Supporting Information

Development of a Raltegravir-based Photoaffinity Labeled Probe for Human Immunodeficiency Virus-1 Integrase Capture

Nicolino Pala,[†] Francesca Esposito,[‡] Enzo Tramontano,[‡] Pankaj Kumar Singh,[†] Vanna Sanna,[§] Mauro Carcelli,[#] Lisa D. Haigh,[°] Sandro Satta,[^,ıı] and Mario Sechi[†,*]

[†]*Department of Chemistry and Pharmacy, Laboratory of Drug Design and Nanomedicine, University of Sassari, Italy;* [‡]*Department of Life and Environmental Sciences, University of Cagliari, Italy;* [§]*Nanomater s.r.l., c/o Porto Conte Ricerche, Alghero, Italy;* [#]*Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Italy;* [°]*Department of Chemistry, Imperial College London, UK;* [^]*Centre for Pharmacology & Therapeutics, Department of Medicine, Imperial College London, UK;* [ıı]*Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, USA*

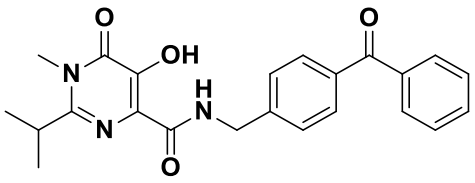
Contents

Experimental Section

Chemistry

All solvents, including anhydrous solvents, and chemicals, were purchased from Aldrich Co., Alfa Aesar, or Carlo Erba, and were used without further purification. All reactions involving air- or moisture-sensitive compounds were performed under a nitrogen atmosphere using oven-dried glassware and syringes to transfer solutions. Analytical thin-layer chromatography (TLC) was carried out on Merck silica gel F-254 plates. Flash chromatography purifications were performed on Merck Silica gel 60 (230–400 mesh ASTM) as a stationary phase. Melting points (mp) were determined using an Electrothermal melting point or a Köfler apparatus and are uncorrected. Nuclear magnetic resonance (^{1}H-NMR and ^{13}C-NMR) spectra were recorded in $CDCl_3$ or DMSO-d_6 on 400 MHz Bruker Avance III. Chemical shifts (δ) are reported in parts per million (ppm) downfield from tetramethylsilane (TMS), used as an internal standard. Splitting patterns are designated as follow: s, singlet; d, doublet; t, triplet; q, quadruplet; m, multiplet; brs, broad singlet; dd, double doublet. The assignment of exchangeable protons (OH and NH) was confirmed by the addition of D_2O. Mass spectra were obtained on a Hewlett-Packard 5989 mass engine spectrometer, or with an ESI-TOF Micromass 4LCZ spectrometer. Mass spectra for peptides and protein were recorded on a Waters LCT Premier operating in ES+ or ES- mode, with a Waters Acquity UPLC system. Infrared (IR) spectra were recorded as thin films or Nujol mulls on NaCl plates with a Perkin-Elmer 781 IR spectrophotometer and are expressed in ν (cm^{-1}). Elemental analyses for purity of compounds **1-3**, and **10** were performed on a Perkin-Elmer Elemental Analyzer 2400-CHN at Laboratory of Microanalysis, Department of Chemistry and Pharmacy, University of Sassari (Italy), and were within ±0.4% of the theoretical values, thus confirming ≥95% purity.

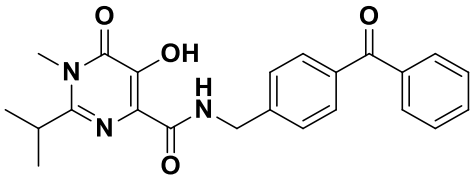
Synthesis of N-(4-benzoylbenzyl)-5-hydroxy-2-isopropyl-1-methyl-6-oxo-1,6-dihydropyrimidine-4-carboxamide (1)

