

## $\beta$ -Diketo Acid Pharmacophore Hypothesis. 1. Discovery of a Novel Class of HIV-1 Integrase Inhibitors

Raveendra Dayam,<sup>†</sup> Tino Sanchez,<sup>†</sup> Omoshile Clement,<sup>‡</sup> Robert Shoemaker,<sup>§</sup> Shizuko Sei,<sup>||</sup> and Nouri Neamati<sup>\*†</sup>

Department of Pharmaceutical Sciences, School of Pharmacy, University of Southern California, 1985 Zonal Avenue, Los Angeles, California 90089, Accelrys, Inc., 9685 Scranton Rd., San Diego, California 92121, Laboratory of Antiviral Drug Mechanisms, SAIC-Frederick, Frederick, Maryland 21702, and Screening Technologies Branch, DTP, DCTD, National Cancer Institute at Frederick, Frederick, Maryland 21702

Received May 27, 2004

HIV-1 Integrase (IN) is an essential enzyme for viral replication. The discovery of  $\beta$ -diketo acids was crucial in the validation of IN as a legitimate target in drug discovery against HIV infection. In this study, we discovered a novel class of IN inhibitors using a 3D pharmacophore guided database search. We used S-1360 (**1**), the first IN inhibitor to undergo clinical trials, and three other analogues to develop a common feature pharmacophore hypothesis. Testing this four-featured pharmacophore against a multiconformational database of 150 000 structurally diverse small molecules yielded 1700 compounds that satisfied the 3D query. Subsequently, all 1700 compounds were docked into the active site of IN. On the basis of docking scores, Lipinski's rule-of-five, and structural novelty, 110 compounds were selected for biological screening. We found that compounds that contain both salicylic acid and a 2-thioxo-4-thiazolidinone (rhodanine) group (e.g. **5–13**) showed significant inhibitory potency against IN, while the presence of either salicylic acid or a rhodanine group alone did not. Although some of the compounds containing only a salicylic acid showed inhibitory potency against IN, none of the compounds containing only rhodanine exhibited considerable potency. Of the 52 compounds reported in this study, 11 compounds (**5**, **6**, **8**, **10–13**, **32–33**, **51**, and **53**) inhibited 3'-processing or strand transfer activities of IN with  $IC_{50} \leq 25 \mu\text{M}$ . This is the first reported use of S-1360 and its analogues as leads in developing a pharmacophore hypothesis for IN inhibition and for identification of new compounds with potent inhibition of this enzyme.

### Introduction

Combination therapy has remarkably improved the outcome of treatment against HIV-1 infection by decreasing the viral loads in infected patients to undetectable levels. However, eradication of the infection has not been achieved because of the persistence of latent HIV-1 in resting memory CD4<sup>+</sup> T-lymphoid cells. Moreover, emergence of viral strains resistant to current drugs necessitates the discovery of drugs with different mechanisms of action.<sup>1,2</sup> The addition of novel drug candidates to the existing combination therapy would certainly improve the outcome of the treatment. HIV-1 integrase (IN) has emerged as an attractive target for antiretroviral drugs because of its crucial role in the viral replication processes. IN catalyzes the integration of viral cDNA into the human genome, which facilitates stable viral replication and sustained infection.<sup>3</sup> Over the years several classes of small molecules such as hydroxylated aromatics and  $\beta$ -diketo acids<sup>4</sup> have been reported as IN inhibitors (for recent reviews see refs 5–7). Among all the IN inhibitors reported, most advanced in development is the  $\beta$ -diketo acid class of compounds,<sup>4,8</sup> of which S-1360 (**1**) was the first compound to enter clinical trials<sup>9</sup> (for a review see ref 10).

Studies with protease and reverse transcriptase inhibitors have proven that the emergence of drug resistance necessitates the development of second- and third-generation drugs. As more IN inhibitors are entering clinical trials, it is important to develop diverse chemical classes of selective inhibitors.

Once a bona fide drug against a known target has been identified, computational approaches, such as pharmacophore-based three-dimensional (3D) database searching, can play a key role in the discovery of novel leads with different chemical scaffolds. A pharmacophore refers to the 3D arrangement of various functional groups (chemical features) that is essential for the molecule to bind effectively against a specific enzyme, protein, or receptor.<sup>11</sup> The large number of successful applications of 3D pharmacophore-based searching in medicinal chemistry clearly demonstrates its utility in the modern drug discovery paradigm (for selected publications, see refs 12–19).

Previously, we have successfully employed the pharmacophore perception in the discovery of novel IN inhibitors. The first reported pharmacophore model developed using a known IN inhibitor, the caffeic acid phenethyl ester (CAPE), yielded several aromatic sulfones from the NCI database.<sup>20</sup> Subsequent three-point pharmacophore models based on depsides and depsidones identified several structurally diverse IN inhibitors.<sup>21–23</sup> Tetracyclines were also discovered using a four point pharmacophore model that was developed on the basis of the low-energy structures of chicoric acid and

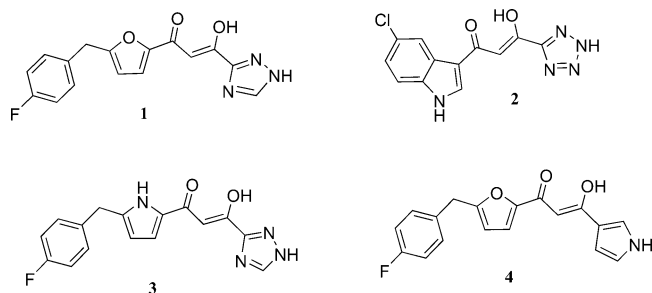
\* To whom correspondence should be addressed. Tel: +1 323-442-2341. Fax: +1 323-442-1390. E-mail: neamati@usc.edu.

<sup>†</sup> University of Southern California.

<sup>‡</sup> Accelrys, Inc.

<sup>§</sup> SAIC-Frederick.

<sup>||</sup> National Cancer Institute at Frederick.



