Potent Small Agonists of Protease Activated Receptor 2

Mei-Kwan Yau,[†] Jacky Y. Suen,[†] Weijun Xu,[†] Junxian Lim,[†] Ligong Liu,[†] Mark N. Adams,[‡] Yaowu He,[‡] John D. Hooper,[‡] Robert C. Reid,[†] and David P. Fairlie*[,†]

[†]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld 4072, Australia

[‡]Translational Research Institute, Mater Research Institute, The University of Queensland, Woolloongabba, Qld 4102, Australia

*Correspondence to: d.fairlie@uq.edu.au

Table of Contents **Page**

Figure S1. Desensitization of agonists **1**, **14** and TFLLR-NH_2 in prostate cancer cells. S2

Compound Characterization Data S2-S4

Cell Culture, Expression Constructs and Transfections S4

Homology Modeling based on Human Nociceptin/Orphanin ORL1 S4-S6

References S6

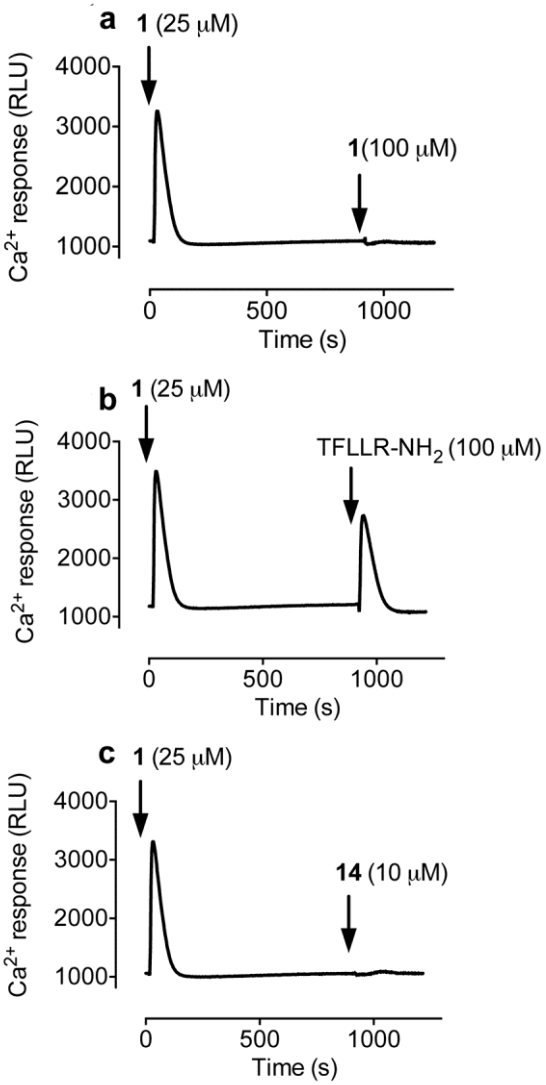

Figure S1. Desensitization of agonists **1**, **14** and TFLLR-NH_2 in prostate cancer cells. (A) Control: PC3 prostate cancer cells were treated with **1** (25 μM) to desensitize PAR2, then with **1** (100 μM) to show PAR2 desensitization. (B) PAR2 was desensitized with **1** (25 μM) before adding PAR1 agonist (TFLLR-NH_2 at 100 μM) to induce a second Ca^{2+} release, confirming that only PAR2 was desensitized. (C) Agonist **14** (10 μM) did not stimulate Ca^{2+} following PAR2 desensitization and thus it is selective for PAR2 over PAR1.

