SUPPORTING INFORMATION

Synthetic Lethality in Pancreatic Cancer: Discovery of a New RAD51-BRCA2 Small Molecule Disruptor That Inhibits Homologous Recombination and Synergizes with Olaparib

Greta Bagnolini,^{§,†,+} Domenico Milano,^{§,+} Marcella Manerba,^{§,+} Fabrizio Schipani,[§] Jose Antonio Ortega,[§] Dario Gioia,[§] Federico Falchi,[§] Andrea Balboni,^{§,†} Fulvia Farabegoli,[†] Francesca De Franco,[†] Janet Robertson,[†] Roberto Pellicciari,[†] Isabella Pallavicini,[⊥] Sebastiano Peri,[⊥] Saverio Minucci,^{‡,⊥} Stefania Girotto,[§] Giuseppina Di Stefano,[#] Marinella Roberti*[†] and Andrea Cavalli*^{§,†}

[§]Computational & Chemical Biology, Istituto Italiano di Tecnologia, via Morego 30, 16163 Genoa, Italy.

[†]Department of Pharmacy and Biotechnology, University of Bologna, Via Belmeloro 6, 40126 Bologna, Italy.

⁷TES Pharma S.r.l., Via Palmiro Togliatti 22bis, I-06073, Corciano, Perugia, Italy.

¹Department of Biosciences, University of Milan, Via Celoria 26, 20100 Milan.

[‡]Department of Experimental Oncology at the IEO, European Institute of Oncology IRCCS, IFOM-IEO Campus, Via Adamello 16, 20100 Milan.

[#]Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Via S. Giacomo 14, 40126 Bologna, Italy.

TABLE OF CONTENT

Computational methods	
Separation and ELISA assay results of enantiomers 4d-I and 4d-II	
Scheme S1. Synthesis of aldehyde 67S6	
Scheme S2. Synthesis of aldehyde 70	
Scheme S3. Synthesis of aldehyde 74	
Scheme S4. Synthesis of aldehyde 75	
Scheme S5. Synthesis of aldehydes 76, 77	
Scheme S6. Synthesis of aldehydes 78, 79 S10	
Scheme S7. Synthesis of aldehydes 80, 81	
Scheme S8. Synthesis of aldehydes 82, 83	
Scheme S9. Synthesis of aldehyde 84	
¹ H-NMR spectrum, ¹³ C-NMR spectrum and HPLC-MS analysis of final compounds 4d-10d , 14d	l-
15d, 18d-57d and HPLC-MS analysis of selected final compounds 4d-10d, 14d-15d, 18-25d, 35d	I,
49d , 57d	

Computational methods

Protein Preparation

The crystal structure of a RAD51-BRCA2 BRC repeat complex was downloaded from the Protein Data Bank (PDB code 1N0W). The structure was then treated with the Schrödinger Suite 2014-4 Protein Preparation Wizard tool. All the selenomethionines were mutated to methionine, water molecules and ions were removed, and an exhaustive sampling of the orientations of groups, whose hydrogen bonding network needs to be optimized, was performed. Finally, the protein structure was refined to relieve steric clashes with a restrained minimization with the OPLS2005 force field until a final RMSD of 0.30 Å with respect to the input protein coordinates.

Database Preparation

A commercially available library of compounds composed of ASINEX and LifeChemicals databases collected from ZINC was prepared with the LigPrep tool of the Schrödinger Suite. The 2D (smi file) structures were converted to 3D structures and for each entry all stereoisomers were generated. The resulting molecules were submitted to Epik and all the tautomers and ionization states at pH 7.0 \pm 2.0 were calculated. Finally, duplicates, compounds with more than 2 chiral centers, Pan-Assay Interference Compounds (PAINS), compounds with Michael acceptor groups, and frequent hitters were deleted. To enrich the database with potential Protein Protein Interaction Inhibitors, the database was filtered with the PPI-HitProfiler tool using the "soft" methods.

High Throughput Docking (Virtual Screening)

All filtered ligands (about 750K) were docked with Glide SP by centering the grid on the position of BRCA **Phe1546**. The 10K top-scoring compounds were re-docked with Glide XP and the 1K top-scoring compounds were selected. Both grid generation and docking calculations were performed with the default settings. The selected compounds were visually inspected to identify compounds able to match the interactions between RAD51 and BRCA and **42** compounds were selected and purchased.

