CHEMMEDCHEM
Chemistry & Drug Discovery

Supporting Information

SUPPORTING INFORMATION

Dipeptide derivatives of AZT: synthesis, chemical stability, activation in human plasma, hPEPT1 affinity and antiviral activity

Cledir Santos, José Morais, Luís Gouveia, Erik De Clercq, Christophe Pannecouque, Carsten Uhd Nielsen, Bente Steffansen, Rui Moreira* and Paula Gomes*

*corresponding authors: pgomes@fc.up.pt; rmoreira@ff.ul.pt

List of Contents **Page**

SPECTROSCOPIC DATA FOR SYNTHETIC INTERMEDIATES

Chemical shifts are reported as parts per million (δ/ppm) from tetramethylsilane (TMS) as internal reference standard. ^{1}H-NMR signals are described as follows: δ_H in ppm (number of protons, peak multiplicity, coupling constant in Hz). Peak multiplicity in ^{1}H-NMR is described by: s (singlet), br s (broad singlet), d (doublet), dd (double doublet) or m (unresolved multiplet). ^{13}C-NMR peaks are described as follows: δ_C in ppm.

Compounds 3a-e

***O*-(*N*-*tert*-Butyloxycarbonyl)glycyl-3'-azido-2',3'-dideoxythymine, *3a*.** Yield, 88%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.38 (9H, s); 1.80 (3H, s); 2.34 (1H, m); 2.44 (1H, m); 3.73 (2H, m); 3.97 (1H, m); 4.30 (2H, m); 4.44 (1H, m); 6.13 (1H, t, *J*=6.3); 7.26 (1H, t, *J*=6.0); 7.45 (1H, s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.02; 28.03; 35.43; 41.84; 59.88; 63.54; 78.27; 80.45; 83.40; 109.95; 135.88; 150.28; 155.79; 163.56; 170.16. $^{m}/_{z}$ ($M+Na^{+}$) = 447.2176 (calcd, 447.1599).

***O*-(*N*-*tert*-Butyloxycarbonyl)alanyl-3'-azido-2',3'-dideoxythymine**, ***3b*.** Yield, 79%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.27 (3H, d, *J*=7.2); 1.36 (9H, s); 1.80 (3H, s); 2.34 (1H, m); 2.35 (1H, m); 2.44 (1H, m); 4.02 (2H, m); 4.27 (2H, m); 4.44 (1H, m); 6.12 (1H, t, *J*=6.6); 7.33 (1H, d, *J*=6.9); 7.43 (1H, s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.01; 16.68; 28.04; 35.40; 49.01; 60.17; 63.81; 78.17; 80.41; 83.47; 109.92; 135.74; 150.26; 155.18; 163.55; 172.83. $^{m}/_{z}$ ($M+Na^{+}$) = 461.2355 (calcd, 461.1755).

***O*-(*N*-*tert*-Butyloxycarbonyl)valyl-3'-azido-2',3'-dideoxythymine**, ***3c*.** Yield, 58%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.86 (3H, d, *J*=4.5); 0.88 (3H, d, *J*=4.8); 1.35 (9H, s); 1.79 (3H, s); 2.02 (1H, m); 2.32 (1H, m); 2.43 (1H, m); 3.85 (1H, t, *J*=7.2); 4.00 (1H, m); 4.27 (2H, m); 4.44 (1H, m); 6.12 (1H, t, *J*=6.6); 7.21 (1H, d, *J*=6.1); 7.44 (1H, s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=11.99; 18.31; 18.84; 28.05; 29.43; 35.32; 59.49; 60.24; 63.74; 78.20; 80.34; 83.49; 109.86; 135.77; 150.27; 155.67; 163.54; 171.72. $^{m}/_{z}$ ($M+Na^{+}$) = 489.2654 (calcd, 489.2068).

***O*-(*N*-*tert*-Butyloxycarbonyl)isoleucyl-3'-azido-2',3'-dideoxythymine**, ***3d*.** Yield, 54%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.81 (6H, m); 1.20 (1H, m); 1.35 (2H, m); 1.37 (9H, s); 1.79 (3H, s); 2.33 (1H, m); 2.44 (1H, m); 3.91 (1H, t, *J*=7.2); 4.01 (1H, m); 4.26 (2H, m); 4.43 (1H, m); 6.12 (1H, t, *J*=6.6); 7.20 (1H, d, *J*=7.5); 7.43 (1H, s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=10.92; 12.01; 15.30; 24.78; 28.06; 35.30; 35.73; 58.32; 60.33; 63.81; 78.19; 80.35; 83.55; 109.84; 135.76; 150.28; 155.60; 163.54; 171.75. $^{m}/_{z}$ ($M+Na^{+}$) = 480.3253 (calcd, 480.2332).

