gel (6 sheets) in chloroform – methanol – acetic acid 7:1: 0.05. Yield:9.8 mg (41.7 %); R_f 0.36 (silica gel, chloroform : methanol 5:1). ^{1}H NMR (300 MHz; $CDCl_3$; Me_4Si): δ_H, ppm 7.45 (d, J=15.5 Hz, 1H, -C**H**=CH-), 7.31 (m, 1H, -CO-N**H**-PEG-NH-CO-), 6.55-7.22 (m, 3H, -C_6H_3-), 6.45 – 6.63 (br.,1H, -CO-NH-PEG-N**H**-CO-), 6.36 (d, J=15.5 Hz, 1H, -CH=C**H**-), 5.20 – 6.0 (br., 2H, -N**H**-CO-N**H**-in biotin), 4.43 (m, 1H, >C**H**-CH< in biotin), 4.23 (m, 1H, >CH-C**H**< in biotin), 3.37-3.7 (m, 40H, -C**H$_2$**- in PEG), 3.09 (m, 1H, -S-C**H**< in biotin), 2.6-2.87 (m, 2H, -S-C**H$_2$**-CH< in biotin), 2.18 (t, 2H, -C**H$_2$**-CO-NH-PEG-), 1.32-1.7 (m, 6H, -C**H$_2$**-C**H$_2$**-C**H$_2$**-CH_2-CO- in biotin); ^{13}C NMR (75Hz, $CDCl_3$): 174.06, 167.29, 164.52, 147.75, 145.66, 143.93, 140.99, 136.76, 127.45, 113.73, 104.98, 77.66, 77.44, 77.24, 76.81, 70.51, 70.43, 70.33, 70.26, 62.09, 60.47, 55.78, 50.74, 45.35, 40.77, 39.56, 39.33, 35.96, 35.88, 28.38, 28.18, 25.76, 23.16; MS (ESI-TOF, $CHCl_3$/MeOH 10:15 v/v): *m/z*=867.4041 $[M+Na]^+$ (calcd for $C_{39}H_{64}N_4O_{14}SNa$ 867.4037).

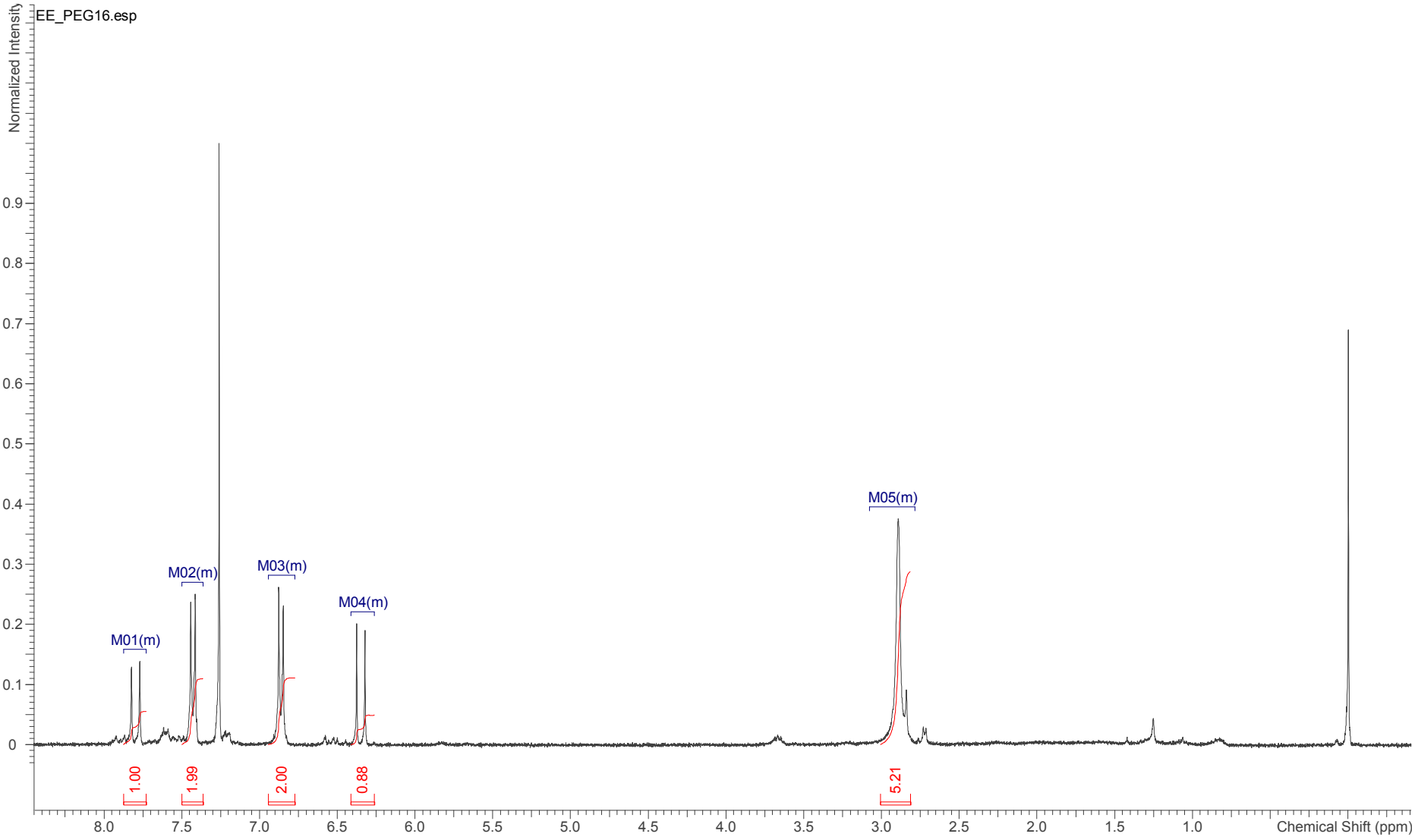


Fig.1. The ^{1}H NMR-spectrum of N-hydroxysuccinimide ester of *trans*-p-coumaric acid in $CDCl_3$.

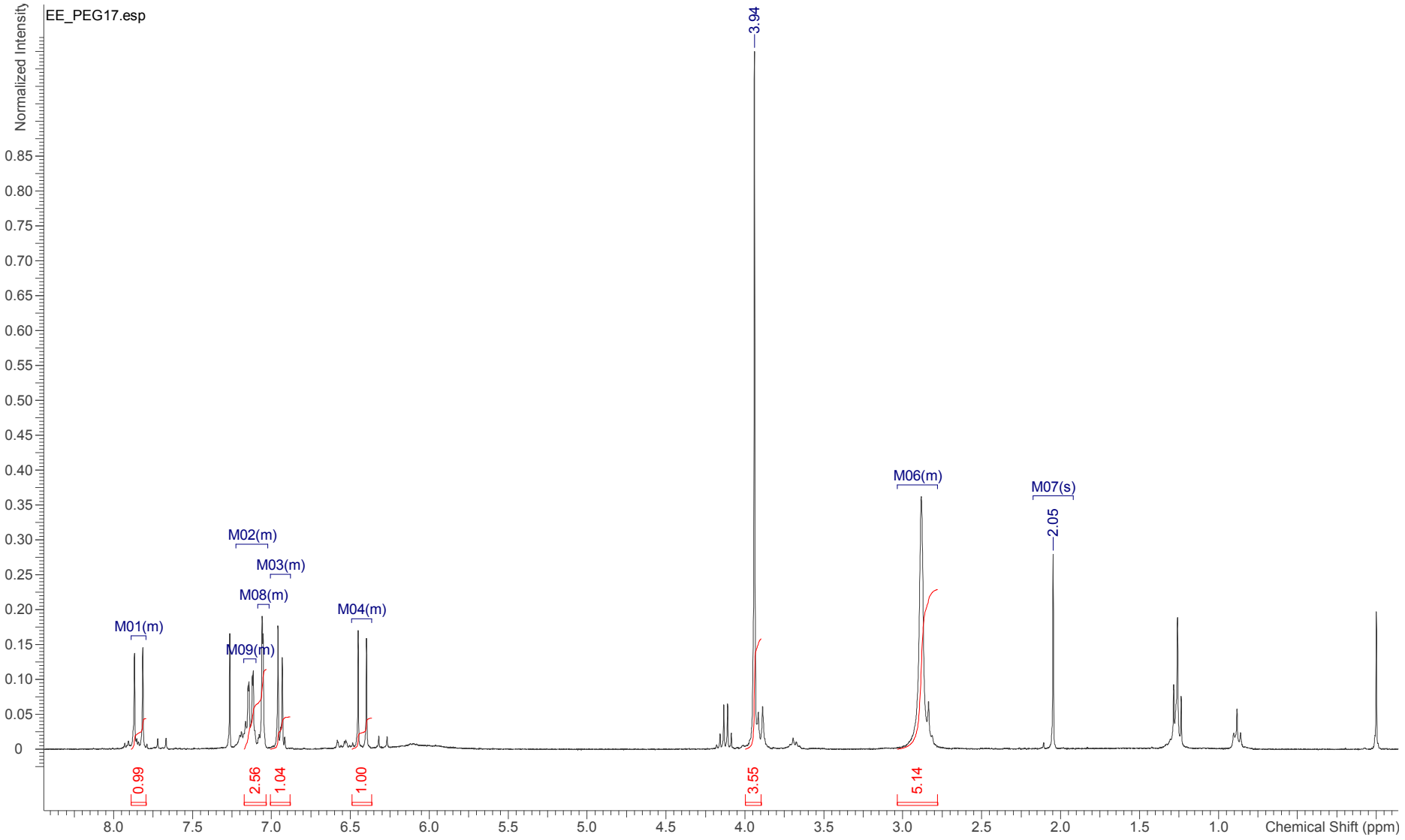


Fig.2. The ^{1}H NMR-spectrum of N-hydroxysuccinimide ester of *trans*-ferulic acid in $CDCl_3$.