Induced-Fit Docking

IFD (Induced-Fit Docking) calculations were performed with the previously prepared protein structure using both enantiomers of each ligand. All the ligands were prepared using Ligprep utility in Schrödinger 2019-2. The IFD protocol involves the use of the Glide docking program to generate a number of initial possible ligand poses followed by a protein side chain optimization using the

Prime protein structure modeling program. After a number of iterations, the process produces a list of final poses, ordered by a proprietary scoring function (that is a combination of the Prime Energy and Glide scoring function).

The Schrödinger Extended Sampling protocol was selected along with the OPLS3e force field. The grid box was centered on the centroid of residues Tyr205, Arg247, and Phe259. Residues within 5.0 Å of ligand poses were refined during the process. The other parameters were set to their default values. Resulting poses were evaluated by visual inspection.

Separation and ELISA assay results of enantiomers 4d-I and 4d-II

The semi-preparative chiral separations of the racemic **4d** by HPLC were performed on a Waters Alliance HPLC instrument consisting of a 1525 Binary HPLC Pump, Waters Fraction Collector III and a 2998 Photodiode Array Detector. The separations were run in isocratic mode on a Daicel ChiralPak AD column (250x10mmID, particle size 10 μ m) with a ChiralPak AD Semi-Prep. Guard pre-column (50x10mmID, particle size 10 μ m). The mobile phase was Heptane-2-Propanol (75:25) with a flowrate = 5mL/min.

To determine the enantiomeric excess (ee) of the enantiomers **4d-I** and **4d-II**, the analytical chiral separations by HPLC were run on a Waters Alliance HPLC instrument consisting of an e2695 Separation Module and a 2998 Photodiode Array Detector. The PDA range was 210-400nm. The analyses were performed in isocratic mode on a Daicel ChiralPak AD column (250x4.6mmID, particle size 10µm). The mobile phase was Heptane-2-Propanol (50:50) with a flow rate = 1mL/min; **4d-I**: $t_R = 6.411min$, >99.5% ee at 240nm; **4d-II**: $t_R = 13.431min$, 96.6% ee at 240nm. The ¹H NMR spectrum was identical to that of racemic **4d** for each enantiomer.

Table S1.

Compound	Retention time (t _R)	QC (UV) @215 nm	Enantiomeric excess (ee)	EC ₅₀ ELISA (µM)
4d-I	6.411 minutes	97%	>99.5%	4 ± 0.5
4d-II	13.431 minutes	99%	96.6%	10 ± 1

Scheme S1: Synthesis of aldehyde 67

The 3-fluoro-4-methoxybenzaldehyde **67** was achieved by the alkylation of the commercially available 3-fluoro-4-hydroxybenzaldehyde **125** with methyl iodide under basic conditions.



Reagents and conditions. Synthesis of **67**. Reagents and conditions: (a) Potassium carbonate (3.3 equiv), anhydrous DMF, methyl iodide **126** (1.2 equiv), 70 °C, 3.5 h.

3-Fluoro-4-methoxybenzaldehyde (67). In a screw capped pressure tube 3-fluoro-4-hydroxybenzaldehyde **125** (461 mg, 3.29 mmol) was dissolved with 3.0 ml of anhydrous DMF then potassium carbonate (1.49 g, 10.8 mmol) was added and the mixture kept under stirring for 5 minutes at rt. Methyl iodide **126** (559 mg, d = 2.28 g/mL, 245 μ L, 3.95 mmol) was then added. The reaction was thus heated to 70 °C and stirred for 3.5 hours. The crude mixture was diluted with CHCl₃ and washed three times with water. The organic layer was dried through a phase separator and the solvent removed under reduced pressure. The crude was purified by normal phase flash column chromatography (SiO₂ 24 g; 0-20% EtOAc/ cyclohexane) to afford **67**. Yield 440 mg, 87%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.87 (d, *J* = 2.1 Hz, 1H), 7.79 (ddd, *J* = 8.4, 2.0, 1.0 Hz, 1H), 7.69 (dd, *J* = 11.4, 2.0 Hz, 1H), 7.39 (t, *J* = 8.4 Hz, 1H), 3.95 (s, 3H). R_t 1.59 min (generic method). ESI-MS for C₈H₇FO₂: calculated 154.0, found *m/z* 155.3 [M+H]⁺.

