SUPPLEMENTARY INFORMATION

KISS, a mammalian *in situ* protein interaction sensor

Sam Lievens[1,2], Sarah Gerlo[1,2], Irma Lemmens[1,2], Dries De Clercq[3], Martijn Risseeuw[3], Nele Vanderroost[1,2], Anne-Sophie De Smet[1,2], Elien Ruyssinck[1,2], Eric Chevet[4], Serge Van Calenbergh[3] and Jan Tavernier[1,2]

[1]Department of Medical Protein Research, VIB, A. Baertsoenkaai 3, 9000 Ghent, Belgium

[2]Department of Biochemistry, Faculty of Medicine and Health Sciences, Ghent University, A. Baertsoenkaai 3, 9000 Ghent, Belgium

[3]Laboratory for Medicinal Chemistry, Faculty of Pharmaceutical Sciences, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium

[4] French National Institute for Health and Medical Research (INSERM) U1053, University of Bordeaux Segalen, 146 Rue Leo Saignat, 33000 Bordeaux, France

Experimental Procedures

Synthesis and characterization of simvastatin methotrexate fusion compound (MFC)

General

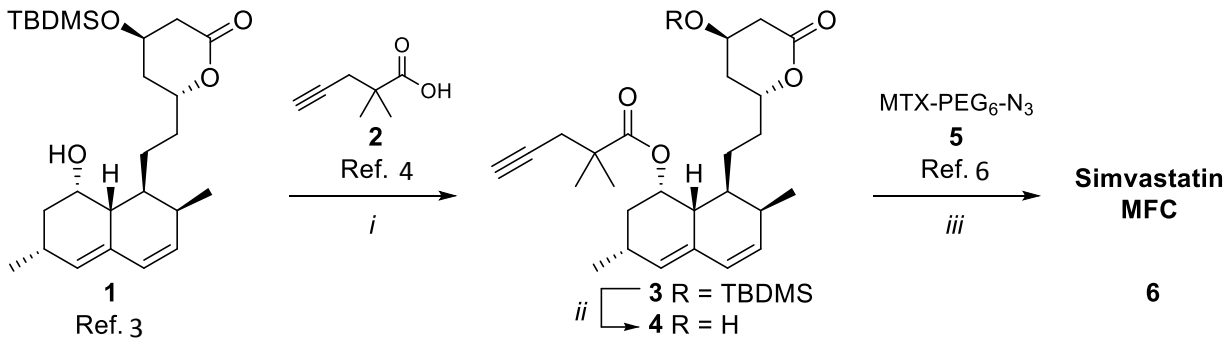
All reactions described were performed under an argon atmosphere and at ambient temperature unless stated otherwise. Reagents and solvents were purchased from Sigma-Aldrich, Acros Organics or TCI Europe and used as received. Reactions were monitored by TLC analysis using TLC aluminum sheets (Macherey-Nagel, Alugram Sil G/UV_{254}) with detection by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot 4H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot 2H_2O$ (10 g/L) in H_2SO_4 (10 %) followed by charring. Column chromatography was performed on Biosolve 60Å silica gel (32-63 μm). LC-MS analyses were carried out on a Waters Alliance 2695 XE separation Module using a Phenomenex Luna reversed phase C18 column (100 x 2.00 mm, 3 μm) and a water/acetonitrile/formic acid gradient system. High-resolution spectra were recorded with a Waters LCT Premier XE Mass spectrometer. ^{1}H- and ^{13}C-NMR spectra were measured on a Varian Mercury-300BB (300/75 MHz) spectrometer. Chemical shifts are given in ppm (δ) relative to tetramethylsilane as an internal standard (^{1}H NMR) or $CDCl_3$, CD_3OD (^{13}C NMR). Coupling constants are given in Hz. Preparative HPLC purifications were carried out using a Laprep preparative RP-HPLC system equipped with a Phenomenex Luna C18 column (21.20 x 250 mm, 5 μm) using a water/acetonitrile/formic acid gradient system. Microwave experiments were performed using a Milestone Microsynth under fiberoptic internal temperature control.

Synthesis of simvastatin MFC

Alkynylated simvastatin was prepared in several steps from lovastatin. Compound **1** and the required 2,2-dimethylpent-4-ynoic acid **2** were obtained according to a protocol by Sorensen *et al*. (3) and Welsch *et al*. (4), respectively. Analogously to a procedure by Hong *et al*. (5) the latter acid was converted in situ to the corresponding acid bromide, which was used to acylate **1** to afford alkyne **4**

after desilylation. Finally, the alkyne-functionalized simvastatin was conjugated with an earlier synthesized MTX-PEG_6-N_3 reagent (6) **5** via CuAAC generating the desired MFC **6**.

Scheme A: Synthesis of the simvastatin MFC.



Reagents and conditions: [i] PPh_3, NBS, $PhNMe_2$, CH_2Cl_2, 0°C→RT, 79.9%; [ii] HCl, THF, dioxane, 0°C→RT, 76.5%; [iii] Na-ascorbate, $CuSO_4$, Et_3N, TBTA, DMF, H_2O, 150°C (μW), 43.6%.

Compound 3

The 2,2-dimethylpent-4-ynoic acid **2** (278 mg, 2.20 mmol) was dissolved in 5.2 mL of CH_2Cl_2 and PPh_3 (1.437 g, 5.48 mmol) was added and cooled to 0°C. Subsequently, NBS (630 mg, 3.54 mmol) was added and the resulting reaction mixture was allowed to stir for 30 minutes at room temperature. After this time, the mixture was cooled to 0°C again and *N*,*N*-dimethylaniline (0.3 mL, 2.38 mmol) and compound **1** (869 mg, 2.00 mmol) were added. The resulting solution was stirred overnight letting the temperature rise to room temperature. The reaction mixture was successively washed with 20 mL of 1.0M HCl, water, sat. $NaHCO_3$, water and sat. brine and then dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by two successive columns (silica gel, 1/9→1/7, v/v, EtOAc/toluene) yielding the starting material **1** (242 mg, 0.56 mmol, 28.0%) as well as the title compound (624 mg, 1.15 mmol, 57.5%, 79.9% based on recovered starting material) as a yellowish oil. ^{1}H NMR (300 MHz, $CDCl_3$) δ = 5.98 (d, 1H, *J* = 9.6 Hz), 5.77 (dd, 1H, *J* = 9.6 Hz; *J* = 6.0 Hz), 5.53-5.49 (m, 1H), 5.38-5.33 (m, 1H), 4.65-4.54 (m, 1H), 4.33-4.27 (m, 1H), 2.62 (dd, 1H, *J* = 17.4 Hz; *J* = 4.5 Hz), 2.54 (ddd, 1H, *J* = 17.4 Hz; *J* = 3.3 Hz; *J* = 1.2 Hz), 2.46-2.39 (m, 3H), 2.39-2.32 (m,

1H), 2.31-2.22 (m, 1H), 2.06-1.91 (m, 3H), 1.91-1.79 (m, 2H), 1.74-1.59 (m, 2H), 1.54-1.29 (m, 3H), 1.26 (s, 3H), 1.25 (s, 3H), 1.08 (d, 3H, J = 7.5 Hz), 0.890 (d, 3H, J = 6.9 Hz), 0.885 (s, 9H), 0.084 (s, 3H), 0.081 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ = 176.1, 170.1, 132.9, 131.5, 129.6, 128.3, 81.0, 76.3, 70.7, 68.7, 63.5, 42.2, 39.2, 37.4, 36.77, 36.75, 33.0, 32.6, 30.6, 29.4, 27.2, 25.7, 24.7, 24.4, 24.1, 23.1, 17.9, 13.8, -4.91, -4.92; HRMS: calcd. for $C_{32}H_{51}O_5Si$ (M+H): 543.3500, found: 543.3502.