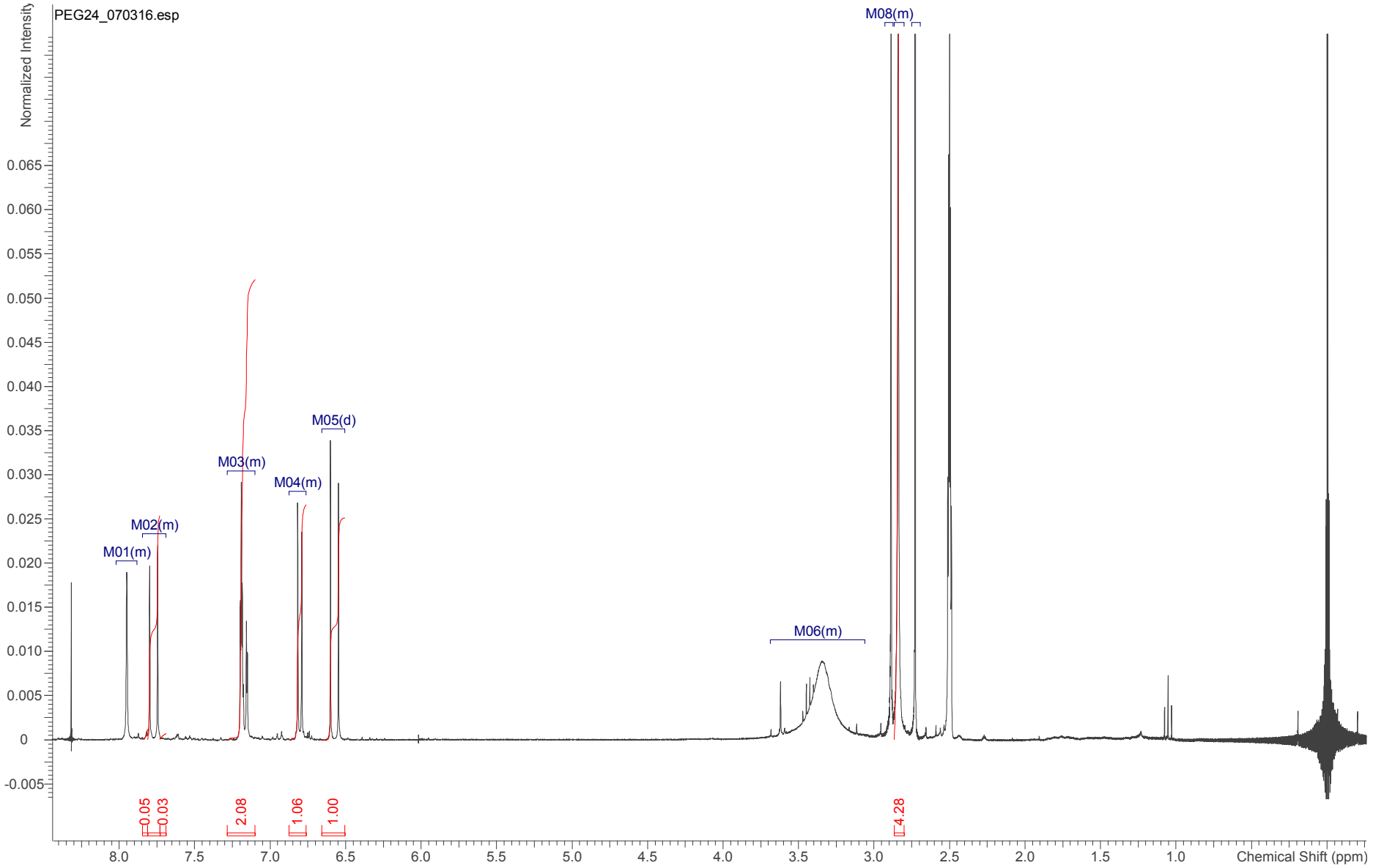

PEG24_070316.esp
Normalized Intensity
M08(m)
M05(d)
M03(m)
M04(m)
M02(m)
M01(m)
M06(m)
0.05
0.03
2.08
1.06
1.00
4.28
Chemical Shift (ppm)

Fig.3. The ^{1}H NMR-spectrum of N-hydroxysuccinimide ester of *trans*-caffeic acid in DMSO-d_6.

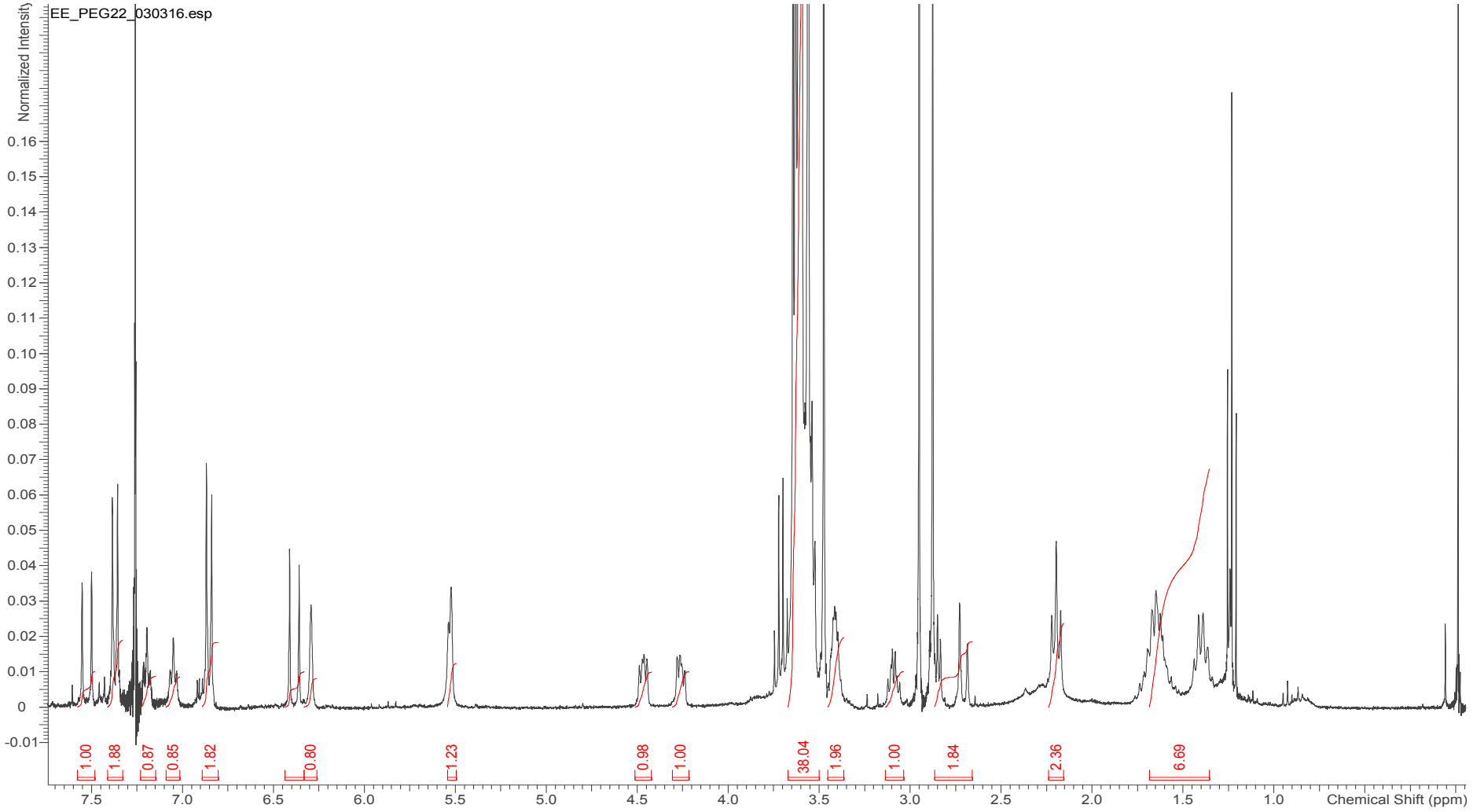


Fig.4. The ^{1}H NMR-spectrum of O-[2-(p-Coumaroylamino)ethyl]-O'-[2-(biotinylamino)ethyl]octaethylene glycol in $CDCl_3$.

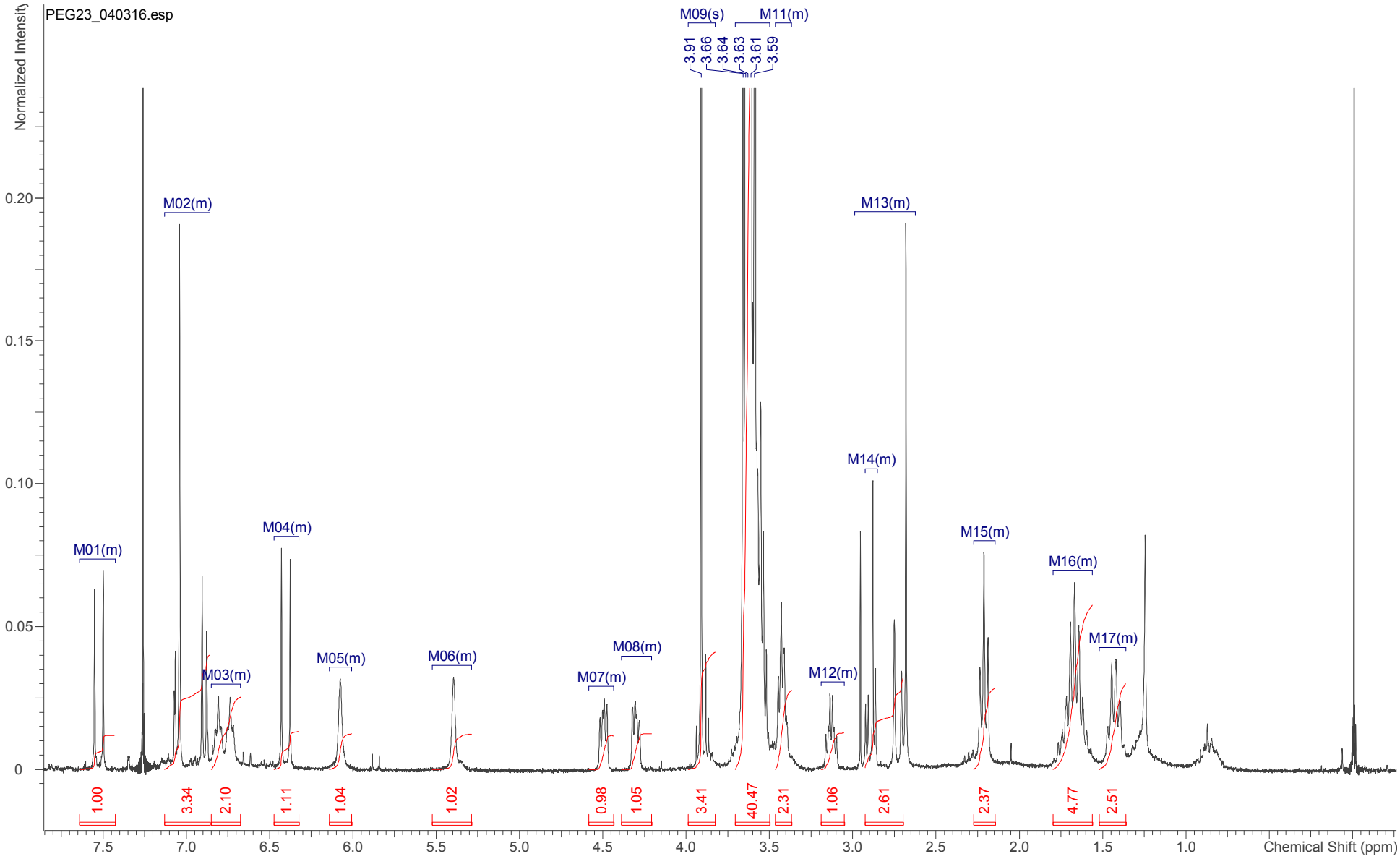


Fig.5. The ^{1}H NMR-spectrum of O-[2-(Feruloylamino)ethyl]-O'-[2-(biotinylamino)ethyl]octaethylene glycol in $CDCl_3$.