**Figure 1.** Structures of the  $\beta$ -diketo acid bioisosteres and analogues used in the generation of common-feature pharmacophore hypotheses.

caffeoylquinic acids.<sup>24</sup> Dynamic pharmacophore models were generated using snapshots from the molecular dynamics trajectories of the core unit of IN in order to incorporate the dynamic nature of the active site region of IN into the pharmacophore model. The dynamic pharmacophore performed well in the identification of known IN inhibitors when compared to a static pharmacophore that was generated from a crystal structure conformation of the core unit of IN.<sup>25</sup> A recent report used the HypoGen module in Catalyst (Accelrys, Inc.) based on the biological activity of 26 compounds to develop a 3D pharmacophore model.<sup>26</sup>

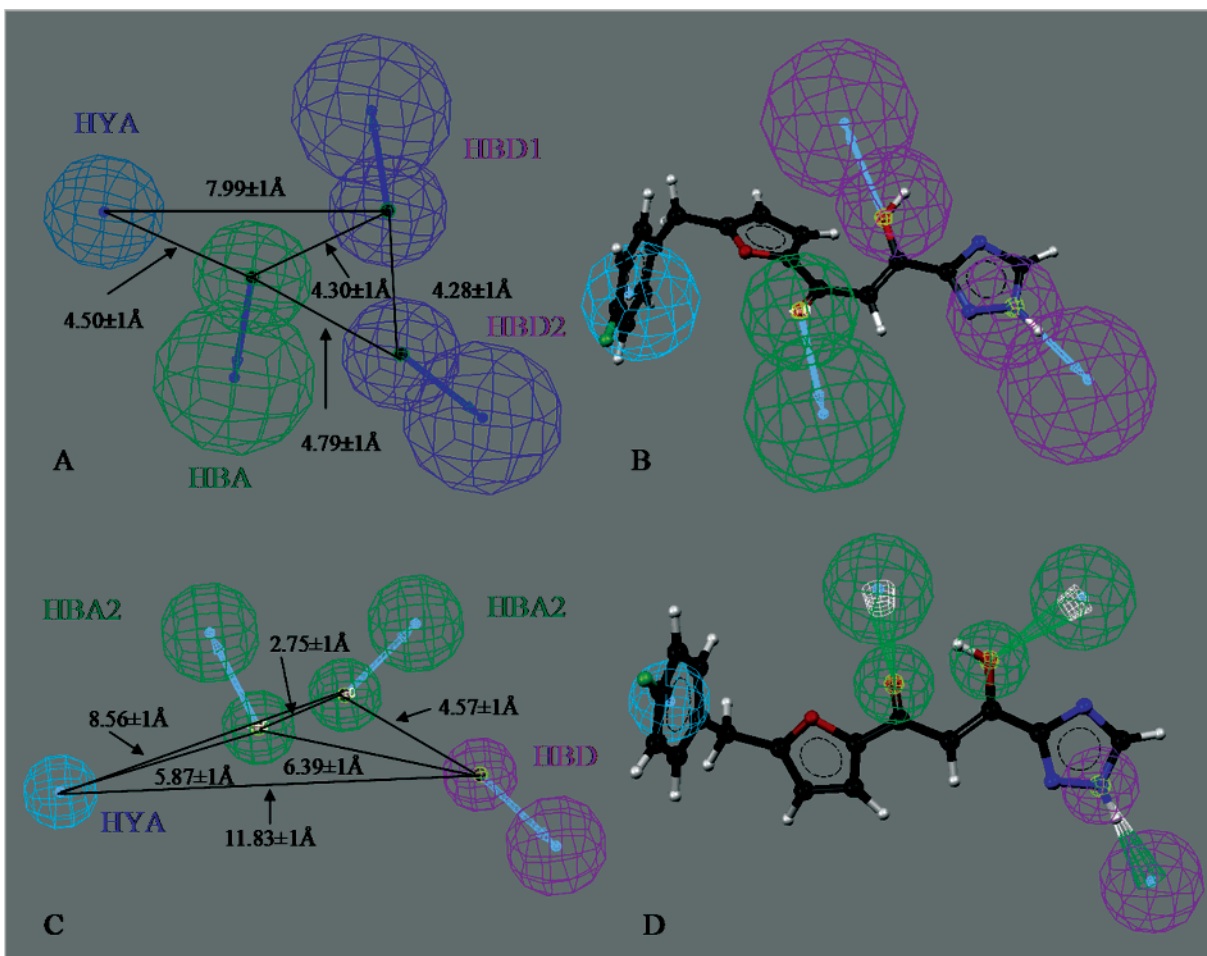
Earlier pharmacophore hypothesis studies were based on lead molecules that did not have proven efficacy in human clinical trials. As part of a study to develop second-generation IN inhibitors, we wanted to take advantage of already available compounds that are in clinical development, such as S-1360. In the present study, we will present the discovery of a novel class of 2-thioxo-4-thiazolidinone-(rhodanine-) and salicylic acid-containing IN inhibitors using a common feature pharmacophore model that was developed on the basis of the  $\beta$ -diketo acid analogues.

## Results and Discussion

**Pharmacophore Hypotheses.** We have generated common-feature pharmacophore hypotheses using a training set consisting of four bioisosteres (1–4) of  $\beta$ -diketo acid-containing inhibitors of IN (Figure 1). The parent compound 5CITEP belongs to the diketo acid-containing compounds, because the tetrazole group is a well-known bioisostere of a carboxylic acid.<sup>27</sup> Ten common-feature pharmacophore hypotheses were generated using HipHop. The HipHop algorithm in Catalyst generates common chemical feature based pharmacophore models (Accelrys, Inc., 2003, Catalyst version 4.8). We selected compounds that selectively inhibit strand transfer reaction of IN with comparable  $IC_{50}$  values and are expected to bind to a similar site on the active site of IN. S-1360 inhibits IN activity with an  $IC_{50}$  value of 20 nM. Compound 2 inhibits IN catalytic activity with an  $IC_{50}$  value of 2.3  $\mu$ M and compounds 3 and 4 inhibit purified IN with  $IC_{50}$  value of <1  $\mu$ M.<sup>6,28,29</sup> For the training set supplied to pharmacophore model generation experiment, it is crucial to include only compounds that exhibit similar activity profile as well as a similar binding mode. Inclusion of compounds with a lack of selectivity for IN or validation thereof can eschew the result of pharmacophore generation. We selected the highest ranked pharmacophore hypothesis (Hypo1) generated by HipHop for the database search (Figure