Scheme S2: Synthesis of aldehyde 70

4'-Fluoro-[1,1'-biphenyl]-4-carbaldehyde **70** was synthesized taking advantage of Suzuki coupling reaction between the commercially available 4-iodobenzaldehyde **127** and (4-fluorophenyl) boronic acid **128** in presence of palladium(0) catalyst.



Reagents and conditions: (a) i. Anhydrous sodium carbonate (3.0 equiv); ii. DMF, H₂O, rt, 15 minutes; iii. Tetrakis(triphenylphosphine)palladium(0) (0.15 equiv), 110 °C, 20 h.

4'-Fluoro-[1,1'-biphenyl]-4-carbaldehyde (**70**). In a screw capped pressure tube 4iodobenzaldehyde **127** (255 mg, 1.1 mmol), (4-fluorophenyl) boronic acid **128** (231 mg, 1.5 mmol) and anhydrous sodium carbonate (350 mg, 3.3 mmol) were added. The tube was deoxygenated with three cycles vacuum/Ar, then DMF (2 mL) and H₂O (400 μ L) were added. The solution was stirred at rt, under Ar flux, for 15 minutes. Tetrakis(triphenylphosphine)palladium(0) (185 mg, 0.16 mmol) was thus added, the mixture was heated to 110 °C and stirred for 20 hours. The crude was then diluted with DCM and washed with water (3 x 50 mL). The organic layer was dried over Na₂SO₄ and the solvent was then removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0-7% EtOAc/ cyclohexane). Yield 226 mg, 66%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 10.06 (s, 1H), 8.03 – 7.97 (m, 2H), 7.94 – 7.88 (m, 2H), 7.87 – 7.81 (m, 2H), 7.40 – 7.32 (m, 2H). R_t 2.30 min (generic method).

Scheme S3: Synthesis of aldehyde 74

1-Methyl-1H-indole-5-carbaldehyde **74** was achieved by the alkylation of the commercially available 1H-indole-5-carboxaldehyde **129** with methyl iodide **126** under basic conditions.



Reagents and conditions: (a) Potassium carbonate (2.0 equiv), anhydrous DMF, methyl iodide **126** (2.0 equiv), 35 °C, 18 h.

1-Methyl-1H-indole-5-carbaldehyde (74). In a screw capped pressure tube 1*H*-indole-5-carboxaldehyde 129 (333 mg, 2.3 mmol) was dissolved with 2.0 mL of anhydrous DMF then

potassium carbonate (630 mg, 4.6 mmol) was added and the mixture kept under stirring for 10 minutes at rt. Methyl iodide **126** (647 mg, d = 2.28 g/mL, 284 μ L, 4.6 mmol) was then added. The reaction was thus heated to 35 °C and stirred for 18 hours. The crude mixture was diluted with DCM and washed three times with water. The organic layer was anhydrified through a phase separator and the solvent removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0- 7.5% EtOAc/ cyclohexane). Yield 286 mg, 81%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.98 (s, 1H), 8.18 (d, *J* = 1.5 Hz, 1H), 7.74 – 7.56 (m, 2H), 7.49 (d, *J* = 3.2 Hz, 1H), 6.67 (dd, *J* = 3.1, 0.8 Hz, 1H), 3.85 (d, *J* = 1.3 Hz, 3H). R_t 1.75 min (generic method). ESI-MS for C₁₀H₉NO: calculated 159.1, found *m/z* 160.1 [M+H]⁺.

Scheme S4: Synthesis of aldehyde 75

1-Ethyl-1H-indole-5-carbaldehyde **75** was achieved by the alkylation of the commercially available 1*H*-indole-5-carboxaldehyde **129** with iodoethane **130** under basic conditions.



Reagents and conditions: (a) Potassium carbonate (2.0 equiv), anhydrous DMF, iodoethane **130** (2.0 equiv), 35 °C, 18 h.