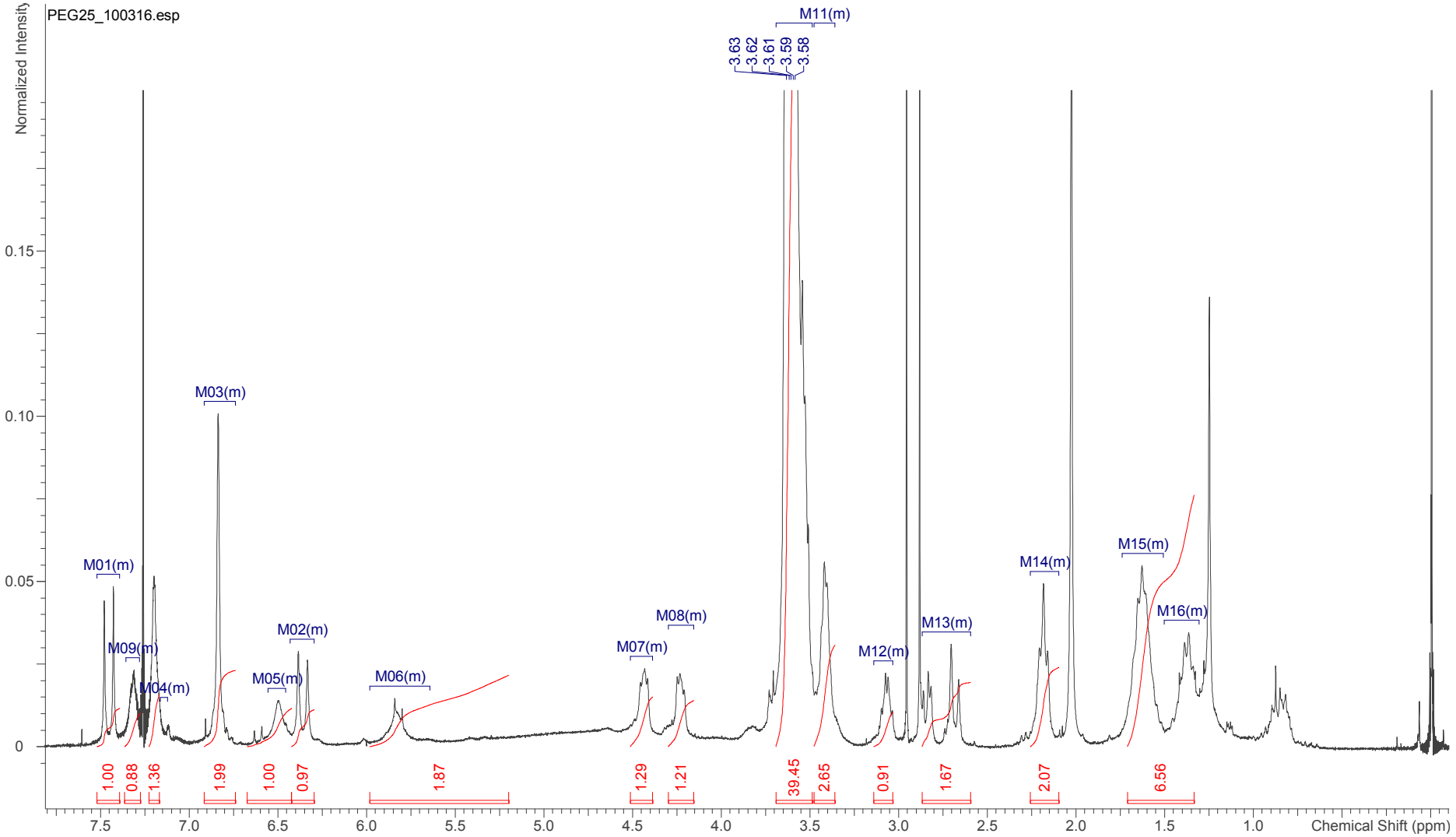


Fig.6. The ^{1}H NMR-spectrum of O-[2-(Caffeoylamino)ethyl]-O'-[2-(biotinylamino)ethyl]octaethylene glycol in $CDCl_3$.

Biopanning

In the panning experiments, all washing, elution and incubation steps, unless noted otherwise, were conducted in final volume of 500 µL in TBT-0.05 (50 mM Tris-HCl, 150 mM NaCl 1% BSA, 0.05% Tween-20, pH 7.5). Experiments were conducted at room temperature unless noted otherwise. For bacterial colony growth Luria Agar (LA) plates were used with appropriate supplementations.

An affinity-based selection was conducted on the scFvM library in three consecutive panning rounds using biotinylated target molecules immobilized on superparamagnetic beads. The magnetic beads were prewashed three times before use and the panning experiments were conducted separately for each of the target antigen in a similar manner. A negative selection was conducted prior to each panning round using biotin-coated beads without the target molecule to reduce unspecific binding to biotin. For the first round, an initial panning input of $6x10^{12}$ cfu/mL scFvM phages were incubated with 0.3 mg/mL of streptavidin (SA)-coated magnetic beads (Dynabeads MyOne Streptavidin T1, Invitrogen, USA) saturated with 100 µM of biotin for 1 h with rotation. The beads were then collected by a magnet and the separated eluate was used for the panning experiments.

The actual panning experiments were conducted on biotinylated target molecules immobilized on a fresh batch of prewashed beads. For immobilization, 0.3 mg/mL beads were incubated with 600 nM of biotinylated target molecules and incubated for 30 min on rotation. The beads were blocked by adding a final concentration of 100 µM of biotin and incubating for 5 min on rotation. Then the beads were washed thrice, the eluate from the negative selection was added on the beads and incubated for 1 h with rotation. Next, the beads were washed twice with of TBT-0.05 and once with of TSAT-0.05 (50 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20, 0.02 % NaN_3 pH 7.5). The tube was changed after each wash. The phages were eluted from the beads by adding 60 µg/ml of trypsin (Sigma-Aldrich, USA) diluted in TBS and incubated for 0.5 h. Trypsin activity was stopped by adding 50 µg/ml of soy bean trypsin inhibitor (Sigma-Aldrich) in to the solution. The beads were collected and the eluate was moved to a new tube.

The phages were recovered by infecting *E. coli* XL1-Blue cells. After growing 3 ml of cells to logarithmic growth phase, the phage eluate was added and incubated statically at 37 °C for 30 min. The panning output was determined by spreading 100 µl of serially diluted infected culture on LA plates (0.5% w/v glucose, 10 µg/ml tetracycline, 25 µg/ml chloramphenicol) and incubated at 37 °C overnight. In addition, the rest of the cultures were plated on large LA plates supplemented as previously. After colony formation, the cells were scraped from the plate in to 15 ml of SB. The scraped cells were used to inoculate 20 ml of fresh SB medium (1% w/v glucose, 10 µg/ml tetracycline, 25 µg/ml chloramphenicol) with a starting OD_{600} of 0.1. After reaching logarithmic growth phase, the cells were infected with $2.5x10^9$ cfu/mL of helper phage VCS M13, mixed gently and incubated statically at 30 °C. Then the cells were pelleted, the supernatant discarded and fresh SB medium (10 µg/ml tetracycline, 25 µg/ml chloramphenicol) without glucose was added. After 1 h of incubation at 30 °C and 300 rpm, 50 µg/ml kanamycin and 100 µM IPTG was added. The incubation was continued overnight at 26 °C and 300 rpm and phage stocks were prepared from these cultures. First, the cell culture supernatant was precipitated by adding 3.3% PEG and 0.4 M NaCl (final conc.) and incubated on ice for 60 minutes. Then, the mixture was centrifuged (15 min, 10 000 g, 4 °C) and the supernatant was removed. The remaining pellet was suspended in 1 ml of TSA/BSA (50 mM Tris-HCl, 150 mM NaCl 1% BSA, pH 7.5) Finally, the stock concentrations were determined by plate titering and used for the next panning round.

The second panning round was performed the same way, except the biotinylated hapten concentration was decreased to 100 nM and 0.1 mg/mL Dynabeads M-280 Tosylactivated coated with neutravidin (Thermo Scientific, USA) were used. The phage concertation was decreased to 10^{11} cfu/mL, trypsin elution was conducted at a final volume of 200 uL, and 1 ml of log-phase culture was used for XL1 infection. The infected cells were plated on 15 ø cm plates and scraped with 5 ml of SB. The phage production, infection and stock preparations were conducted as described for the first round. The third round was conducted as the second except 50 µg/mL of SA beads (M-280 Streptavidin, Invitrogen)

were used and the phage concentrations were $2.2x10^{10}$ for ferulate binders and $3.8x10^{10}$ for coumarate and caffeate binders.

Cloning, expression, lysate preparation and protein purification

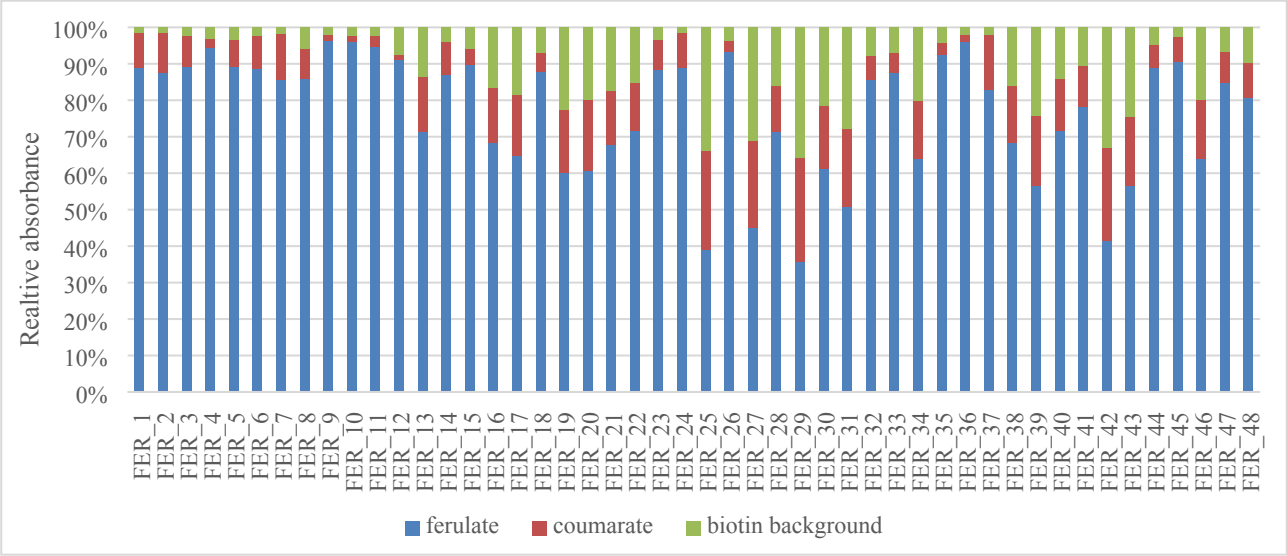
A random pool of genes coding for the scFvs from the pEB32x vector collected from the 3rd panning round were cloned to an expression vector pLK06H**.** To isolate a random pool of the scFv genes, 1 ml of infected XL1 cells were collected and the plasmids extracted with GeneJET Plasmid Miniprep Kit (Fermentas, Lithuania). The scFvs were digested with SfiI (Thermo Scientific) and cloned to the expression plasmid pLK06H to form a scFv-AP fusion. To obtain non-fusion forms of the scFvs, the fragments were cloned to pLK04H (without an AP site). The ligated plasmids were transformed to XL1 cells via electroporation and selected on LA plates (tetracycline 10 µg/mL, ampicillin 100 µg/mL). The clones were verified with restriction analyses.