2A). Hypo1 has four features: a hydrophobic aromatic (AR), an H-bond acceptor (HBA) and two H-bond donors (HBD1 and HBD2) (Figure 2). The distance between AR and HBD1 and HBA were found to be  $7.99 \pm 1$  Å and  $4.5 \pm 1$  Å, respectively. The distance between HBD1 and HBD2 and HBA were found to be  $4.28 \pm 1$  and  $4.30 \pm 1$  Å, respectively. The distance between HBD2 and HBA was found to be  $4.79 \pm 1$  Å. Figure 2B shows the alignment of S-1360 against Hypo1. This alignment represents a good match of features present in the ligand to the pharmacophore model (Fit score = 3.98/4.0). The mapping of Hypo1 onto S-1360 was performed using the “BEST FIT” method in Catalyst. During the fitting process, conformations of S-1360 were calculated within the 20 kcal/mol energy threshold to minimize the distance between Hypo1 features and mapped atoms of S-1360. The possible reason for the orientation of “hydrogen” is formation of an intramolecular H-bond between OH and N(5) triazole (OH $\cdots$ N, 2.61 Å) in this particular conformation, which gave a maximum fit value of 3.98 with 10.47 kcal/mol of energy. A conformation of S-1360 that has “hydrogen” oriented in the H-bond donor feature direction is also observed, but it gave a fitness score of 2.89 with conformational energy of 14.64 kcal/mol (not shown). Hypo1 was then used as a search query to retrieve compounds that fit all the features of the query from a chemical database. A collection of ~150 000 compounds (ChemBridge, San Diego, CA) was converted into a Catalyst multiconformer 3D database. We also generated a different pharmacophore hypothesis (Hypo2), shown in Figure 2C,D, using a set of conformations similar to the conformation of 5CITEP in the IN–5CITEP complex crystal structure<sup>30</sup> and a single-crystal structure conformation of one of its analogues.<sup>31</sup> In this particular conformation, the furan, the keto–enol, and the triazole groups were in the same plane, and an intramolecular H-bond was observed between the hydroxyl of enol and the oxygen atom of the keto group. It is important to note that the mapped conformation (Hypo1) is different from the crystal structure conformation (Hypo2), and the importance of each conformation to the biologically relevant conformation on the active site of IN in the presence of  $Mg^{2+}$  and DNA will require additional data. Studies are in progress using Hypo2 to search our expanded database, and the results will be reported in due course.

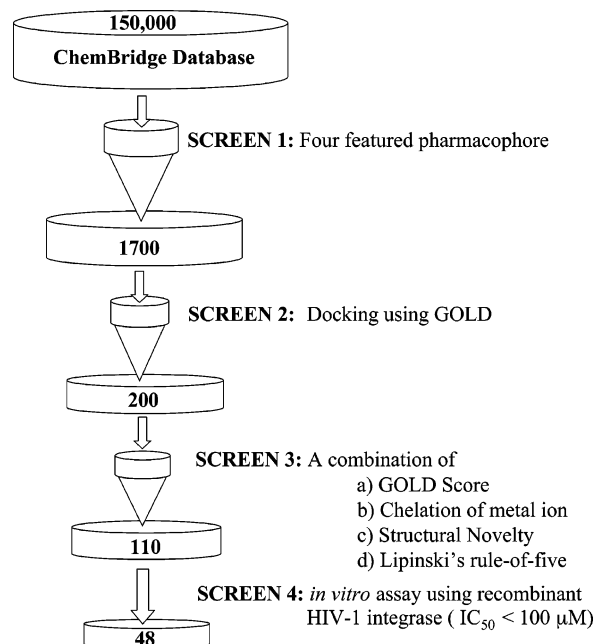
**Compound Selection.** The compound selection process that was implemented in the present study is shown as a flowchart in Figure 3. The database search using Hypo1 as the search query retrieved 1700 hits. All retrieved compounds from the Catalyst database search were then docked into the active site of IN using GOLD (CCDC 2001, GOLD v1.2). Docking was performed in order to rank the compounds on the basis of their ability to form favorable interactions within the active site of IN. On the basis of docking scores from GOLD, the top 200 compounds were selected for further analysis. The docked poses of the top 200 compounds inside the IN active site were visually examined. Finally, 110 compounds were chosen for in vitro assay against purified IN on the basis of the docking scores from GOLD, structural novelty, ability to chelate  $Mg^{2+}$  ion in the active site of IN, and satisfying Lipinski’s rule-



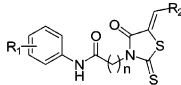
**Figure 2.** (A) The best ranked HipHop-generated four-featured pharmacophore hypothesis (Hypo1). Hypo1 was used in the database mining to identify novel integrase inhibitors. (B) S-1360 mapped onto Hypo1. (C) Hypo2 was generated using a conformation similar to the crystal structure conformation of 5-CITEP. (D) S-1360 mapped onto Hypo2. The orientation of the keto–enol group is the major difference between the two models. Both models were selected because the biologically relevant conformation on the active site of IN in the presence of  $Mg^{2+}$  and DNA is not currently known. Pharmacophore features are shown: hydrophobic aromatic (HYA), light blue; H-bond acceptor (HBA), green; H-bond donor (HBD), magenta.

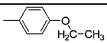
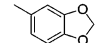
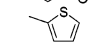
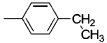
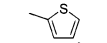
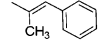
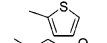
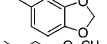
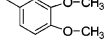
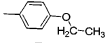
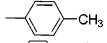
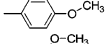
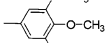
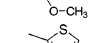
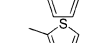
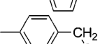
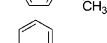
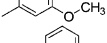
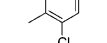
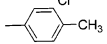
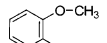
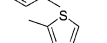
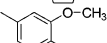
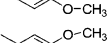
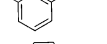
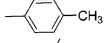
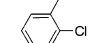
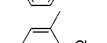
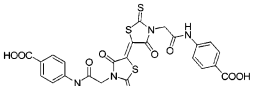
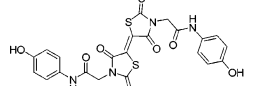
of-five.<sup>32,33</sup> Of the 110 compounds tested against purified IN, 48 compounds inhibited IN with  $IC_{50}$  values ranging from 7 to 100  $\mu M$ . Thirty structurally related compounds are presented in this report. The remaining structurally unrelated compounds will be reported elsewhere. To establish a coherent structure–activity relationship, 22 additional compounds were also tested on the basis of a 2D substructure search (see below).

