Supporting Information

Design, Synthesis, Biological and Structural Evaluations of Novel HIV-1 Protease Inhibitors to Combat Drug Resistance

*Maloy Kumar Parai,[1] David J. Huggins,[2] Hong Cao,[3] Madhavi N. L. Nalam,[4] Akbar Ali,[3] Celia A. Schiffer,[4] Bruce Tidor,[2] and Tariq M. Rana[1, 3]**

[1]Program for RNA Biology, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037; [2]Department of Biological Engineering and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139; [3]Chemical Biology Program, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, [4]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605.

Table of Contents

General Procedure for the Synthesis of Activated Carbonates 7a-j.

(A): **(*R*)-methyl 3-(((4-nitrophenoxy)carbonyl)oxy)butanoate.** To a solution of **6a** (0.20 g, 1.69 mmol) and *N*-methylmorpholine (0.5 mL, 4.86 mmol) in dry THF (20 mL), *p*-nitrophenyl chloroformate (0.68 g, 3.38 mmol) was added and the resulting mixture was stirred at 23 °C for 1 h. After completion of the reaction, water was added to the reaction mixture and the solvent was removed under reduced pressure. Then aqueous phase was extracted with CH_2Cl_2. The organic extracts were dried over anhydrous Na_2SO_4 and the solvent was removed. The crude mixture was purified by flash-chromatography (1:8 EtOAc-hexane) to afford 0.37 g (78 %) of **7a** as a pale-yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.22-8.18 (m, 2H), 7.33-7.30 (m, 2H), 5.25-5.19 (m, 1H), 3.65 (s, 3H), 2.72 (dd, J = 9.6, 4.8 Hz, 1H), 2.56 (dd, J = 9.6, 3.3 Hz, 1H), 1.39 (d, J = 3.9 Hz, 3H).

Compounds **7a-c**, **7f-j** were prepared following this general procedure.

4-nitrophenyl (2-(2,2,2-trifluoroacetamido)ethyl) carbonate (7b).

The title compound was synthesized from **6b** as described for **7a** in 62% yield. Flash-chromatography was performed using a 1:5 mixture of EtOAc and hexane as the eluent. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30-8.25 (m, 2H), 7.40-7.36 (m, 2H), 7.07 (bs, 1H), 4.44 (t, J = 4.8 Hz, 2H), 3.81-3.76 (m, 2H).

2-(2,2-dichloroacetamido)ethyl (4-nitrophenyl) carbonate (7c).

The title compound was synthesized from commercially available 2,2-dichloro-*N*-(2-hydroxyethyl)acetamide **6c** as described for **7a** in 69% yield. Flash-chromatography was performed using a 1:6 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.28 (d, J = 8.7 Hz, 2H), 7.39 (d, J = 8.7 Hz, 2H), 7.04 (bs, 1H), 5.89 (s, 1H), 4.16-4.09 (m, 2H), 3.84-3.78 (m, 2H).

(*R*)-methyl-2-methyl-3-(((4-nitrophenoxy)carbonyl)oxy)propanoate (7f).

The title compound was synthesized from commercially available (*S*)-methyl 3-hydroxy-2-methylpropanoate **6f** as described for **7a** in 73% yield. Flash-chromatography was performed using a 1:6 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.27 (d, J = 9.2 Hz, 2H), 7.35 (d, J = 9.2 Hz, 2H), 4.45 (dd, J = 10.5, 6.8 Hz, 1H), 4.30 (dd, J = 10.5, 5.9, Hz, 1H), 3.66 (s, 3H), 2.89-2.81 (m, 1H), 1.21 (d, J = 6.9 Hz, 3H).

4-nitrophenyl (2-(2-oxoimidazolidin-1-yl)ethyl) carbonate (7g).

The title compound was obtained from commercially available 1-(2-hydroxyethyl)imidazolidin-2-one **6g** as described for **7a** in 75% yield. Flash-chromatography was performed using a 1:5 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (d, J = 9.2 Hz, 2H), 7.36 (d, J = 9.2 Hz, 2H), 5.38 (bs, 1H), 4.26-4.19 (m, 2H), 3.57-3.50 (m, 2H), 3.47-3.36 (m, 4H).