Compound 4

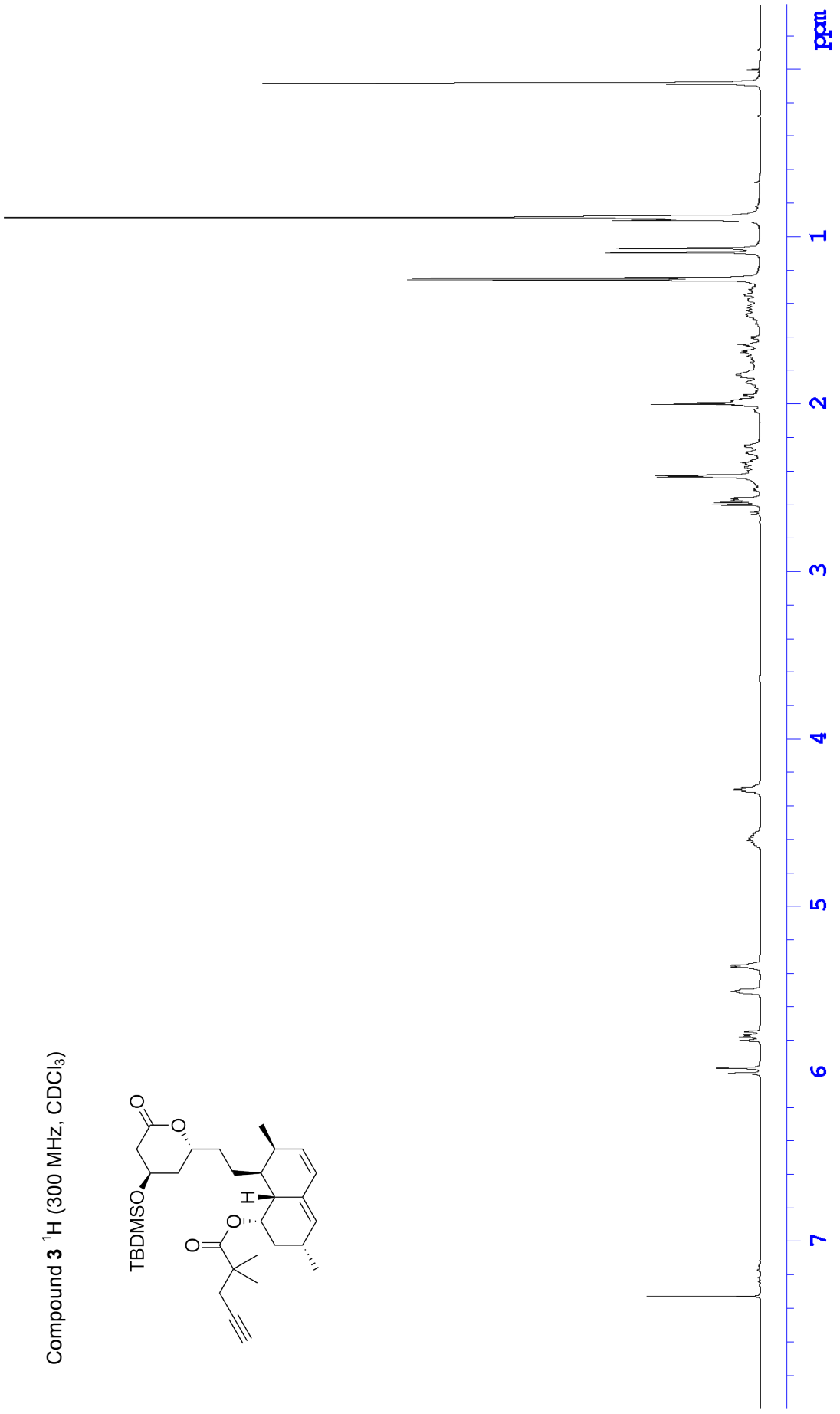
Compound **3** (390 mg, 0.72 mmol) was dissolved in 6.5 mL of THF and 0.34 mL of 1,4-dioxane. After the mixture was cooled to 0°C, concentrated HCl (482 µL) was added and stirred for 5 hours at room temperature. The reaction mixture was subsequently poured in sat. $NaHCO_3$ (30 mL) and after separation of both phases, the aqueous layer was washed with CH_2Cl_2 (3 x 30 mL). All organic fractions were pooled, dried over Na_2SO_4 and taken to dryness. The residue was purified by silica gel chromatography (1/1, v/v, EtOAc/toluene) yielding the title compound (234 mg, 0.55 mmol, 76.5%) as a yellowish/beige foam. ^{1}H NMR (300 MHz, $CDCl_3$) δ = 5.99 (d, 1H, J = 9.6 Hz), 5.78 (dd, 1H, J = 9.3 Hz; J = 6.0 Hz), 5.53-5.49 (m, 1H), 5.40-5.35 (m, 1H), 4.69-4.58 (m, 1H), 4.41-4.33 (m, 1H), 2.74 (dd, 1H, J = 17.6 Hz; J = 5.0 Hz), 2.62 (ddd, 1H, J = 17.6 Hz; J = 3.8 Hz; J = 1.5 Hz), 2.49-2.41 (m, 3H), 2.39-2.33 (m, 1H), 2.31-2.23 (m, 1H), 2.05 (t, 1H, J = 2.7 Hz), 2.01-1.93 (m, 3H), 1.90-1.77 (m, 2H), 1.76-1.64 (m, 2H), 1.55-1.30 (m, 3H), 1.26 (s, 3H), 1.25 (s, 3H), 1.08 (d, 3H, J = 7.2 Hz), 0.89 (d, 3H, J = 6.9 Hz); ^{13}C NMR (75 MHz, $CDCl_3$) δ = 176.6, 170.6, 133.1, 131.6, 129.8, 128.5, 81.3, 76.5, 71.0, 69.0, 62.8, 42.5, 38.8, 37.6, 36.8, 36.2, 33.1, 32.9, 30.8, 29.7, 27.4, 24.9, 24.7, 24.5, 23.3, 14.0; HRMS: calcd. for $C_{26}H_{36}O_5Na$ (M+Na): 451.2455, found: 451.2466.

