Supporting Information

Virtual screening and biological validation of novel influenza virus PA endonuclease inhibitors

Nicolino Pala,^{†,||,*} Annelies Stevaert,^{‡,||} Roberto Dallocchio,[§] Alessandro Dessì,[§] Dominga Rogolino,[⊥] Mauro Carcelli,[⊥] Vanna Sanna,[†] Mario Sechi,[†] Lieve Naesens^{‡,*}

[†]Dipartimento di Chimica e Farmacia, Università di Sassari, Via Vienna 2, 07100 Sassari, Italy; [‡]Rega Institute for Medical Research, KU Leuven, B-3000 Leuven, Belgium; [§]CNR-Consiglio Nazionale delle Ricerche - Istituto di Chimica Biomolecolare, Sassari, 07100 Li Punti, Italy; [⊥]Dipartimento di Chimica, Università di Parma, Parco Area delle Scienze 17/A, 43124 Parma, Italy

Author Contributions: ^{||}These authors contributed equally to this work.

*Corresponding authors:

Nicolino Pala, PhD Dipartimento di Chimica e Farmacia Università di Sassari Via Vienna 2, 07100 Sassari, Italy Phone: +39 079 228 654; Fax: +39 079 229559 E-mail: <u>nikpal@uniss.it</u>

Lieve Naesens, PhD Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy KU Leuven, B-3000 Leuven, Belgium Minderbroedersstraat 10 Phone: +32-16-337345 E-mail: <u>lieve.naesens@rega.kuleuven.be</u>

Materials and methods

The compounds **10**, **11** and **15** have been prepared using a previously described procedure.¹ compounds **9** and **12-14** belonged to our collection, whereas the remaining compounds were purchased from different suppliers (i.e. Maybridge Ltd. for **1**; Life Chemicals Inc. for **2**; Princeton BioMolecular Research Inc. for **3**; InterBioScreen Ltd. for **4**; ChemDiv Inc. for **5**; ENAMINE Ltd. for **6**; and Vitas-M laboratory Ltd. for **7** and **8**).

Chemistry. Anhydrous solvents and all reagents were purchased from Aldrich, Merck or Carlo Erba. All reactions involving air- or moisture-sensitive compounds were performed under nitrogen atmosphere using oven-dried glassware and syringes to transfer solutions. Melting points (mp) were determined using an Electrothermal or a Köfler apparatus and are uncorrected. Infrared (IR) spectra were recorded in nujol with a Perkin-Elmer 2400 spectrophotometer and are expressed in $v \text{ cm}^{-1}$. The ESI-MS spectra were collected by using a quadrupole-time-of-flight micro mass spectrometer (Micromass, Manchester, UK) equipped with a pneumatically assisted ESI interface. The system was controlled by MassLynx software version 4.0 (Micromass). The nebulizing gas (nitrogen, 99.999% purity) and the desolvation gas (nitrogen, 99.998% purity) were delivered at a flow-rate of 10 and 600 L/h respectively. Continuum mode full-scan mass spectra were acquired using an acquisition time of 1 s and an interscan delay of 0.1 s. QqTOF external calibration was performed using a 0.1% phosphoric acid solution and a fifth-order nonlinear calibration curve was usually adopted. Nuclear magnetic resonance (1H-NMR) spectra were determined in CDCl₃ and DMSO-*d*₆ on a Varian XL-200 (200 MHz) spectrometer. Chemical shifts (δ scale) are reported in parts per million (ppm) downfield from tetramethylsilane (TMS) used as an internal standard. The assignment of exchangeable protons (OH and NH) was confirmed by the addition of D_2O . Analytical thin-layer chromatography (TLC) was performed on Merck silica gel F-254 plates. For flash chromatography, Merck Silica gel 60 was used with a particle size 0.040-0.063 mm (230-400 mesh ASTM).