4-nitrophenyl (2-(2-oxooxazolidin-3-yl)ethyl) carbonate (7h).

The title compound was synthesized from commercially available 3-(2-hydroxyethyl)oxazolidin-2-one **6h** as described for **6a** in 79% yield. Flash-chromatography was performed using a 1:5 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.26 (d, J = 8.8 Hz, 2H), 7.37 (d, J = 8.8 Hz, 2H), 4.28-4.01(m, 4H), 3.76-3.68 (m, 4H).

(*R*)-4-nitrophenyl (2-oxotetrahydrofuran-3-yl) carbonate (7i).

The title compound was synthesized from commercially available (*R*)-3-hydroxydihydrofuran-2(3H)-one **6i** as described for **6a** in 58% yield. Flash-chromatography was performed using a 1:6 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30-8.25 (m, 2H), 7.43-7.37 (m, 2H), 5.46 (dd, J = 9.2, 3.2 Hz, 1H), 4.09-4.01 (m, 1H), 3.99-3.93 (m, 1H), 2.67 (dd, J = 9.5, 4.3 Hz, 1H), 2.59 (dd, J = 9.5, 3.8 Hz, 1H).

(*S*)-4-nitrophenyl ((5-oxopyrrolidin-2-yl)methyl) carbonate (7j).

The title compound was obtained from commercially available (*S*)-5-(hydroxymethyl)pyrrolidin-2-one **6j** as described for **7a** in 72% yield. Flash-chromatography was performed using a 1:5 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (d, J = 8.7 Hz, 2H) 7.40 (d, J = 8.7 Hz, 2H), 5.49 (bs, 1H), 4.07-3.96 (m, 1H), 3.77-3.69 (m, 2H), 2.38-2.29 (m, 2H), 2.21-2.14 (m, 2H).

(B): (*R*)-2,5-dioxopyrrolidin-1-yl (5-oxohexan-2-yl) carbonate (7d). To a solution of **6d** (0.20 g, 1.72 mmol) in dry acetonitrile (10 mL), *N,N'*-disuccimidyl carbonate (0.57 g, 2.24 mmol) and Et_3N (0.5 mL, 3.47 mmol) were added and the resulting mixture was stirred at 25 °C. After 8 h, the solvent was removed under reduced pressure, the residue was dissolved in EtOAc and small amount of saturated solution of $NaHCO_3$ was added to it, the aqueous phase was extracted with EtOAc. The organic extracts were dried over anhydrous Na_2SO_4, and the solvent was removed in vacuo. Purification of the crude residue using a mixture of 1:4 EtOAc and hexane afforded **6d** (0.075 g) in 41% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.96-4.88 (m, 1H), 2.88 (s, 4H), 2.58 (t, J = 7.3 Hz, 2H), 2.16 (s, 3H), 2.04-1.85 (m, 2H), 1.39 (d, J = 5.9 Hz, 3H).

***tert*-butyl (2-((((2,5-dioxopyrrolidin-1-yl)oxy)carbonyl)oxy)ethyl)carbamate (7e)**.

The title compound was synthesized from commercially available *tert*-butyl (2-hydroxyethyl)carbamate **6e** as described for **7d** in 35% yield. Flash-chromatography was performed using a 1:4 mixture of EtOAc and hexane as the eluant. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.09-5.03 (m, 1H), 4.31 (t, J = 5.0 Hz, 1H), 4.12 (t, J = 5.3 Hz, 1H) 3.44-3.26 (m, 2H), 2.79 (s, 4H), 1.37 (s, 9H).

Synthesis of Core Scaffolds of the Designed Protease Inhibitors:

General Procedure for the Ring Opening of Epoxide with Amines.

To a solution of the chiral epoxide **8** *tert-butyl* ((*S*)-oxiran-2-yl)-2-phenylethyl)carbamate (1.0 g, 3.79 mmol) in EtOH (25 mL) was added (*S*)-2-methylbutyl amine **9a** (0.5 mL, 4.23 mmol) and the

mixture was heated at 65 °C for 4 h. After cooling to room temperature, solvents were removed under reduced pressure. Product was purified by recrystallization from EtOAc-hexanes mixture to provide the product as white solid in excellent yield.