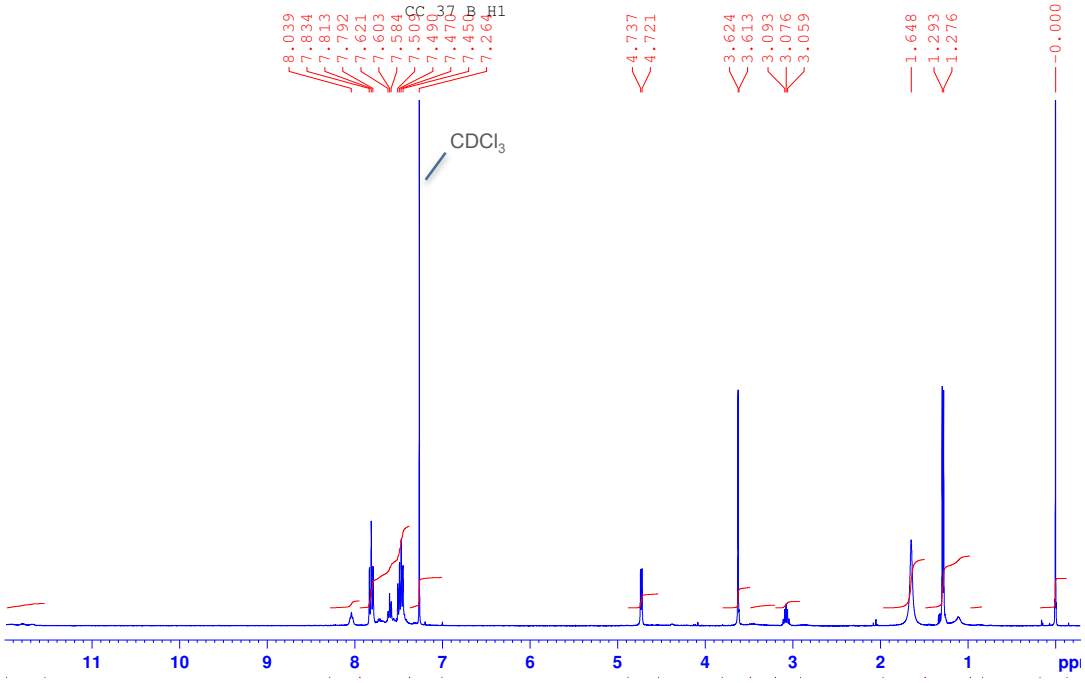


To a solution of 0.102 g (0.4807 mmol, 1.7 eq) of 4-(aminomethyl)(phenyl)(phenyl) methanone (**3**) in 3.5 mL of CH_2Cl_2, 0.081 g (0.4242 mmol, 1.5 eq) of N-(3-dimethylpropyl)-N'-ethylcarbodiimide hydrochloride (EDC·HCl), 0.060 g (0.2828 mmol, 1.5 eq) of 1-hydroxybenzotriazole hydrate (HOBt) and 0.060 g (0.2828 mmol, 1 eq) of 5 (hydroxy)-2-isopropyl-1-methyl-6-bone-1,6-dihydropyrimidine-4-carboxylate (**2**) were added sequentially, and the solution was left under stirring at room temperature for 4.5 h. Subsequently, the mixture was diluted with a saturated solution of $NaHCO_3$ (7 mL) and extracted with ethyl acetate (4 x 20 mL). Then, the organic phase was then counter-extracted in succession with 1N HCl (2 x 20 mL) and a saturated solution of $NaHCO_3$. Finally, a further washing is carried out with a saturated NaCl solution and dried over Na_2SO_4. After evaporation of the solvent, 0.055 g of the desired compound (**1**) was obtained in the form of a light brown solid, which, after trituration with diethyl ether, does not require further purification.

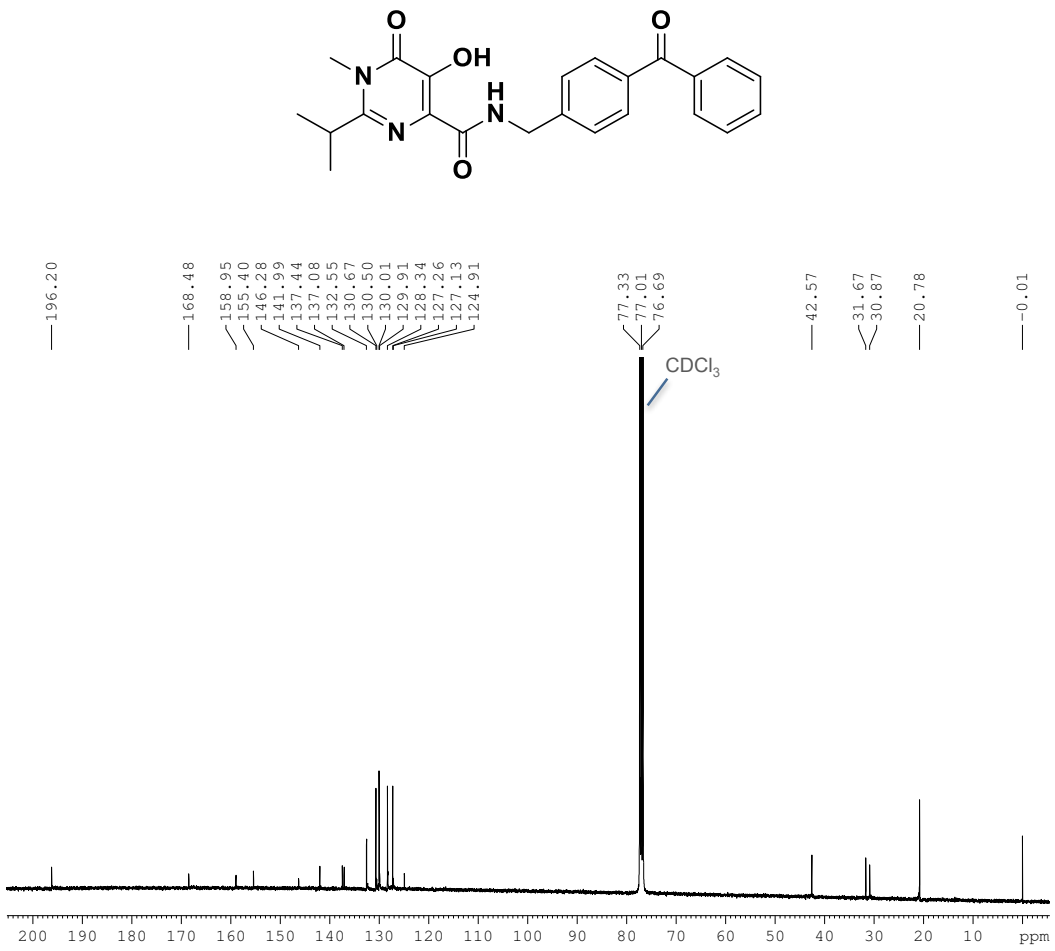
Yield: 48%. **Melting point:** 158-160°C. **^{1}H-NMR 400 MHz ($CDCl_3$):** δ 11.90 (brs, 1H, OH); 8.03 (s, 1H, NH); 7.81-7.70 (m, 4H, Ar-H); 7.59-7.45 (m, 5H, Ar-H); 4.73 (s, 2H, -CH_2); 3.63 (s, 3H -CH_3N); 3.07 (m, 1H, -CH); 1.28 (d, 6H, -CH_3). **^{13}C-NMR ($CDCl_3$):** δ 196.2, 168.5, 158.9, 155.4, 146.3, 142.0, 137.4, 137.1, 132.5, 130.7, 130.0, 129.9. 128.3, 127.2, 127.1, 124.9. 42.5, 31.7, 30.9, 20.8. **ESI/MS (+, m/z):** 404 [M^+]; *m/z* 426 (M^++Na). Anal. Calc. for ($C_{23}H_{23}N_3O_4$): C, 68.13; H, 5.72; N, 10.36. Found: C, 68.27; H, 5.88; N, 10.11.