**In Vitro Inhibitory Activities of a Novel Class of IN Inhibitors.** The structures and IN inhibitory activities of compounds 5–34 are given in Table 1. Twenty out of 30 compounds (66%) retrieved from the database using the highest ranked pharmacophore hypothesis, Hypo1, inhibited both 3'-processing and strand transfer activities of IN with  $IC_{50}$  values less than 100  $\mu M$ . In Figure 4A is shown one of the active compounds, 5, mapped onto Hypo1 to better understand the correlation between the pharmacophoric features of Hypo1 and the chemical features important for activity. All compounds possessed a five-membered rhodanine ring as a common structural unit linked to various aromatic groups via a three to five carbon aliphatic amide chain (Table 1). Active compounds 5–13 ( $IC_{50} < 44 \mu M$ ) contained a salicylic acid group in addition to



**Figure 3.** In silico screening protocol implemented for the discovery of integrase inhibitors.

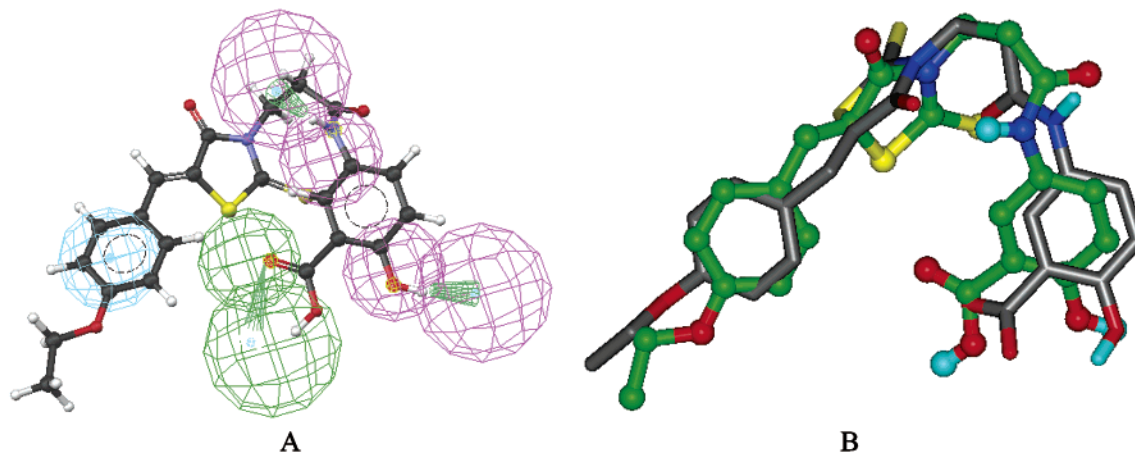
**Table 1.** Inhibition of HIV-1 Integrase Catalytic Activities and GOLD Scores of a Novel Class of Compounds Retrieved from the Database Using the Best Pharmacophore Hypo


Compound	R <sub>1</sub>	R <sub>2</sub>	n	Inhibition of IN Catalytic Activities		GOLD Score
				IC <sub>50</sub> (μM)		
				3'-Processing	Strand Transfer	
5	3-COOH, 4-OH		2	17±13	11±6	52.18
6	3-COOH, 4-OH		2	15±10	17±11	57.90
7	3-COOH, 4-OH		3	44±20	35±16	52.59
8	3-COOH, 4-OH		2	15±3	11±5	51.57
9	3-COOH, 4-OH		2	38±16	38±16	56.39
10	3-COOH, 4-OH		2	36±23	23±3	57.25
11	3-OH, 4-COOH		3	32±14	25±5	54.57
12	3-OH, 4-COOH		2	33±23	23±13	57.62
13	3-OH, 4-COOH		2	61±34	17±4	57.46
14	3-COOH		2	93±12	>100	51.98
15	3-COOH		2	100	>100	56.15
16	3-COOH		2	>100	>100	53.43
17	4-COOH		2	97±6	85±15	50.05
18	4-COOH		2	86±15	86±17	51.26
19	3-COOH		2	>100	>100	55.78
20	3-COOH		1	>100	>100	55.06
21	3-COOH		1	>100	>100	52.07
22	3-COOH		1	>100	>100	58.95
23	3-COOH		1	98±4	>100	56.04
24	4-OH		2	99±3	100	43.93
25	3-OH		2	>100	>100	47.81
26	3-OH		2	>100	90±10	53.28
27	2-OH, 4-NO <sub>2</sub>		1	>100	>100	49.84
28	2-OH, 5-NO <sub>2</sub>		1	84±24	80±18	47.49
29	2-OH, 4-NO <sub>2</sub>		1	>100	>100	46.44
30	2-OH, 5-NO <sub>2</sub>		1	20±7	18±11	49.77
31	2-OH, 5-Cl		1	32±11	17±2	46.78
32	2-OH, 4-NO <sub>2</sub>		1	>100	>100	51.99
33				59±31	75	58.64
34				59±37	69	51.30

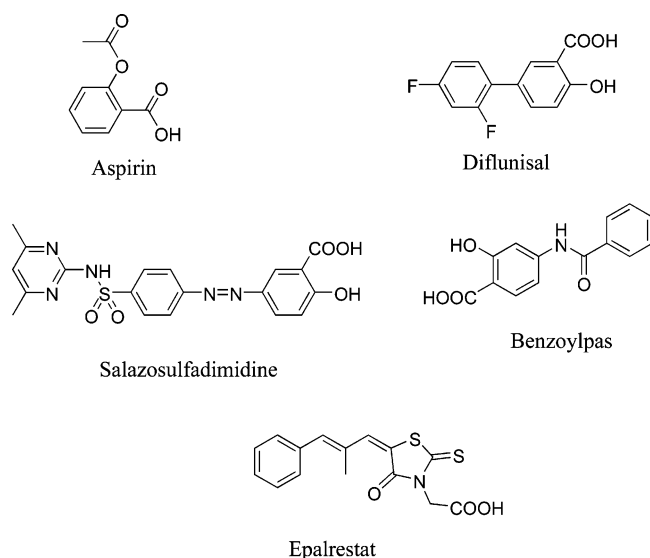
the rhodanine ring. Compounds **14–23** with a 3- or 4-benzoic acid group in place of a salicylic acid group showed moderate activity or no activity at the highest tested concentration. Compounds **30** and **31** with a nitro phenol in place of the salicylic acid group were active, while compounds **24–26** with a phenol were inactive. The dimeric compounds **33** and **34** with two rhodanine rings and a 4-carboxybenzoic acid (**33**) or a phenol (**34**) in place of the salicylic acid group of the active compounds inhibited 3'-processing of IN with IC<sub>50</sub> values of 59 and 59 μM and strand transfer activities with IC<sub>50</sub> values of 75 and 69 μM, respectively. Compounds containing a variety of substitutions (R<sub>2</sub> in Table 1) on the fourth position of the rhodanine ring were retrieved using Hypo1. It appears that substituents on the fourth position of the rhodanine ring have considerable impact on activity. For example, compounds **5**, **6**, **8**, and **11** with substituents such as 4-ethoxyphenyl, phenyl-3,4-dioxymethylene, 4-ethylphenyl, and 2-thiophene were very potent. All compounds with a salicylic acid group were active, while some of the compounds with 3- or 4-benzoic acid showed moderate activity. Compounds with 5-nitro- or 5-chlorophenol showed high potency, suggesting that electron withdrawing groups on the 3- and 4- or 2- and 5-positions on the aromatic ring (R<sub>1</sub> in Table 1) along with certain substituents at R<sub>2</sub> may be important for inhibitory potency against IN.