Synthesis of *N*-(3,4-dihydroxyphenethyl)-5,6-dihydroxy-1*H*-indole-2-carboxamide [10]

A solution of **21** (1.0 mmol) in anhydrous dichloromethane (60 mL) was cooled to -70 °C and boron tribromide 1 M solution in dichloromethane (8.0 mmol) was slowly added under nitrogen atmosphere. The reaction mixture was stirred at about -40 °C for 4 hours, then quenched with methanol. After removal of the solvent in vacuum, the crude residue was purified by flash chromatography (ethyl acetate) to obtain a white powder. Yield: 58%; R_f: 0.57 (ethyl acetate); mp: 223 °C dec. (lit. 221-223 °C dec.¹); IR (nujol) ν cm⁻¹ = 3320 (OH), 3280 (NH), 1620 (C=O);

¹H-NMR (CDCl₃:DMSO-d₆): δ 10.53 (bs, 1H, NH), 8.40 (bs, 4H, OH), 7.76 (bs, 1H, NH), 6.91-6.70 (m, 4H, Ar-H), 6.69 (s, 1H, Ar-H), 6.55-6.51 (m, 1H, Ar-H), 3.44 (m, 2H, CH₂), 2.73 (t, 2H, CH₂); ESI/MS (+, *m/z*): 328 [M⁺]; Anal. Calcd for $C_{17}H_{16}N_2O_5$: C 62.18, H 4.91, N 8.54. Found: C 62.03, H 4.96, N 8.37.

Synthesis of *N*,*N*'-(propane-1,3-diyl)bis(5,6-dihydroxy-1*H*-indole-2-carboxamide) [11]

To a -70 °C cooled solution of **22** (1.4 mmol) in anhydrous dichloromethane (280 mL), boron tribromide 1 M solution in dichloromethane (11.2 mmol) was added slowly and the resulting mixture was stirred at 0 °C for 4 hours. After quenching with methanol, the solvents were removed under vacuum. The crude residue was washed with water and tritured with diethyl ether to obtain a gray powder. Yield: 46%; R_f : 0.43 (ethyl acetate); mp: 212-212.5 °C dec. (lit. 208-210 °C dec.¹); IR (nujol) ν cm⁻¹ = 3380 (OH), 3280 (NH), 1630 (C=O); ¹H-NMR (CDCl₃:DMSO-d₆): δ 10.90 (bs, 2H, NH), 8.24 (bs, 4H, OH), 7.78 (bs, 2H, NH), 6.95-6.84 (m, 6H, Ar-H), 3.40 (m, 4H, CH₂), 1.84 (m, 2H, CH₂); ESI/MS (+, *m/z*): 424 [M⁺]; Anal. Calcd for C₂₁H₂₀N₄O₆•0.5H₂O: C 58.18, H 4.89, N 12.93. Found: C 57.88, H 5.01, N 12.82.

Synthesis of *N*,*N*'-(butane-1,4-diyl)bis(5,6-dihydroxy-1*H*-indole-2-carboxamide) [15]

To a -70 °C cooled solution of **23** (1.0 mmol) in anhydrous dichloromethane (200 mL), boron tribromide 1M solution in dichloromethane (8.0 mmol) was added slowly, and the resulting mixture was stirred at 0 °C for 4 hours. After quenching with methanol, the solvents were removed under vacuum. The crude residue was washed with water and tritured with diethyl ether to obtain a gray powder. Yield: 65%; R_f: 0.44 (ethyl acetate); mp: 304 °C dec. (lit. 304-306 °C dec.¹); IR (nujol) ν cm⁻¹ = 3400 (OH), 3200 (NH), 1630 (C=O); ¹H-NMR (CDCl₃:DMSO-d₆): δ 10.72 (bs, 2H, NH), 8.80-8.20 (bs, 4H, OH), 8.04 (bs, 2H, NH), 6.90 (s, 2H, Ar-H), 6.86 (s, 4H, Ar-H), 3.48-3.23 (m, 4H, CH₂), 1.80-1.55 (m, 4H, CH₂); ESI/MS (+, *m/z*): 438 [M⁺]; Anal. Calcd for C₂₂H₂₂N₄O₆: C 59.98, H 5.34, N 12.55. Found: C 57.88, H 5.01, N 12.82.