^{1}H-NMR spectrum of (1)

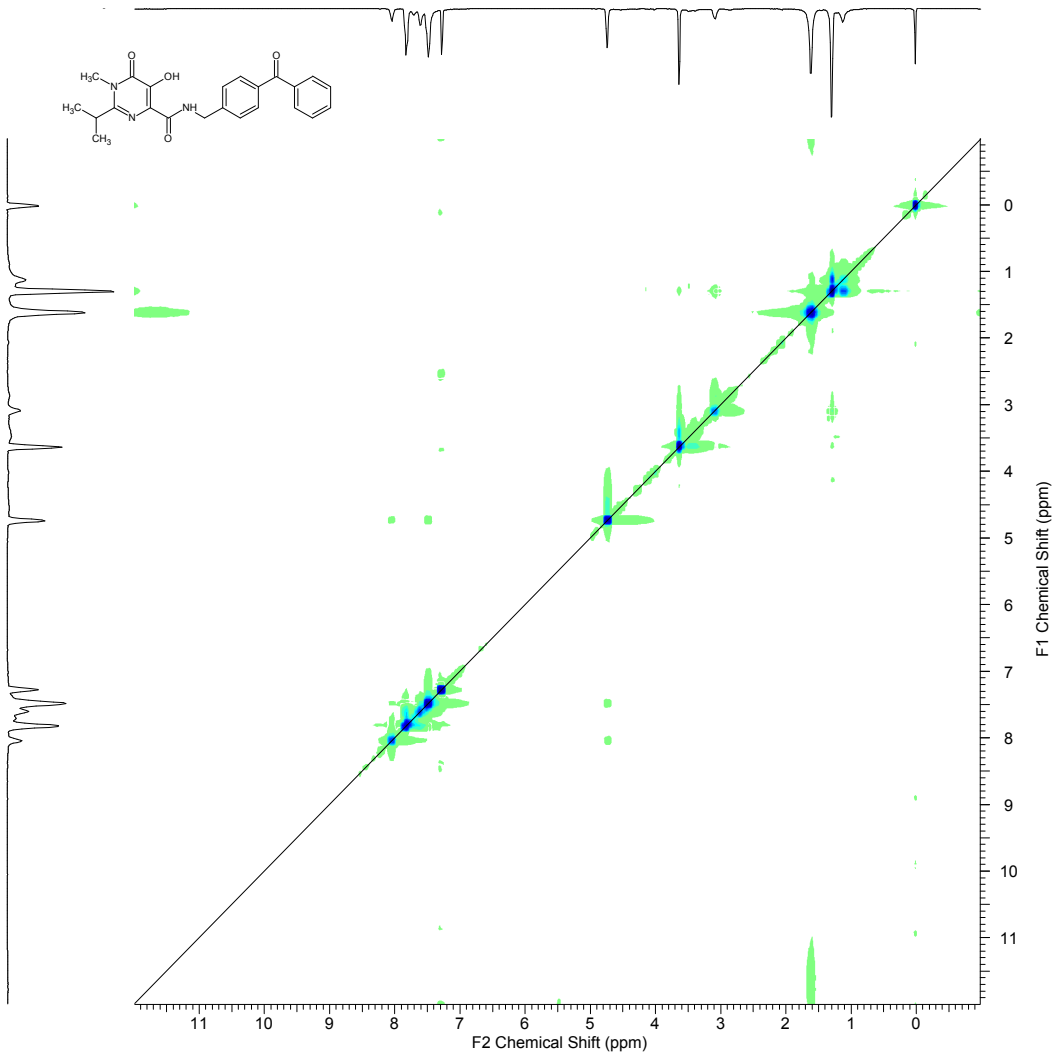




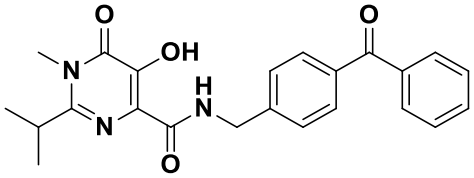
^{13}C-NMR spectrum of (1)

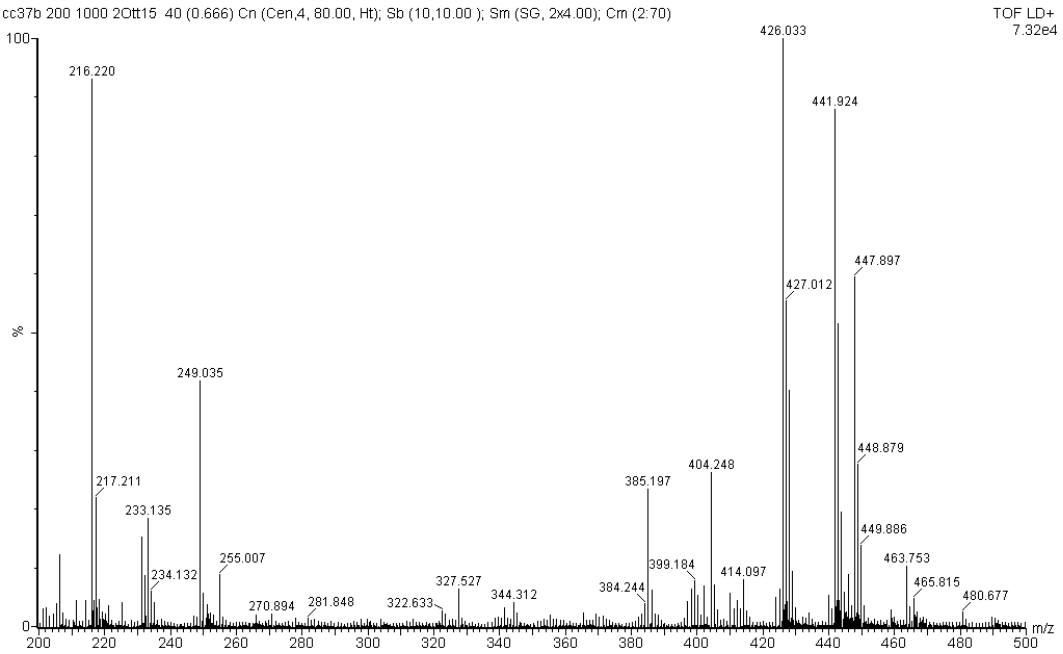


NOESY spectrum of (1)