The most potent compounds had both the rhodanine and the salicylic acid groups. Because the rhodanine ring and salicylic acid moiety are structural components of several biologically active molecules, we thought it would be interesting to develop them as IN inhibitors. Previously, many compounds with the rhodanine scaffold were reported as antimicrobial, antifungal, insecticidal, antimalarial, antidiabetic, hepatitis C virus protease inhibiting, and antitumor agents.<sup>34–36</sup> A rhodanine-containing aldose reductase inhibitor, epalrestat, has been used in the treatment of diabetic neuropathy (Figure 5).<sup>37,38</sup> An example of rhodanine-ring-containing compounds previously reported to have IN inhibitory activity contained a rhodanine ring attached to a phenylsulfonamide moiety. This compound showed strong inhibitory activity against IN.<sup>20</sup> Like the rhodanine moiety, the salicylic acid group is also known for its druglike characteristics. For example, aspirin as well as several known drugs have salicylic acid as an essential structural feature. Compounds containing the salicylic acid substructure were reported as analgesic, antiinflammatory, antibacterial, and antituberculosis agents. Some of the therapeutic agents that have salicylic acid or rhodanine group are shown in Figure 5.

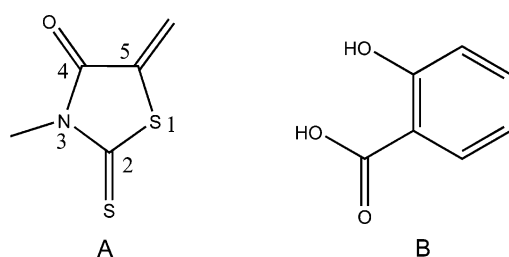
To evaluate the influence of the salicylic acid group and the rhodanine ring together or independently on IN inhibitory activity, two sets of compounds were retrieved from the database that had either a salicylic group or a rhodanine ring using the two substructure queries shown in Figure 6. The structures and IN inhibitory activities of the rhodanine-ring-containing compounds are given in Table 2, while the data for the salicylic acid-containing compounds are given in Table 3. The 2D database search using two substructures was carried out mainly to achieve two goals: (i) To evaluate the influence of the salicylic acid group and the rhoda-



**Figure 4.** (A) Compounds **5** mapped on to Hypo1. The pharmacophore features are shown: hydrophobic aromatic, light blue; H-bond acceptor, green; H-bond donor, magenta. (B) Both conformations of compound **5**; the conformation that was mapped onto Hypo1 (green ball-and-stick model) and the bound conformation inside the IN active site (gray stick model) are superimposed. The orientation of the rhodanine group was different in two conformations, but the positions of atoms that mapped key pharmacophoric features of Hypo1 were similar.



**Figure 5.** Therapeutic agents containing the salicylic acid or the rhodanine moiety.



**Figure 6.** Substructures A (A) and B (B) used to retrieve compounds from the database.

nine ring together or independently on IN inhibitory activity, as these two drug fragments were the main structural components of the active compounds, and (ii) to generate more data for a more coherent structure–activity relationship and further optimization processes. The rhodanine-ring-containing compounds **35–45**, with various aromatic, nonaromatic, and heterocyclic substitutions on the N of the rhodanine ring, showed only modest IN inhibitory activity. Compounds **36**, **38–40**, **42**, and **44** inhibited IN with  $IC_{50}$  values less than 100

$\mu M$ , and the remaining compounds were inactive. Among all the rhodanine-ring-containing compounds, **42**, with  $IC_{50}$  values 52 and 27  $\mu M$  for 3'-processing and strand transfer, respectively, was the most potent. On the other hand, **43**, with a relatively small substituent (phenyl) on the N atom of the rhodanine ring (replaces large substituents in all the other compounds), was inactive. The rhodanine-containing compounds **36**, **40**, **42**, and **44** fit Hypo1 and did show IN inhibitory activity. Remaining compounds **35**, **37–39**, **41**, **43**, and **45** do not fit Hypo1. It appears that the rhodanine ring is an important structural unit in this series of compounds, and its presence would lead to increased IN inhibitory potency. However, the inactive compounds of this series demonstrate that the presence of the rhodanine moiety alone does not ensure potency.

The salicylic acid-containing compounds **46–56** were retrieved using the substructure query B (Figure 6B). Biological testing of these compounds against the IN assay showed that compounds **51** and **53** were the most potent and inhibited both 3'-processing and strand transfer activities of IN with  $IC_{50}$  values of 13, 16  $\mu M$  and 12, 8  $\mu M$ , respectively. Some of the salicylic acid-containing compounds, **46**, **48**, **55**, and **56**, poorly fit (fitness score < 1) the Hypo1 pharmacophore model, while compounds **47** and **49–54** do not fit Hypo1. Despite the fact that these compounds do not fit Hypo1, salicylic acid-containing compounds **51** and **53** showed good IN inhibitory activity comparable to that of the active compounds. Compounds **49**, **50**, **52**, **54**, and **56** were moderately active ( $IC_{50} = 32–75 \mu M$ ), while compounds **46–48** and **55** were inactive ( $IC_{50} > 100 \mu M$ ). It is clear that the salicylic acid group is therefore important for IN inhibitory activity. Careful analysis of the activity profiles of these three sets of compounds demonstrates that the therapeutically known rhodanine and salicylic acid groups linked together exhibit the desired potency against IN. The database search using Hypo1 was confined to a collection of ~150 000 compounds and did not yield more potent inhibitors than the parent compounds. Most of the active compounds identified in this study are structurally novel and amenable to optimization. However, studies are under-