The following compounds were prepared by this general procedure:

***tert*-butyl ((2*S*,3*R*)-3-hydroxy-4-(((*S*)-2-methylbutyl)amino)-1-phenylbutan-2-yl)carbamate (10a)**. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.30-7.17 (m, 5H), 4.78-4.75 (m, 1H), 3.80 (bs, 1H), 3.48-3.43 (m, 1H), 3.00-2.94 (m, 1H), 2.88-2.80 (m, 1H), 2.67 (d, J = 4.8 Hz, 2H), 2.53-2.47 (m, 1H), 2.41-2.34 (m, 1H), 1.54-1.38 (m, 12H), 1.35-1.06 (m, 1H), 0.96-0.85 (m, 6H); MS (ESI) *m/z* 351.18 $(M + H)^+$.

***tert*-butyl ((2*S*,3*R*)-3-hydroxy-4-(pentylamino)-1-phenylbutan-2-yl)carbamate (10b)**.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29-7.18 (m, 5H), 4.74 (bs, 1H), 3.81 (bs, 1H), 3.45-3.42 (m, 1H), 3.04-2.97 (m, 1H), 2.89-2.82 (m, 1H), 2.65 (d, J = 4.5 Hz, 2H), 2.55-2.50 (m, 1H), 1.57-1.49 (m, 2H), 1.29 (s, 9H), 1.28-0.97 (m, 6H), 0.73 (t, J = 7.3 Hz, 3H); MS (ESI) *m/z* 351.15 $(M + H)^+$.

General Procedure for the Synthesis of (Hydroxyethylamino)sulfonamides, 12a-d.

To an ice-cooled solution of the secondary amine **10a** (1.0 g, 2.85 mmol) in CH_2Cl_2 (20 mL) was added an aqueous solution of Na_2CO_3 (0.50 g, 4.71 mmol in 3 mL H_2O) followed by the slow addition of 4-methoxyphenyl sulfonyl chloride **11a** (0.60 g, 2.90 mmol) solution in CH_2Cl_2 (5 mL). After 15 minutes the reaction mixture was warmed to temperature and stirred till no starting material was detected by TLC. Reaction mixture was diluted with CH_2Cl_2 and layers were separated. Organic phase was washed with saturated aqueous NaCl solution, dried (Na_2SO_4), filtered and evaporated. Crude mixture was purified by flash chromatography on silica gel using mixture of EtOAc and hexanes as eluent to afford pure product.

The following compounds were prepared by this general procedure:

***tert*-butyl ((2*S*,3*R*)-3-hydroxy-4-(4-methoxy-*N*-((*S*)-2-methylbutyl)phenylsulfonamido)-1-phenylbutan-2-yl)carbamate (12a)**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.64-7.62 (m, 2H), 7.24-7.14 (m, 5H), 6.91-6.88 (m, 2H), 4.67 (d, J = 5.1 Hz, 1H), 3.83-3.70 (s, 6H), 3.01-2.87 (m, 5H), 2.74-2.67 (m, 1H), 1.52-1.41 (m, 2H), 1.37 (s, 9H), 1.01-0.98 (m, 1H), 0.79-0.75 (m, 6H); MS (ESI) *m/z* 543.36 $(M + Na)^+$.

***tert*-butyl ((2*S*,3*R*)-3-hydroxy-4-(*N*-((*S*)-2-methylbutyl)benzo[*d*]thiazole-6-sulfonamido)-1-phenylbutan-2-yl)carbamate (12b)**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.13 (s, 1H), 8.37 (d, J = 1.2 Hz, 1H), 8.16 (d, J = 8.5 Hz, 1H), 7.81 (dd, J = 8.5, 1.8 Hz, 1H), 7.27-7.09 (m, 5H), 5.24 (d, J = 8.5 Hz, 1H), 4.60 (d, J = 7.3 Hz, 1H), 3.83 (bs, 1H), 3.74-3.64 (m, 2H), 3.14-2.91 (m, 3H), 2.91-2.77 (m, 2H), 1.62-1.50 (m, 1H), 1.47-1.34 (m, 1H), 1.27 (s, 9H), 1.06-0.90 (m, 1H), 0.81-0.72 (m, 6H); MS (ESI) *m/z* 570.49 $(M + Na)^+$.