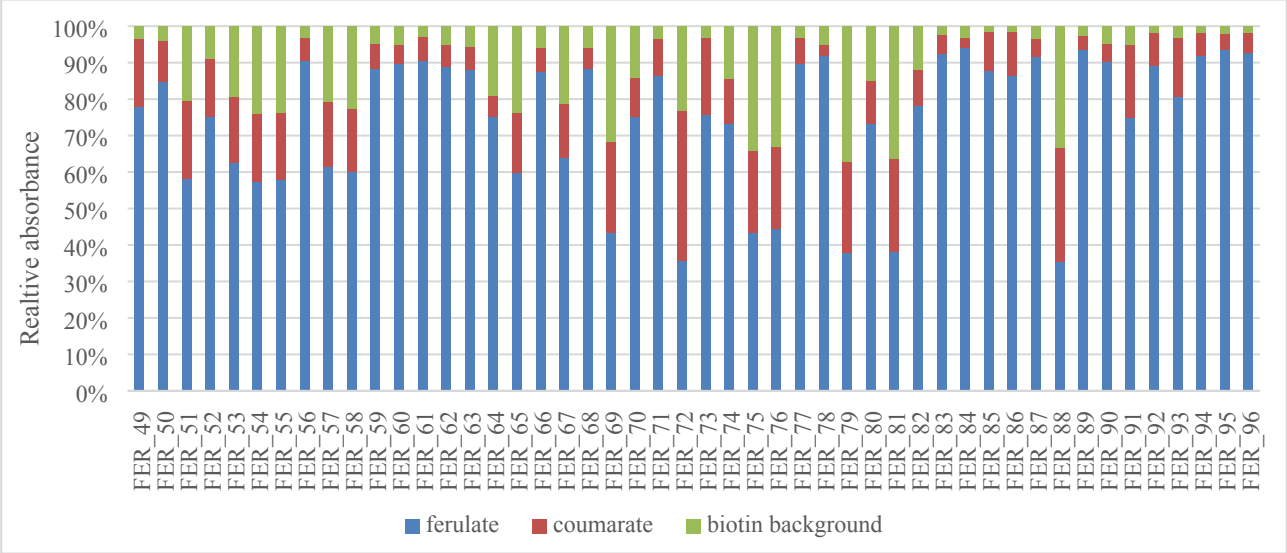
For primary screening of the produced recombinant antibodies, single colonies (100-200 colonies per target hapten) were picked from the transformant plates and grown in 96-well plates containing 200 µL of SB medium with glucose (1%) and appropriate antibiotics supplementation. The plates were sealed with breathable sealing tape and incubated for 16-18 hours at 37 °C with 900 rpm and 70% moisture. Then, the medium was exchanged for glucose free SB, the cells were grown up to log-phase and induced with 0.1-0.2 mM IPTG. After induction, the temperature was decreased to RT or 22 °C and incubation continued for another 16-18 hours. After incubation, the cells were lysed directly with lysis solution (1 mg/mL lysosome from chicken egg white lysosome, Sigma, 2.5 U/mL Benzonase (Merck, USA) in TBS [5 mM Tris-HCl, 15 mM NaCl] pH 7.5). Lysis incubation was set at room temperature with shaking for 30 min. Next, the cells were freeze-thawed thrice, lysate was centrifuged and the supernatant containing the AP-scFvs was used for primary screening assays. Lysate for secondary screening was obtained as above, except the cultivation volume was increased to 5 ml,

shaking adjusted to 300 rpm, and no moisture control was used. Additionally, lysis solution was added on pelleted cells in a final volume of 1 ml.

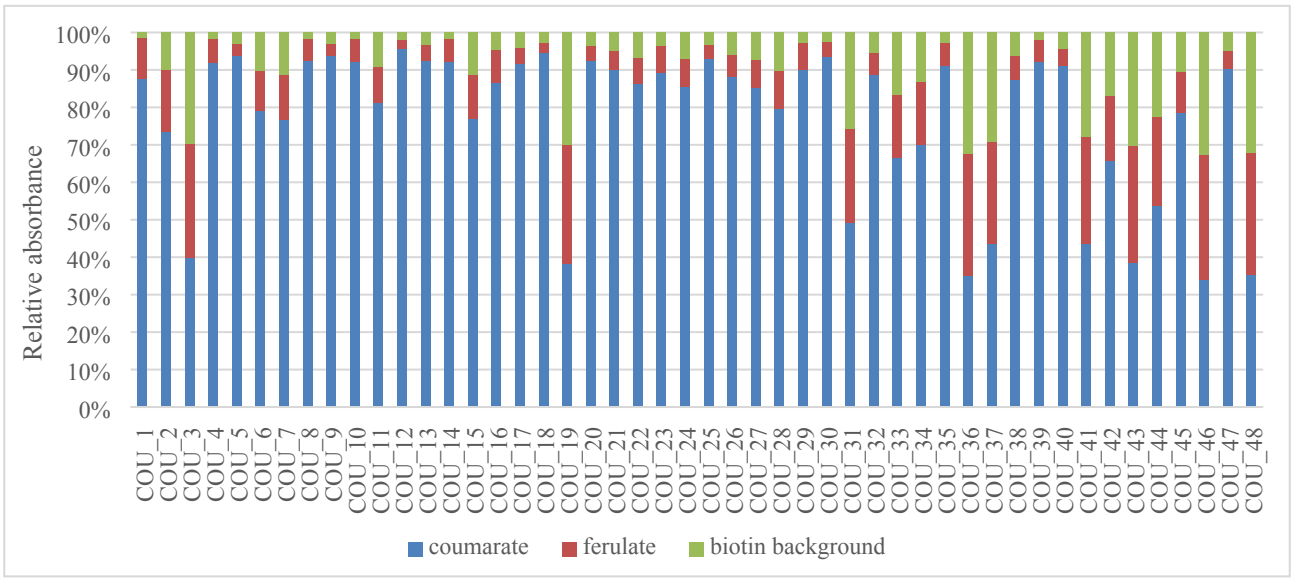
Recombinant antibody expression of selected clones for protein purification was conducted similarly to the secondary screening, except in 50 ml of SB supplemented with 0.4% glucose and 100 µg/mL ampicillin. For cell lysis, 1 ml of binding buffer containing 5 mM imidazole, 20 mM Tris-HCl, and 0.5 M NaCl was added per 1 g of wet weight of cell pellet. After 30 min of incubation on ice, the samples were sonicated 2 x 30 s and centrifuged for 30 min at 30000 g. The lysate was purified using immobilized metal affinity chromatography (IMAC) (His GraviTrap, GE Healthcare, Sweden) using 40 mM and 1 M imidazole for washing and elution steps, respectively. Purity of the proteins was verified with SDS-PAGE and protein concentrations were determined with Qubit 3.0 fluorometer and assay kit (ThermoFisher Scientific) or NanoDrop 2000 UV-Vis Spectrophotometer (ThermoFischer Scientific) at 280 nm. The samples were filter-sterilized and 0.05% NaN_3 and 0.1% Triton X-100 were added.

