Bioorganic Chemistry **28,** 140–155 (2000)
doi:10.1006/bioo.2000.1166, available online at http://www.idealibrary.com on IDEAL®

Carbonyl J Derivatives: A New Class of HIV-1 Integrase Inhibitors

Karl Maurer,*,† Ann H. Tang,‡ George L. Kenyon,*,§ Andrew D. Leavitt‡,¶

‡*Department of Laboratory Medicine,* ¶*Department of Internal Medicine, and* **School of Pharmacy, University of California, San Francisco, California;* §*College of Pharmacy, University of Michigan, Ann Arbor, Michigan; and* †*Combimatrix, Burlingame, California*

Received October 8, 1999

Integration of a DNA copy of the HIV-1 genome is required for viral replication and pathogenicity, and this highly specific molecular process is mediated by the virus-encoded integrase protein. The requirement for integration, combined with the lack of a known analogous process in mammalian cells, makes integrase an attractive target for therapeutic inhibitors of HIV-1 replication. While many reports of HIV-1 IN inhibitors exist, no such compounds have yet emerged to treat HIV-1 infection. As such, new classes of integrase inhibitors are needed. We have combined molecular modeling and combinatorial chemistry to identify and develop a new class of HIV-1 integrase inhibitors, the Carbonyl J [*N*,*N*′-bis(2-(5-hydroxy-7-naphthalenesulfonic acid)urea] derivatives. This new class includes a number of compounds with sub-micromolar IC_{50} values for inhibiting purified HIV-1 integrase *in vitro*. Herein we describe the chemical characteristics that are important for integrase inhibition and cell toxicity within the Carbonyl J derivatives.

INTRODUCTION

Retroviral integration, the covalent ligation of a complete copy of the viral genome into the host cell DNA, is an essential step in the retroviral life cycle and is mediated by the virus-encoded integrase (IN) protein (1). Following receptor-mediated cell entry, virus-encoded reverse transcriptase (RT) uses the viral RNA as a template for synthesis of a DNA copy of the viral genome. IN then removes two nucleotides from the 3′-end of each strand of the viral DNA, referred to as 3′-processing. While still in a nucleoprotein complex, also called the preintegration complex, the viral genome migrates into the nucleus where IN acts to link covalently the DNA copy of the viral genome to the host cell DNA. The latter step, called strand transfer, generates the provirus that serves as a template for future virus production and ensures that all daughter cells contain a complete copy of the virus genome.

IN is one of three enzymes common to all retroviruses, the other two being RT and protease (PR). RT and PR are well characterized, their crystallographic structures have been determined, and each is a target of clinically available anti-retroviral agents (*2*,*3*). While anti-RT and anti-PR therapeutics have made significant advances in controlling HIV-1 infection and improving the lives of people with AIDS, problems with toxicity and resistance continue to plague the current anti-retroviral armamentarium (*2*,*3*). As such, new anti-retroviral agents, directed against new targets, remain an important goal of AIDS research.

The essential role of IN in viral replication, and the lack of a known functional analog in human cells, make IN an attractive anti-retroviral target (*1*). Multiple classes of

0045-2068/00 $35.00

IN inhibitors have been reported to date, including nucleotides/nucleosides (*4,5*), oligonucleotides (*5–10*), DNA binding molecules (*11,12*), proteins (*13,14*), and peptides (*15*). The majority of reports, however, have involved hydroxylated and polyhydroxylated aromatic compounds including aurintricarboxylic acid and its derivatives (*16,17*), cosalines (*18,19*), and topoisomerase inhibitors (*12*). While the initial searches for IN inhibitors were rather broad in scope, many subsequent ones have been driven by the search for better hydroxylated aromatic compounds (*20–23*) and by searching the NCI drug database for compounds that share a pharmacophore with published IN inhibitors (*24–28*). While there are numerous reported inhibitors, only one has suggested a direct inhibition of IN activity in culture (*29,30*). As such, new anti-IN lead compounds are needed, and this will most likely happen by taking new approaches to identifying such compounds.