O*-(*N*-*tert*-Butyloxycarbonyl)phenylalanyl-3'-azido-2',3'-dideoxythymine**, ***3e. Yield, 89%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.32 (9H, s); 1.80 (3H, s); 2.29 (2H, m); 2.95 (2H, m); 3.95 (1H, m); 4.20 (4H, m); 6.11 (1H, t, *J*=6.6); 7.24 (5H, m); 7.39 (2H, m); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.02; 27.99; 35.51; 36.33; 55.40; 60.15; 63.97; 78.30; 80.30; 83.37; 110.00; 126.43; 128.15; 128.96; 135.54; 137.27; 150.25; 155.27; 163.52; 171.82. $^{m}/_{z}$ ($M+H^{+}$) = 515.4576 (calcd, 515.2249).

Compounds 4a-e

***O*-Glycyl-3'-azido-2',3'-dideoxythymine trifluoroacetic acid salt, *4a*.** Yield, 89%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.86 (3H, s); 2.36 (1H, m); 2.54 (1H, m); 3.85 (2H, m); 3.98 (1H, m); 4.40 (2H, m); 4.52 (1H, m); 6.15 (1H, d, *J*=6.9); 7.48 (1H, s); 8.38 (3H, br s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.02; 35.22; 59.49; 64.40; 80.19; 83.61; 109.96; 136.25; 150.33; 163.60; 167.50. $C_{14}H_{17}F_3N_6O_7$•H_2O (456.13 gmol^{-1}): N, 18.42; C, 36.69; H, 4.19%; Found: N, 18.29; C, 36.55; H, 4.31%.

***O*-Alanyl-3'-azido-2',3'-dideoxythymine trifluoroacetic acid salt, *4b*.** Yield, 83%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.43 (3H, d, *J*=7.2); 1.80 (3H, s); 2.37 (1H, m); 2.52 (1H, m); 3.99 (1H, m); 4.12 (1H, m); 4.40 (2H, m); 4.53 (1H, m); 6.13 (1H, t, *J*=6.6); 7.48 (1H, s); 8.50 (3H, br s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=11.50; 15.10; 34.67; 47.61; 59.07; 64.04; 79.60; 83.15; 109.51; 135.76; 149.86; 163.13; 169.25. $C_{15}H_{19}F_3N_6O_7$•H_2O (470.15 gmol^{-1}): N, 17.87; C, 38.32; H, 4.50%; Found: N, 17.21; C, 38.03; H, 4.47%.

***O*-Valyl-3'-azido-2',3'-dideoxythymine trifluoroacetic acid salt, *4c*.** Yield, 72%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.94 (3H, d, *J*=6.9 Hz); 0.97 (3H, d, *J*=6.9); 1.79 (3H, s); 2.17 (1H, m); 2.34 (1H, m); 2.53 (1H, m); 3.91 (1H, m); 4.01 (1H, m); 4.43 (2H, m); 4.50 (1H, m); 6.11 (1H, t, *J*=6.6); 7.48 (1H, s); 8.46 (3H, br s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.00; 17.44; 18.00; 29.33; 34.98; 57.32; 59.91; 64.78; 80.06; 83.74; 109.91; 136.23; 150.33; 163.58; 168.72. $C_{17}H_{23}F_3N_6O_7$•1.5 H_2O (507.18 gmol^{-1}): N, 16.57; C, 40.25; H, 5.17%; Found: N, 16.38; C, 40.22; H, 4.95%.

***O*-Isoleucyl-3'-azido-2',3'-dideoxythymine trifluoroacetic acid salt, *4d*.** Yield, 59%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.87 (6H, m); 1.27 (1H, m); 1.43 (1H,

m); 1.79 (3H, s); 1.90 (1H, m); 2.33 (1H, m); 2.54 (1H, m); 3.97 (1H, m); 4.02 (1H, m); 4.42 (2H, m); 4.50 (1H, m); 6.13 (1H, t, *J*=6.6); 7.48 (1H, s); 8.51 (3H, br s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=11.31; 11.98; 14.08; 24.99; 34.90; 35.90; 56.17; 60.08; 64.88; 80.08; 83.84; 109.91; 136.25; 150.36; 163.58; 168.56. $C_{18}H_{25}F_3N_6O_7$•H_2O (515.19 gmol^{-1}): N, 16.31; C, 41.96; H, 5.28%; Found: N, 15.87; C, 42.40; H, 5.38%.