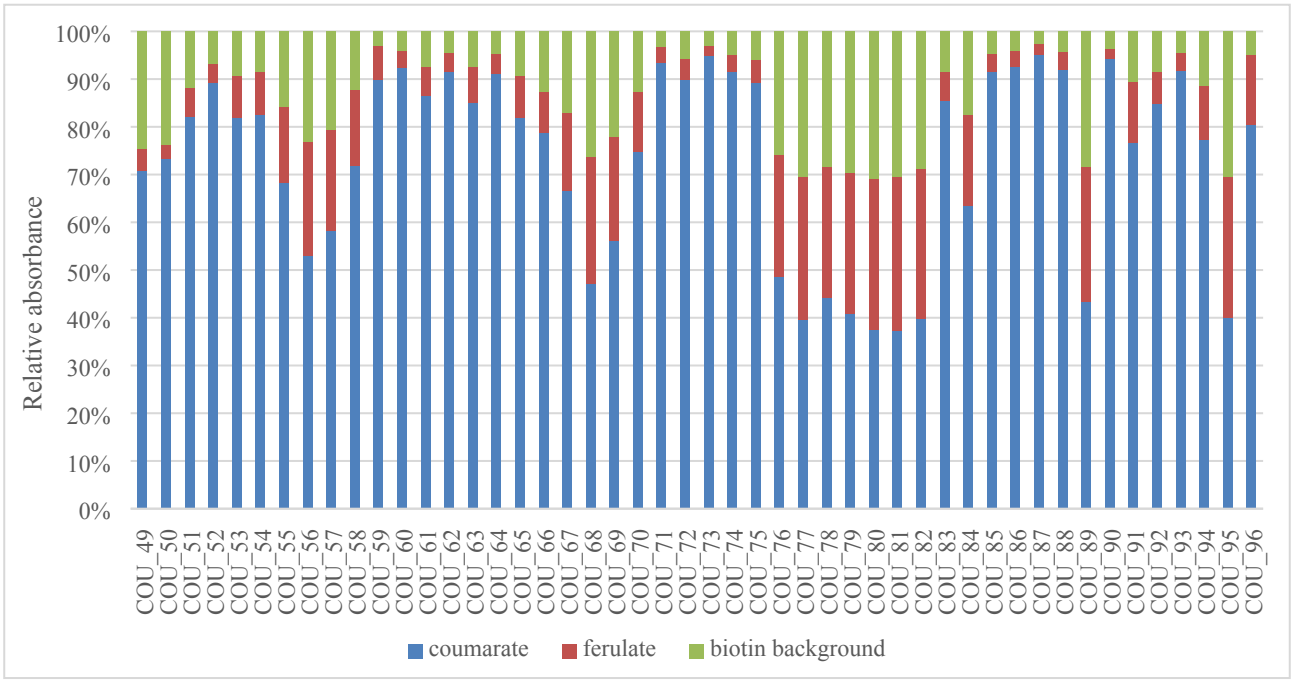
ESI figures for initial screening



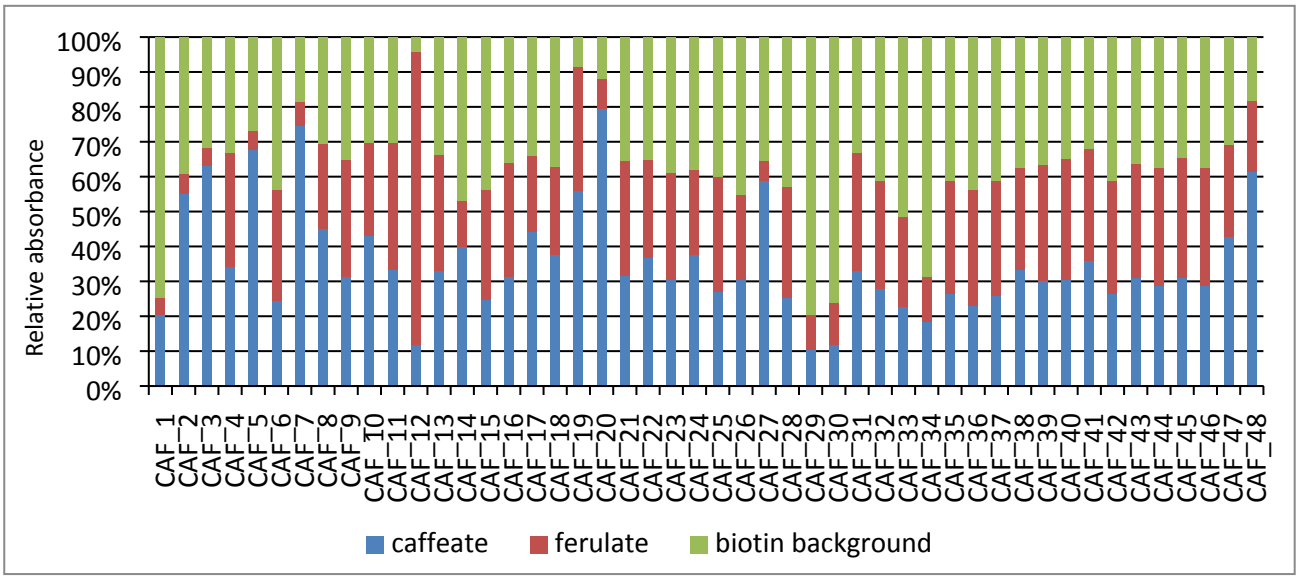


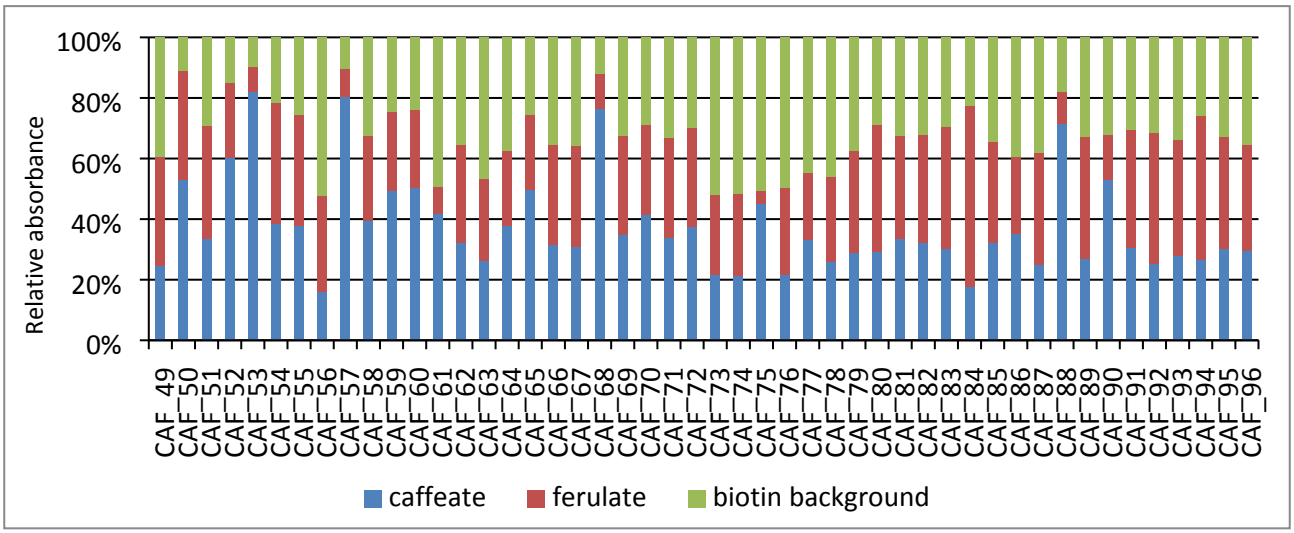
ESI Figure 7: (fragmented to 2 separate figures for clarity) The initial screening result of binders targeted against ferulate. Blue = wells saturated with biotinylated ferulate, red = wells saturated with biotinylated coumarate, green = wells saturated with biotinylated biotin (background). The figures show the relative proportion of the obtained signal for the target molecule, cross binding molecule and background signal.

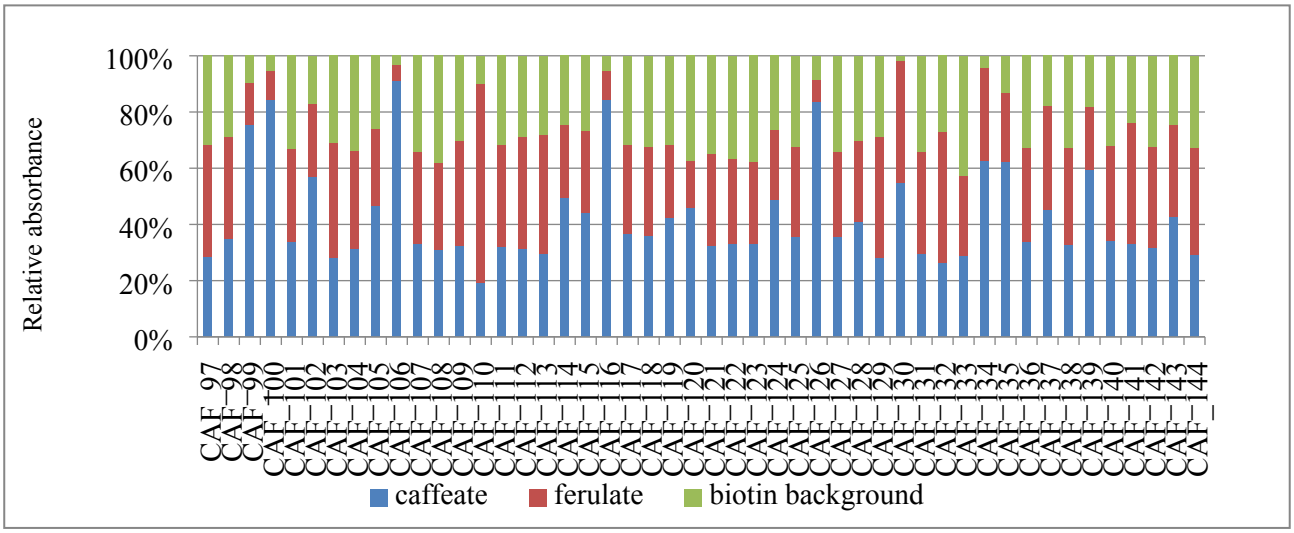


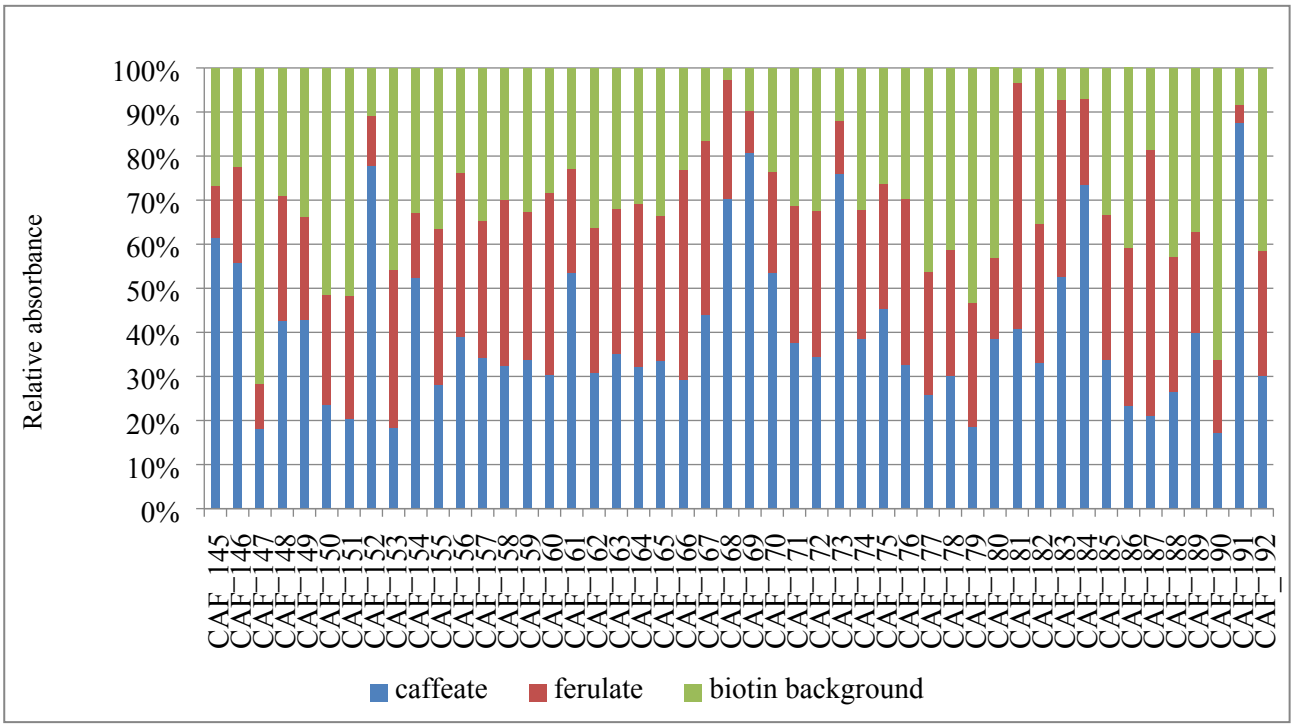


ESI Figure 8: (fragmented to 2 separate figures for clarity) The initial screening result of binders targeted against coumarate. Blue = wells saturated with biotinylated coumarate, red = wells saturated with biotinylated ferulate, green = wells saturated with biotinylated biotin (background). The figures show the relative proportion of the obtained signal for the target molecule, cross binding molecule and background signal.


Relative absorbance
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
CAF_1 CAF_2 CAF_3 CAF_4 CAF_5 CAF_6 CAF_7 CAF_8 CAF_9 CAF_10 CAF_11 CAF_12 CAF_13 CAF_14 CAF_15 CAF_16 CAF_17 CAF_18 CAF_19 CAF_20 CAF_21 CAF_22 CAF_23 CAF_24 CAF_25 CAF_26 CAF_27 CAF_28 CAF_29 CAF_30 CAF_31 CAF_32 CAF_33 CAF_34 CAF_35 CAF_36 CAF_37 CAF_38 CAF_39 CAF_40 CAF_41 CAF_42 CAF_43 CAF_44 CAF_45 CAF_46 CAF_47 CAF_48
caffeate
ferulate
biotin background

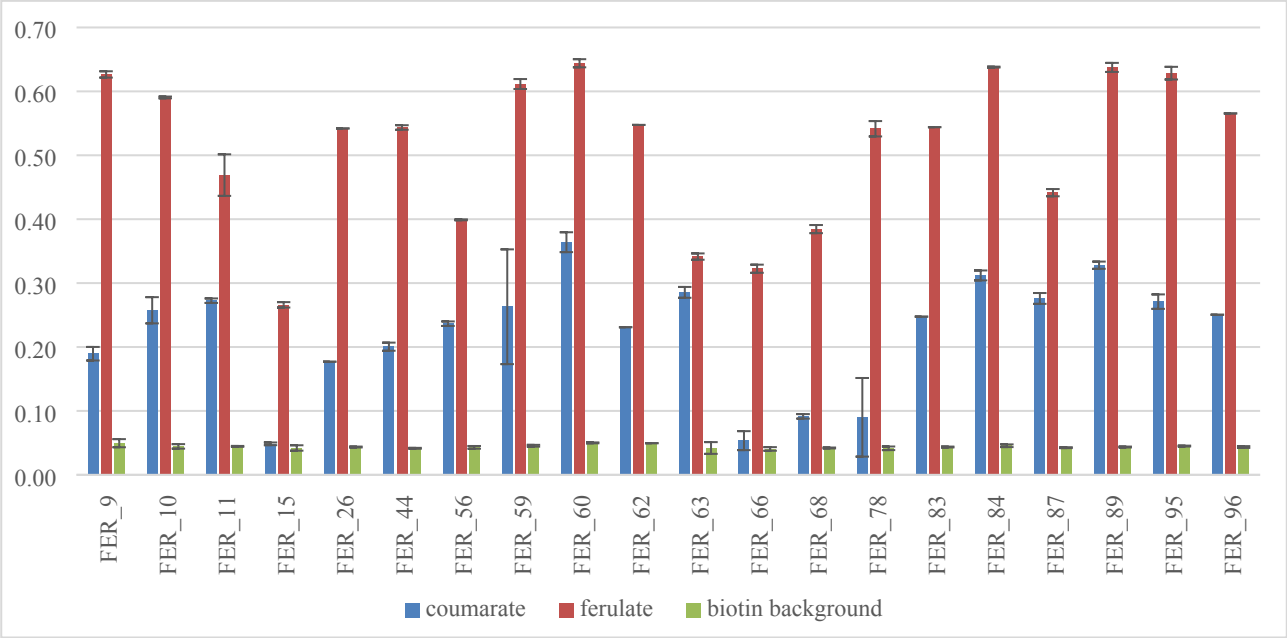

Relative absorbance
100%
80%
60%
40%
20%
0%
CAF_49 CAF_50 CAF_51 CAF_52 CAF_53 CAF_54 CAF_55 CAF_56 CAF_57 CAF_58 CAF_59 CAF_60 CAF_61 CAF_62 CAF_63 CAF_64 CAF_65 CAF_66 CAF_67 CAF_68 CAF_69 CAF_70 CAF_71 CAF_72 CAF_73 CAF_74 CAF_75 CAF_76 CAF_77 CAF_78 CAF_79 CAF_80 CAF_81 CAF_82 CAF_83 CAF_84 CAF_85 CAF_86 CAF_87 CAF_88 CAF_89 CAF_90 CAF_91 CAF_92 CAF_93 CAF_94 CAF_95 CAF_96
caffeate
ferulate
biotin background