**Table 2.** Inhibition of HIV-1 Integrase Catalytic Activities and GOLD Scores of Rhodanine-Containing Compounds Retrieved from the Database Using Substructure A

Compound	Structure	Inhibition of IN Catalytic Activities IC <sub>50</sub> (μM)		GOLD Score
		3'-Processing	Strand Transfer	
35		>100	>100	46.03
36		83±21	44±16	50.69
37		>100	>100	50.89
38		98±5	100	54.21
39		97±6	93±12	56.55
40		88±22	84±2	59.21
41		>100	>100	46.41
42		52±11	27±6	47.27
43		>100	>100	36.51
44		70±27	38±14	46.41
45		>100	100	50.30

way to apply this pharmacophore to search a larger database, which is expected to yield more promising results. It is also important to bear in mind that S-1360 is a clinical candidate and it was optimized in a systematic way by the scientists at Shionogi & Co. Our primary aim was to identify structurally diverse leads for IN using the Hypo1. Further optimization of these leads is expected to give more potent IN inhibitors.

**Antiviral Activity in Cell Culture.** Thirteen compounds (**5–12**, **30–31**, **51**, and **53**) were selected for determination of their antiviral activity in HIV-1 infected CEM cells. Compounds **9** and **53** with 50% effective concentrations (EC<sub>50</sub>) of 21 ± 5 and 59 ± 5, respectively and 50% cytotoxic concentrations (CC<sub>50</sub>) of 48 ± 6 and 133 ± 10 μM, respectively, were the most potent. A maximum protection of 55–76% was achieved at 20 μM of **53**. Rhodanine-containing compounds **7** and

**Table 3.** Inhibition of HIV-1 Integrase Catalytic Activities and GOLD Scores of Salicylic Acid-Containing Compounds Retrieved from the Database Using Substructure B

Compound	R	Inhibition of IN Catalytic Activities IC <sub>50</sub> (μM)		Gold Score
		3'-Processing	Strand Transfer	
46		>100	>100	44.21
47		100	100	48.80
48		>100	>100	43.25
49		40±10	43±6	54.73
50		61±10	32±1	50.37
51		13±6	12±8	54.07
52		66±6	44±4	41.98
53		16±5	8±1	52.06
54		96±7	73±24	36.46
55		>100	>100	47.78
56		78±39	75±43	49.36

**12** yielded a maximum protection of 22–36% at 63 μM and of 18–21% at 200 μM, respectively. Other compounds tested did not show a significant antiviral activity. Studies are in progress to optimize the potency of the active compounds.

**Druglike Properties of Active Compounds.** Oral bioavailability is a desirable property of compounds under investigation in the drug discovery process. Lipinski's rule-of-five is a simple model to forecast the absorption and intestinal permeability of a compound.<sup>32,33</sup> In the rule-of-five model, compounds are considered likely to be well-absorbed when they possess ClogP < 5, molecular weight < 500, number of H-bond donors < 5, and number H-bond acceptors < 10. As shown in Table 4, some of the active compounds satisfy these requirements. Modification of the rule-of-five based on the number of rotatable bond and polar surface area was implemented within the Accord software suite (Accelrys, Inc.) as described.<sup>39–41</sup> The calculated atom-based log P (AlogP98) of some of these active compounds ranges from 2.3 to 3.5, and H-bond donor and acceptor counts

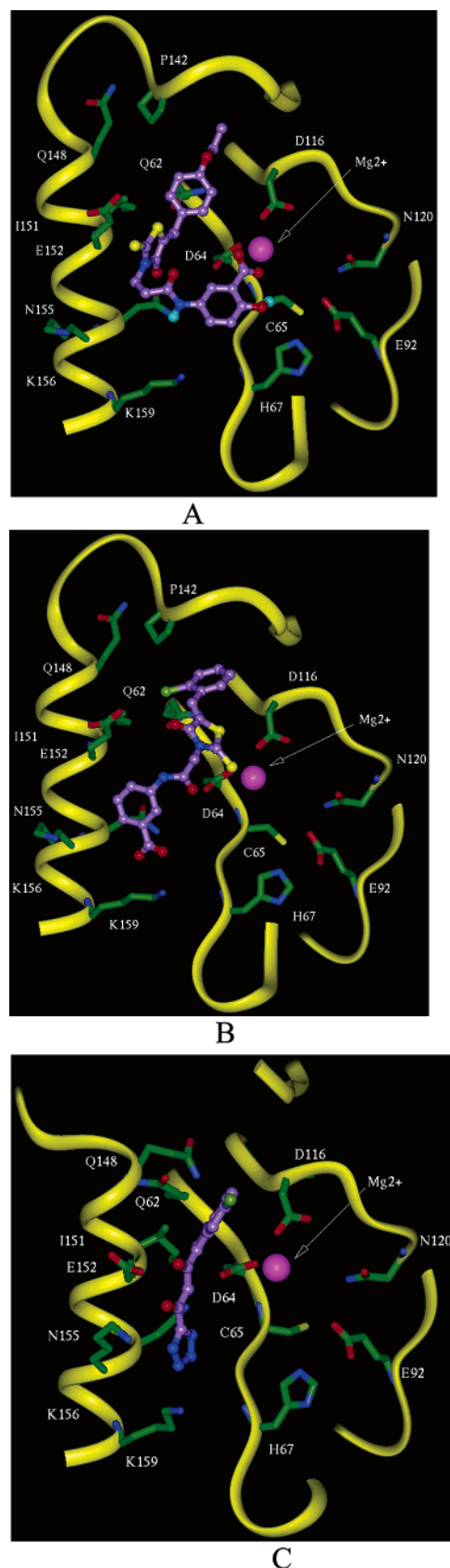
**Table 4.** Druglike Physicochemical Properties of Some of the Active Compounds

compd	MW <sup>a</sup>	ALogP98 <sup>b</sup>	HBA <sup>c</sup>	HBD <sup>d</sup>	R <sub>bonds</sub> <sup>e</sup>	fPSA <sup>f</sup>
5	472.52	2.92	8	3	9	115.11
6	472.49	2.36	9	3	7	124.04
7	448.51	2.53	7	3	8	106.18
8	456.52	3.53	7	3	8	106.18
9	434.49	2.21	7	3	7	106.18
10	468.53	3.51	7	3	8	106.18
11	448.52	2.53	7	3	8	106.18
12	472.48	2.36	9	3	7	124.04
13	488.52	2.55	9	3	9	124.04

<sup>a</sup> Molecular weight. <sup>b</sup> Calculated atom-based log *P*. <sup>c</sup> Number of H-bond acceptors. <sup>d</sup> Number of H-bond donors. <sup>e</sup> Number of rotatable bonds. <sup>f</sup> Fast polar surface area (Å<sup>2</sup>).

are <5 and <10, respectively.<sup>40,41</sup> The calculated fast polar surface area (fPSA) of the compounds spans from 106 to 124 Å<sup>2</sup>, which indicates that these compounds would be well-absorbed in the human intestine, as absorption problems arise when the calculated fPSA value of a compound is >140 Å<sup>2</sup> (Table 4). Therefore, these compounds are potentially interesting for optimization, and studies are underway to design and synthesize more potent analogues.