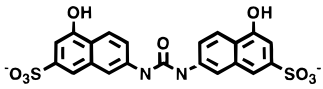
We have used a combination of structure-based computer modeling and combinatorial chemistry to identify new inhibitors of HIV-1 RT and HIV-1 IN. During one of our structure-based searches for new IN inhibitors, we noticed a structural motif similarity between a new set of potential IN inhibitor compounds and a class of compounds that were already under evaluation for their anti-RT properties, the Carbonyl J class of compounds (*31*). The lead RT inhibitor compound at that time, Carbonyl J (Fig. 1; compound **1**), was therefore tested for its anti-IN activity. It had an IC_{50} of 4 μM for *in vivo* IN-mediated 3′-processing and strand transfer. Subsequent structural similarity searches identified Calcomine orange (Fig. 1; compound **2**) with much improved anti-IN activity. Herein we describe our structure–function analysis of this new class of HIV-1 IN inhibitors, the Carbonyl J derivatives, which have the benefit of also possessing significant anti-RT activity (*31*). The combined anti-IN and anti-RT effect in a single compound makes the Carbonyl J derivatives an intriguing class of compounds for further study and development.

EXPERIMENTAL

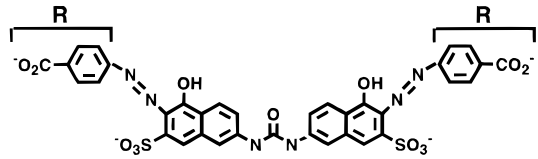
Materials. Carbonyl J (*N,N′*-bis(2-(5-hydroxy-7-naphthalenesulfonic acid)) urea) was purchased from Pfaltz and Bauer, Inc. (Waterbury, CT). Calcomine orange was purchased from Sigma (St. Louis, MO). All other chemicals were purchased from Aldrich Chemical (Milwaukee, WI).

Synthesis. A general procedure was used for production of the di(diazo) inhibitors reported. This procedure may be scaled from approximately 25 to 500 mg of aniline or an aniline derivative. A sample of the aniline to be coupled (1.9 mM) was slurried in 2 ml of water and 1.25 ml of 20% H_2SO_4 was added followed by the addition of 1.8 mM $NaNO_2$ in 1 ml water. The reaction was mixed until >95% of the precipitate dissolved (generally <5 min) and was then added to 15 ml of water. A saturated Na_2CO_3 solution was added (typically close to 6 ml) to achieve a pH of 9. A solution of compound **1** or analog (0.95 mM fully dissolved in 10 ml warm water) was added and the reaction mixture immediately turned dark orange-red. Saturated Na_2CO_3 solution (3 ml) was added to maintain basicity. The resulting solution was stirred at room temperature for 2.5–4 h, acidified with 20% H_2SO_4 to a pH of approximately 0 (approximately 5 ml), and then diluted with water to 140 ml. The precipitate was isolated by centrifugation, washed with water (2×80 ml) to remove salts, methanol, or ethanol (2×40 ml) to remove residual organics, and then vacuum dried overnight. Electron spray and liquid secondary ion mass spectroscopy analyses were performed at the UCSF Mass Spectrometry Facility (A.L. Burlingame, Director).

A



CARBONYL J (1)



CALCOMINE ORANGE (2)

B

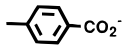
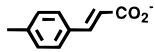
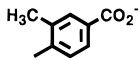
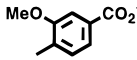
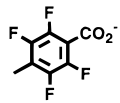
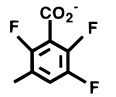
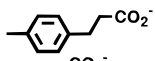
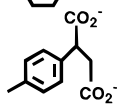
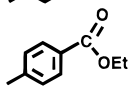
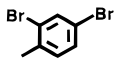
COMPOUND #	R	IC_{50} (μM) 3'-P	IC_{50} (μM) ST	TC_{50} (μM)	T.I. (TC_{50}/IC_{50})
1	N/A	4.0	4.0	ND	—
2		0.35	0.50	85	170
3		2.7	3.2	150	47
4		0.4	0.70	90	129
5		0.5	0.70	70	100
6		0.20	0.33	165	507
7		1.0	1.0	335	335
8		0.50	0.50	140	280
9		0.50	1.0	300	300
10		0.30	0.70	ND	—
11		2.0	4.0	ND	—
12		2.7	1.3	ND	—
13		4.0	0.60	ND	—

FIG. 1. Structure–function analysis and toxicity data for Carbonyl J derivatives. (A) The two drawings at the top show the chemical structure of Carbonyl J (**1**) and Calcomine Orange (**2**). The "R" group of Calcomine Orange is relevant to compounds **2–13** in B. (B) Carbonyl J (**1**) has no "R" group as shown in A. The IC_{50} values are from *in vitro* assays and represent the concentration of inhibitor that yields a 50% reduction in IN activity compared to control reactions performed in the absence of inhibitor. The TC_{50} represents the concentration of inhibitor that yields a 50% reduction in live cells after 5 days of continuous exposure to the compound as compared to control cells grown for 5 days in the absence of inhibitor. Compounds **10–13** were not tested (ND) for TC_{50} due to poor solubility. A T.I. (therapeutic index: TC_{50}/IC_{50}) value is therefore provided for compounds **2–9.**