1-Ethyl-1H-indole-5-carbaldehyde (**75**). In a screw capped pressure tube 1*H*-indole-5-carboxaldehyde **129** (338 mg, 2.3 mmol) was dissolved with 2.0 ml of anhydrous DMF then potassium carbonate (638 mg, 4.6 mmol) was added and the mixture kept under stirring for 10 minutes at rt. Iodoethane **130** (503 mg, d = 1.94 g/mL, 344 μ L, 4.6 mmol) was then added. The reaction was thus heated to 35 °C and stirred for 18 hours. The crude mixture was diluted with DCM and washed three times with water. The organic layer was dried through a phase separator and the solvent removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0- 5%, EtOAc/ cyclohexane). Yield 363 mg, 91%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.98 (s, 1H), 8.18 (t, *J* = 1.1 Hz, 1H), 7.71 – 7.62 (m, 2H), 7.57 (d, *J* = 3.2 Hz, 1H), 6.67 (dd, *J* = 3.2, 0.7 Hz, 1H), 4.27 (q, *J* = 7.2 Hz, 2H), 1.37 (t, *J* = 7.2 Hz, 3H). R_t 1.94 min (generic method). ESI-MS for C₁₁H₁₁NO: calculated 173.1, found *m/z* 174.0 [M+H]⁺.

Scheme S5: Synthesis of aldehydes 76, 77

Ethylindazole-5-carbaldehydes **76** and **77** were synthesized by the alkylation of the commercially available 1*H*-indazole-5-carboxaldehyde **131** with ethyl bromide **132** under basic conditions.



Reagents and conditions: (a) Potassium carbonate (3.0 equiv), anhydrous DMF, ethyl bromide **132** (1.1 equiv), 40 °C, 5 h.

Ethylindazole-5-carbaldehydes (76, 77). In a screw capped pressure tube 1*H*-indazole-5carboxaldehyde **131** (300 mg, 2.84 mmol) was dissolved with 3.0 mL of anhydrous DMF then potassium carbonate (1.13 g, 8.2 mmol) was added and the mixture kept under stirring for 10 minutes at rt. Ethyl bromide **132** (335 mg, d = 1.46 g/mL, 230 µL, 3.1 mmol) was then added. The reaction was thus heated to 40 °C and stirred for 5 hours. The crude mixture was diluted with DCM and washed with water (3x50 mL). The organic layer was anhydrified through a phase separator and the solvent removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0- 50% EtOH/ DMC). Mono- and bi-dimensional ¹H- and ¹³C-NMR (HMBC) analyses confirmed the structure of the title compounds named **76** (1-ethyl-1Hindazole-5-carbaldehyde), which showed a positive ¹H-¹³C correlation between C<u>H</u>₂ at 4.49 ppm and a quaternary C at 141.7 ppm, and **77** (2-ethyl-2H-indazole-5-carbaldehyde), which showed a positive ¹H-¹³C correlation between C<u>H</u>₂ at 4.51 ppm and the C₃ at 127.5 ppm, as shown in the reaction scheme above according to their elution order. Yields: (**76**) 179 mg, 50%; (**77**) 121 mg, 34%.

76) ¹H NMR (400 MHz, DMSO- d_6) δ 10.04 (s, 1H), 8.43 (t, J = 1.1 Hz, 1H), 8.34 (d, J = 0.8 Hz, 1H), 7.92 – 7.81 (m, 2H), 4.50 (q, J = 7.2 Hz, 2H), 1.42 (t, J = 7.2 Hz, 3H). ¹³C NMR (101 MHz, DMSO- d_6) δ 192.67, 141.43, 135.64, 130.49, 128.16, 124.80, 123.78, 110.94, 43.85, 15.32. R_t 1.56 min (generic method). ESI-MS for C₁₀H₁₀N₂O: calculated 174.1, found *m*/*z* 175.3 [M+H]⁺;

77) ¹H NMR (400 MHz, DMSO- d_6) δ 9.98 (s, 1H), 8.75 (s, 1H), 8.45 (d, J = 1.2 Hz, 1H), 7.69 (qd, J = 9.0, 1.3 Hz, 2H), 4.52 (q, J = 7.3 Hz, 2H), 1.54 (t, J = 7.3 Hz, 3H). ¹³C NMR (101 MHz,

DMSO- d_6) δ 192.69, 150.17, 130.99, 130.81, 127.49, 122.46, 121.23, 118.24, 48.63, 16.02. R_t 1.36 min (generic method). ESI-MS for C₁₀H₁₀N₂O: calculated 174.1, found *m*/*z* 175.3 [M+H]⁺.