***O*-Phenylalanyl-3'-azido-2',3'-dideoxythymine trifluoroacetic acid salt, *4e*.** Yield, 79%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.81 (3H, s); 2.32 (1H, m); 3.07 (1H, dd, *J*=13.9, 5.6); 3.19 (1H, dd, *J*=13.9, 5.9); 3.91 (1H, m); 4.29 (4H, m); 6.11 (1H, t, *J*=6.6); 7.27 (5H, m); 7.42 (1H, s); 8.61 (3H, br s); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.06; 35.15; 35.93; 53.31; 59.53; 64.79; 79.92; 83.56; 109.98; 127.27; 128.56; 129.31; 134.47; 136.09; 150.30; 163.59; 168.74. $C_{21}H_{23}F_3N_6O_7$•H_2O (546.18 gmol^{-1}): N, 15.38; C, 46.18; H, 4.61%; Found: N, 15.15; C, 46.07; H, 4.45%.

Compounds 5a-h

***O*-(*N*-*tert*-Butyloxycarbonyl)glycylglycyl-3'-azido-2',3'-dideoxythymine, *5a*.** Yield, 25%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.38 (9H, s); 1.80 (3H, s); 2.33 (1H, m); 2.44 (1H, m); 3.57 (2H, d, *J*=6.0); 3.90 (2H, m); 3.98 (1H, m); 4.29 (2H, m); 4.46 (1H, m); 6.13 (1H, t, *J*=6.6); 7.00 (1H, t, *J*=6.0); 7.46 (1H, s); 8.22 (1H, t, *J*=5.7); 11.3 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.01; 28.09; 35.43; 40.53; 42.93; 59.82; 63.58; 77.99; 80.42; 83.37; 109.94; 135.90; 150.28; 155.67; 163.56; 169.54; 169.98. $^{m}/_{z}$ ($M+Na^{+}$) = 504.1924 (calcd, 504.1813).

***O*-(*N*-*tert*-Butyloxycarbonyl)glycylvalyl-3'-azido-2',3'-dideoxythymine, *5b*.** Yield, 37%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.88 (6H, m); 1.37 (9H, s); 1.80 (3H, s); 2.06 (1H, m); 2.33 (1H, m); 2.45 (1H, m); 3.60 (2H, d, *J*=5.4); 4.02 (1H, m); 4.23 (1H, m); 4.29 (2H, m); 4.45 (1H, m); 6.12 (1H, t, *J*=6.6); 6.94 (1H, t, *J*=5.4); 7.44 (1H, s); 8.01 (1H, d, *J*=8.1); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.00; 17.93; 18.80; 28.08; 28.89; 35.36; 42.74; 57.23; 60.24; 63.95; 77.92; 80.37; 83.66; 109.86; 135.81; 150.27; 155.70; 163.55; 169.64; 171.23. $^{m}/_{z}$ ($M+H^{+}$-Boc) = 424.2318 (calcd, 424.1939); $^{m}/_{z}$ ($M+Na^{+}$) = 546.2841 (calcd, 546.2283); $^{m}/_{z}$ ($M+K^{+}$) = 562.2581 (calcd, 562.2022).

***O*-(*N*-*tert*-Butyloxycarbonyl)glycylphenylalanyl-3'-azido-2',3'-dideoxythymine, *5c*.** Yield, 66%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.37 (9H, s); 1.80 (3H, s); 2.30 (2H, m); 3.00 (2H, m); 3.55 (2H, d, *J*=5.4); 3.96 (1H, m); 4.21 (3H, m); 4.49 (1H, m); 6.10 (1H, t, *J*=6.6); 6.91 (1H, t, *J*=5.7); 7.22 (5H, m); 7.42 (1H, s); 8.27 (1H, d, *J*=7.2); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.04; 28.08; 35.53; 36.60; 42.69; 53.72; 60.03; 64.06; 77.92; 80.29; 83.48; 109.98; 126.58; 128.23; 129.01; 135.63; 136.79; 150.25; 155.63; 163.56; 169.47; 171.18. $^{m}/_{z}$ ($M+Na^{+}$) = 594.2361 (calcd, 594.2283).

***O*-(*N*-*tert*-Butyloxycarbonyl)valylglycyl-3'-azido-2',3'-dideoxythymine, *5d*.** Yield, 69%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.83 (3H, d, *J*=6.9); 0.87 (3H, d, *J*=6.9); 1.37 (9H, s); 1.80 (3H, s); 1.94 (1H, m); 2.33 (1H, m); 2.46 (1H, m); 3.83 (1H, m); 3.90 (2H, m); 3.97 (1H, m); 4.30 (2H, m); 4.46 (1H, m); 6.13 (1H, t, *J*=6.6); 6.65 (1H, d, *J*=9.0); 7.46 (1H, s); 8.32 (1H, t, *J*=5.4); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=11.99; 17.94; 19.04; 28.07; 30.33; 35.40; 40.50; 59.37; 59.88; 63.61; 77.90; 80.46; 83.42; 109.92; 135.92; 150.28; 155.31; 163.55; 169.48; 171.98. $^{m}/_{z}$ (M+H^{+}-Boc) = 424.1887 (calcd, 424.1939); $^{m}/_{z}$ (M+Na^{+}) = 546.2386 (calcd, 546.2283).