Matrix-assisted laser desorption ionization mass spectroscopy (MALDIMS) was performed on a PerSeptive Biosystems Voyager-DE instrument and were internally calibrated by close proximity spotting. NMR spectra (^{1}H and ^{13}C) were taken on a GE 300 MHz instrument, and shifts were measured against tetramethylsilane. Data presented come from compounds of >95% purity as judged by NMR spectra, without any unidentified peaks or peaks attributable to synthetic by-products.

*N,N′-bis-(2-(5-hydroxy-7-naphthalenesulfonic acid))oxalamide (**18**).* A mixture of oxallyl di-imidazole (1.98 g, 10.4 mM), J-acid (6-amino-1-hydroxy-3-naphthalenesulfonic acid; 5.00 g, 2(1 mM, dried on pump overnight) and imidazole (1.60 g, 23.5 mM, recrystallized from benzene) was diluted with DMF (20 ml, dry). After stirring for a few hours, the reaction became cloudy and remained so for the duration of the reaction. After 4 days the reaction mixture was centrifuged, and the pelleted material was washed with ethanol (4 × 10–15 ml) until the liquid layer was only lightly colored. The material was dried in vacuum to yield 2.67 g of an off-white powder containing trace DMF. Negative ion liquid secondary ion mass spectroscopy found (MH-) 531^{1}H NMR (300 MHz, DMSO) 11.10 (s, 2H) 9.03 (s, 2H) 8.45 (s, 2H) 8.10 (d, *J* = 9 Hz, 2H) 7.93 (d, *J* = 9 Hz, 2H) 7.67 (s, 4H) 7.54 (s, 2H) 7.14 (s, 2H) ^{13}C NMR (75 MHz, DMSO) 158.86, 152.75, 146.58, 135.96, 134.44, 133.61, 122.68, 121.74, 119.68, 119.46, 117.55, 114.79, 105.52

*N,N′-bis-(2-(5-hydroxy-6-(azo-(4-benzoic acid))-7-naphthalenesulfonic acid))urea (**2**).* Negative ion electron spray mass spectroscopy (ESMS) found (MH-) 799, (MNa-) 821 ^{1}H NMR (300 MHz, DMSO) 15.81 (s, 2H) 12.5–13.0 (broad s, 2 H) 9.53 (s, 2H) 8.21 (d, *J* = 8.7 Hz, 2H) 7.99 (d, *J* = 8.1 Hz 4H) 7.82 (d, *J* = 8.1 Hz, 4H) 7.76 (s, 2H) 7.50 (s, 2H) ^{13}C NMR (75 MHz, DMSO) 177.81, 166.88, 151.85, 146.14, 144.82, 143.50, 137.00, 130.84, 129.40, 128.65, 126.90, 125.04, 122.29, 117.91, 116.62, 116.35

*N,N′-bis-(2-(5-hydroxy-6-(azo-(4-cinnaminic acid))-7-naphthalenesulfonic acid)) urea (**3**).* Negative ion ESMS found (MH-) 851, (MNa-) 873 ^{1}H NMR (300 MHz, DMSO) 15.94 (s, 2H) 12.08 (broad s, 2H) 9.58 (s, 2H) 8.20 (d, *J* = 8.4 Hz, 2H) 7.77 (m, 12H) 7.62 (d, *J* = 15.9 Hz, 2H) 7.51 (s, 2H) 6.56 (d, *J* = 15.9 Hz, 2H) ^{13}C NMR (75 MHz, DMSO) 177.00, 167.34, 151.55, 144.25, 143.87, 143.71, 143.12, 136.72, 131.08, 129.37, 128.76, 128.16, 124.70, 121.05, 117.89, 117.37, 117.11, 115.89