Scheme S6: Synthesis of aldehydes 78, 79

Propylindazole-5-carbaldehydes **78** and **79** were synthesized by the alkylation of the commercially available 1H-indazole-5-carboxaldehyde **131** with 1-bromopropane **133** under basic conditions.



Reagents and conditions: (a) Potassium carbonate (2.6 equiv), anhydrous DMF, 1-bromopropane **133** (2.0 equiv), rt, 14 h.

Propylindazole-5-carbaldehydes (78, 79). In a screw capped pressure tube 1H-indazole-5carboxyaldehyde 131 (152 mg, 1.0 mmol) was dissolved with 1.1 mL of anhydrous DMF then potassium carbonate (359 mg, 2.6 mmol) was added and the mixture kept under stirring for 30 minutes at rt. 1-Bromopropane 133 (256 mg, d = 1.353 g/mL, 189 µL, 2.0 mmol) was then added. The reaction was thus stirred for 14 hours at rt. The crude mixture was diluted with DCM and washed with water (3x50 mL). The organic layer was dried through a phase separator and the solvent removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0-70% EtOH/ DCM). Mono- and bi-dimensional ¹H- and ¹³C-NMR (HMBC) analyses confirmed the structure of the title compounds named 78 (1-propyl-1H-indazole-5carbaldehyde), which showed a positive ${}^{1}\text{H}{}^{-13}\text{C}$ correlation between CH₂ at 4.42 ppm and a quaternary C at 142.0 ppm, and 79 (2-propyl-2H-indazole-5-carbaldehyde), which showed a positive ${}^{1}\text{H}-{}^{13}\text{C}$ correlation between CH₂ at 4.44 ppm and the C₃ at 128.2 ppm, as shown in the reaction scheme above according to their elution order. Yields: (78) 99 mg, 50%; (79) 66 mg, 34%. **78**) ¹H NMR (400 MHz, DMSO- d_6) δ 10.02 (s, 1H), 8.42 (d, J = 1.2 Hz, 1H), 8.33 (s, 1H), 7.88 – 7.80 (m, 2H), 4.42 (t, J = 6.9 Hz, 2H), 1.85 (h, J = 7.2 Hz, 2H), 0.85 – 0.77 (m, 3H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 192.18, 141.58, 135.19, 129.99, 127.69, 124.33, 123.15, 110.55, 49.84, 22.80, 11.02. R_t 1.82 min (generic method). ESI-MS for C₁₁H₁₂N₂O: calculated 188.1, found m/z 189.1 $[M+H]^{+}$.

79) ¹H NMR (400 MHz, DMSO- d_6) δ 9.97 (s, 1H), 8.73 (s, 1H), 8.44 (d, J = 1.4 Hz, 1H), 7.78 – 7.62 (m, 2H), 4.43 (t, J = 7.0 Hz, 2H), 1.95 (h, J = 7.2 Hz, 2H), 0.85 (t, J = 7.4 Hz, 3H). ¹³C NMR (101 MHz, DMSO- d_6) δ 192.21, 149.74, 130.53, 130.36, 127.71, 121.97, 120.65, 117.78, 54.60, 23.20, 10.83. R_t 1.61 min (generic method). ESI-MS for C₁₁H₁₂N₂O: calculated 188.1, found m/z 189 [M+H]⁺.

Scheme S7: Synthesis of aldehydes 80, 81

Cyclohexylindazole-5-carbaldehydes **80** and **81** were synthesized by the alkylation of the commercially available 1H-indazole-5-carboxaldehyde **131** with 1-bromocyclohexane **134** under basic conditions.



Reagents and conditions: (a) Potassium carbonate (2.1 equiv), anhydrous DMF, 1-bromocyclohexane **134** (2.1 equiv), rt, 6 days.