Synthesis of 5H-[1,3]dioxolo[4,5-f]indole-6-carbonyl chloride [17]

A suspension of **16** (11 mmol) and phosphorous pentachloride (13.5 mmol) in anhydrous diethyl ether (75 mL) was stirred at room temperature for 2 hours. The solvent was removed under vacuum and the crude residue was treated twice with diethyl ether and thrice with chloroform. The resulting product was redissolved in diethyl ether (100 mL) and used without further purification.

Synthesis of *N*-(3,4-dihydroxyphenethyl)-5*H*-[1,3]dioxolo[4,5-*f*]indole-6-carboxamide [21]

To a solution of **17** (4.38 mmol) in diethyl ether (70 mL) a solution of **18** (6.60 mmol) in dioxane (6 mL) was added dropwise in anhydrous conditions. The reaction mixture was stirred at room temperature for 2 hours. The precipitate formed was filtered and washed with water. Yield: 60 %; R_f: 0.64 (petroleum ether/ethyl acetate = 1:1); mp: 200-202.5 °C (lit. 201-202 °C¹); IR (nujol) ν cm⁻¹ = 3290 (NH), 1620 (C=O); ¹H-NMR (CDCl₃:DMSO-d₆): δ 10.80 (bs, 1H, NH), 7.55 (bs, 1H, NH), 6.93-6.91 (m, 2H, Ar-H), 6.80 (m, 3H, Ar-H), 5.93 (s, 2H, CH₂), 3.84 (s, 3H, OCH₃), 3.82 (s, 3H, OCH₃), 3.64 (m, 2H, CH₂), 2.87 (t, 2H, CH₂); ESI/MS (+, *m/z*): 368 [M⁺]; Anal. Calcd for C₂₁H₂₀N₄O₆•0.1H₂O: C 64.88, H 5.50, N 7.57. Found: C 64.69, H 5.27, N 7.49.

Synthesis of N,N'-(propane-1,3-diyl)bis(5H-[1,3]dioxolo[4,5-f]indole-6-carboxamide) [22]

To a freshly prepared solution of **17** (2.92 mmol) in anhydrous diethyl ether (50 mL) a solution of **19** (4.40 mmol) in dioxane (4 mL) was added dropwise. The resulting mixture was stirred at room temperature for 2 hours. The precipitate that formed was filtered and washed with water. Yield: 35 %; R_f: 0.75 (petroleum ether/ethyl acetate = 3:7); mp: 269 °C dec. (lit. 268-269 °C dec.¹); IR (nujol) ν cm⁻¹ = 3280 (NH), 1620 (C=O); ¹H-NMR (CDCl₃:DMSO-d₆): δ 11.30 (bs, 2H, NH), 8.30 (bs, 2H, NH), 6.97-6.90 (m, 6H, Ar-H), 5.93 (s, 4H, CH₂), 3.50 (m, 4H, CH₂), 1.82 (m, 2H, CH₂); ESI/MS (+, *m/z*): 448 [M⁺]. Anal. Calcd for C₂₁H₂₀N₄O₆•0.5H₂O: C 60.37, H 4.63, N 12.25. Found: C 60.42, H 4.58, N 12.29.

Synthesis of N,N'-(butane-1,4-diyl)bis(5H-[1,3]dioxolo[4,5-f]indole-6-carboxamide) [23]

To a freshly prepared solution of **17** (1.75 mmol) in anhydrous diethyl ether (30 mL) a solution of **20** (2.64 mmol) in dioxane (2.5 mL) was added dropwise. The resulting mixture was stirred at room temperature for 2 hours. The precipitate that formed was filtered and washed with water. Yield: 58 %; R_f: 0.79 (petroleum ether/ethyl acetate = 3:7); mp: 300 °C dec. (lit. 297-299 °C dec.¹); IR (nujol) ν cm⁻¹ = 3220 (NH), 1630 (C=O); ¹H-NMR (CDCl₃:DMSO-d₆): δ 11.00 (bs, 2H, NH), 8.16 (bs, 2H, NH), 6.97 (s, 2H, Ar-H), 6.92 (s, 2H, Ar-H), 6.89 (s, 2H, Ar-H), 5.92 (s, 4H, OCH₂O), 3.44-3.29 (m, 4H, CH₂), 1.76-1.52 (m, 4H, CH₂); ESI/MS (+, *m/z*): 462 [M⁺]; Anal. Calcd for C₂₄H₂₂N₄O₆: C 62.33, H 4.79, N 12.12. Found: C 62.28, H 4.90, N 11.90.