*N,N′-bis-(2-(5-hydroxy-6-(azo-(2-methyl-4-benzoic acid))-7-naphthalenesulfonic acid)) urea (**4**).* Negative ion ESMS found (MH-) 827, (MNa-) 849 ^{1}H NMR (300 MHz, DMSO) 16.17 (s, 2H) 9.64 (s, 2H) 8.19 (d, *J* = 7.8 Hz, 2H) 8.09 (d, *J* = 7.5 Hz, 2H) 7.85 (s, 6H) 7.62 (d, *J* = 7.8 Hz, 2H) 7.57 (s, 2H) 2.45 (s, 6H) ^{13}C NMR (75 MHz, DMSO) 177.77, 167.06, 151.77, 144.78, 144.07, 143.33, 136.92, 132.09, 130.14, 128.71, 128.55, 126.64, 125.17, 124.93, 122.39, 117.91, 116.20, 115.33, 56.09 (EtOH) 18.60 (EtOH) 16.72

*N,N′-bis-(2-(5-hydroxy-6-(azo-(2-methoxy-4-benzoic acid))-7-naphthalenesulfonic acid))urea (**5**).* Negative ion ESMS found (MH-) 799, (MNa-) 821 ^{1}H NMR (300 MHz, DMSO) 15.94 (s, 2H) 9.69 (s, 2H) 8.21 (d, *J* = 8.4 Hz, 2H) 8.04 (d, *J* = 7.8 Hz, 2H) 7.85 (s, 2H) 7.65 (m, 6H) 7.56 (s, 2H) 4.06 (s, 6H) ^{13}C NMR (75 MHz, DMSO) 178.00, 166.93, 151.85, 147.44, 144.78, 143.45, 136.91, 134.94, 130.15, 128.80, 127.36, 125.11, 123.86, 122.28, 117.86, 116.28, 115.38, 112.13, 56.23

*N,N′-bis-(2-(5-hydroxy-6-(azo-(3-benzoic acid))-7-naphthalenesulfonic acid)) urea (**8**).* Negative ion ESMS found (MH-) 799, (MNa-) 821 ^{1}H NMR (300 MHz, DMSO) 16.03 (s, 2H) 9.55 (s, 4H) 8.23 (d + s, 2H) 8. 08 (d, *J* = 8.1 Hz, 2H) 7.77 (s, 6H) 7.60 (t, *J* = 7.8 Hz 2H) 7.52 (s, 2H) ^{13}C NMR (75 MHz, DMSO) 176.74, 167.02, 151.99, 151.91, 144.80,

143.35, 143.17, 136.88, 132.25, 129.94, 128.69, 126. 39, 125.09, 121.58, 121.29, 118.16, 117.88, 116.19, 56.16 (EtOH), 18.64 (EtOH)

*N,N′-bis-(2-(5-hydroxy-6-(azo-(4-(3-propanoic acid)phenyl)-7-naphthalenesulfonic acid)) urea (**9**)*. Negative ion MALDIMS found (MH-) 855, (MNa-) 877 ^{1}H NMR (300 MHz, DMSO) 16.22, (s, 2H) 12.07 (broad s, 2H) 9.53 (s, 2H) 9.21 (d, 8.4 Hz, 2H) 7.78–7.68 (m, 8H) 7.50 (s, 2H) 7.33 (d, *J* = 7.8 Hz, 4H)2.87 (t, poorly resolved, 4H) 2.57 (t, *J* = 7.2 Hz, 4H) ^{13}C NMR (75 MHz, DMSO) 175.22, 173.80, 170.39, 151.96, 144.30, 143.48, 141.10, 139.16, 136.64, 129.41, 128.09, 124.99, 120.48, 117.75, 117.52, 115.92, 59.81(EtOAc), 56.09(EtOH), 35. 29, 30.06, 20.81 (EtOAc), 18.59 (EtOH), 14.14 (EtOAc).

*N,N′-bis-(2-(5-hydroxy-6-(azo-(4-(2-succinic acid)phenyl)-7-naphthalenesulfonic acid)) urea (**10**)*. Negative ion MALDIMS found (MH-) 944, (MNa-) 965 ^{1}H NMR (300 MHz, DMSO) 16.05 (broas S, 2H) 9.49 (s, 2H) 8.21 (d, *J* = 8.4 Hz, 2H) 7.77–7.01 (m, 8H) 7.48 (s, 2H) 7.38 (d, *J* = 7.8 Hz, 4H) 3.94 (dd, *J* = 4.5, 10.2 Hz, 2H) 2.99 (dd, *J* = 10.2, 16.5 Hz, 2H) 2.58 (dd, *J* = 4.5, 16.8 Hz, 2H) ^{13}C NMR (75 MHz, DMSO) 175.96, 174.00, 172.74, 152.01, 144.46, 143.40, 141.95, 136.81, 136.40, 128.97, 128.32, 125.11, 120.96, 117.96, 117.60, 116.16 46.55, 37.49