Cyclohexylindazole-5-carbaldehydes (80, 81). In a screw capped pressure tube 1*H*-indazole-5-carboxaldehyde **131** (165 mg, 2.13 mmol) was dissolved with 1.0 mL of anhydrous DMF then potassium carbonate (624 mg, 4.52 mmol) was added and the mixture kept under stirring for 30 minutes at rt. Bromocyclohexane **134** (737 mg, d = 1.335 g/mL, 552 μ L, 4.52 mmol) was then added. The reaction was thus stirred at rt for 6 days. The crude mixture was diluted with DCM and washed with water (3x50 mL). The organic layer was dried through a phase separator and the solvent removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0-50% EtOH/ DCM). Mono- and bi-dimensional ¹H- and ¹³C-NMR (HMBC) analyses confirmed the structure of the title compounds named **80** (1-cyclohexyl-1H-indazole-5-carbaldehyde) and **81** (2-cyclohexyl-2H-indazole-5-carbaldehyde) according to their elution order, as shown in the reaction scheme above. **81** showed a positive ¹H-¹³C correlation between C<u>H</u> at 4.54 ppm and the C₃ at 126.1 ppm. Yields: (**80**) 111 mg, 21%; (**81**) 75 mg, 14%.

80) ¹H NMR (400 MHz, DMSO- d_6) δ 10.02 (s, 1H), 8.41 (t, J = 1.0 Hz, 1H), 8.32 (s, 1H), 7.91 – 7.81 (m, 2H), 4.67 (tt, J = 10.0, 5.0 Hz, 1H), 1.89 (ddt, J = 24.4, 11.6, 3.7 Hz, 6H), 1.71 (dt, J = 12.7, 3.3 Hz, 1H), 1.50 (dtt, J = 16.7, 8.1, 4.4 Hz, 2H), 1.27 (qt, J = 12.8, 3.5 Hz, 1H). ¹³C NMR (101 MHz, DMSO- d_6) δ 192.18, 140.59, 134.95, 130.06, 127.69, 124.09, 123.14, 110.54, 56.92, 32.23, 24.94. R_t 2.30 min (generic method). ESI-MS for C₁₄H₁₆N₂O: calculated 228.1, found m/z 229.0 [M+H]⁺.

81) ¹H NMR (400 MHz, DMSO- d_6) δ 9.97 (s, 1H), 8.76 (d, J = 0.9 Hz, 1H), 8.43 (t, J = 1.2 Hz, 1H), 7.74 – 7.62 (m, 2H), 4.53 (tt, J = 11.4, 3.8 Hz, 1H), 2.18 – 2.09 (m, 2H), 1.97 – 1.81 (m, 4H), 1.78 – 1.65 (m, 1H), 1.46 (qt, J = 12.9, 3.5 Hz, 2H), 1.27 (qt, J = 12.8, 3.5 Hz, 1H). ¹³C NMR (101 MHz, DMSO- d_6) δ 192.23, 149.27, 130.48, 130.40, 125.65, 121.89, 120.46, 117.88, 62.06, 33.08, 24.83, 24.77. R_t 2.05 min (generic method). ESI-MS for C₁₄H₁₆N₂O: calculated 228.1, found *m*/*z* 228.8 [M+H]⁺.

Scheme S8: Synthesis of aldehydes 82, 83

The synthesis of the 4-(ethylindazol-5-yl)benzaldehydes **82** and **83** start from the Suzuki coupling reaction between the commercially available 4-iodobenzaldehyde **127** and 1*H*-indazole-5-boronic acid **135** in presence of palladium(0) catalyst to give the corresponding intermediate **136**. The alkylation of **136** with ethyl bromide **132** under basic conditions afforded the aldehydes **82** and **83**.



Reagents and conditions: (a) i. Anhydrous sodium carbonate (3.0 equiv); ii. DMF, H₂O, rt, 15 minutes; iii. Tetrakis(triphenylphosphine)palladium(0) (0.15 equiv), 110 °C, 22 h; (b) Potassium carbonate (2.0 equiv), anhydrous DMF, ethyl bromide **132** (2.5 equiv), rt, 40 h.