Simvastatin MFC 6

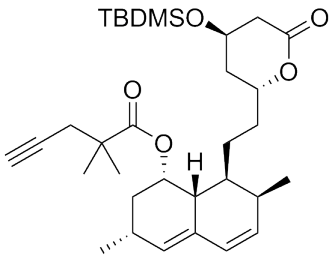
The earlier prepared α-*tert*-butylester protected MTX-PEG_6-N_3 reagent(6) (160 mg, 0.2 mmol) was taken up in a mixture of TFA and CH_2Cl_2 (5 mL, 1:1, v/v) and stirred for 40 minutes at room temperature. The reaction mixture was then taken to dryness, coevaporated twice with toluene and concentrated under high vacuum for 1 hour. The resulting deprotected MTX-PEG_6-N_3 reagent **5** residue was dissolved in 2.8 mL of DMF, taken up in a mixture of water and DMF (3 mL, 1:1, v/v)

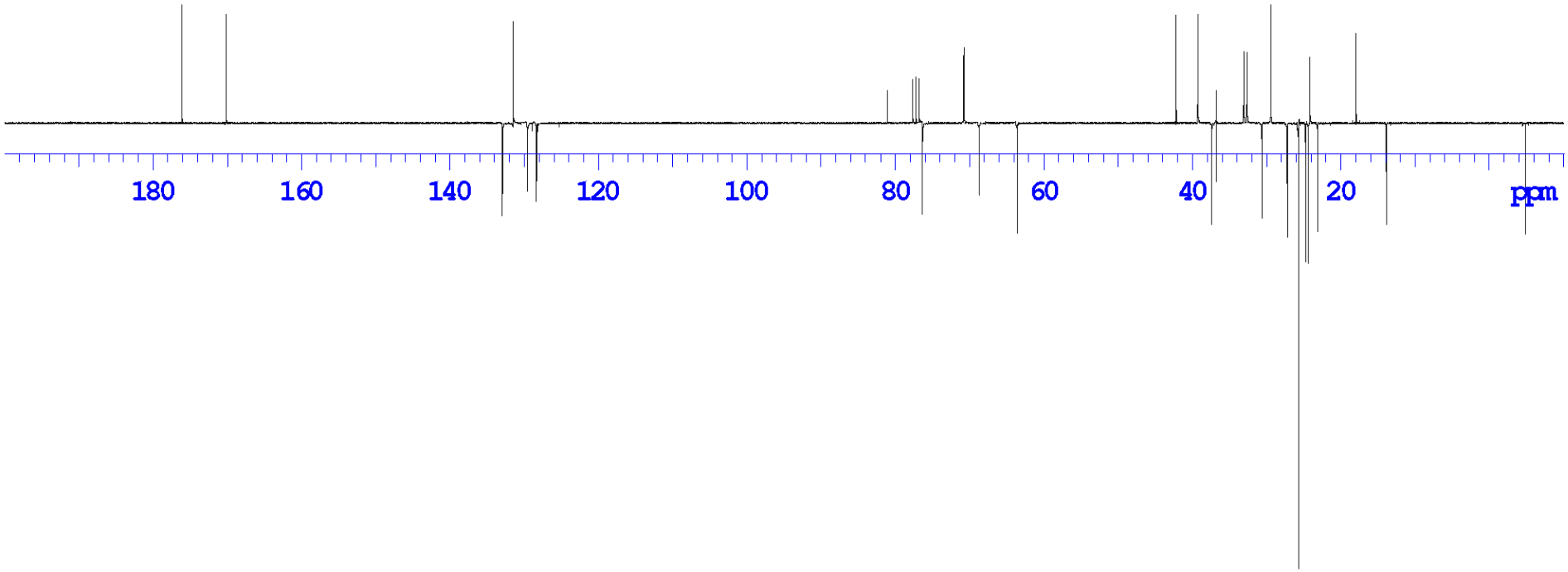
and alkyne **4** (43 mg, 0.1 mmol), $CuSO_4$ (400 µL, 0.05M, 0.2 eq.) and Na-ascorbate (200 µL, 0.5M, 1.0 eq.) were added. Finally, the resulting reaction mixture was charged with a catalytic amount of both Et_3N and TBTA (7) and heated to 150°C for 15 minutes (microwave). After cooling to room temperature, the solution was concentrated *in vacuo*. The residue was purified by preparative RP-HPLC (MeCN/water/formic acid 10→70%) yielding the title compound (51 mg, 44 µmol, 43.6%) as a pale yellow amorphous solid. LC-HRMS: Rt = 4.69 min (92.17% AUC, 30→100% MeCN, 15 min run). Note that the peak eluting at a Rt of 4.09 min (6.42% AUC) corresponds to the lactone ring opened simvastatin MFC analogue. We reasoned that ring opening of the β-hydroxy-δ-valerolactone pharmacophore could not be responsible for reduced bait-prey interactions, since both lactones as their hydrolyzed linear analogues constitute the statin class of drugs (cf. simvastatin vs. atorvastatin). Furthermore, we estimated that hydrolysis of the lactone would presumably occur to some extent anyhow, due to esterases operating in the intracellular environment; HRMS: calcd. for $C_{58}H_{84}N_{12}O_{14}$ (M+2H): 586.3110, found: 586.3058. Lactone ring opened simvastatin MFC analogue: calcd. for $C_{58}H_{86}N_{12}O_{15}$ (M+2H): 595.3162, found: 595.3105.

Selected NMR spectra

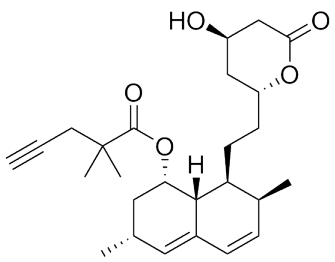


Compound **3** ^{13}C APT (75 MHz, $CDCl_3$)