Compound characterization data

Compound purity and identity were assessed by rpHPLC, high-resolution mass spectroscopy (HRMS) and NMR spectroscopy. All assayed compounds were ≥ 95% pure, as determined by rpHPLC (UV detection at 214, 230 and 254 nm, Phenomenex Luna C18 column, 5 μm, 4.6 x 250 mm). Standard HPLC conditions were used, unless otherwise indicated, at a flow rate of 1 mL/min using solvent A and B: 20% B to 100% B linear gradient over 10 min, followed by a further 10 min at 100% B where solvent A was H_2O + 0.1% TFA and solvent B was 90% MeCN, 10% H_2O + 0.1% TFA with UV detection at 214, 230 and 254 nm. High-resolution mass spectra (HRMS) measurements were obtained on a Bruker microTOF mass spectrometer

equipped with a Dionex LC system (Chromeleon) in positive ion mode by direct infusion in MeCN at 100 μL/h using sodium formate clusters as an internal calibrant. Data was processed using Bruker Daltonics DataAnalysis 3.4 software. Mass accuracy was better than 1 ppm error. Rat plasma samples were analyzed by UPLC-MS on a Shimadzu Nexera system using a Zorbax eclipse plus C18 100 x 2.1 mm 1.8 μm column and MeCN/water (both containing 0.1% formic acid) gradients at flow rate 0.5 mL/min. ^{1}H and ^{13}C NMR spectra were recorded on Bruker Avance 600 spectrometer at 298 K in deuterated solvents indicated and were referenced to residual ^{1}H signals; DMSO-d_6 2.50 ppm.

Isoxazole-Cha-Ile-GRL-NH_2 (**7**). t_R = 8.0 min; HRMS: $[MH]^+$ 705.4406 (calc. for $C_{33}H_{57}N_{10}O_7^+$) 705.4404 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.78–0.95 (m, 14H), 1.04–1.22 (m, 4H), 1.30 (m, 1H), 1.41–1.76 (m, 16H), 3.08 (q, *J* = 6.4 Hz, 2H), 3.66–3.81 (2 sets of dd, *J* = 5.7 Hz, 2H), 4.14–4.23 (m, 2H), 4.29 (m, 1H), 4.56 (m, 1H), 6.97 (br s, 1H), 7.16 (d, *J* = 1.9 Hz, 1H), 7.29 (br s, 1H), 7.49 (t, *J* = 5.8 Hz, 1H), 7.90–7.99 (m, 3H), 8.25 (t, *J* = 5.8 Hz, 1H), 8.76 (d, *J* = 1.9 Hz, 1H), 8.98 (d, *J* = 8.5 Hz, 1H).

Isoxazole-Cha-Ile-GR-NH_2 (**8**). t_R = 7.3 min; HRMS: $[MH]^+$ 592.3566 (calc. for $C_{27}H_{46}N_9O_6^+$) 592.3568 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.77–0.96 (m, 8H), 1.04–1.21 (m, 4H), 1.31 (m, 1H), 1.39–1.76 (m, 13H), 3.08 (q, *J* = 6.6 Hz, 2H), 3.69–3.79 (2 sets of dd, *J* = 5.7 Hz, 2H), 4.14–4.23 (m, 2H), 4.56 (m, 1H), 7.11 (br s, 1H), 7.16 (d, *J* = 1.9 Hz, 1H), 7.39 (br s, 1H), 7.53 (t, *J* = 5.9 Hz, 1H), 7.93 (d, *J* = 8.4 Hz, 1H), 7.97 (d, *J* = 8.6 Hz, 1H), 8.23 (t, *J* = 5.6 Hz, 1H), 8.76 (d, *J* = 1.9 Hz, 1H), 8.99 (d, *J* = 8.4 Hz, 1H).

Isoxazole-Cha-Ile-G-NH_2 (**9**). t_R = 8.5 min; HRMS: $[MH]^+$ 436.2554 (calc. for $C_{21}H_{34}N_5O_5^+$) 436.2551 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.77–0.96 (m, 8H), 1.03–1.22 (m, 4H), 1.31 (m, 1H), 1.46 (m, 1H), 1.52–1.78 (m, 8H), 3.56–3.70 (2 sets of dd, *J* = 5.7 Hz, 2H), 4.14 (t, *J* = 7.8 Hz, 1H), 4.57 (m, 1H), 7.05 (br s, 1H), 7.17 (d, *J* = 1.9 Hz, 1H), 7.20 (br s, 1H), 8.03 (d, *J* = 8.5 Hz, 1H), 8.13 (t, *J* = 5.8 Hz, 1H), 8.75 (d, *J* = 1.9 Hz, 1H), 8.98 (d, *J* = 8.3 Hz, 1H).