Database optimization

The database used in the present study was created by merging the Clean Lead Database deposited in ZINC² (4,869,601 structures) and an in-house Database containing small molecules designed as metalloenzymes inhibitors (861 structures). The resulting Database contained 4,870,462 unique structures that were washed using MOE (Molecular Operating Environment³) as follow: (a) salts of Group I were disconnected and only the organic fraction was retained; (b) explicit hydrogen atoms were added; (c) protons were added to neutralize the structures; (d) finally, strong acids and bases were considered as ionized.

Pharmacophore model generation and validation

Thirteen pharmacophore hypotheses were generated by the pharmacophore elucidation module implemented in MOE, and using as training set a series of 12 influenza virus endonuclease inhibitors (Chart 1). Settings were kept as follows: Stochastic search with unique clusterization was selected, and at least two metal ligator and one aromatic feature were imposed. Emphasize Donor and Acceptor Atoms was used as alignment constraints. All pharmacophore hypotheses generated were evaluated using a test set of 50 unique structure (Chart 2) composed of 10 PA-Nter inhibitors (**T1-10**), and 40 structures randomly selected from the Zinc Lead Database (**T11-T50**). Pharmacophore search was carried out enabling a partial match status of at least 4 pharmacophore features. As output preferences, Conformations and Best per Molecule were selected.

Chart 1. Chemical structures of the 12 compounds used as training set for the pharmacophore model generation.



TS9

TS10

TS11

TS12

COOF соон F₃CC соон соон Т6 T7 Т8 T10 Ľ,º↓ C T13 T12 T14 T11 T15 Ĭ T17 H. T21 T22 T23 T25 N T29 T30 T27 `CI т33 T35 .N T39 T40 T43 т42 таа T46 T47 T48 T49 T50

Chart 2. Chemical structures of the 50 compounds of the test set for the pharmacophore model validation.

Protein preparation

The crystal structure of inhibitor-free PA-Nter protein of influenza virus (subtype A/Victoria/3/1975, H3N2) was retrieved from the RCSB Protein Data Bank (PDB ID: 2W69⁴). Chain A (residues 1-209) containing two Mn²⁺ ions as metal cofactors was chosen for this study. The protein target was prepared using AutoDock 4.2⁵ and AMBER 11,⁶ as reported by us previously.⁷ Briefly, water molecules and sulfate ions were removed while hydrogen atoms were added using the ADT module implemented in AutoDock. The missing residues at positions 72 (Leu) and 142 (Asn) in this chain A were incorporated from chain B of the same protein after superimposition of the backbones of residues His41, Glu80, Pro107, Asp108, Leu109, Glu119, Ile120 (rms = 0.2541) and substitution by fitting on Leu71, Lys73 and His74 for Leu72; and Thr143 and His144 for Asn142. The partially solved residues Phe105 and Arg185 were also refined according to the data for chain B, and the protein structure was analyzed with tleap module of AMBER 11, using Amber *ff03.r1*. To include Mn^{2+} ions, the force field file *frcmod* was modified as follows: Mass Mn 54.938, nonbonded 6-12 interaction Mn 1.69 (Å) and 0.014 (kcal/mol). Na⁺ ions were added to obtain charge neutrality, and charges were adjusted using Gasteiger charges module for proteins implemented in Autodock, paying particular attention to the protonation states of the catalytic triad Glu80, Asp108 and Glu119.