***tert*-butyl ((2*S*,3*R*)-3-hydroxy-4-(4-methoxy-*N*-pentylphenylsulfonamido)-1-phenylbutan-2-yl)carbamate (12c)**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.64 (d, J = 9.2 Hz, 2H), 7.23-7.09 (m, 5H), 6.89 (d, J = 9.2 Hz, 2H), 4.76 (bs, 1H), 3.79 (s, 3H), 3.77-3.69 (m, 2H), 3.13-2.85 (m, 3H), 2.92-2.79 (m, 2H), 1.51-1.48 (m, 2H), 1.28 (s, 9H), 1.29-0.98 (m, 6H), 0.73 (t, J = 7.3 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.92, 156.30, 137.84, 130.33, 129.56 (2C), 129.33 (2C), 128.43 (2C), 126.38, 114.28 (2C), 79.42, 72.57, 55.58, 54.59, 52.45, 50.46, 35.35, 28.76, 28.22 (3C), 22.31, 22.19, 13.92; MS (ESI) *m/z* 543.33 $(M + Na)^+$.

***tert*-butyl ((2*S*,3*R*)-3-hydroxy-4-(*N*-pentylbenzo[*d*]thiazole-6-sulfonamido)-1-phenylbutan-2-yl)carbamate (12d)**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.12 (s, 1H), 8.34 (d, J = 1.8 Hz, 1H), 8.15 (d, J = 8.5 Hz, 1H), 7.82 (dd, J = 8.5, 1.8 Hz, 1H), 7.24-7.09 (m, 5H), 4.68 (d, J = 7.3 Hz, 1H), 3.85 (bs, 1H), 3.78-3.67 (m, 2H), 3.16-2.90 (m, 3H), 2.96 (dd, J = 14.0, 4.9 Hz, 1H), 2.91-2.83 (m, 1H), 1.49-1.36 (m, 1H), 1.29 (s, 9H), 1.23-0.99 (m, 6H), 0.76 (t, J = 7.3 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 156.89, 155.12, 154.38, 137.74, 135.17, 133.22, 128.45 (2C), 127.41 (2C), 125.40,

123.67, 123.24, 121.05, 78.74, 71.48, 53.70, 51.26, 49.28, 35.35, 28.68, 28.17 (3C), 22.10 (2C), 13.83; MS (ESI) *m/z* 570.50 $(M + Na)^+$.

Purity of Target Compounds Determined by HPLC

Compound	System		Compound	System	
	R_t (min)	Purity (%)		R_t (min)	Purity (%)
12b	4.76	97.32	**13f**	4.69	98.35
13a	4.60	96.76	**14f**	4.20	97.68
14a	3.97	97.41	**15f**	4.83	98.39
15a	4.43	97.35	**16f**	4.22	99.09
16a	4.35	97.35	**13g**	4.97	98.81
13b	4.01	99.57	**14g**	3.76	97.61
14b	3.84	98.01	**15g**	3.69	98.29
15b	3.98	98.42	**16g**	3.34	98.49
16b	3.70	98.82	**13h**	3.50	99.05
13c	4.37	99.10	**14h**	3.59	99.27
14c	4.51	98.41	**15h**	3.65	98.24
16c	4.45	99.30	**16h**	3.29	98.93
13d	4.35	97.21	**13i**	3.86	99.22
14d	4.24	98.15	**14i**	3.57	98.34
15d	4.54	99.25	**15i**	3.01	99.06
16d	3.96	97.53	**16i**	3.63	98.90
13e	5.36	97.34	**13j**	3.54	96.97
14e	5.18	98.82	**14j**	3.33	96.83
15e	5.33	98.84	**15j**	3.80	95.36
16e	4.97	98.78	**16j**	3.37	99.15