Relative absorbance
100%
80%
60%
40%
20%
0%
CAF_97 CAF_98 CAF_99 CAF_100 CAF_101 CAF_102 CAF_103 CAF_104 CAF_105 CAF_106 CAF_107 CAF_108 CAF_109 CAF_110 CAF_111 CAF_112 CAF_113 CAF_114 CAF_115 CAF_116 CAF_117 CAF_118 CAF_119 CAF_120 CAF_121 CAF_122 CAF_123 CAF_124 CAF_125 CAF_126 CAF_127 CAF_128 CAF_129 CAF_130 CAF_131 CAF_132 CAF_133 CAF_134 CAF_135 CAF_136 CAF_137 CAF_138 CAF_139 CAF_140 CAF_141 CAF_142 CAF_143 CAF_144
caffeate
ferulate
biotin background

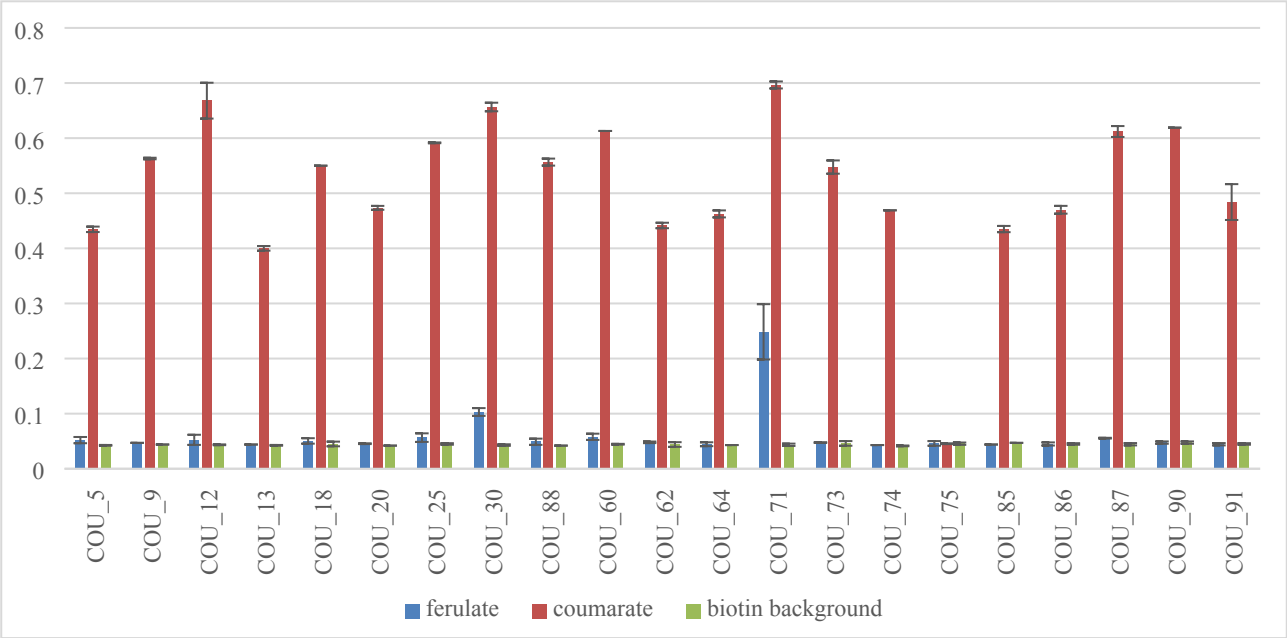


ESI Figure 9: (fragmented to 4 separate figures for clarity) The initial screening result of binders targeted against caffeate: Blue = wells saturated with biotinylated caffeate, red = wells saturated with biotinylated ferulate, green = wells saturated with biotinylated biotin (background). The figures show the relative proportion of the obtained signal for the target molecule, cross binding molecule and background signal.

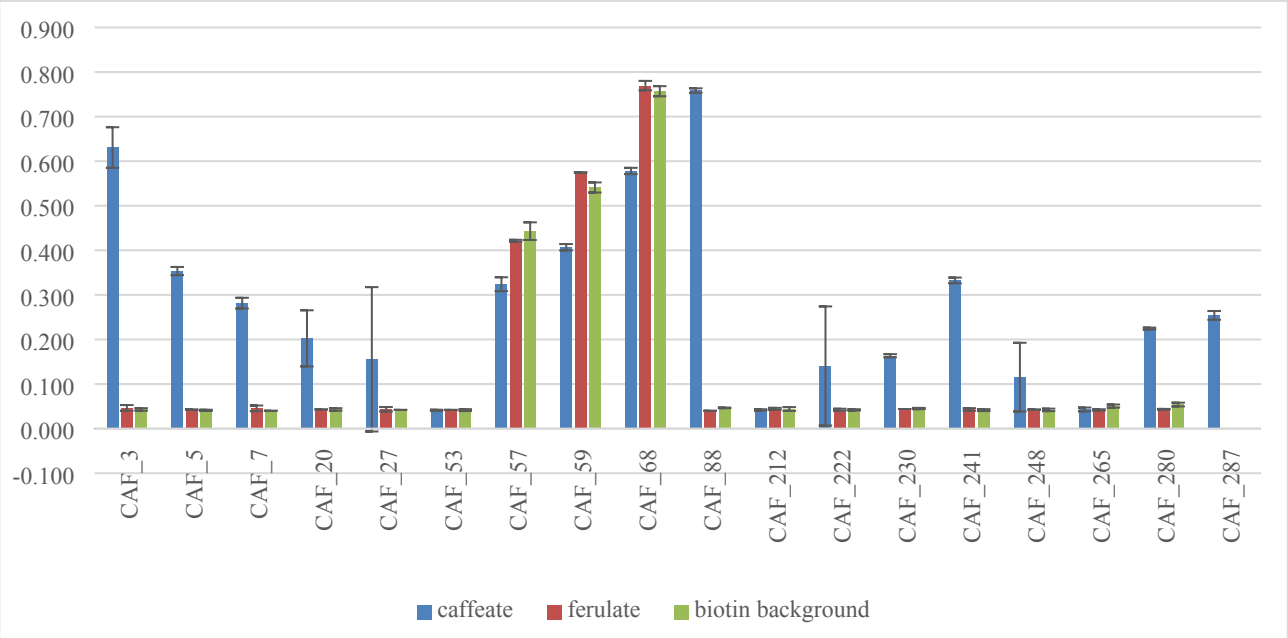
ESI figures of secondary screening of AP fusions with the target molecule and a similar molecule.



ESI Figure 10: Secondary screening results for the binders targeted against ferulate. Red bars = biotinylated target molecule (ferulate), blue bars = biotinylated cross binding molecule (coumarate), green bars = background (biotin). Standard deviation of duplicates is shown as error bars.



ESI Figure 11: Secondary screening results for the binders targeted against coumarate. Red bars = biotinylated target molecule (coumarate), blue bars = biotinylated cross binding molecule (ferulate), green bars = background (biotin). Standard deviation of duplicates is shown as error bars.



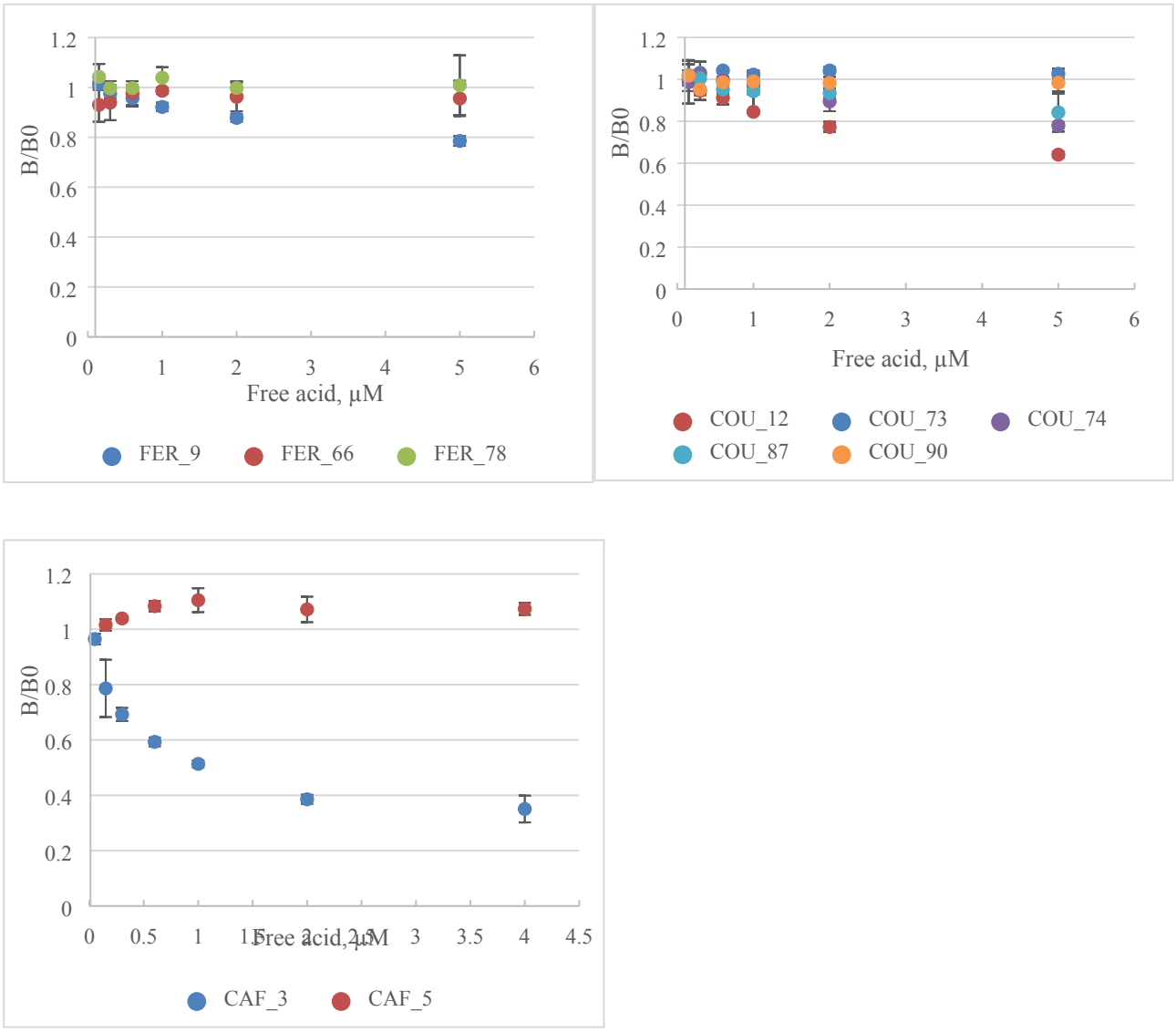
ESI Figure 12: Secondary screening results for the binders targeted against caffeate. Blue bars = biotinylated target molecule (caffeate), red bars = biotinylated cross binding molecule (ferulate), green bars = background (biotin). Standard deviation of duplicates is shown as error bars.