Isoxazole-Cha-Ile-NH_2 (**10**). t_R = 8.9 min; HRMS: $[MH]^+$ 379.2340 (calc. for $C_{19}H_{31}N_4O_4^+$) 379.2340 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.78–0.84 (m, 6H), 0.84–0.95 (m, 2H), 1.02–1.13 (m, 3H), 1.13–1.22 (m, 1H), 1.27–1.35 (m, 1H), 1.39–1.47 (m, 1H), 1.51–1.60 (m, 2H), 1.60–1.73 (m, 6H), 4.12 (dd, *J* = 7.6, 8.8 Hz, 1H), 4.52–4.57 (m, 1H), 7.02 (s, 1H), 7.15 (d, *J* = 1.9 Hz, 1H), 7.37 (s, 1H), 7.82 (d, *J* = 9.0 Hz, 1H), 8.74 (d, *J* = 1.9 Hz, 1H), 8.97 (d, *J* = 7.9 Hz, 1H).

Isoxazole-Cha-Chg-GRL-NH_2 (**11**). t_R = 8.4 min; HRMS: $[MH]^+$ 731.4563 (calc. for $C_{35}H_{59}N_{10}O_7^+$) 731.4561 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.79–1.21 (m, 16H), 1.30 (m, 1H), 1.40–1.73 (m, 20H), 3.07 (q, *J* = 6.8 Hz, 2H), 3.66–3.80 (2 sets of dd, *J* = 5.7 Hz, 2H), 4.14 (t, *J* = 7.8 Hz, 1H), 4.20 (m, 1H), 4.29 (m, 1H), 4.57 (m, 1H), 6.97 (br s, 1H), 7.16 (d, *J* = 1.9 Hz, 1H), 7.29 (br s, 1H), 7.48 (t, *J* = 5.7 Hz, 1H), 7.87–7.99 (m, 3H), 8.26 (t, *J* = 5.6 Hz, 1H), 8.75 (d, *J* = 1.9 Hz, 1H), 8.99 (d, *J* = 8.3 Hz, 1H).

Isoxazole-Cha-Chg-GR-NH_2 (**12**). t_R = 7.7 min; HRMS: $[MH]^+$ 618.3722 (calc. for $C_{29}H_{48}N_9O_6^+$) 618.3723 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.82–1.21 (m,

10H), 1.30 (m, 1H), 1.39–1.74 (m, 17H), 3.08 (q, *J* = 6.7 Hz, 2H), 3.72 (m, 2H), 4.13 (t, *J* = 8.0 Hz, 1H), 4.19 (m, 1H), 4.56 (m, 1H), 7.11 (br s, 1H), 7.16 (d, *J* = 1.9 Hz, 1H), 7.39 (br s, 1H), 7.51 (t, *J* = 5.6 Hz, 1H), 7.87–7.95 (2 sets of d, *J* = 8.6 Hz, 2H), 8.24 (t, *J* = 5.6 Hz, 1H), 8.75 (d, *J* = 1.9 Hz, 1H), 8.99 (d, *J* = 8.4 Hz, 1H).

NH_2-Cha-Chg-GR-NH_2 (**15**). t_R = 5.9 min; HRMS: $[MH]^+$ 523.3715 (calc. for $C_{25}H_{47}N_8O_4^+$) 523.3715 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.78–1.21 (m, 10H), 1.28 (m, 1H), 1.40–1.77 (m, 17H), 3.09 (q, *J* = 6.2 Hz, 2H), 3.69–3.83 (2 sets of dd, *J* = 5.7 Hz, 2H), 3.89 (m, 1H), 4.17–4.27 (m, 2H), 7.09 (br s, 1H), 7.39 (br s, 1H), 7.62 (t, *J* = 5.6 Hz, 1H), 7.93 (d, *J* = 8.0 Hz, 1H), 8.08 (d, *J* = 4.25 Hz, 3H), 8.30 (t, *J* = 5.9 Hz, 1H), 8.47 (d, *J* = 8.7 Hz, 1H).