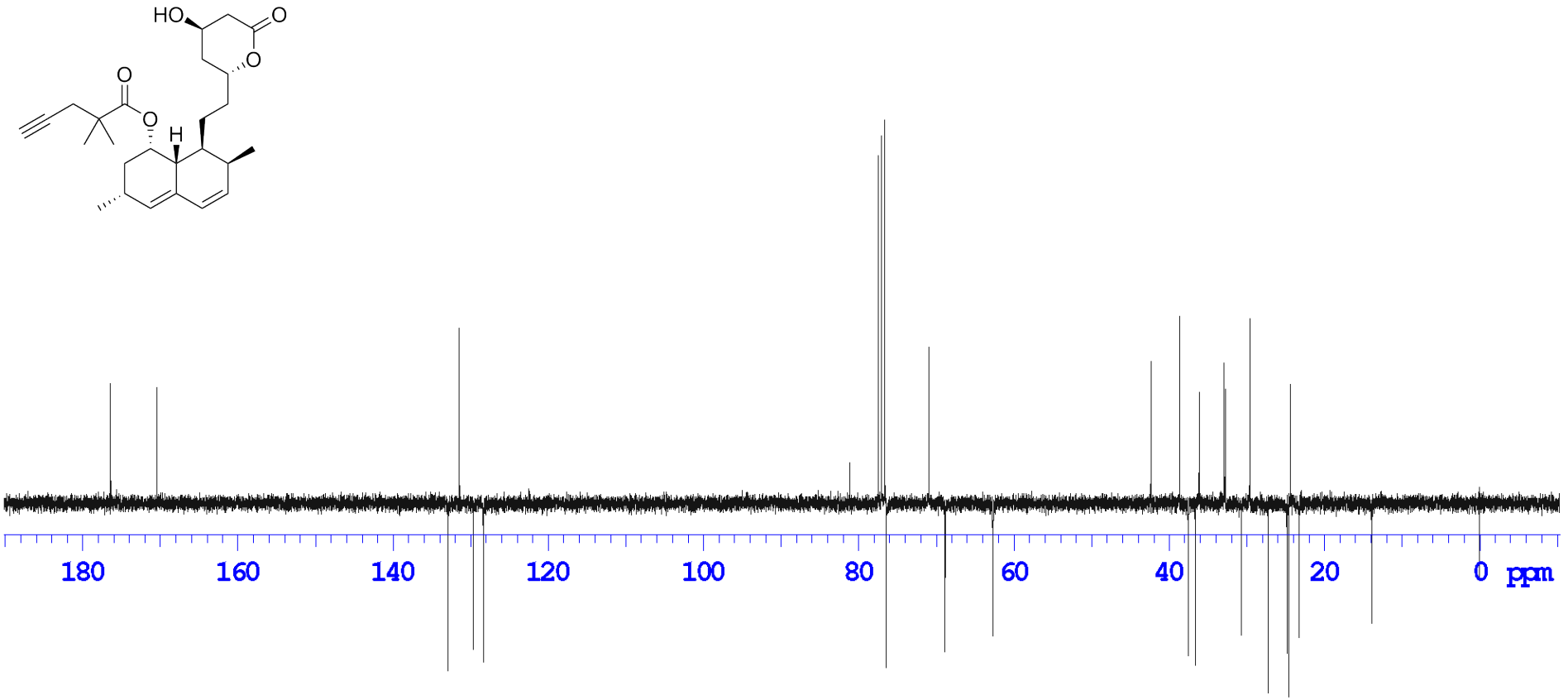




Compound **4** ^{1}H (300 MHz, $CDCl_3$)

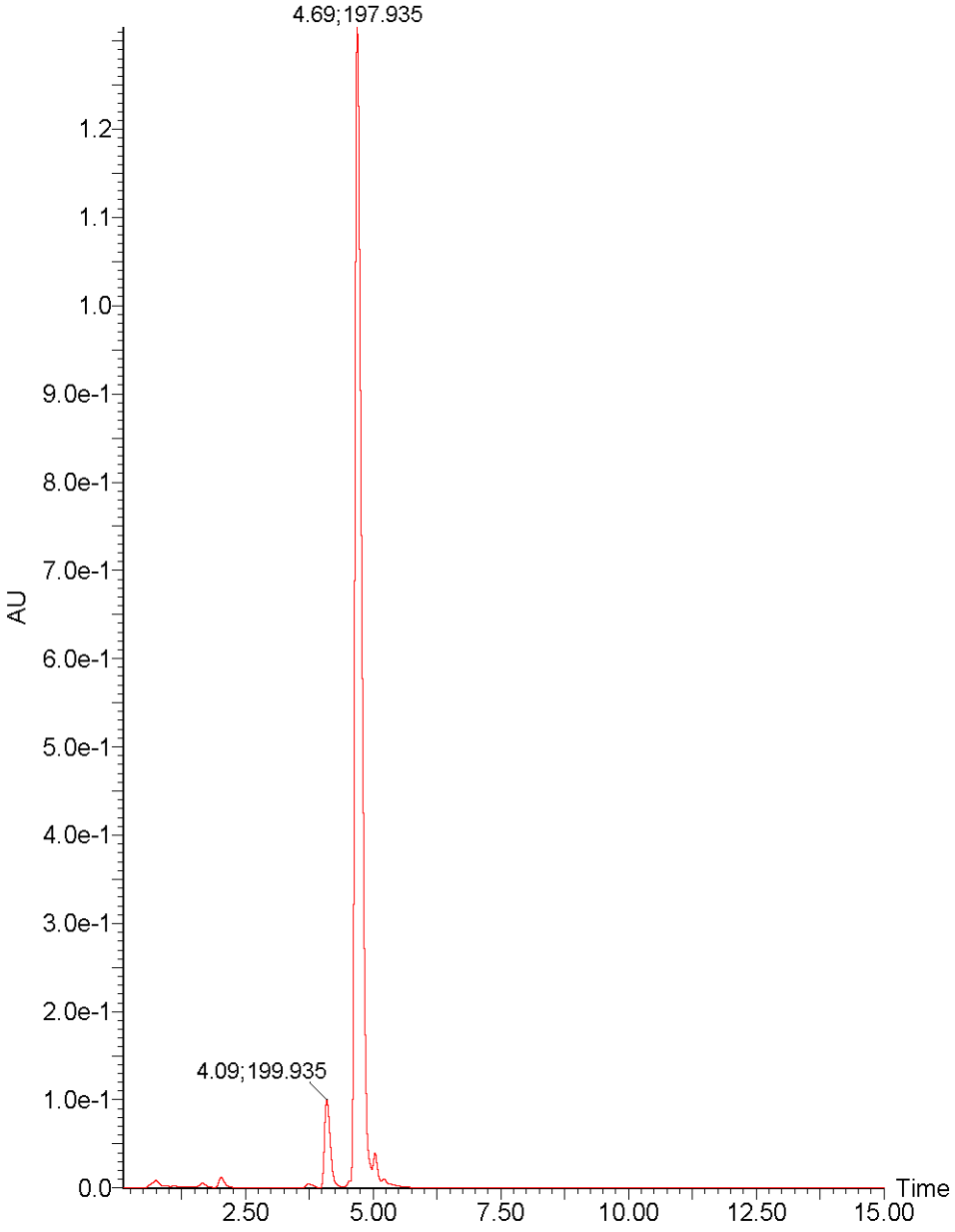


Compound **4** ^{13}C APT (75 MHz, $CDCl_3$)



Selected LC-HRMS spectrum

Simvastatin MFC **6** (30→100% MeCN, 15 min run)