ESI table of sequence data

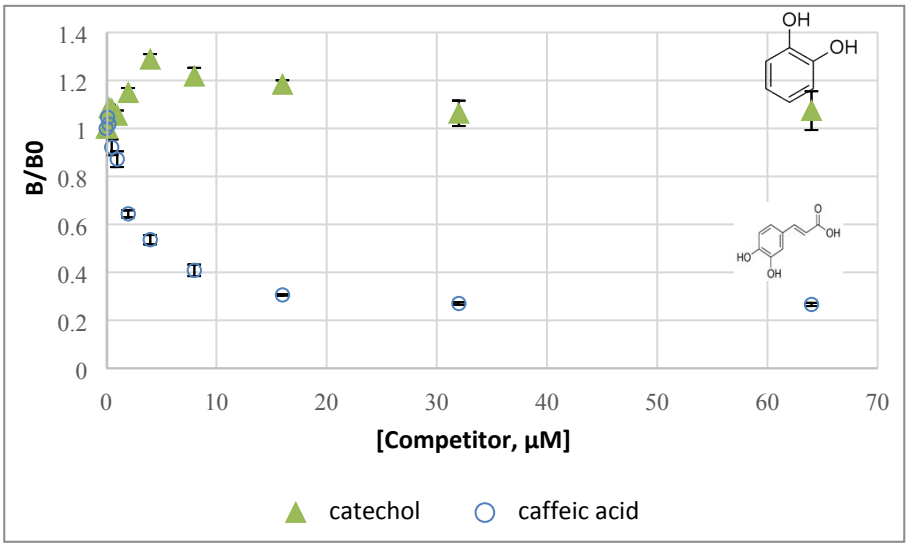
ESI Table 1: Sequencing data of the selected clones.

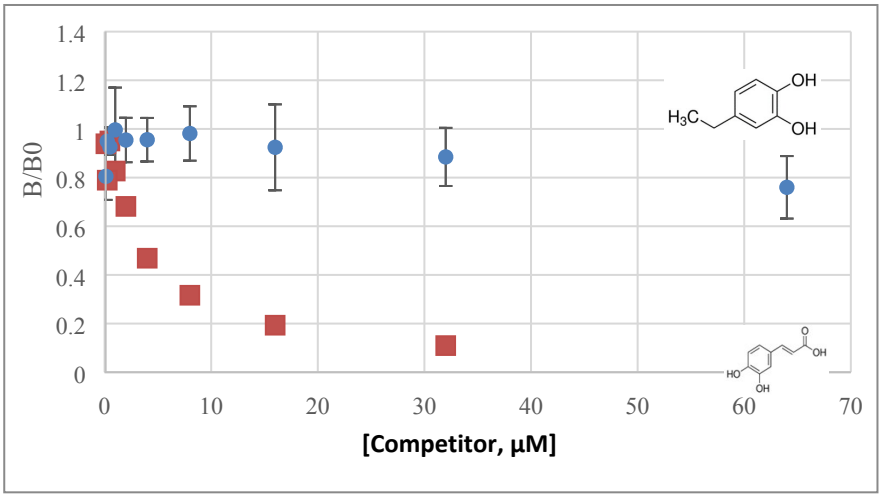
	Loop 1 light chain (CDR-L1)	Loop 3 light chain (CDR-L3)	Loop 3 heavy chain (CDR-H1)	Loop 3 heavy chain (CDR-H2)	AA deletions in VH
FER_9, FER_15	**SSSLH**	HQNNYTPH	**SYA**	**SSITSSGGSTD**	TTYYYIWGQGTL------
FER_26, FER_44, FER_ 66, FER_68,	**SSYLN**	LQYTYDPH	**SYA**	**SGIAPSGGSTY**	RQGFGPWGQGTL------
FER_78	**SSYLH**	HQNNSIPY	**SYG**	**SGIAPSGGSTH**	TSYYYIWGQGTL------
COU_9, COU_90	**SSYLH**	LQHYYFPH	**SYL**	**SSITPSGGST**	SD-------FGIW
COU_12	**SSYLN**	LQHSYYPH	**SYA**	**SSIYPSGGSTN**	AG--------GYGW
COU_13	**SSYLN**	LQYNSIPH	**SYS**	**SSITSSGGSTH**	GG--------AGGW
COU_73	**SSYLH**	LQHTSFPH	**SYS**	**SWITPSGGSTN**	TP-------IGYW
COU_74	**SSSLN**	LQYSYTPH	**SYA**	**SSIAPSGGSTN**	AG--------GYLW
COU_87	**SSYLN**	LQYSYAPH	**SYA**	**SQINPSGGSTN**	AG--------AYLW
CAF_3	**SSSLS**	LQNNSIPF	**SYG**	**SWIDSSGGSTD**	--------VIW
CAF_5, CAF_7, CAF_20, , CAF_57, CAF_241, CAF_280, CAF_287	**SSSLN**	LQGNYNPR	**SYG**	**SRITPSGGSTD**	GRGG------AYW

ESI figures for initial competitive data



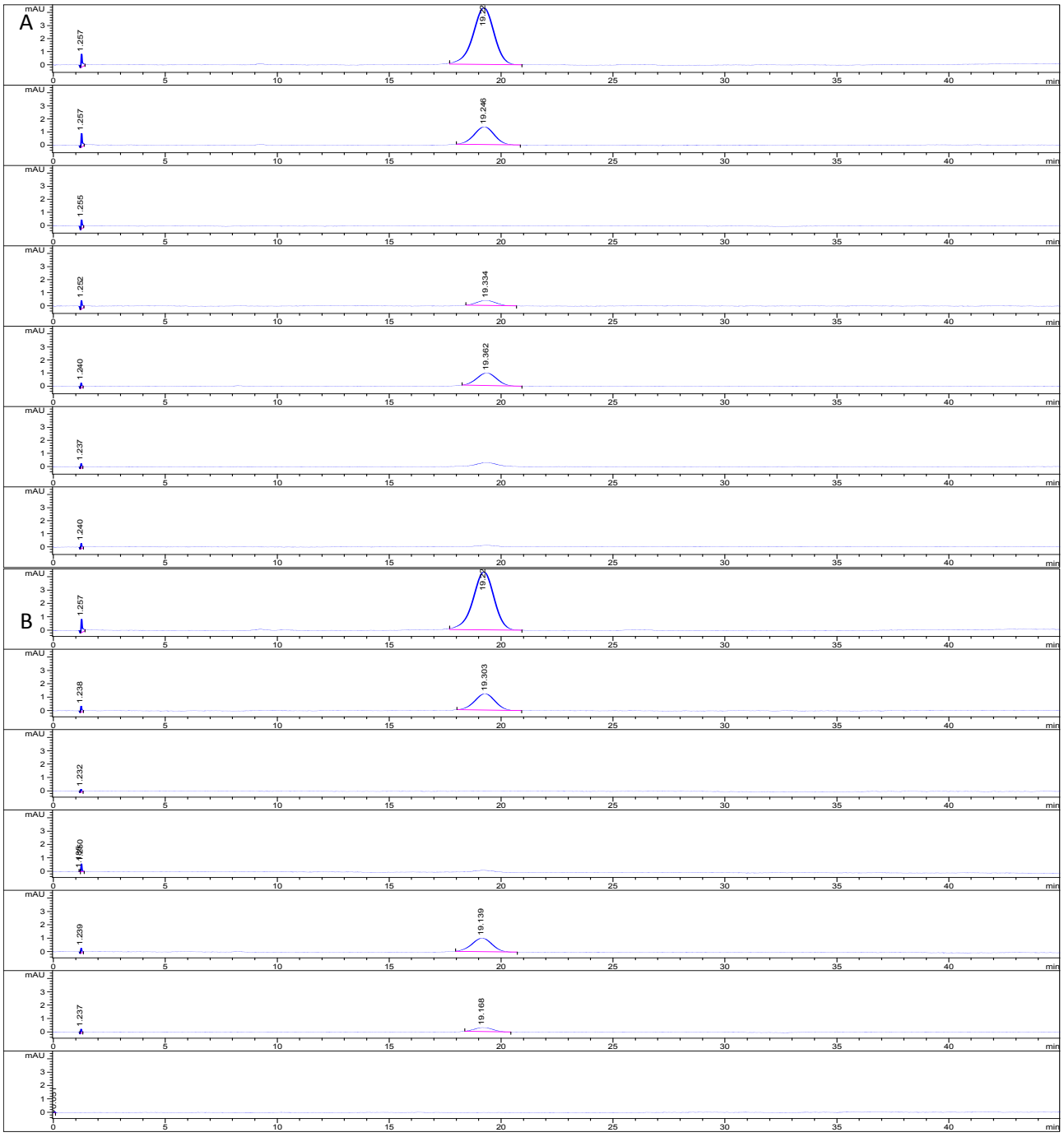
ESI Figure 13: Initial competitive screening with lysates containing ferulate (top left), coumarate (top right) and caffeate (bottom) scFv-APs. The scFv-APs that were further analyzed are shown with a connected line.