*N,N′-bis-(2-(5-hydroxy-6-(azo-(3-(carboxyethoxy)phenyl))-7-naphthalenesulfonic acid))urea (**11**)*. Negative ion MALDIMS found (MH-) 855, (MNa-) 877 ^{1}H NMR (300 MHz, DMSO) 15.91 (s, 2H) 9.59 (s, 2H) 8.19 (d + s, 4H) 8.10 (d, *J* = 7.8 Hz, 2H) 7.77 (m, 6H) 7.59 (t, *J* = 7.5 Hz, 2H) 7.52 (s, 2H) 4.34 (q, *J* = 6.9, 4H) 1.35 (t, *J* = 7.5 Hz, 6H) ^{13}C NMR (75 MHz, DMSO) 176.68, 165.33, 151.80, 144.66, 143.19, 136.77, 131.19, 129.91, 128.76, 128.37, 125.92, 124.95, 121.75, 121.26, 118.03, 117.82, 116.09, 61.04, 56.12 (EtOH), 18.58 (EtOH), 14.18

*N,N′-bis-(2-(5-hydroxy-6-(azo-(4-(carboxyethoxy)phenyl))-7-naphthalenesulfonic acid))urea (**12**)*. Negative ion MALDIMS found (MH-) 855, (MNa-) 877 ^{1}H NMR (300 MHz, DMSO) 15.77 (s, 2H) 9.55 (s, 2H) 8.18 (d, *J* = 8.4 Hz, 2H) 7.98 (d, *J* = 8.4 Hz, 4H) 7.82 (d, *J* = 8.4 Hz, 4H) 7.76 (s, 2H) 7.71 (d, *J* = 9.0 Hz, 2H) 7.51 (s, 2H) 4.31 (q, *J* = 6.9 Hz, 4H) 1.34 (t, *J* = 6.9 Hz, 6H) ^{13}C NMR (75 MHz, DMSO) 177.84, 165.25, 151.65, 146.27, 144.81, 143.08, 136.82, 130.59, 129.48, 128.57, 125.79, 124.92, 122.71, 117.83, 116.58, 116.28, 60.61, 14.21

*N,N′-bis-(2-(5-hydroxy-6-(azo-(2,4-dibromophenyl))-7-naphthalenesulfonic acid))urea (**13**)*. Negative ion MALDIMS found (MH-) 1025, 1027 (MNa-) 1048 ^{1}H NMR (300 MHz, DMSO) 16.00 (s, 2H) 9.56 (s, 2H) 8.16 (d, *J* = 8.4 Hz, 2H) 8.00 (d, *J* = 8.7 Hz, 2H) 7.93 (s, 2H) 7.75 (s, 2H) 7.67 (m, 2H) 7.55 (s, 2H) ^{13}C NMR (75 MHz, DMSO) 177.33, 151.62, 144.78, 143.34, 140.09, 136.91, 134.45, 131.94, 130.03, 128.71, 124.59, 122.69, 119.10, 117.81, 117.23, 116.29, 111.39

*N,N′-bis-(2-(5-hydroxy-6-(azo-(2,3,5,6-tetrafluoro-4-benzoic acid)-7-naphthalenesulfonic acid))urea (**6**)*. Negative ion MALDIMS found (M(−3H)-) 939, (M(−4H)Na-) 961 ^{1}H NMR (300 MHz, DMSO) 15.82(s, 2H) 10.88 (s, 2H) 9.58 (s, 2H) 8.24, (d, *J* = 8.7, 2H) 7.84 (s, 2H) 7.79 (d, *J* = 9. 0 Hz, 2H) 7.57 (s, 2H) ^{13}C NMR (75 MHz, DMSO, 50°C) 172.67, 160.82, 151.35, 146.42 (m), 144.34, 143.14 (m), 141.73, 141.10 (m), 139.44 (m), 137.71 (m), 136.57, 134.67, 134.50, 128.64, 127.95, 123.24, 121.96, 118.26, 115.94, 180.92 (m) (multiplets due to ^{19}F splitting).