NH_2-Chg-GR-NH_2 (**16**). t_R = 2.9 min; HRMS: $[MH]^+$ 370.2561 (calc. for $C_{16}H_{32}N_7O_3^+$) 370.2562 (found); ^{1}H NMR (600 MHz, DMSO–d_6), δ 0.99–1.21 (m, 5H), 1.40–1.77 (m, 10H), 3.09 (q, *J* = 6.7 Hz, 2H), 3.62 (m, 1H), 3.79–3.95 (2 sets of dd, *J* = 5.7 Hz, 2H), 4.21 (m, 1H), 7.10 (br s, 1H), 7.42 (br s, 1H), 7.63 (t, *J* = 5.8 Hz, 1H), 8.05–8.16 (m, 4H), 8.61 (t, *J* = 5.6 Hz, 1H).

Cell culture, expression constructs and transfections

Cell culture reagents were obtained from Invitrogen (Carlsbad, CA) or Sigma Aldrich (St. Louis, MO). Chinese Hamster (CHO) cells were cultured in medium at 37°C and 5% CO_2 according to ATCC instructions. CHO cells were grown in F12 supplemented with 200 µg/mL hygromycin B. PC3 cells were grown in DMEM. All media were supplemented with 10% FBS, 100 units/mL penicillin and 100 units/mL streptomycin. During cell culture passage, cell dissociation solution (Sigma Aldrich, St Louis, MO) was used to dissociate cells from culture flask surfaces and cells were counted manually using a haemocytometer or automated cell counter. MDA-MB-231 cells were maintained in L-15 with 10% FBS and 1% penicillin/streptromycin. Cells were cultured at 37°C in a humidified incubator. Site directed mutagenesis to generate PAR2 expression constructs was performed using a previously described.[1]PAR2 construct as template and the QuickChange kit (Stratagene) according to manufacturer's instructions. Primer sequences are available on request. Sequences of all constructs were confirmed by at the Australian Genomic Research Facility (St Lucia, Australia). CHO cells were transfected using Lipofectamine 2000, and stable transfectants selected in media containing hygromycin B (600 µg/mL).

Homology modeling based on human Nociceptin/Orphanin ORL1

Human PAR2 protein sequence (P55085) was obtained from Uniprot and aligned using sALign to the sequence of the crystal structure of the Nociceptin/orphanin FQ receptor (4EA3, sequence identity = 28%, resolution = 3.01 Å.[2] Although human PAR1 shares a higher sequence identity with PAR2 (41%), we did not choose the PAR1 crystal structure as the template for a PAR2 homology model due to the large bound natural product ligand vorapaxar, which in our opinion is distorted from the likely binding site for smaller ligands. However, no such complication existed for the nociceptin/orphanin FQ receptor crystal structure, which we instead chose as the template for our PAR2 model. The produced alignment was then manually inspected and adjusted in Jalview program. 20 homology models were built using Modeller 9v10. The model with the lowest discrete optimization protein energy (DOPE) score was selected for further refinement. ECL2 was further refined using the loop refine-

ment option in Prime and the final model that incorporated the optimized ECL2 was minimized in Prime using the truncated-Newton energy minimization (OPLS_2005 force field with restrained helical backbone). The final model was finally refined using protein preparation wizard in Schrodinger to optimize hydrogen bond networks and for a restrained energy minimization (OPLS_2005 force field and heavy atom movement <0.5 Å). The quality of the homology models was assessed for different physical and chemical parameters. "Structure Assessment" module available at Swiss-Model portal (http://swissmodel.expasy.org/).