References

1. Eyckerman S*, et al.* (2001) Design and application of a cytokine-receptor-based interaction trap. *Nature Cell Biology* 3(12):1114-1119.
2. Lievens S, Peelman F, De Bosscher K, Lemmens I, & Tavernier J (2011) MAPPIT: a protein interaction toolbox built on insights in cytokine receptor signaling. *Cytokine Growth Factor Rev* 22(5-6):321-329.
3. Sorensen JL, Auclair K, Kennedy J, Hutchinson CR, & Vederas JC (2003) Transformations of cyclic nonaketides by Aspergillus terreus mutants blocked for lovastatin biosynthesis at the lovA and lovC genes. *Organic & Biomolecular Chemistry* 1(1):50-59.
4. Welsch T, Tran HA, & Witulski B (2010) Total Syntheses of the Marine Illudalanes Alcyopterosin I, L, M, N, and C. *Organic Letters* 12(24):5644-5647.
5. Hong CI, Kim JW, Shin HJ, Kang TW, & Cho DO (2001).
6. Risseeuw MD*, et al.* (2013) A "Clickable" MTX Reagent as a Practical Tool for Profiling Small-Molecule-Intracellular Target Interactions via MASPIT. *ChemMedChem* 8(3):521-526.
7. Chan TR, Hilgraf R, Sharpless KB, & Fokin VV (2004) Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Organic Letters* 6(17):2853-2855.

Figure legends

Figure S1. Identification of TYK2-containing preys in MAPPIT cDNA library screens. (**a**) Schematic overview of the MAPPIT principle. A bait protein of interest ("X") is fused to (a portion of) a homomeric cytokine receptor (e.g. the leptin receptor, or a chimera containing the extracellular domain of the erythropoietin receptor fused to the transmembrane and intracellular domains of the leptin receptor) containing a mutated STAT3 recruitment site in its cytoplasmic tail (grey dot); a prey protein ("Y") is tethered to a gp130 cytokine receptor fragment with functional STAT3 docking sites (black dots); upon administration of the appropriate cytokine ("cyt"), constitutively associated JAK2 kinases ("JAK") are activated by cross-phosphorylation ("P"); when bait and prey interact, JAKs phosphorylate the STAT3 docking sites on the prey chimeras ("P"), leading to the recruitment of STATs; these STATs, in turn, are phosphorylated by the JAKs ("P"), which results in their activation; activated STATs migrate to the nucleus where dimers of these transcription factors activate transcription of a (firefly luciferase) reporter gene. For details see (1, 2). (**b**) Tyrosine kinase 2 (TYK2)-containing prey fusion proteins were identified in multiple MAPPIT cDNA library screens with unrelated bait proteins. In these preys, the gp130 cytokine receptor fragment that relays the MAPPIT signal was fused to a C-terminal portion of TYK2, invariably containing the intact kinase domain. Cells expressing such a prey exhibited a luciferase reporter signal that was independent of cytokine stimulation or the presence of the MAPPIT bait. This suggested that the TYK2 kinase domain in these fusion proteins is able to phosphorylate ("P") both the STAT3 docking sites in the attached gp130 cytokine receptor moiety and the STAT3 transcription factors that are being recruited to these sites. (**c**) TYK2-containing preys identified as artifacts in MAPPIT cDNA library screens consisted of an in-frame fusion of a C-terminal portion of TYK2 with the gp130 receptor fragment relaying the STAT3-dependent signal in MAPPIT. Arrows indicate the fusion sites in the TYK2 sequence of such preys. (**d**) MAPPIT evaluation of a TYK2-containing prey. Cells expressing the E3 ubiquitin ligase protein SKP1 as a bait together with one of its known interaction partners FBXW11 (FBX) as a prey generated a luciferase signal only upon stimulation of the cells with the MAPPIT-activating cytokine (S) and not in unstimulated conditions (NS). This signal is specific for the

interaction, as in the presence of just one of the interaction partners (SKP1 bait combined with unfused gp130 as a control or empty vector combined with FBX prey) no signal was detected. In the case of a TYK2 prey, a luciferase signal was detected independent of cytokine stimulation and regardless of the presence of the bait. Prey expression was evaluated through anti-Flag western blot; actin expression was revealed to confirm equal loading.

Figure S2. Analysis of interactions among HIV-1 RT p51 and p66 subunits. (**a**) Luciferase data correspond to that in Fig. 3b, here displaying raw instead of normalized luciferase signals. Western blots confirm expression of prey (anti-Flag) and bait (anti-HA) fusion proteins. Equal loading was evaluated by revealing actin expression. (**b**) Detection of prey phosphorylation upon binding to interacting bait. As in Figure 3b and panel a of Figure S2, a KISS assay was performed with combinations of HIV-1 RT p51 bait and either a control prey (consisting of unfused gp130 fragment) or a RT p66 prey, generating a similar p51-p66-dependent luciferase signal. In parallel, the prey in cells from the same transfection was precipitated by means of the gp130-fused Flag tag and assayed in an immunoblot using a phosphotyrosine-specific antibody, indicating that the p51-interacting p66 prey fusion was phosphorylated. Control blots show expression of bait (anti-HA) and prey (anti-Flag) in each of the lysates. Equal loading was evaluated through revealing actin expression levels.

Figure S3. Detection of ligand-dependent interaction between GPCRs and beta arrestin 2. Cells expressing somatostatin receptor 2 or angiotensin receptor 1 bait together with beta arrestin 2 prey chimeras exhibit a ligand-dependent luciferase signal. Cells were stimulated with saturating concentrations (S; 10μM) of the appropriate ligand (somatostatin or angiotensin II for somatostatin receptor 2 or angiotensin receptor 1, respectively). No signal was observed in unstimulated cells (NS) or in cells expressing unfused gp130 control prey. Prey and bait expression was confirmed through anti-Flag and anti-HA western blot. Actin expression was revealed to confirm equal loading.

Figure S4. Evaluation of ER stress-induced ERN1 clustering. (**a**) Model for ERN1 activation by ER stress. In unstressed cells, the ER chaperone BiP is bound to the ERN1 luminal portion (L), preventing ERN1 luminal domain interaction. Upon accumulation of misfolded proteins in the ER, these compete with ERN1 for binding to BiP (1), hence promoting clustering of ERN1 luminal domains (2), followed by trans-autophosphorylation of adjacent cytoplasmic ERN1 kinase domains (K). Activation of the ERN1 kinase triggers conformational changes that in turn promote activation of the C-terminal endonuclease domain (E) of ERN1 that is associated with higher-order oligomerization (3). D123P mutation of ERN1 prevents dimerization of the luminal domains (*a), K599A mutation prevents trans-autophosphorylation of the ERN1 kinase domain (*b). (**b**) Subcellular localization of ERN1cyt prey expression. Expression of an ERN1 prey consisting of only the cytoplasmic ERN1 domain fused to gp130 was visualized by anti-gp130 staining. Co-staining with an antibody against the ER-resident calnexin protein indicated that expression of this partial ERN1 prey is not restricted to the ER. DAPI staining was used to visualize cell nuclei (blue). Scale bars indicate 20μm. (**c, d, e**) Western blots confirming expression of bait, prey and unfused expression constructs used in Fig. 3f, g and h, respectively.