*N,N′-bis-(2-(5-hydroxy-6-(azo-(2,4,5-trifluoro-3-benzoic acid)-7-naphthalenesulfonic acid))urea (**7**)*. Negative ion MALDIMS found (M(−3H)-) 903, (M(−4H)Na-) 925 ^{1}H

NMR (300 MHz, DMSO) 16.04 (broad s, 2H) 9.50 (s, 2H) 8.24, (d, *J* = 8.7 Hz, 2H) 7.93 (dd, *J* = 6.9, 10.8 Hz, 2H) 7.84 (s, 2H) 7.75 (d, *J* = 8.7 Hz, 2H) 7.57 (s, 2H) ^{13}C NMR (75 MHz, DMSO) 171.22, 166.10, 151.90, 150.10, 148.76, 146.91, 145.41, 145.02 (m), 144.02, 143.16, 136.59, 129.09, 127.85, 124.45 (m), 123.84, 121.11, 118.26, 115.86, 110.41 (m), 106.90 (m) (multiplets due to ^{19}F splitting).

N,N′-bis-(2-(5-hydroxy-6-(azo-(4-benzoic acid))-7-naphthalenesulfonic acid))oxalamide ***(14)***. Negative ion MALDIMS found (MH-) 827, (MNa-) 849 ^{1}H NMR (300 MHz, DMSO) 15.81 (broad S, 2H) 11.36 (s, 2H) 8.23 (m, 4H) 8.05 (d, *J* = 9.0 Hz, 2H) 7.99 (d, *J* = 8.1 Hz, 4H) 7.83 (d, *J* = 8.4 Hz, 4H) 7.47 (s, 2H) ^{13}C NMR (75 MHz, DMSO) 177.19, 166.57, 158.71, 146.04, 144.10, 142.28, 136.55, 130. 61, 129.39, 127.97, 127.11, 126.78, 121.45, 119.65, 119.02, 116.67

N,N′-bis-(2-(5-hydroxy-6-(azo-(4-(2-cinnaminic acid)phenyl))-7-naphthalenesulfonic acid))oxalamide ***(15)***. Negative ion ESMS found (MH-) 879, (MNa-) ^{1}H NMR (300 MHz, DMSO, note spectra is poorly resolved) 15.94 (s, 2H), 11.28 (s, 2H) 8.23 (s, 4H) 8.01 (d, *J* = 6 Hz, 2H) 7.78 (s, 4H) 7.58 (d, *J* = 15.9 Hz, 2H) 7.48 (s, 2H) 6.53 (d, *J* = 15.3 Hz, 2H) ^{13}C NMR (75 MHz, DMSO) 176.50, 167.69, 158.79, 144.14, 143.38, 142.22, 136.56, 131.69, 129.68, 129.02, 128.02, 126.83, 121.01, 119.71, 119.10, 118.96, 118.34, 117.68

N,N′-bis-(2-(5-hydroxy-6-(azo-(3-benzoic acid))-7-naphthalenesulfonic acid))oxalamide ***(19)***. Negative ion MALDIMS found (MH-) 827, (MNa-) 849 ^{1}H NMR (300 MHz, DMSO) 16.03 (s, 2H) 11.36 (s, 2H) 8.26 (m, 6H) 8.07 (m, 4H) 7.79 (d, *J* = 7.2 Hz, 2H) 7.60 (t, *J* = 7.8 Hz, 2H) 7.47 (s, 2H)) ^{13}C NMR (75 MHz, DMSO) 175.90, 166.83, 158.86, 144.26, 143.19, 142.15, 136.55, 132.16, 129.79, 128.80, 127.91, 126.75, 126.50, 121.36, 120.68, 119.70, 118.94, 118.36

N,N′-bis-(2-(5-hydroxy-6-(azo-(4-succinic acid))-7-naphthalenesulfonic acid))oxalamide (racemic mixture) ***(20)***. Negative ion MALDIMS found (MH-) 972, (MNa-) 993 ^{1}H NMR (300 MHz, DMSO) 16.2 (very broas S, 2H) 11.35 (s, 2H) 8.28 (d + s, 4H) 8.05 (d, *J* = 8.7 Hz, 2H) 7.75 (d, *J* = 8.1 Hz, 4H) 7.48 (s, 2H) 7.40 (d, *J* = 8.4 Hz, 4Hz) 3.95 (dd, *J* = 4.8, 9.6 Hz, 2H) 3.00 (dd, *J* = 10.2, 16.5, 2H) 2.59 (dd, *J* = 4.8, 16.8 Hz, 2H)) ^{13}C NMR (75 MHz, DMSO) 175.44, 174.31, 173.05, 159.18, 143.62, 142.28, 137.17, 136.59, 129.37, 128.62, 128.19, 127.30, 121.03, 120.38, 119.51, 118.26, 46.90, 37.74