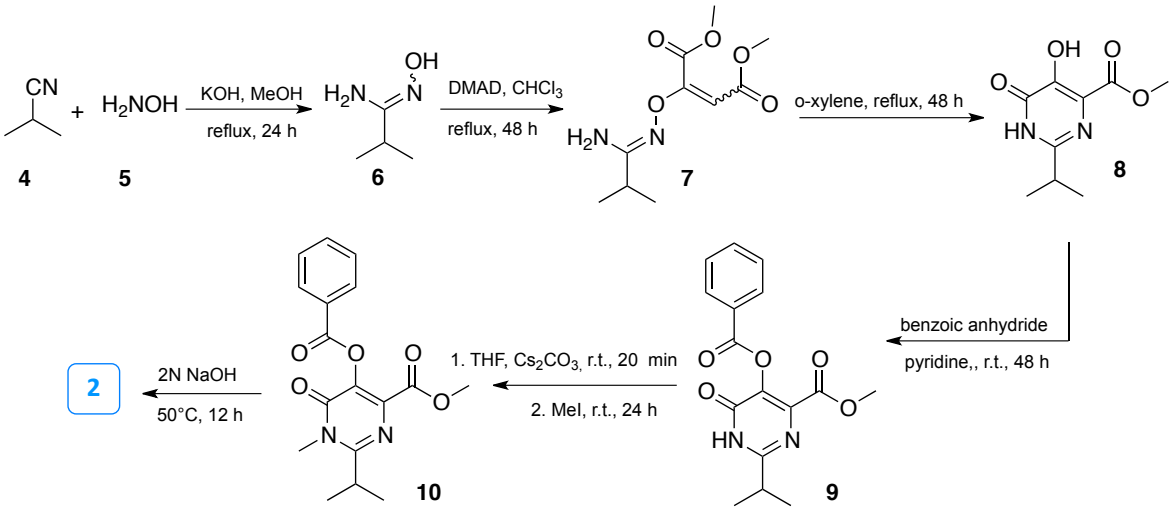
Mass spectrum of (1)



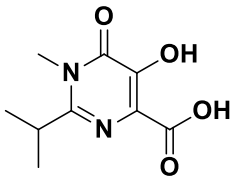


Scheme 2. Detailed synthetic route for the preparation of the synthon 2.

Briefly, the N'-hydroxy-2-methylpropanimideamide (**6**) was obtained by reaction between isobutyrronitrile (**4**) and hydroxylamine hydrochloride (**5**) with KOH in methanol under reflux. The oxime **6** was next reacted with dimethylacetylene dicarboxylate (DMAD) in chloroform, to obtain quantitatively the dimethylester **7**, which was cyclized to generate the dihydroxypyrimidine derivative **8** by reflux condition in *o*-xylene.



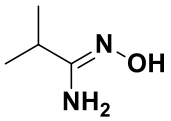
Synthesis of 5-hydroxy-1-methyl-6-oxo-2-(propan-2-yl)-1,6-dihydropyrimidin-4-carboxylic acid (2)



To a solution of methyl 5-(benzoyloxy)-2-isopropyl-1-methyl-6-oxo-1,6-dihydropyrimidin-4-carboxylate (**10**) (0.325 g, 0.9838 mmol) in 5 mL of metanol, 1.72 mL of 2N NaOH (3.5 eq) were added. The mixture is left under stirring at 50 °C for 12 h and then, after cooling, neutralized with 1N HCl. A white-beige precipitate was formed, which is filtered and dried.

Yield: 70% **Melting point:** 110°C. **^{1}H-NMR 400 MHz ($CDCl_3$):** δ 10.0-9-90 (brs, 1H, COOH), 3.65 (s, 3H -CH_3N); 3.11 (m, 1H, -CH); 1.32 (d, 6H, -CH_3). **^{13}C-NMR ($CDCl_3$):** 188.3, 158.2, 157.0, 146.8, 122.5, 31.7, 31.2, 20.6. **ESI/MS (+, m/z):** 212 [M^+]. Anal. Calc. for ($C_9H_{12}N_2O_4$): C, 50.94; H, 5.70; N, 13.20. Found: C, 50.78; H, 5.49; N, 13.04.

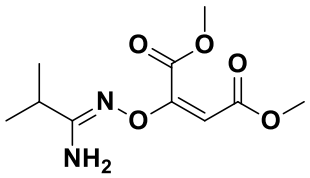
Synthesis of N'-hydroxy-2-methylpropyl-imid-amide (6)[2,3]



To a solution of hydroxylamine hydrochloride (6.00 g, 86.34 mmol) in 60 mL of methanol, KOH powder (5.32 g, 94.97 mmol, 1.1 eq) was added. The mixture is stirred for 30 minutes, and the formation of a KCl precipitate was observed, which is eliminated by filtration, and washed with methanol. Isobutyrronitrile (5.97 g, 86.34 mmol) was added to the clear solution and the mixture was refluxed for 24 h. Upon completion of reaction, the solvent was removed under vacuum to obtain a yellow oil, which does not require further purification.

Yield: 88%. **^{1}H-NMR 400 MHz ($CDCl_3$):** δ 8.60 (bs, 1H, OH), 4.48 (s, br, 2H, NH_2), 2.44 (m, 1H, CH), 1.18 (dd, 6H, CH_3). **ESI/MS (+, m/z):** 102 [M^+].

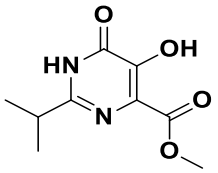
Synthesis of dimethyl 2- (1-amino-2-methylpropylideneaminoxy) but-2-endioate (7)[2,3]



N'-hydroxy-2-methylpropyl-imid-amide (**6**) (6.0 g, 58.74 mmol) was dissolved in 30 mL of anhydrous $CHCl_3$. To this solution a second solution of dimethylacetylene dicarboxylate (DMAD) (10.37 g, 73.42 mmol, 1.25 eq) was added drowise in the same solvent. The mixture turned yellow and is left under reflux for 48 h, then the solvent was removed under vacuum. The brown oil obtained did not require further purification.

Yield: 100%. **^{1}H NMR 400 MHz ($CDCl_3$): main isomer:** δ 5.87 (s, 1H, =CH); 4.74 (s, br, 2H, NH_2); 3.97 (s, 3H, CH_3); 3.72 (s, 3H, CH_3); 2.50 (CH); 1.20 (CH_3). **Secondary isomer:** δ 5.76 (s, 1H, =CH); 5.52 (s, br, 1H, NH); 5.22 (s, br, 1H, NH); 3.84 (s, 3H, CH3); 3.74 (s, 3H, CH_3). **ESI/MS (+, m/z):** 244 [M^+].