**Docking Studies.** As part of the compound selection process, all the compounds 5–56 were docked into the active site of IN and the bound conformations inside the active site of IN were visually examined. Although a significant correlation between GOLD scores and inhibitory activities of the compounds was not routinely observed, in general, a trend was seen in which many of the active compounds scored high. The bound conformation of one of the highly active compounds, 5, inside the active site of IN is shown in Figure 7A, while the bound conformation of an inactive compound, 29, is shown in Figure 7B. The salicylic acid group in 5 occupied an area surrounded by amino acid residues D64, C65, N155, and K159, while the rhodanine moiety of the compound was placed between D64 and E152. The 4-ethoxyphenyl group of 5 was placed near Q62, D116, I151, and P142. The oxygen atoms of the carboxylate group of 5 formed coordinate bonds with Mg<sup>2+</sup> ion. Three H-bond interactions were observed between the hydroxyl and the carboxylate groups of 5 and the active site amino acid residues C65 and H67 (Table 5). As a comparison, the bound conformation of compound 5 (Figure 7A) and the conformation that was mapped onto Hypo1 (Figure 4A) were superimposed as shown in Figure 4B. It is important to note that although the orientation of the rhodanine group is different in these two conformations, positions of atoms that are mapped by the key pharmacophoric features of Hypo1 are similar (Figure 4B). Like 5, compound 29 also bound to a similar area of the IN active site, but it adopted a somewhat different conformation within the active site. The phenol moiety of 29 occupied a space surrounded by N155, K156, and K159. An H-bond interaction was observed between the hydroxyl oxygen of 29 and  $\epsilon$ -NH of K159 (HO...HN K159). The rhodanine moiety of 29 occupied an area between D64 and E152, while the thiophene group was placed near Q62 and P142. The predicted binding area of 5 and 29 inside the IN active site (Table 5) was similar to the binding area of 5CITEP (Table 5) and the binding area of 5CITEP in IN–5CITEP complex crystal structure (Figure 7C). In its



**Figure 7.** The bound conformations of compounds 5 (A) and 29 (B), and the crystallographic conformation of 2 (C) inside the IN active site (PDB 1QS4). The violet ball-and-stick models represent 5, 29, and 2, while the yellow ribbon model represents backbone of the IN active site region. The prominent amino acid residues are rendered as stick models. The magenta sphere represents Mg<sup>2+</sup>.

**Table 5.** H-Bond Interactions and the Active Site Amino Acid Residues Interacting with Compounds **5** and **29**

compd	H-bonding interactions (Å)	interacting amino acid residues
<b>5</b> <sup>a</sup>	HO $\cdots$ HN H67 (2.97) <sup>b</sup> OH $\cdots$ O=C C65 (2.22) <sup>c</sup> C=O $\cdots$ HN C65 (2.89) <sup>c</sup>	Q62, D64, C65, T66, H67, D116, P142, I151, E152, N155, K159
<b>29</b>	HO $\cdots$ HN K159 (2.19)	Q62, D64, P142, I151, E152, N155, K156, K159

<sup>a</sup> The carboxylate oxygen atoms of **5** coordinate with active site Mg<sup>2+</sup> ion; CO $\cdots$ Mg<sup>2+</sup>(2.12 Å), CO $\cdots$ Mg<sup>2+</sup>(1.65 Å). <sup>a</sup> Side chain. <sup>b</sup> Backbone.

bound conformation inside the IN active site, **5** established strong interactions with catalytically vital amino acid residues D64 and E152 and Mg<sup>2+</sup> ion, while the inactive compound **29** did not form such interactions inside the active site of IN. The bound conformation of **5** revealed that the salicylic acid group is involved in strong interactions with amino acid residue D64 and Mg<sup>2+</sup> ion, while the rhodanine moiety interacted with E152. These specific interactions of the salicylic acid group and the rhodanine moiety with catalytically important amino acid residues inside IN active site were responsible for its inhibitory activity against IN.

## Conclusions

A novel class of IN inhibitors was successfully discovered using a pharmacophore hypothesis derived from a template of the clinical drug S-1360 and its analogues. The potent compounds identified via this approach contained two therapeutically known drug fragments, a salicylic acid and a rhodanine ring, as major structural components, which make these compounds promising for drug development targeting IN. Further structure-based optimization is in progress to enhance their potency and antiviral activities.

## Experimental Section

**Generation of Pharmacophore Hypotheses.** Common-feature pharmacophore hypotheses were generated using four bioisosteres of  $\beta$ -diketo acid-containing inhibitors (**1–4**) of IN (Figure 1). The structures and conformations of the four compounds were built within Catalyst on a multiprocessor Linux PC in parallel to a 24-cpu Silicon Graphics Onyx workstation as described.<sup>31,42</sup> The Poling algorithm implemented within Catalyst was used to generate conformations for all the compounds.<sup>43–45</sup> For each compound, all feasible unique conformations were generated over a 20 kcal/mol range using the BEST flexible conformation generation option in Catalyst. The Catalyst HipHop module was used to generate the common feature hypotheses. HipHop evaluates a collection of conformational models of molecules and a selection of chemical features and identifies configurations of features (pharmacophore) common to these molecules. The top-ranking pharmacophores are expected to identify the hypothetical orientation of the active compounds and the common binding features interacting with the target. All training set inhibitors were somewhat structurally diverse but possessed some common chemical features, and comparable inhibitory potencies. On the basis of the structural information from these known inhibitors and the active site of IN, a set of features was selected to be present in the pharmacophore-generation experiment. The chemical features considered in the pharmacophore model generation run were H-bond donor (HBD), H-bond acceptor (HBA), and aromatic ring center (AR). HipHop was forced to incorporate these features in the 10 top-ranking pharmacophore hypotheses generated.