Integrase activity and inhibition assays. Oligonucleotide-based assays for *in vitro* HIV-1 IN 3′-processing and strand transfer were performed essentially as described previously (32). The IN protein assayed was expressed in yeast and purified as described (*32*). Briefly, 15-μl reaction volumes included 20 mM morpholinepropanesulfonic acid (MOPS) pH 7.0, 5 mM DTT, 10 mM $MnCl_2$, 5 pmol of purified, full-length HIV-1 IN, and 0.5 pmol of radiolabeled, duplex *att* site oligonucleotide (U5-29+: 5′-TTTAGTCAGTGTGGAAAATCTCTAG**CA**GT-3′, and U5-29-: 5′-ACTGCTAGAGATTTTCCACACTGACTAAA-3′; the bold nucleotides indicate the invariant CA dinucleotide near the viral end). The U5-29+ was radiolabeled at its 5′-end to follow the fate of the duplex during a reaction (*32*). Approximately 20 mM NaCl is present as a "carry over" from the protein storage buffer. Inhibitors were preincubated for 10 min at 30°C in complete reaction mix lacking the *att* site oligonucleotide. The reaction was initiated by adding the oligonucleotide substrate, and the reaction was stopped after 20 min at 37°C by the adding gel loading dye (95% formamide) and heating to 100°C for 2 min. Reaction mixtures were electrophoresed through 7 M urea 15% acrylamide denaturing gels, the gels were dried on 3 mm paper, and the signal was quantified using a PhosphorImager (Molecular Dynamics).

Inhibitor concentrations were always performed in duplicate and positive controls (no inhibitor) were always performed in triplicate. Results presented are an average of two or more independent experiments.

Cell toxicity assays. Toxicity assays were performed using a slight modification to an assay frequently used to test for the toxicity of possible HIV-1 inhibitors (*33*). We changed from the standard assay substrate tetrazolium XTT (Polysciences) to the fluorescent substrate Calcein AM (Molecular Probes, Inc.) because many of our compounds are colored and therefore interfered with the absorbence characteristics of the standard substrate. CEM-SS cells, a human T-cell line, were grown in RPMI media supplemented with 10% fetal bovine serum, 100 units of penicillin G sodium and 0.1 mg of streptomycin sulfate per ml. Assays were performed using 100 μl of cells at 50,000/ml and inhibitor at final concentrations of 10, 100, 200, 500, 1000 μM. Cells were incubated in the presence of the compounds for 5 days. One hundred microliters of Calcein AM (Molecular Probes, Inc.) at 4 μM in phosphate-buffered saline was then added to each well. Calcein AM enters cells with intact membranes and is converted by intracellular esterase from a virtually nonfluorescent compound to an intensely fluorescent one (*34*). The plate was incubated at room temperature for 30–45 min and read with PerSeptive Biosystems' Fluorescence Multi-well Plate reader at 485 nm for excitation and 530 nm for emission, using the Cyto-Flour II software. Toxicity assays are performed in quadruplicate, and all data presented represent the average of two or more experiments.

Restriction enzyme inhibition assays. One unit of ApaL I (New England Biolabs), the manufacturer's recommended buffer, and inhibitor at either one- or fivefold the 3′-processing IC_{50} concentration were pre-incubated in 15 μl for 10 min at 30°C. Five microliters of plasmid DNA at 80 ng/μl was added to the reaction followed by a 90-min incubation at 37°C. The reaction was stopped by addition of 6X loading dye and the restriction digests were electrophoresed in a 0.8% agarose gel in TAE buffer. The DNA bands indicative of enzyme activity were quantified with the IS1000 Digital Imaging System (Alpha Innotech Corporation). The amount of enzyme and the time of the reaction were both within the linear range of enzyme activity for the reaction conditions.

RESULTS

Carbonyl J inhibits HIV-1 IN in vitro. Carbonyl J (**1**), the initial compound screened, demonstrated an IC_{50} value of 4 μM for 3′-processing and for strand transfer when tested for anti-IN activity using in vitro assays that employ purified HIV-1 IN and synthetic mimics of the HIV-1 *att* site (*32*). A series of Carbonyl J derivatives, largely compounds in which the hydroxyl or sulfonic acid groups were repositioned or removed, either lessened, or failed to improve, the anti-IN activity seen with Carbonyl J (data not shown).