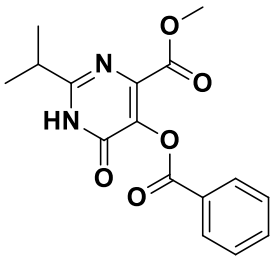
Synthesis of 5-hydroxy-2-isopropyl-6-oxo-1,6-dihydropyrimidin-4-methyl carboxylate (8)[2,3]



A solution of 10 g (40.94 mmol) of dimethyl 2-(1-amino-2-methylpropylideneaminoxy)but-2-endioate (**7**) in 120 mL of *o*-xylene was refluxed for 48 h, monitoring the progress of the reaction by thin layer chromatography (CH_2Cl_2:MeOH = 9.5:0.5). After evaporation of the solvent, a dark oil was obtained, which was dissolved in warm ethyl acetate (25 mL). After cooling to -18°C, a beige precipitate was formed which was subsequently recrystallized with ethyl acetate.

Yield: 55%. **Melting point:** 235-236°C (decomposition). **^{1}H NMR 400 MHz (DMSO-d_6):** δ 12.66 (s, br, 1H, OH); 10.17 (s, br, 1H, NH); 3.81 (s, 3H, CH_3); 2.77 (m, 1H, CH); 1.16 (d, 6H, CH_3). **ESI/MS (+, m/z):** 212 [M^+].

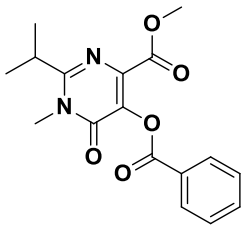
Synthesis of methyl 5-(benzoyloxy)-2-isopropyl-6-oxo-1,6-dihydropyrimidin-4-carboxylate (9)[2,3]



To a solution of 5-hydroxy-2-isopropyl-6-oxo-1,6-dihydropyrimidin-4-methyl carboxylate (**8**) (3.0 g (14.14 mmol) dissolved in anhydrous pyridine (40 mL), benzoic anhydride (3.19 g, 14.14 mmol) was added, and it is left under stirring for 48 h at room temperature. At the end, 1M HCl was added and the mixture is extracted with ethyl acetate (4 x 30 mL). The combined organic phases were washed with 1 N HCl, with brine, and dried over anhydrous Na_2SO_4. After evaporation of the solvent, a solid was obtained, which was triturated with ethyl ether to give a white powder.

Yield: 87%. **Melting point:** 176-178°C. **IR (cm^{-1}):** ν NH 3175; ν C=O 1740, 1680. **^{1}H NMR 400 MHz (DMSO-d_6):** δ 13.29 (s, br, 1H, NH); 8.07 (d, 2H, Ar-H); 7.78 (t, 1H, Ar-H), 7.63 (t, 2H, Ar-H); 3.75 (s, 3H, CH_3); 2.91 (m, 1H, CH); 1.23 (dd, 6H, CH_3). **ESI/MS (+, m/z):** 316 [M^+].

Synthesis of methyl 5-(benzoyloxy)-2-isopropyl-1-methyl-6-oxo-1,6-dihydropyrimidin-4-carboxylate (10)[2,3]

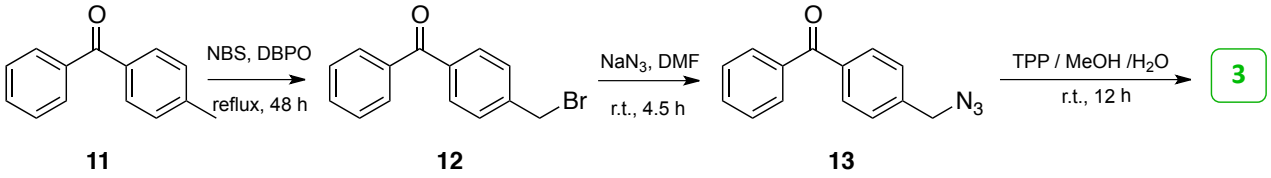


In a solution of methyl 5-(benzoyloxy)-2-isopropyl-6-oxo-1,6-dihydropyrimidin-4-carboxylate (**9**) (3.0 g, 9.48 mmol) in anhydrous tetrahydrofuran (90 mL), Cs_2CO_3 (6.18 g, 18.96 mmol, 2 eq) was added, and the mixture was stirred for 20 minutes at room temperature. Then, methyl iodide (3.90 g, 27.49 mmol, 2.9 eq) was added dropwise, and the reaction mixture was then stirred for 24 hours at room temperature. At the end, after dilution with HCl 1N, the mixture was extracted with ethyl acetate, and the organic phase was washed with brine and dried over anhydrous Na_2SO_4. The solvent was removed in vacuo to obtain a dark hygroscopic solid, which is used without further purifications.

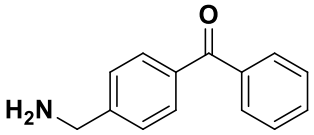
Yield: 98%. **IR (cm−1):** ν C=O = 1737, 1687. **^{1}H NMR 400 MHz ($CDCl_3$):** δ 8.20 (d, 2H, Ar-H); 7.64 (t, 1H, Ar-H); 7.51 (t, 2H, Ar-H); 3.85 (s, 3H, CH_3); 3.68 (s, 3H, CH_3); 3.17 (m, 1H, CH); 1.39 (d, 6H, CH_3). **ESI/MS (+, m/z):** 330 [M^+]. Anal. Calc. for ($C_{17}H_{18}N_2O_5$): C, 61.81; H, 5.49; N, 8.48. Found: C, 61.98; H, 5.29; N, 8.42.

Scheme 3. Detailed synthetic route for the preparation of the synthon 3.