Figure S5. Three-hybrid KISS. Western blot confirming expression of bait, prey and unfused expression constructs used in Fig. 5b.

Figure S1

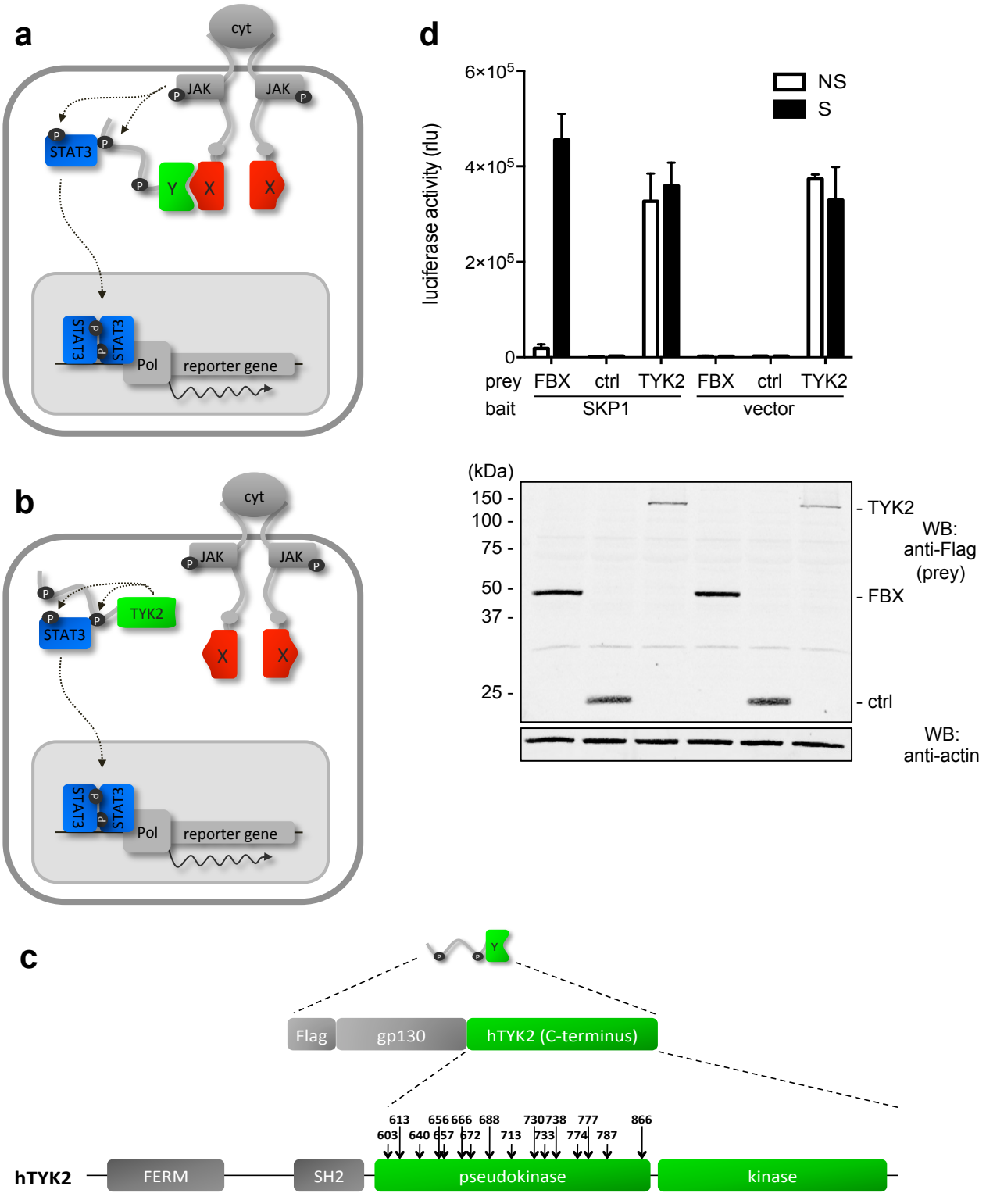

a
cyt
JAK
JAK
STAT3
Y
X
X
STAT3
STAT3
Pol
reporter gene
b
cyt
JAK
JAK
TYK2
STAT3
X
X
STAT3
STAT3
Pol
reporter gene
d
luciferase activity (rlu)
6×10^5
4×10^5
2×10^5
0
NS
S
prey
FBX
ctrl
TYK2
FBX
ctrl
TYK2
bait
SKP1
vector
(kDa)
150 -
100 -
75 -
50 -
37 -
25 -
- TYK2
WB:
anti-Flag
(prey)
- FBX
- ctrl
WB:
anti-actin
c
Flag
gp130
hTYK2 (C-terminus)
613
603
640
656
657
666
672
688
713
730
733
738
774
777
787
866
hTYK2
FERM
SH2
pseudokinase
kinase