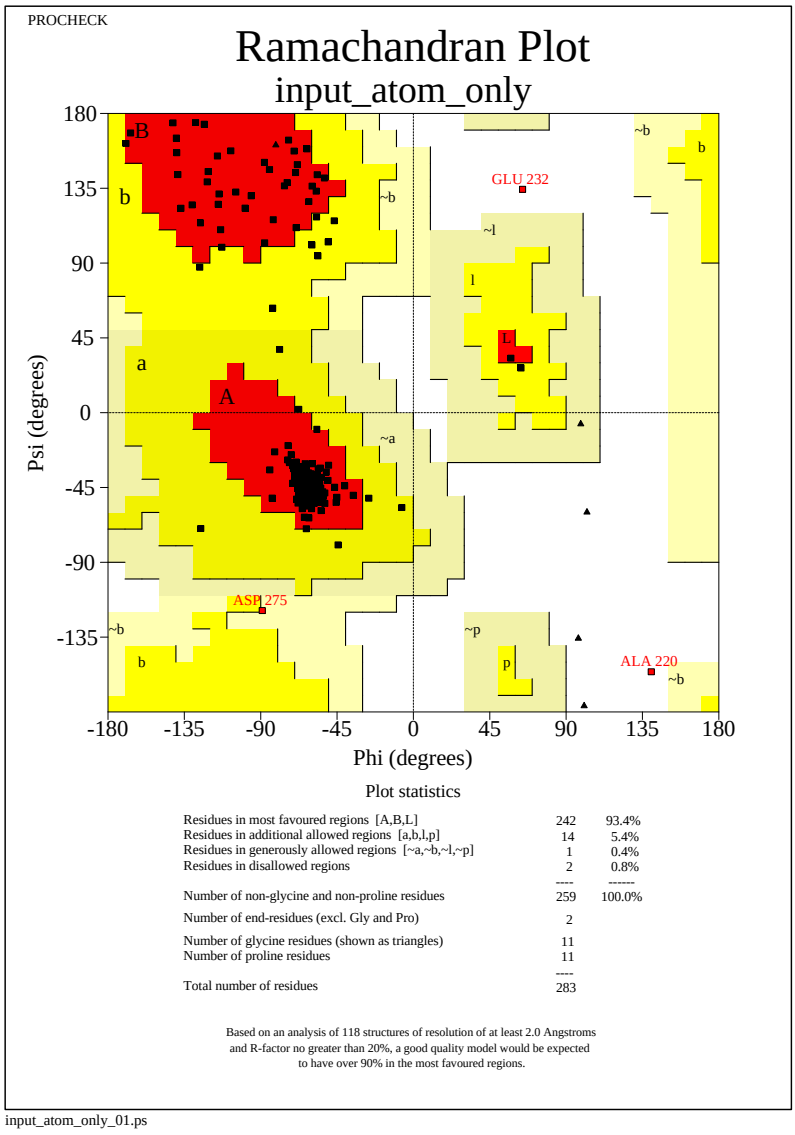

Figure S1. Ramachandran plot of PAR2 homology model derived from the crystal structure of the Nociceptin/orphanin FQ receptor (4EA3).

Ligand Preparation. The structure of **14** was drawn in ChemBioDraw13.0 as a neutral species with correct stereochemistry and saved as a 2D sdf file. LigPrep in Schrodinger Suite software (v 9.4) was used to convert the 2D sdf file into 3D maestro and sdf files. LigPrep generated a single 3D structure per ligand, minimized using the OPLS2005 force field and protonation state corrected to pH 7.4 using Epik.

Docking protocol. Agonist **14** was docked using GOLD 5.2.1 (CCDC, Cambridge, UK). Docking was done with default parameters without constraints. Final analysis and visualization was done using PyMOL Molecular Graphics System, Version 1.6.0.0, Schrödinger, LLC.

References

(1) Ramsay, A. J.; Dong, Y.; Hunt, M. L.; Linn, M.; Samaratunga, H.; Clements, J. A.; Hooper, J. D. Kallikrein-related peptidase 4 (KLK4) initiates intracellular signaling via protease-activated receptors (PARs). KLK4 and PAR-2 are co-expressed during prostate cancer progression. *J. Biol. Chem.* **2008**, *283*, 12293-12304.
(2) Thompson, A. A.; Liu, W.; Chun, E.; Katritch, V.; Wu, H.; Vardy, E.; Huang, X. P.; Trapella, C.; Guerrini, R.; Calo, G.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* **2012**, *485*, 395-399.