The preparation of aminodiphenylketone **3** (Scheme 2) was accomplished by starting from the 4-methylbenzophenone (**11**), which was refluxed with N-bromosuccinimide (NBS) and benzoyl peroxide in carbon tetrachloride (CH_4Cl_4) to give the bromoderivative **12**. Next, the azide intermediate **13** was obtained quantitatively by treating **12** with sodium azide in dimethylformamide (DMF) at room temperature. **13** was then converted into the corresponding aminomethyl derivative (**3**) by reaction with triphenylphosphine (TPP) in methanol at room temperature.



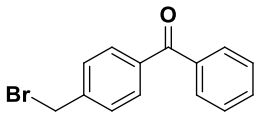
Synthesis of 4-(aminomethyl)(phenyl)(phenyl)methanone (3)[1]



To a solution of 4-(azidomethyl)(phenyl)(phenyl)methanone **(8)** (0.300 g, 1.2644 mmol) in a mixture of methanol and water 2:1 (1.5 mL), 0.5 g of triphenylphosphine (1.8966 mmol, 1.5 eq) was added. The mixture is kept under stirring at room temperature for 18 h, and then diluted with water (3 mL), acidified with 1N HCl and washed with ethyl acetate. The mother liquors are alkalized with a solution of $NaHCO_3$ and subjected to extraction with ethyl acetate (3 x 5 mL); the combined organic phases are then dried over Na_2SO_4 and concentrated in a rotary evaporator. The final product is then obtained in the form of a pale yellow oil that does not require further purification.

Yield: 92%. **^{1}H-NMR 400 MHz ($CDCl_3$):** δ 7.79 (d, 2H, Ar-H); 7.59 (t, 1H, Ar-H); 7.46 (m, 4H, Ar-H); 3.97 (s, 2H -CH_2NH_2). **ESI/MS (+, m/z):** 211 [M^+]. Anal. Calc. for ($C_{14}H_{13}NO$): C, 79.59; H, 6.20; N, 6.63. Found: C, 79.65; H, 6.35; N, 6.51.

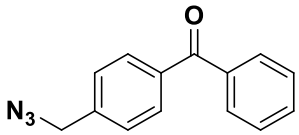
Synthesis of 4-(bromomethyl)(phenyl)(phenyl)methanone (12)[4]



To a solution of 4-methylbenzophenone (5.00 g, 18.2 mmol) in CCl_4 (70 mL), 3.89 g (21.8 mmol) of N-bromosucinimide is added. Then, 0.35 g of dibenzoyl peroxide (1.46 mmol) is added and the reaction mixture was refluxed for 48 hours. Next, after filtration of the precipitate that formed, the solvent was removed in vacuo to obtain an oil which solidified in a short time. The crude product was purified by silica gel column chromatography (petroleum ether:ethyl acetate = 9.5:0.5) to afford a yellow solid.

Yield: 80%. **Melting point:** 110°C. **^{1}H-NMR 400 MHz ($CDCl_3$):** δ 7.83-7.77 (m, 4H, Ar-H), 7.61-7.45 (m, 5H, Ar-H), 4.54 (s, 2H, CH_2). **ESI/MS (+, m/z):** 275 [M^+].

Synthesis of 4-(azidomethyl)(phenyl)(phenyl)methanone (13)[4]



A solution of 4-(bromomethyl)(phenyl)(phenyl)methanone (**7**) (0.350 g, 1.270 mmol) and NaN_3 (0.088 g, 1.3484 mmol, 1.06 eq) in 3.5 mL of dimethylformamide is kept under stirring at room temperature for 4.5 h. At the end the reaction mixture is diluted with water and extracted with diethyl ether (3 x 3 mL). The organic phase is washed with water (3 x 3 mL), dried over Na_2SO_4 and concentrated in vacuo. The final product is obtained in the form of a pale yellow oil which does not require further purification.

Yield: 100%. **^{1}H-NMR 400 MHz ($CDCl_3$):** δ 7.80 (t, 4H, Ar-H), 7.59 (t, 1H, Ar-H), 7.41-7.50 (m, 4H, Ar-H), 4.43 (s, 2H, -CH_2-). **ESI/MS (+, m/z):** 237 [M^+].

Expression and purification of recombinant HIV-1 IN

Recombinant 6xHis tagged wt HIV-1 IN protein was expressed and purified as described).[5,6] Briefly, the pKBIN6Hthr plasmids containing wt IN sequence was used to express the proteins in Escherichia coli strain BL21 (DE3).[7] The protein purification was done using a Ni-Sepharose column with an imidazole gradient from 20 mM to 500 mM concentration in a 50 mM HEPES (pH 7.5) buffer containing 1M NaCl, 7.5 mM CHAPS, 2 mM β-mercaptoethanol, followed by a heparin column purification with a NaCl gradient from 0 to 1M concentrations. Fractions containing IN were pooled and stored in 15% glycerol at -80 °C.[8]

HIV-1 IN inhibition Homogeneous Time Resolved Fluorescence (HTRF) assays

The HIV-1 IN inhibition assay allows measuring the inhibition of 3'-processing and strand-transfer IN reactions in the absence of recombinant LEDGF/p75 protein.[9] Briefly, 50 nM IN was pre-incubated with increasing concentration of compounds for 1 hour at room temperature in reaction buffer containing 20 mM HEPES pH 7.5, 1 mM DTT, 1% Glycerol, 20 mM MgCl2, 0.05% Brij-35 and 0.1 mg/ml BSA. To this mixture, 9 nM DNA donor substrate (5'-ACAGGCCTAGCACGCGTCG-Biotin-3' annealed with 5'-CGACGCGTGGTAGGCCTGT-Biotin3') and 50 nM DNA acceptor substrate (5'-Cy5-ATGTGGAAAATCTCTAGCAGT-3' annealed with 5'-Cy5- TGAGCTCGAGATTTTCCACAT-3') were added and incubated at 37 °C for 90 minutes. After this time, 4 nM of Europium-Streptavidin were added at the reaction mixture and the HTRF signal was recorded using a Perkin Elmer Victor 3 plate reader using a 314 nm for excitation wavelength and 668 and 620 nm for the wavelength of the acceptor and the donor substrates emission, respectively.