MOE docking procedure

MOE docking protocol was applied to the receptor model, which was optimized selecting AMBER99 as force field and fixing hydrogens and charges. Then, the dockable space was set by selecting the residues around to the metal cofactors in a radius of 9 Å. Keeping this selection as docking site, the compound libraries were used as Ligand, Triangle Matcher as placement feature, and London dG as rescoring function.

AutoDock docking refinement

To refine the hit search, a docking procedure previously optimized for PA-Nter was used.⁷ Briefly, docking was performed with AutoDock version 4.2. using the empirical free energy function and the Lamarckian protocol. The atomic charges for the protein were assigned using the Gasteiger-Marsili method. Mass-centered grid maps were generated with 80 grid points for every direction and with 0.375 Å spacing by the AutoGrid program for the whole protein target. Random starting position on the entire protein surface, random orientations and torsions were used for the ligands. The distance-dependent dielectric permittivity of Mehler and Solmajer was used for the calculation of the electrostatic potential grid maps. One hundred independent

docking runs were carried out for each ligand. The cluster analyses were computed with a cluster tolerance of less than 1.5 Å in positional root-mean-square deviation with AutoDock Tools. The structures were then clusterized by scaffold similarity, thus permitting to identify the 15 top ranked hits selected for biological assays.

Figure 1. Superimposition of compounds 10 (white), 11 (yellow), and 15 (orange) on the pharmacophore model PH4-3.



Plasmid-based endonuclease assay

The procedure used was a slight modification to a previously published method.⁸ Briefly, recombinant PA-Nter [residues 1-217 from the PA protein of influenza virus strain A/X-31] was expressed in *E. coli* and purified.⁹ One microgram of the enzyme was incubated with 1 μ g (16.7 nM) of single-stranded circular DNA plasmid M13mp18 (Bayou Biolabs, Metairie, Louisiana) in the presence of the test compounds and at a final volume of 25 μ L. The assay buffer contained 50 mM Tris-HCl pH 8, 100 mM NaCl, 10 mM β -mercaptoethanol and 1 mM MnCl₂. After 2 h incubation at 37 °C, the reaction was stopped by heat inactivation (80 °C, 20 min). Endonucleolytic digestion of the plasmid was visualized by gel electrophoresis on a 1% agarose gel with ethidium bromide staining and the amount of remaining intact plasmid was quantified by ImageQuant TL software (GE Healthcare). The percentage inhibition of PA endonuclease activity was plotted against the compound concentration on a semi-logarithmic plot, using GraphPad Prism software (GraphPad Software, La Jolla, CA). The 50% inhibitory concentrations (IC₅₀) were obtained by nonlinear least- squares regression analysis of the results from three independent experiments.

Influenza virus yield assay

We previously published in full detail the virus yield assay to determine the anti-influenza virus activity in MDCK (Madin-Darby canine kidney) cell cultures.⁷ Briefly, serial dilutions of the test compounds were added to subconfluent cultures of MDCK cells in 96-well plates. Immediately after compound addition, influenza virus A/PR/8/34 (A/H1N1) was added. After 24 h incubation at 35 °C, the supernatants were collected and the virus amount in these samples was estimated by determining the viral genome copy number in a one-step quantitative real-time reverse transcription (RT-PCR) assay described in detail elsewhere.¹⁰ The EC₉₉ and EC₉₀ values were calculated by interpolation from data of at least three experiments and defined as the compound concentration causing respectively a 2-log₁₀ and 1-log₁₀ reduction in vRNA copy number, as compared to the virus control receiving no compound. In parallel, uninfected MDCK cells were used to determine the 50% cytotoxic concentration (CC₅₀) values after 24 h incubation with compounds, by using the MTS cell viability assay (CellTiter 96 AQueous One Solution Cell Proliferation Assay; Promega). The antiviral drug ribavirin and selective PA-Nter inhibitor L-742,001¹¹ were included as reference compounds.