Figure S2

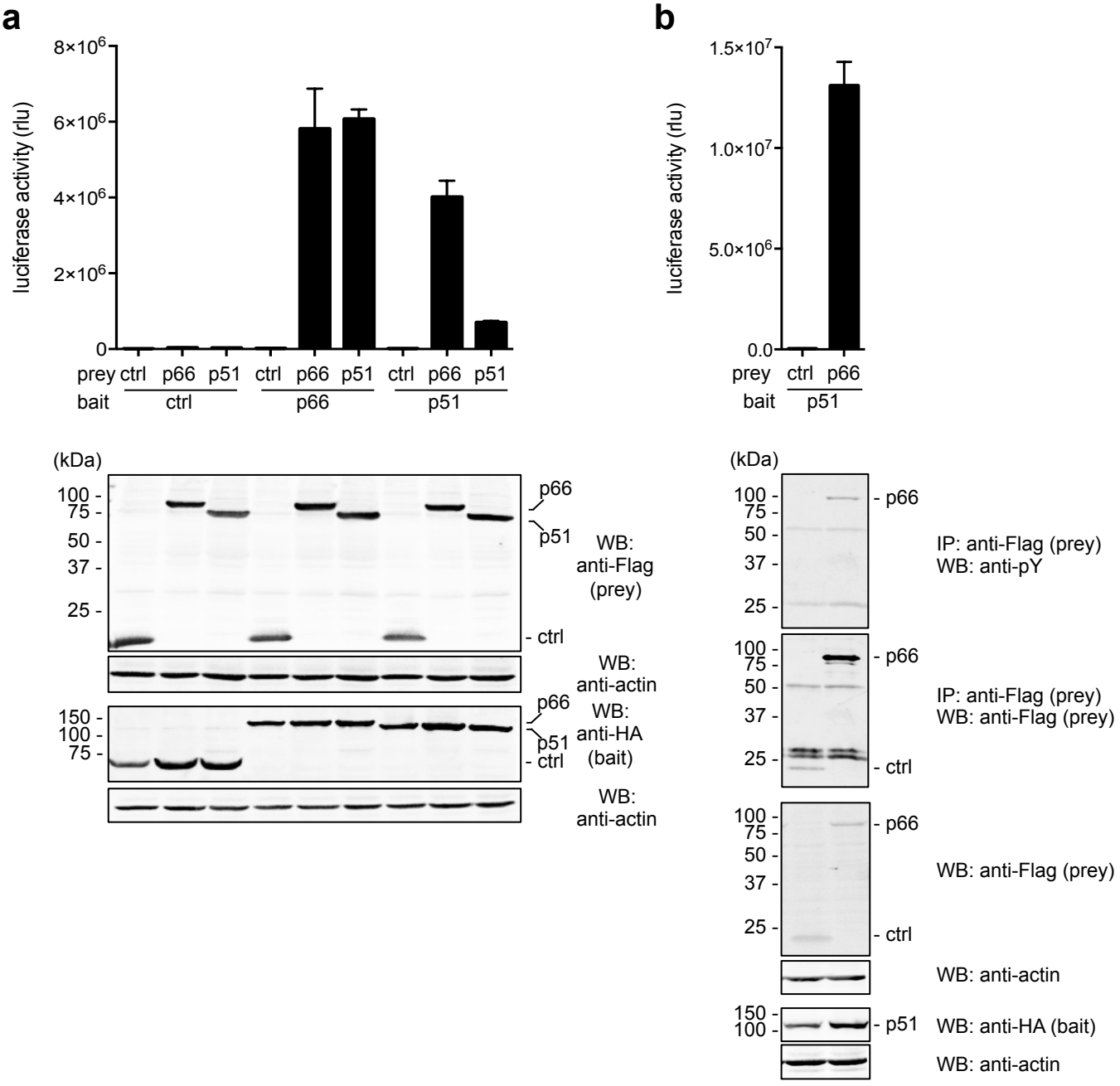

a
luciferase activity (rlu)
8×10^6
6×10^6
4×10^6
2×10^6
0
prey ctrl p66 p51 ctrl p66 p51 ctrl p66 p51
bait ctrl p66 p51
(kDa)
100 -
75 -
50 -
37 -
25 -
p66
p51
- ctrl
WB: anti-Flag (prey)
WB: anti-actin
150 -
100 -
75 -
p66
p51
- ctrl
WB: anti-HA (bait)
WB: anti-actin
b
luciferase activity (rlu)
1.5×10^7
1.0×10^7
5.0×10^6
0.0
prey ctrl p66
bait p51
(kDa)
100 -
75 -
50 -
37 -
25 -
- p66
IP: anti-Flag (prey)
WB: anti-pY
- p66
- ctrl
IP: anti-Flag (prey)
WB: anti-Flag (prey)
- p66
- ctrl
WB: anti-Flag (prey)
WB: anti-actin
150 -
100 -
- p51
WB: anti-HA (bait)
WB: anti-actin

Figure S3

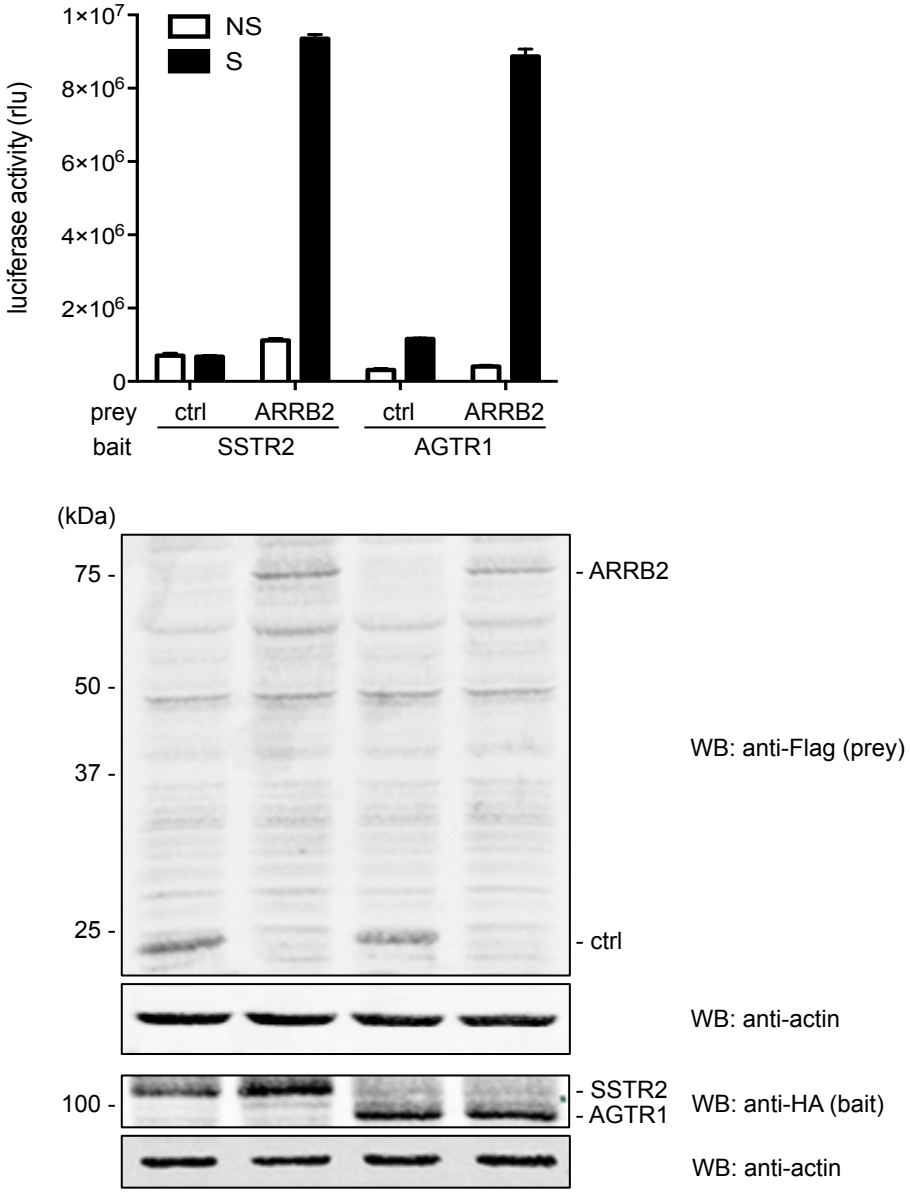

1×10^7
8×10^6
6×10^6
4×10^6
2×10^6
0
luciferase activity (rlu)
NS
S
prey
ctrl
ARRB2
ctrl
ARRB2
bait
SSTR2
AGTR1
(kDa)
75 -
50 -
37 -
25 -
- ARRB2
- ctrl
WB: anti-Flag (prey)
WB: anti-actin
100 -
- SSTR2
- AGTR1
WB: anti-HA (bait)
WB: anti-actin