Photoaffinity labelling of HIV IN

An aliquot of 5 µM of HIV-IN protein (Abcam, recombinant HIV-1 IN protein, ab49082) dissolved in 50% glycerol, 0.2% Triton-X-100, 1.5M Urea, 25mM Tris HCl at pH 8 was incubated overnight at 4°C with 25 µM or 7.5 µM of **1** (1:5, or 1:1.5, protein:photoprobe, respectively) in DMSO. Crosslinking reaction was performed on ice in a 0.6 µL Eppendorf by irradiating for 15 minutes at 360 nm with a UV lamp (3 cm distance).

HIV-1 IN sequence (Abcam)

M[1]HHHHHHFLD GIDKAQEEHE KYHSNWRAMA SDFNLPPVVA KEIVASCDKC QLKGEAMHGQ VDCSPIWQL DCTHLEGKVI LVAVHVASGY IEAEVIPAET GQETAYFILK LAGRWPVKTI HTDNGSNFTS TTVKAACWWA GIKQEFGIPY NPQSQGVIES MNKELKKIIG QVRDQAEHLK TAVQMAVFIH NFKRKGGIGG YSAGERIVDI IATDIQTKEL QKQITKIQNF RVYYRDSRDP LWKGPAKLLW KGECAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDEDHHHHHH[300]

SDS-page esperiments

10 µg of HIV-1 IN alone and in complex of **1** after photoactivation were mixed with laemmli buffer (ratio 5:1). The samples were heated for 5 minutes at 95 °C and immediately placed on ice before electrophoresis. The samples and the ladder (Spectra Multicolor Broad Range Protein Ladder - Thermo Scientific) were loaded into a pre-made Invitrogen 12% polyacrylamide gel and run for 1h and 30 minutes at 150V. After the cassette was dismounted, the proteins samples and the ladder were transfer onto a PVDF membrane by using Trans-Blot Turbo Transfer System (Bio-rad) at 20V for 7 minutes. The membrane was stained with Ponceau S solution for 30 minutes under Promega blue to enhance band visualisation. In order to remove the excess of Comassie brilliant blue, the PVDF membrane was immerse in 100% MeOH for 1 minutes (2 times). Image was taken with ChemiDoc Imaging Systems (Bio-rad).

Trypsin HIV-1 IN digestion

The native protein or the complex HIV-1 IN + **1** subjected to tryptic digestion were dissolved in 50 mM Tris-HCl buffer with a pH 7.6. Trypsin Gold (Promega) was added to the complex 1:20 (w/w) ratio, for a final concentration of 0.5 mg/mL. The complex with the trypsin was incubated overnight at 37°C. The trypsin digestion was stopped by freezing and the digested samples was stored at -20°C before mass spectrometry analysis.

Site of cleavage and peptides generated after tryptic digestion

#	Frag.	Mass	pI	Peptide
1	1	1760.95	6.56	MHHHHHHFLDGIDK
2	2	869.89	4.75	AQEEHEK
3	3	861.92	8.75	YHSNWR
4	4	1459.72	5.88	AMASDFNLPPVVAK
5	5	863.98	4.37	EIVASCDK
6	6	490.62	8.22	CQLK
7	7	2654.97	4.72	GEAMHGQVDCSPIWQLDCTHLEGK
8	8	3431.97	4.48	VILVAVHVASGYIEAEVIPAETGQETAYFILK
9	9	415.49	9.75	LAGR
10	10	528.65	8.75	WPVK
11	11	1722.83	6.41	TIHTDNGSNFTSTTVK
12	12	1005.20	8.27	AACWWAGIK
13	13	2266.51	4.53	QEFGIPYNPQSQGVIESMNK
14	**14**	**388.46**	**6.10**	**ELK**
15	15	146.19	8.75	K
16	16	684.84	9.75	IIGQVR
17	17	839.90	5.32	DQAEHLK
18	18	1505.80	8.44	TAVQMAVFIHNFK
19	19	174.20	9.75	R
20	20	146.19	8.75	K
21	21	1023.07	6.00	GGIGGYSAGER
22	22	1329.56	4.21	IVDIIATDIQTK
23	23	516.60	6.10	ELQK
24	24	488.58	8.75	QITK
25	25	676.77	9.75	IQNFR
26	26	599.69	8.56	VYYR
27	27	376.37	5.84	DSR
28	28	657.77	5.84	DPLWK
29	29	371.44	8.75	GPAK
30	30	558.72	8.75	LLWK
31	31	1490.65	4.03	GECAVVIQDNSDIK
32	32	469.58	9.72	VVPR
33	33	174.20	9.75	R
34	34	146.19	8.75	K
35	35	217.27	8.80	AK
36	36	400.52	9.75	IIR
37	37	481.51	5.83	DYGK
38	38	1152.26	4.21	QMAGDDCVASR
39	39	1328.29	5.99	QDEDHHHHHH

Alignment of HIV-1 IN sequences

M[1*]HHHHHHF[1]LD GIDKAQEEHE KYHSNWRAMA SDFNLPPVVA KEIVASCDKC QLKGEAM[50]
HGQVDCSPIWQL D[64]CTHLEGKVI LVAVHVASGY IEAEVIPAET GQETAYFILK LAGRWPVKTI
HTD[116]NGSNFTS TTVKAACWWA GIKQEFGIPY NPQSQGVIE[152]S MNKELKKIIG QVRDQAEHLK
TAVQMAVFIH NFKRKGGIGG YSAGERIVDI IATDIQTKE[212]L QKQITKIQNF RVYYRDSRDP
LWKGPAKLLW KGECAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED[288]
HHHHHH[300*]

Legenda:
M[1*] - H[300*] = HIV-1 IN sequence from Abcam
F[1]-D[288] = HIV-1 IN sequence in regular protein (few residues are missing from the X-ray structure derived sequences)
F[1]-D[50] = N-terminal domain (NTD) in regular protein
H[51]-Q[212] = catalytic core domain (CCD) in regular protein
D[64]D[116]E[152] = catalytic triad
EIVASCDK (fragment **5**), **ELK** (fragment **14**) = identified labelled peptides