4-(1*H***-Indazol-5-yl)benzaldehyde (136)**. In a Schlenk tube 4-iodobenzaldehyde **127** (300 mg, 1.3 mmol), 1*H*-indazole-5-boronic acid **135** (316 mg, 1.9 mmol) and anhydrous sodium carbonate (413 mg, 3.9 mmol) were added. The tube was deoxygenated with three cycles vacuum/Ar, then DMF (2 mL) and H₂O (400 µL) were added. The solution was stirred at rt, under Ar flux, for 15 minutes. Tetrakis(triphenylphosphine)palladium(0) (220 mg, 0.19 mmol) was thus added, the mixture was heated to 110 °C and stirred for 22 hours. The crude was diluted with DCM and washed with water (3 x 50 mL), the solvent thus removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0-5% EtOH/ DCM). Yield 205 mg, 71%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 13.20 (s, 1H), 10.05 (s, 1H), 8.22 – 8.12 (m, 2H), 7.98 (q, *J* = 8.5 Hz, 4H), 7.84 – 7.61 (m, 2H). R_t 1.77 min (generic method). ESI-MS for C₁₄H₁₀N₂O: calculated 222.1, found *m*/*z* 223.0 [M+H]⁺, 221.2 [M-H]⁻

4-(Ethylindazol-5-yl)benzaldehydes (82, 83). In a screw capped pressure tube 4-(1*H*-indazol-5-yl)benzaldehyde **136** (204 mg, 0.92 mmol) was dissolved with 1.0 mL of anhydrous DMF then potassium carbonate (254 mg, 1.8 mmol) was added and the mixture kept under stirring for 15 minutes at rt. Ethyl bromide **132** (251 mg, d = 1.46 g/mL, 172 μ L, 2.3 mmol) was then added. The reaction was thus stirred at rt for 40 hours. The crude mixture was diluted with DCM and washed with water (3x50 mL). The organic layer was dried through a phase separator and the solvent

removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0- 5% EtOH/ DMC). Mono- and bi-dimensional ¹H- and ¹³C-NMR analyses confirmed the structure of the title compounds named **82** (4-(1-ethyl-1H-indazol-5-yl)benzaldehyde) and **83** (4-(2-ethyl-2H-indazol-5-yl)benzaldehyde) according to their elution order, as shown in the reaction scheme above. Yields: (**82**) 125 mg, 54%; (**83**) 82 mg, 39%.

82) ¹H NMR (400 MHz, DMSO- d_6) δ 10.05 (s, 1H), 8.17 (t, J = 1.3 Hz, 1H), 8.15 (s, 1H), 8.02 – 7.94 (m, 4H), 7.80 (d, J = 1.3 Hz, 2H), 4.48 (q, J = 7.2 Hz, 2H), 1.42 (t, J = 7.2 Hz, 3H). ¹³C NMR (101 MHz, DMSO- d_6) δ 192.65, 146.47, 138.64, 134.57, 133.34, 131.34, 130.17, 127.36, 125.54, 124.24, 119.64, 110.32, 43.20, 14.93. R_t 2.19 min (generic method). ESI-MS for C₁₆H₁₄N₂O: calculated 250.1, found m/z 251.1 [M+H]⁺.

83) ¹H NMR (400 MHz, DMSO- d_6) δ 10.05 (s, 1H), 8.49 (d, J = 0.9 Hz, 1H), 8.13 (dd, J = 1.8, 0.9 Hz, 1H), 8.02 – 7.91 (m, 4H), 7.78 – 7.60 (m, 2H), 4.49 (q, J = 7.3 Hz, 2H), 1.53 (t, J = 7.3 Hz, 3H). ¹³C NMR (101 MHz, DMSO- d_6) δ 192.63, 147.68, 146.74, 134.52, 131.49, 130.14, 127.19, 124.97, 124.34, 121.86, 119.45, 117.72, 47.88, 15.72. R_t 1.97 min (generic method). ESI-MS for C₁₆H₁₄N₂O: calculated 250.1, found *m*/*z* 251.1 [M+H]⁺.

Scheme S9: Synthesis of aldehyde 84

4-(1-Propyl-1H-pyrazol-4-yl)benzaldehyde **84** was achieved taking advantage of Suzuki coupling reaction between the commercially available 4-iodobenzaldehyde **127** and (1-propyl-1*H*-pyrazol-4-yl)boronic acid **137** in presence of palladium(0) catalyst.



Reagents and conditions: (a) i. Anhydrous sodium carbonate (3.0 equiv); ii. DMF, H₂O, rt, 15 minutes; iii. Tetrakis(triphenylphosphine)palladium(0) (0.15 equiv), 110 °C, 18 h.