Similarity search yields more potent Carbonyl J derivative. The failure to improve upon the anti-IN activity of Carbonyl J by derivatives described above led us to search the available chemical data base for compounds structurally similar to Carbonyl J. One such compound, Calcomine Orange (**2**), resembles Carbonyl J but contains two symmetrically attached phenyl-azo groups (Fig 1). Calcomine Orange demonstrated an IC_{50} of 350 nM for 3′-processing and 500 nM for strand transfer (Figs. 1 and 2).

Synthesis of Carbonyl J derivatives. Having identified Calcomine Orange as a potent lead compound for the development of IN inhibitors, we made a number of symmetrical

analogues by varying the appended aryl groups. Calcamine Orange can be seen as arising from the addition of an aryl diazonium salt to Carbonyl J, with the azo addition being directed by the phenol group in the adjacent (ortho) position (*35, 36*). While aryl diazonium salts are not generally commercially available, they are readily produced from a wide variety of commercially available anilines via reaction with nitrous acid (37,38).

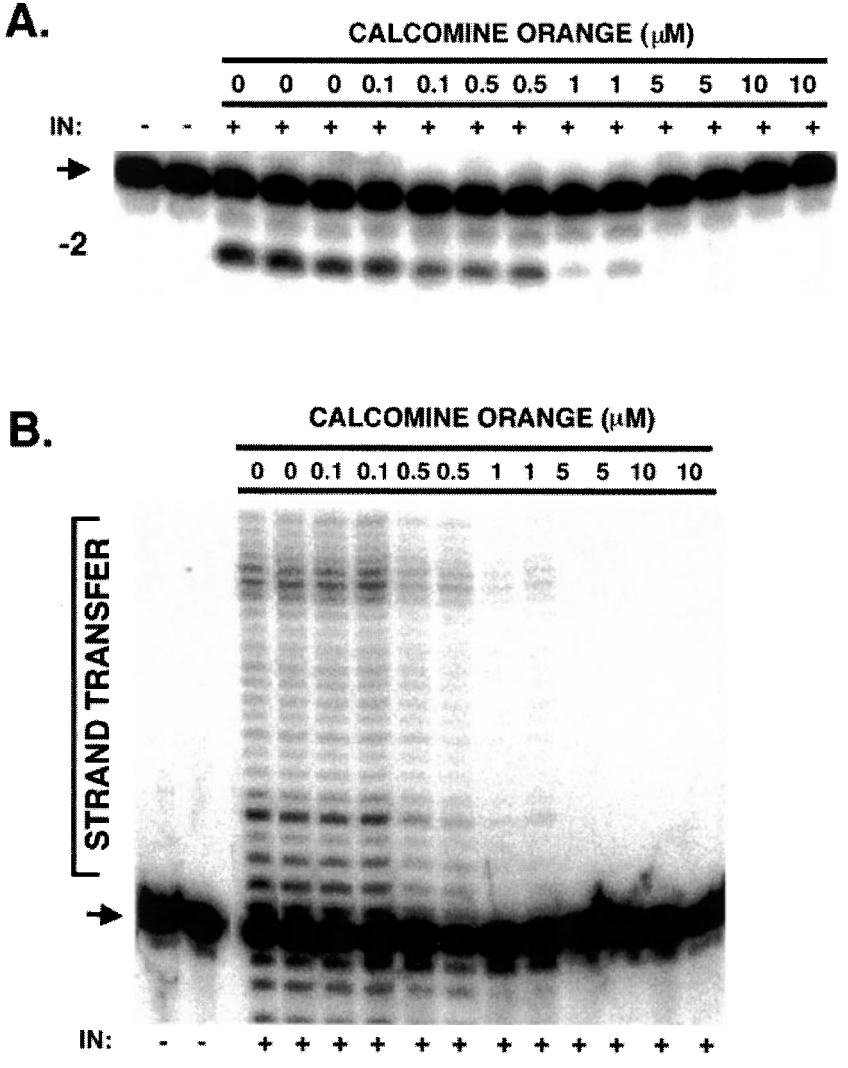


FIG. 2. Dose–response curve for Calcomine Orange inhibition of HIV-1 IN activity *in vitro.* The Calcomine Orange dose–response curve for 3′-processing (top) and strand transfer (bottom) represent the type of data obtained for all compounds described. Arrows indicate the starting material for each experiment, –2 indicates the products of 3′-processing that migrate faster than the starting material in A. The products of strand transfer migrate slower than the starting material and are indicated in B. Inhibitor concentration (μM) is indicated at the top of each panel.