***O*-(*N*-*tert*-Butyloxycarbonyl)valylalanyl-3'-azido-2',3'-dideoxythymine, *5e*.** Yield, 67%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.81 (3H, d, *J*=6.6); 0.86 (3H, d, *J*=6.6); 1.37 (12H, m); 1.80 (3H, s); 1.92 (1H, m); 2.33 (1H, m); 2.47 (1H, m); 3.83 (1H, m); 3.98 (1H, m); 4.27 (3H, m); 4.45 (1H, m); 6.11 (1H, t, *J*=6.6); 6.55 (1H, d, *J*=9.0); 7.45 (1H, s); 8.29 (1H, d, *J*=6.3); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=11.96; 16.72; 17.90; 19.02; 28.07; 30.47; 35.34; 47.54; 58.97; 60.14; 63.84; 77.85; 80.41; 83.54; 109.89; 135.87; 150.26; 155.28; 163.54; 171.24; 172.15. $^{m}/_{z}$ (M+H^{+}-Boc) = 438.2084 (calcd, 438.2096); $^{m}/_{z}$ (M+Na^{+}) = 560.2596 (calcd, 560.2439); $^{m}/_{z}$ (M+K^{+}) = 576.2386 (calcd, 576.2179).

***O*-(*N*-*tert*-Butyloxycarbonyl)valylvalyl-3'-azido-2',3'-dideoxythymine, *5f*.** Yield, 45%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.86 (12H, m); 1.37 (9H, s); 1.80 (3H, s); 1.91 (1H, m); 2.07 (1H, m); 2.33 (1H, m); 2.46 (1H, m); 3.88 (1H, m); 3.99 (1H, m); 4.22 (1H, m); 4.28 (2H, m); 4.45 (1H, m); 6.12 (1H, t, *J*=6.6); 6.70 (1H, d, *J*=9.0); 7.45 (1H, s); 8.02 (1H, d, *J*=7.8); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=11.97; 17.97; 18.14; 18.73; 19.06; 28.06; 29.73; 30.20; 35.28; 57.24; 59.36; 60.23; 63.86; 77.89; 80.35; 83.64; 109.85; 135.89; 150.27; 155.29;

163.53; 171.02; 171.78. $^{m}/_{z}$ (M+H^+-Boc) = 466.2418 (calcd, 466.2414); $^{m}/_{z}$ (M+Na^+) = 588.2879 (calcd, 588.2752); $^{m}/_{z}$ (M+K^+) = 604.2621 (calcd, 604.2492).

***O*-(*N*-*tert*-Butyloxycarbonyl)valylisoleucyl-3'-azido-2',3'-dideoxythymine, *5g*.** Yield, 31%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=0.83 (12H, m); 1.20 (1H, m); 1.37 (11H, m); 1.79 (3H, s); 1.90 (1H, m); 2.31 (1H, m); 2.46 (1H, m); 3.86 (1H, m); 3.99 (1H, m); 4.27 (3H, m); 4.44 (1H, m); 6.12 (1H, t, *J*=6.6); 6.67 (1H, d, *J*=8.7); 7.45 (1H, s); 8.02 (1H, d, *J*=7.5); 11.4 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=10.92; 11.97; 15.21; 18.13; 19.04; 24.52; 28.05; 30.21; 35.23; 36.06; 56.24; 59.32; 60.32; 63.92; 77.89; 80.36; 83.70; 109.84; 135.90; 150.28; 155.30; 163.53; 171.04; 171.70. $^{m}/_{z}$ (M+H^+-Boc) = 480.2359 (calcd, 480.2571); $^{m}/_{z}$ (M+Na^+) = 602.2781 (calcd, 602.2909).