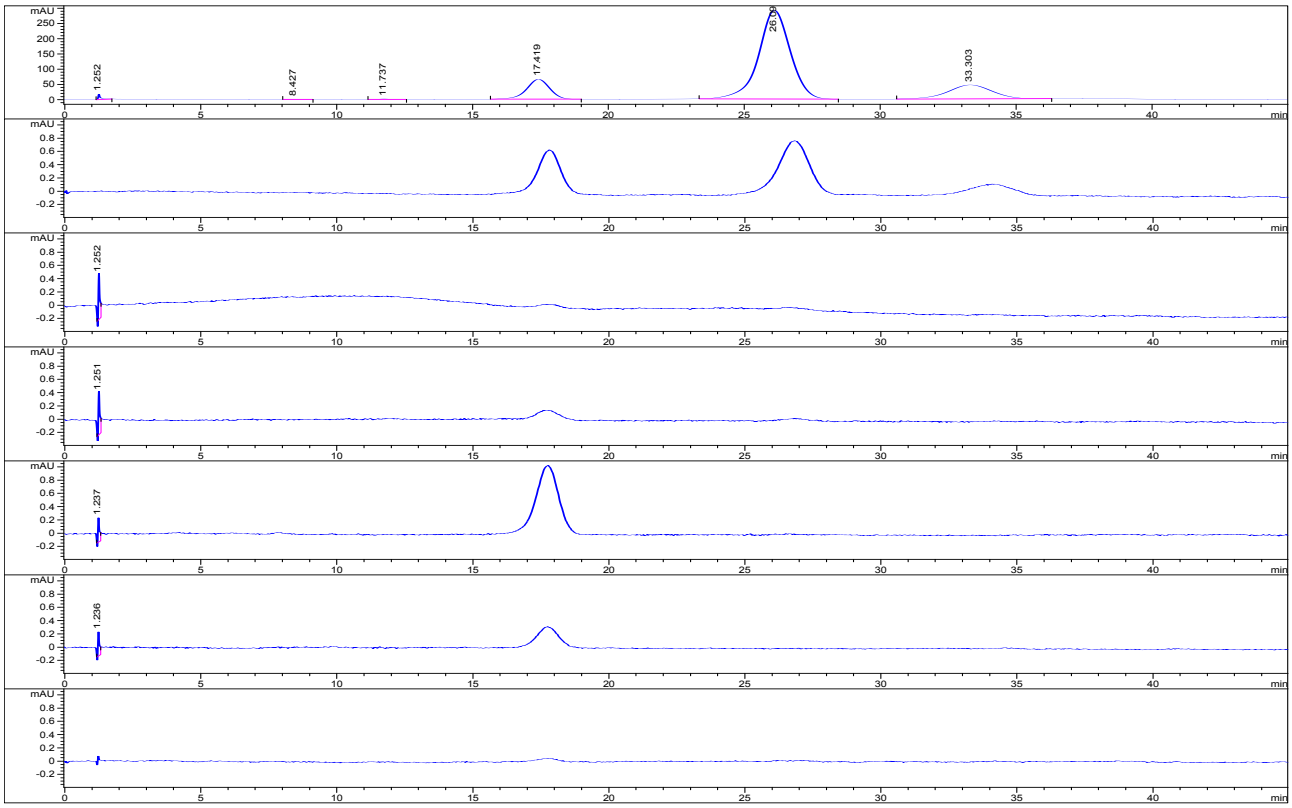


ESI Figure 14: Competitive binding with purified scFv CAF_3. A) 1.2-dihydroxybenene (catechol) and caffeic acid B) 4-ethlycatechol and caffeic acid. Signal intensity is shown as relative absorbance against maximum signal (B/B0). The maximum signal was obtained from an assay without a competitor.

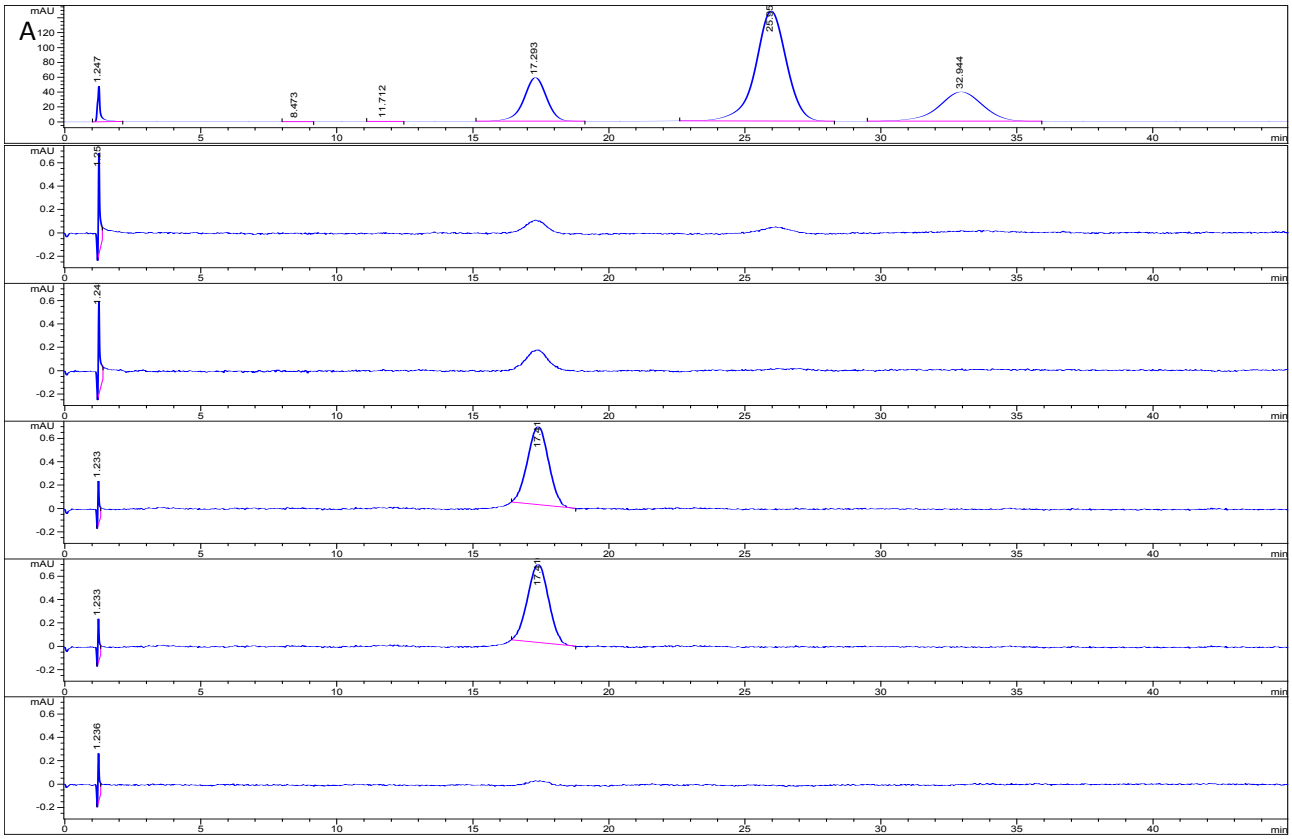
HPLC data for column purification



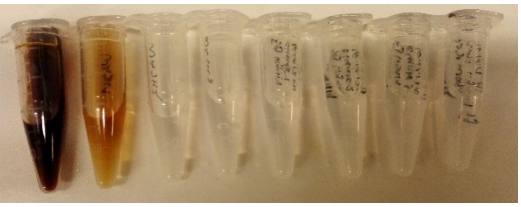
ESI figure 14: Recovery of caffeate from caffeate solution. Top (A) and bottom (B) figures show the results of two different recovery experiments made with the same column. Top rows: sample with caffeate (RT 17 min). Second rows: last flow through fraction. Third rows: last wash fraction. Rows 4-6: elution fractions. Last row: 1 ml fraction after elution.



ESI Figure 15: Recovery of caffeate from synthetic LDM mixture. Top row: spiked sample with coumarate (RT 27 min), ferulate (RT 35 min) and caffeate (RT 17 min). Second row: last flow fraction. Third row: last wash fraction. Rows 4-6: elution fractions. Last row: 1 ml fraction after elution. Note: first row not in scale.

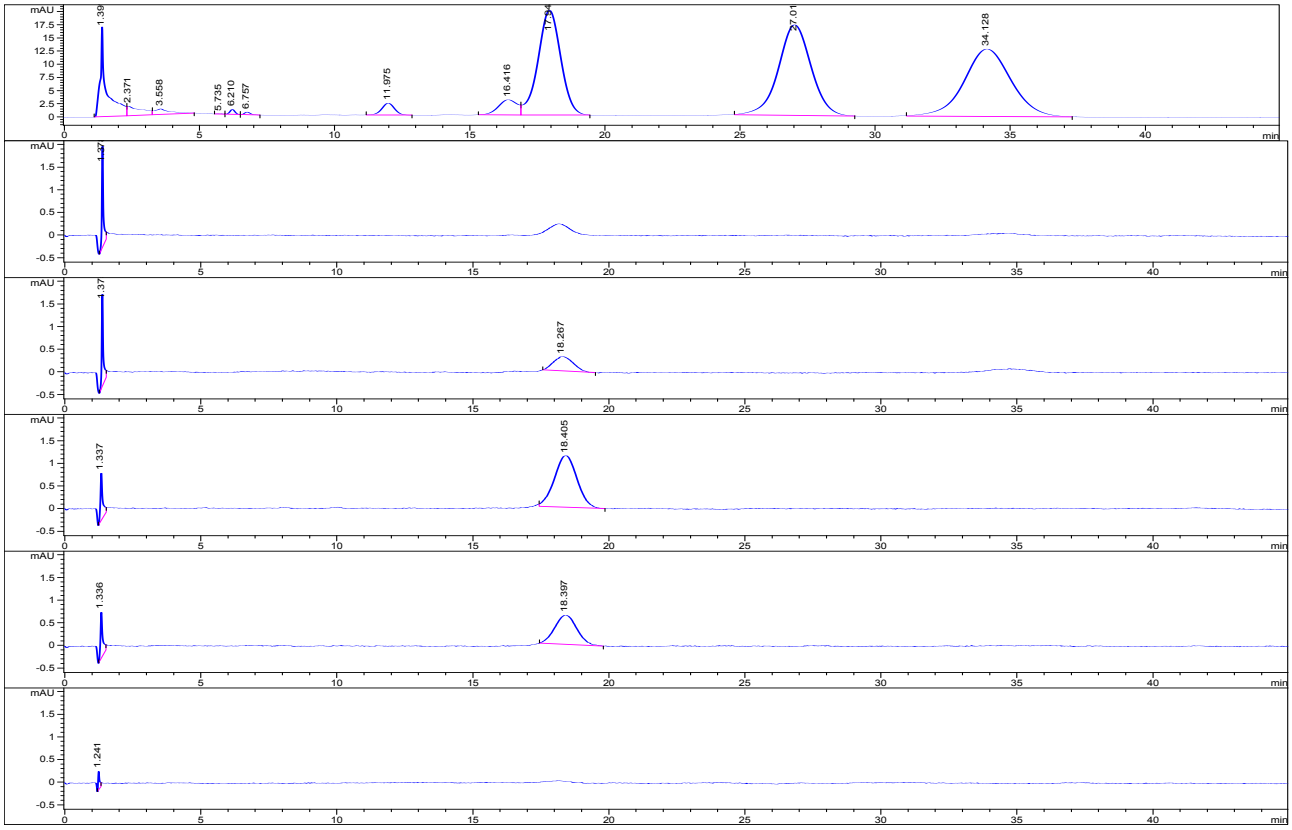


B



ESI Figure 16: A) HPLC data for Kraft lignin (310 nm). A) Kraft lignin spiked with coumarate (RT 27 min), ferulate (RT 35 min) and caffeate (RT 17 min). Top row: initial sample. Second row: wash fraction. Rows 3-5: elution 1, elution 2 and elution 3. Row 6: 1 ml fraction after elution. B) Collected

fractions from purification column. From left to right: initial sample, three last wash fractions, elution 1, 2, 3 and 1 ml fraction after elution.



ESI Figure 17: HPLC data for rice straw hydrolysate (310 nm). Top row: initial sample spiked with coumarate (RT 27 min), ferulate (RT 35 min) and caffeate (RT 17 min). Second row: last wash fraction. Row 3-5: elution 1, elution 2 and elution 3. Row 6: 1 ml fraction after elution.