Molecular Modelling

Molecular modeling was performed on a personal Macbook installed with IOS operating system. For the docking analysis of the photoprobe in the catalytic domain of HIV-IN, "Glide" was used. Glide consist of multiple modes for docking processing. For the present study, the XP mode of docking was used. The photoprobe **1** was sketched and cleaned by utilizing the "builder tool" option available in Maestro.[10] The molecules were then optimized with OPLS3 force-field using the "Ligprep" module of Maestro at a pH range of 7 ± 2.0. For HIV-IN, the co-crystallized X-ray structure, PDB ID:1QS4[11] available in the RCSB database, was retrieved. The protein structure was then optimized using the "protein preparation wizard" module implemented in the Maestro program and the optimization included completion of bond order, addition of hydrogen atoms, construction of missing side chains and loops. Finally, the complex was minimized using an OPLS3 force field to RMSD of 0.15 Å employing the "minimize" option of the protein preparation wizard. Grid was generated using the energy minimized protein structure followed by docking of the photoprobe. Molecular dynamic (MD) simulation was performed on the selected best docked pose of photoprobe in the catalytic domain of HIV-IN. This analysis was carried out using "MOE" software.[12] Partial charges were calculated and energy minimizations were done. MD simulation was carried for the time duration of 2ns. The NVT statistical ensemble was conducted at 310 K. The Nose-Poincare-Anderson equations were used to solve the equations of motion. The simulations were conducted with a time step of 0.5 fs and the coordinate data were stored in the database. RMSD value was calculated to determine the stability of the complex after the time period of 2ns.

References

1) Xie, H.; Ng, D.; Savinov, S. N.; Dey, B.; Kwong, P. D.; Wyatt, R.; Smith, A. B., 3rd; Hendrickson, W. A. Structure-activity relationships in the binding of chemically derivatized CD4 to gp120 from human immunodeficiency virus. *J. Med. Chem.* **2007**, *50,* 4898-4908
2) Bacchi, A.; Carcelli, M.; Compari, C.; Fisicaro, E.; Pala, N.; Rispoli, G.; Rogolino, D.; Sanchez, T. W.; Sechi, M.; Sinisi, V.; Neamati, N. Investigating the role of metal chelation in HIV-1 integrase strand transfer inhibitors. *J. Med. Chem.* **2011**, *54*, 8407-8420
3) Gardelli, C.; Nizi, E.; Muraglia, E.; Crescenzi, B.; Ferrara, M.; Orvieto, F.; Pace, P.; Pescatore, G.; Poma, M.; Ferreira Mdel, R.; Scarpelli, R.; Homnick, C. F.; Ikemoto, N.; Alfieri, A.; Verdirame, M.; Bonelli, F.; Paz, O. G.; Taliani, M.; Monteagudo, E.; Pesci, S.; Laufer, R.; Felock, P.; Stillmock, K. A.; Hazuda, D.; Rowley, M.; Summa, V. Discovery and synthesis of HIV integrase inhibitors: development of potent and orally bioavailable N-methyl pyrimidones. *J. Med. Chem.* **2007**, *50*, 4953-4975
4) Sechi, M.; Carta, F.; Sannia, L.; Dallocchio, R.; Dessì, A.; Neamati, N. Design, synthesis, molecular modeling and anti-HIV 1 integrase activity of a series of photoactivatable diketo acid-containing inhibitors as affinity probes. *Antiviral Res.* **2009**, *81*, 267-276
5) Kessl, J.J., Jena, N., Koh, Y., Taskent-Sezgin, H., Slaughter, A., Feng, L., de Silva, S., Wu, L., Le Grice, S.F.J., Engelman, A.; Fuchs, J.R.; Kvaratskhelia, M. Multimode, cooperative mechanism of action of allosteric HIV-1 integrase inhibitors. *J. Biol. Chem.* **2012**, *287*, 16801–16811
6) Esposito, F.; Ambrosio, F.A.; Maleddu, R.; Costa, G.; Rocca, R.; Maccioni, E.; Catalano, R.; Romeo, I.; Eleftheriou, P.; Karia, D.C.; Tsirides, P.; Godvani, N.; Pandya, H.; Corona, A.; Alcaro, S.; Artese, A.; Geronikaki, A.; Tramontano, E. Chromenone derivatives as a versatile scaffold with dual mode of inhibition of HIV-1 reverse transcriptase-associated Ribonuclease H function and integrase activity. *Eur. J. Med. Chem.* **2019**, *182*, 111617
7) Sanna, C.; Scognamiglio, M.; Fiorentino, A.; Corona, A.; Graziani, V.; Caredda, A.; Cortis, P.; Montisci, M.; Ceresola, E.R.; Canducci, F.; Poli, F.; Tramontano, E.; Esposito F. Prenylated phloroglucinols from Hypericum scruglii, an endemic species of Sardinia (Italy), as new dual HIV-1 inhibitors effective on HIV-1 replication. *PlosOne* **2018**, 13, e0195168
8) Sanna, C.; Rigano, D., Corona. A.; Piano, D.; Formisano, C.; Farci, D.; Franzini, G.; Ballero, M.; Chianese, G.; Tramontano, E.; Taglialatela-Scafati, O.; Esposito, F. Dual HIV-1 reverse transcriptase and integrase inhibitors from Limonium morisianum Arrigoni, an endemic species of Sardinia (Italy). *Nat. Prod. Res.* **2018**, *4*, 1-6
9) Esposito, F.; Tintori, C.; Martini, R.; Christ, F.; Debyser, Z.; Ferrarese, R.; Cabiddu, G.; Corona, A.; Ceresola, E.R.; Calcaterra, A.; Iovine, V.; Botta, B.; Clementi, M.; Canducci, F.; Botta, M.; Tramontano, E. Kuwanon-L as a new allosteric HIV-1 integrase inhibitor: molecular modeling and biological evaluation. *ChemBioChem* **2015**, *16*, 2507-2512
10) The authors are grateful to Schrödinger Team for the opportunity to evaluate a trial of the last version of software Maestro.
11) Goldgur, Y.; Craigie, R.; Cohen, G.H.; Fujiwara, T.; Yoshinaga, T.; Fujishita, T.; Sugimoto, H.; Endo, T.; Murai, H.; Davies, D.R. Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13040–13043
12) Molecular Operating Environment, MOE 2009.10; Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2009**.