***O*-(*N*-*tert*-Butyloxycarbonyl)phenylalanylglycyl-3'-azido-2',3'-dideoxythymine, *5h*.** Yield, 74%; ^{1}H NMR (300 MHz, [d_6]DMSO, 25 ºC, TMS): δ=1.32 (9H, s); 1.80 (3H, s); 2.84 (2H, m); 3.02 (2H, m); 3.67 (2H, m); 3.86 (1H, m); 4.04 (1H, m); 4.13 (3H, m); 6.58 (1H, m); 6.93 (1H, d, *J*=9.0); 7.24 (5H, m); 7.81 (1H, s); 10.77 (1H, s). ^{13}C NMR (75.4 MHz, [d_6]DMSO, 25 ºC, TMS): δ=12.02; 28.02; 30.33; 37.34; 40.63; 55.39; 59.90; 63.68; 77.89; 80.48; 83.46; 109.92; 126.05; 127.89; 129.05; 135.91; 138.11; 150.28; 155.14; 163.55; 169.61; 172.35. $^{m}/_{z}$ (M+H^+-Boc) = 472.1571 (calcd, 472.1939).

SOME ADDITIONAL EXAMPLES OF TIME PROFILES FOR THE DECOMPOSITION OF COMPOUNDS 2 IN 80% HUMAN PLASMA

Figure S1. Time profile for the decomposition of **2e** (🗹) into **4a** (■) and AZT ( ) in 80% human plasma at pH 7.4 and 37ºC; solid lines represent the best fit of experimental data to a three parallel first-order reactions model (see text).

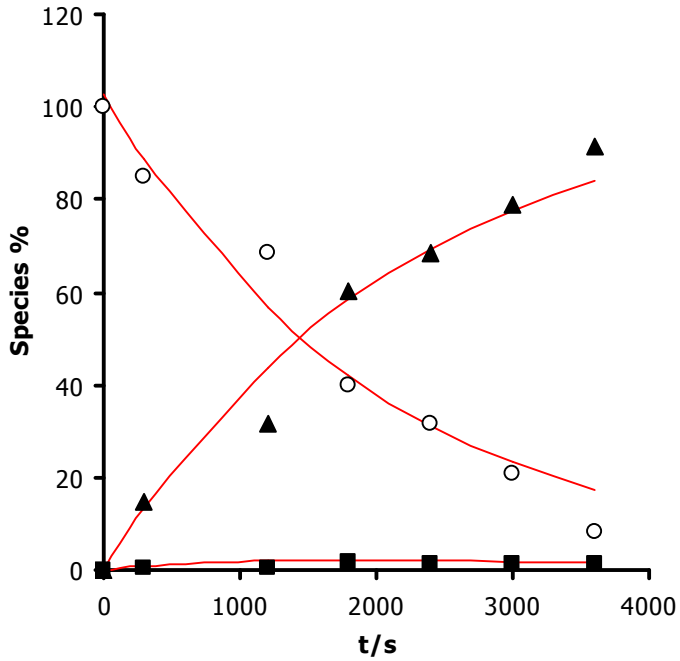

Figure S2. Time profile for the decomposition of **2f** (▩) into **4c** (■) and AZT ( ) in 80% human plasma at pH 7.4 and 37ºC; solid lines represent the best fit of experimental data to a three parallel first-order reactions model (see text).

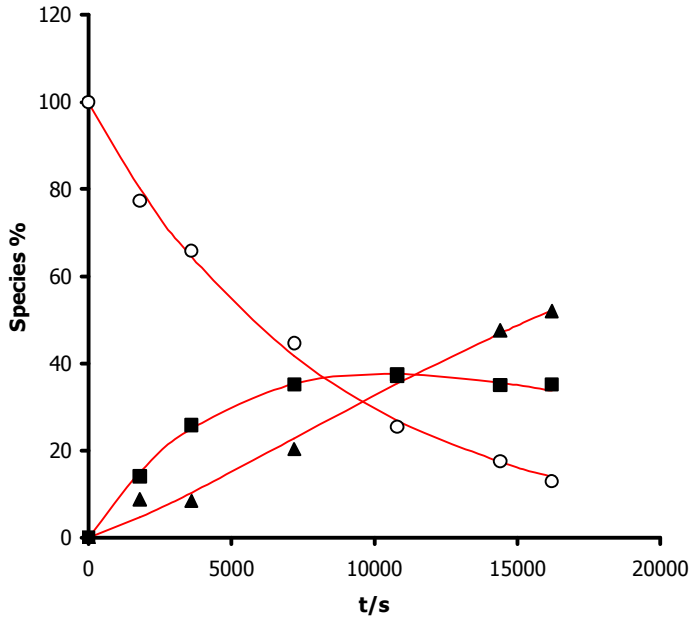

Figure S3. Time profile for the decomposition of **2g** (▩) into **4d** (■) and AZT ( ) in 80% human plasma at pH 7.4 and 37ºC; solid lines represent the best fit of experimental data to a three parallel first-order reactions model (see text).