Influenza vRNP reconstitution assay

The procedure to determine the inhibitory effect of the compounds on influenza virus vRNPs reconstituted in HEK293T cells, is described in full detail elsewhere.⁷ Briefly, the four relevant expression plasmids¹² (containing the PA, PB1, PB2 and nucleoprotein sequences from influenza A/PR/8/34) were combined with a firefly luciferase reporter plasmid, and cotransfected into HEK293T cells using Lipofectamine 2000. After addition of the test compounds, the plates were incubated for 24 h at 37 °C. Luciferase activity was then determined using the ONE-Glo luciferase assay system (Promega). The percentage of vRNP inhibition, resulting from four independent experiments, was plotted against the compound concentration on a semilogarithmic plot. The 50% effective concentration (EC₅₀) was defined as the compound concentration causing 50% reduction in the vRNP-driven firefly luciferase signal, as compared to cells receiving medium instead of compound, and was calculated by nonlinear least- squares regression analysis using GraphPad Prism software. In parallel, the compound cytotoxicity, expressed as CC_{50} , was determined after 24 h incubation, using the spectrophotometric MTS cell viability assay described above. Ribavirin and L-742,001 were included as reference compounds.

References

- Sechi, M.; Angotzi, G.; Dallocchio, R.; Dessì, A.; Carta, F.; Sannia, L.; Mariani, A.; Fiori, S.; Sanchez, T.; Movsessian, L.; Plasencia, C.; Neamati, N. Design and synthesis of novel dihydroxyindole-2-carboxylic acids as HIV-1 integrase inhibitors. *Antiviral Chem. Chemother.* 2004, *15*, 67-81.
- (2) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: a free tool to discover chemistry for biology. *J Chem. Inf. Model.* **2012**, *52*, 1757-1768.
- (3) Molecular Operating Environment, MOE 2009.10; Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2009.
- (4) Dias, A.; Bouvier, D.; Crépin, T.; McCarthy, A. A.; Hart, D. J.; Baudin, F.; Cusack, S.; Ruigrok, R. W. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **2009**, *458*, 914-918.
- (5) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785-2791.
- (6) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. Amber 11. 2010.
- (7) Stevaert, A.; Dallocchio, R.; Dessì, A.; Pala, N.; Rogolino, D.; Sechi, M.; Naesens, L. Mutational analysis of the binding pockets of the diketo acid inhibitor L-742,001 in the influenza virus PA endonuclease. *J. Virol.* **2013**, *87*, 10524-10538.
- (8) Carcelli, M.; Rogolino, D.; Bacchi, A.; Rispoli, G.; Fisicaro, E.; Compari, C.; Sechi, M.; Stevaert, A.; Naesens, L. Metal-chelating 2-hydroxyphenyl amide pharmacophore for inhibition of influenza virus endonuclease. *Mol. Pharm.* **2014**, *11*, 304-316.
- (9) Stevaert, A.; Nurra, S.; Pala, N.; Carcelli, M.; Rogolino, D.; Shepard, C.; Domaoal, R. A.; Kim, B.; Alfonso-Prieto, M.; Marras, S. A. E.; Sechi, M.; Naesens, L. An integrated biological approach to guide the development of metal-chelating inhibitors of influenza virus PA endonuclease. *Mol. Pharmacol.* **2015**, *87*, 323-337.
- (10) Vanderlinden, E.; Göktas, F.; Cesur, Z.; Froeyen, M.; Reed, M. L.; Russell, C. J.; Cesur, N.; Naesens, L. Novel inhibitors of influenza virus fusion: structure-activity relationship and interaction with the viral hemagglutinin. *J. Virol.* **2010**, *84*, 4277-4288.
- (11) Hastings, J. C.; Selnick, H.; Wolanski, B.; Tomassini, J. E. Anti-influenza virus activities of 4-substituted 2,4-dioxobutanoic acid inhibitors. *Antimicrob. Agents Chemother.* **1996**, *40*, 1304-1307.
- (12) Kim, M.; Kim, S. Y.; Lee, H. W.; Shin, J. S.; Kim, P.; Jung, Y. S.; Jeong, H. S.; Hyun, J. K.; Lee, C. K. Inhibition of influenza virus internalization by (-)-epigallocatechin-3-gallate. *Antiviral Res.* **2013**, *100*, 460-472.