4-(1-Propyl-1H-pyrazol-4-yl)benzaldehyde (84). In a Schlenk tube 4-iodobenzaldehyde 127 (326 mg, 1.4 mmol), (1-propyl-1*H*-pyrazol-4-yl)boronic acid 137 (323 mg, 2.1 mmol) and anhydrous sodium carbonate (445 mg, 4.2 mmol) were added. The tube was deoxygenated with three cycles

vacuum/Ar, then anhydrous DMF (1.1 mL) and H₂O (300 µL) were added. The solution was stirred at rt, under Ar flux, for 15 minutes. Tetrakis(triphenylphosphine)palladium(0) (243 mg, 0.2 mmol) was thus added, the mixture was heated to 110 °C and stirred for 18 hours. The reaction mixture was diluted with DCM and washed with water (3x50 mL), the solvent was thus removed under reduced pressure. Purified by normal phase flash column chromatography (SiO₂ gold 24 g, 0-30% EtOAc/cyclohexane). Yield 238 mg, 97%. ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.95 (s, 1H), 8.37 (s, 1H), 8.03 (s, 1H), 7.94 – 7.75 (m, 4H), 4.10 (t, *J* = 6.9 Hz, 2H), 1.83 (h, *J* = 7.2 Hz, 2H), 0.85 (t, *J* = 7.4 Hz, 3H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 192.17, 138.95, 136.76, 133.78, 130.30, 128.33, 125.10, 120.48, 53.14, 23.05, 10.90. Rt 1.86 min (generic method). ESI-MS for C₁₃H₁₄N₂O: calculated 214.1, found *m/z* 215.5 [M+H]⁺.

¹H-NMR spectrum, ¹³C-NMR spectrum and HPLC-MS analysis of final compounds 4d-10d, 14d-15d, 18d-57d

¹H-NMR spectrum (400 MHz, DMSO- d_6) of **4d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **4d**



HPLC-MS analysis of 4d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **5d**



13 C-NMR spectrum (101 MHz, DMSO- d_6) of **5d**



HPLC-MS analysis of **5d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **6d**



 13 C-NMR spectrum (101 MHz, DMSO- d_6) of **6d**





S23





¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **7d**



HPLC-MS analysis of **7d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **8d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **8d**



HPLC-MS analysis of 8d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **9d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **9d**



HPLC-MS analysis of 9d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **10d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **10d**



HPLC-MS analysis of 10d




¹H-NMR spectrum (400 MHz, DMSO- d_6) of **14d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **14d**



HPLC-MS analysis of 14d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **15d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **15d**



HPLC-MS analysis of 15d









¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **18d**.



HPLC-MS analysis of 18d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **19d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **19d**









¹H-NMR spectrum (400 MHz, DMSO- d_6) of **20d**

¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **20d**



HPLC-MS analysis of 20d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **21d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **21d**











¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **22d**



HPLC-MS analysis of 22d









¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **23d**



HPLC-MS analysis of 23d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **24d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **24d**



HPLC-MS analysis of 24d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **25d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **25d**



HPLC-MS analysis of **25d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **26d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **26d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **27d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **27d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **28d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **28d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **29d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **29d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **30d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **30d**







¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **31d**


¹H-NMR spectrum (400 MHz, DMSO- d_6) of **32d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **32d**





¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **33d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **34d**

¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **34d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **35d**

¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **35d**



HPLC-MS analysis of 35d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **36d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **36d**







¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **37d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **38d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **38d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **39d**

¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **39d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **40d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **40d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **41d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **41d**







¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **42d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **43d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **43d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **44d**

¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **44d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **45d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **45d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **46d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **46d**







¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **47d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **48d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **48d**





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **49d**

¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **49d**



HPLC-MS analysis of 49d





¹H-NMR spectrum (400 MHz, DMSO- d_6) of **50d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **50d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **51d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **51d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **52d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **52d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **53d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **53d**







¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **54d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **55d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **55d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **56d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **56d**



¹H-NMR spectrum (400 MHz, DMSO- d_6) of **57d**



¹³C-NMR spectrum (101 MHz, DMSO- d_6) of **57d**



HPLC-MS analysis of 57d