Figure S4

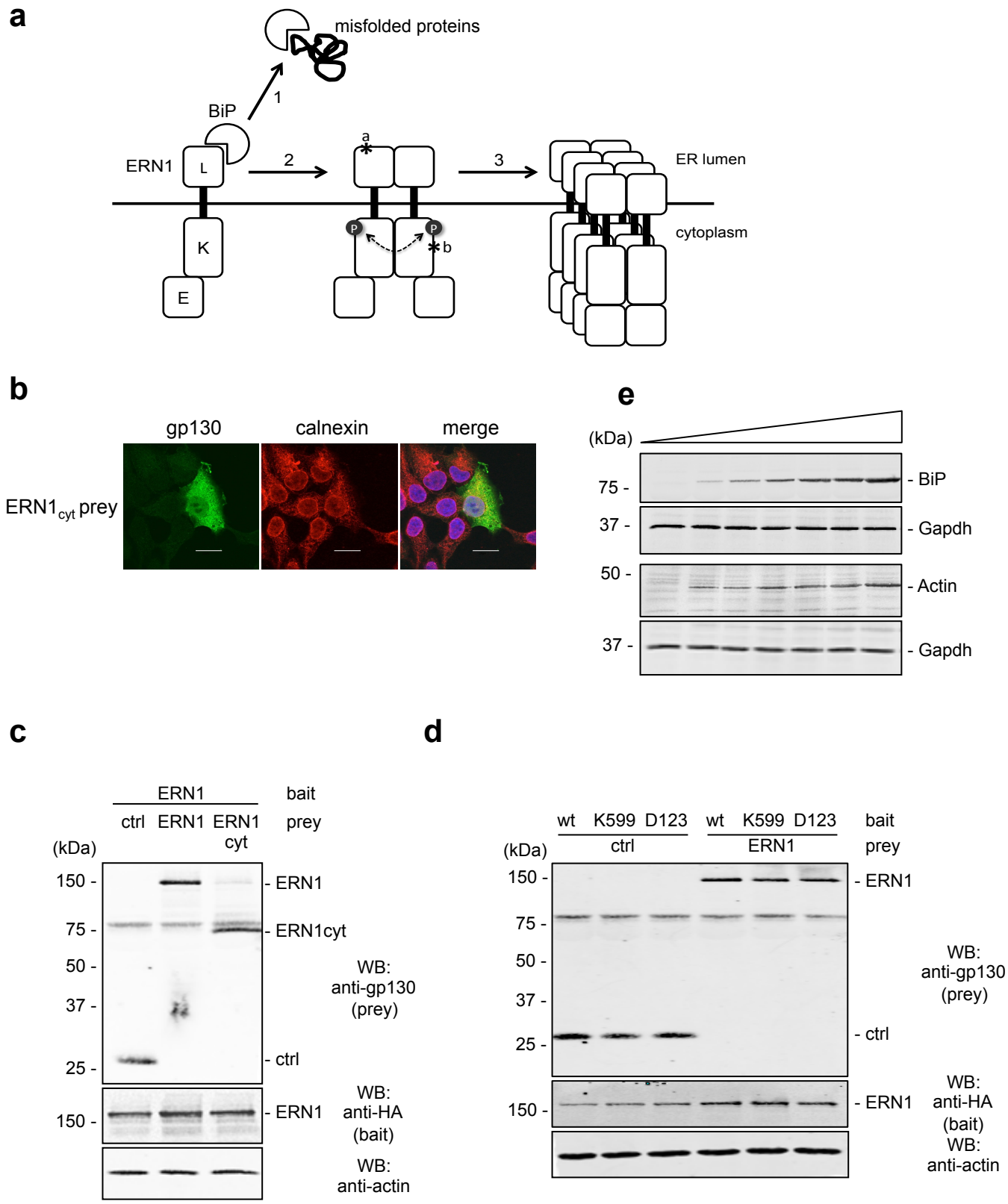
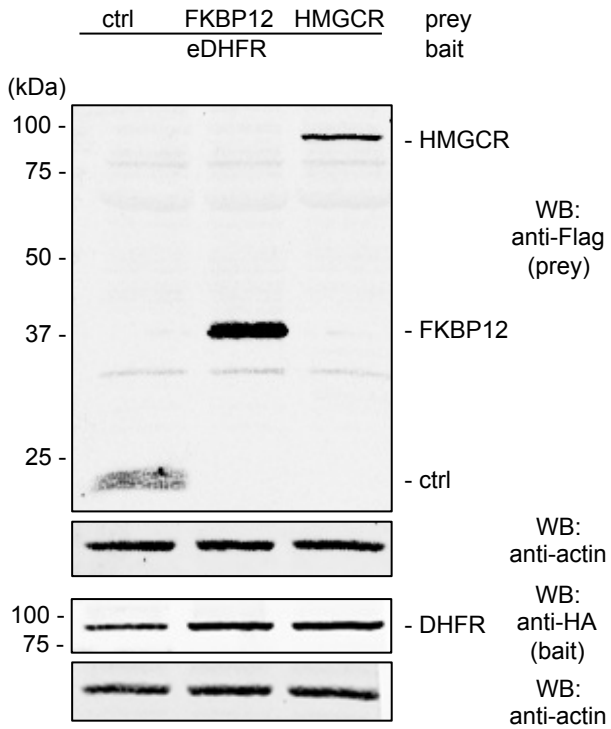

a
misfolded proteins
BiP
1
ERN1
L
2
a
3
ER lumen
K
E
P
P
b
cytoplasm
b
gp130
calnexin
merge
ERN1cyt prey
e
(kDa)
75 -
- BiP
37 -
- Gapdh
50 -
- Actin
37 -
- Gapdh
c
ERN1
bait
ctrl
ERN1
ERN1 cyt
prey
(kDa)
150 -
- ERN1
75 -
- ERN1cyt
50 -
37 -
WB: anti-gp130 (prey)
25 -
- ctrl
150 -
- ERN1
WB: anti-HA (bait)
WB: anti-actin
d
wt
K599
D123
wt
K599
D123
bait
ctrl
ERN1
prey
(kDa)
150 -
- ERN1
75 -
50 -
37 -
WB: anti-gp130 (prey)
25 -
- ctrl
150 -
- ERN1
WB: anti-HA (bait)
WB: anti-actin

Figure S5


ctrl
FKBP12
HMGCR
prey
eDHFR
bait
(kDa)
100 -
75 -
50 -
37 -
25 -
- HMGCR
WB:
anti-Flag
(prey)
- FKBP12
- ctrl
WB:
anti-actin
100 -
75 -
- DHFR
WB:
anti-HA
(bait)
WB:
anti-actin