**Database Search.** The highest ranked common-feature pharmacophore (Figure 2A) was used as a search query to retrieve molecules with novel chemical structure and desired chemical features from an in-house multiconformer Catalyst-formatted database consisting of ~150 000 compounds. The Best Flexible Search Databases/Spread Sheets method in Catalyst was used to search the database. Druglike properties of the retrieved hits from the database search were calculated using Accord for Excel (Accelrys, Inc.).

**Docking.** The subunit B of the core domain X-ray structure of IN (PDB 1BIS) in which all the active site amino acid residues were resolved was chosen for our docking purpose.<sup>46</sup> A Mg<sup>2+</sup> ion was placed in the active site between carboxylate oxygen atoms of amino acid residues D64 and D116 considering the geometry of the Mg<sup>2+</sup> ion that was present in the subunit A of IN in PDB 1BIS and subunit A in IN–5CITEP complex X-ray structure (PDB 1SQ4).<sup>30</sup> All the water molecules present in protein were removed and proper protonation states were assigned for acidic and basic residues of the protein. Docking was performed using version 1.2 of the GOLD (Genetic Optimization for Ligand Docking) software package.<sup>47–49</sup> GOLD is an automated ligand-docking program that uses a genetic algorithm to explore the full range of ligand conformational flexibility with partial flexibility of the receptor. The algorithm was tested on a dataset of over 300 complexes extracted from the Brookhaven Protein DataBank. GOLD succeeded in more than 70% cases in reproducing the experimental bound conformation of the ligand.<sup>47</sup> GOLD requires a user defined binding site. It searches for a cavity within the defined area and considers all the solvent accessible atoms in the defined area as active site atoms. A 20 Å radius active site was defined considering the carboxylate oxygen (OD1) atom of residue D64 as the center of the active site. All the compounds retrieved by the pharmacophore model (Hypo1) were docked into the active site of IN. On the basis of the GOLD fitness score, for each molecule a bound conformation with high fitness score was considered as the best bound conformation. The fitness function that was implemented in GOLD consisted basically of H-bonding, complex energy, and ligand internal energy terms. A population of possible docked orientations of the ligand is set up at random. Each member of the population is encoded as a chromosome, which contains information about the mapping of ligand H-bond atoms onto (complementary) protein H-bond atoms, mapping of hydrophobic points on the ligand onto protein hydrophobic points, and the conformation around flexible ligand bonds and protein OH groups. A number of parameters control the precise operation of the genetic algorithm, viz, population size, selection pressure, number of operations, number of islands, niche size, and operator weights, such as migrate, mutate, and crossover. All docking runs were carried out using standard default settings with a population size of 100, a selection pressure of 1.1, a maximum of 100 000 operations, number of islands as 5, a niche size of 2, and a mutation and crossover rate of 95.

**Materials, Chemicals, and Enzymes.** All compounds were dissolved in DMSO and the stock solutions were stored at –20 °C. The  $\gamma$ -[<sup>32</sup>P]ATP was purchased from either Amersham Biosciences or ICN. The expression systems for the wild-type IN and soluble mutant IN<sup>F185KC280S</sup> were generous gifts of Dr. Robert Craigie, Laboratory of Molecular Biology, NIDDK, NIH, Bethesda, MD.

**Preparation of Oligonucleotide Substrates.** The oligonucleotides 21top (5'-GTGTGGAAAATCTCTAGCAGT-3') and 21bot (5'-ACTGCTAGAGATTTTCCACAC-3') were purchased from Norris Cancer Center Microsequencing Core Facility (University of Southern California) and purified by UV shadowing on polyacrylamide gel. To analyze the extent of 3'-processing and strand transfer using 5'-end-labeled substrates, 21top was 5'-end labeled using T<sub>4</sub> polynucleotide kinase (Epicenter, Madison, WI) and  $\gamma$ -[<sup>32</sup>P]ATP (Amersham Biosciences or ICN). The kinase was heat-inactivated and 21bot was added in 1.5 molar excess. The mixture was heated at 95 °C, allowed to cool slowly to room temperature, and run



through a spin 25 minicolumn (USA Scientific) to separate annealed double-stranded oligonucleotide from unincorporated material.

**Integrase Assays.** To determine the extent of 3'-processing and strand transfer, wild-type IN was preincubated at a final concentration of 200 nM with the inhibitor in reaction buffer [50 mM NaCl, 1 mM HEPES, pH 7.5, 50  $\mu$ M EDTA, 50  $\mu$ M dithiothreitol, 10% glycerol (w/v), 7.5 mM MnCl<sub>2</sub>, 0.1 mg/mL bovine serum albumin, 10 mM 2-mercaptoethanol, 10% dimethyl sulfoxide, and 25 mM MOPS, pH 7.2] at 30 °C for 30 min. Then, 20 nM of the 5'-end <sup>32</sup>P-labeled linear oligonucleotide substrate was added, and incubation was continued for an additional 1 h. Reactions were quenched by the addition of an equal volume (16  $\mu$ L) of loading dye (98% deionized formamide, 10 mM EDTA, 0.025% xylene cyanol, and 0.025% bromophenol blue). An aliquot (5  $\mu$ L) was electrophoresed on a denaturing 20% polyacrylamide gel (0.09 M Tris-borate, pH 8.3, 2 mM EDTA, 20% acrylamide, 8 M urea).

Gels were dried, exposed in a PhosphorImager cassette, analyzed using a Typhoon 8610 Variable Mode Imager (Amersham Biosciences), and quantitated using ImageQuant 5.2. Percent inhibition (% I) was calculated using the following equation:

$$\% I = 100 \times [1 - (D - C)/(N - C)]$$

where *C*, *N*, and *D* are the fractions of 21-mer substrate converted to 19-mer (3'-processing product) or strand transfer products for DNA alone, DNA plus IN, and IN plus drug, respectively. The IC<sub>50</sub> values were determined by plotting the logarithm of drug concentration versus percent inhibition to obtain the concentration that produced 50% inhibition.

**Anti-HIV Assays in Cultured Cells.** The anti-HIV activity was evaluated in human T cell line CEM-SS infected with HIV-1 as described by Weislow et al.<sup>50</sup> In brief, cells were plated in 96-well plates at 5 × 10<sup>3</sup> cells/well and infected with HIV-1<sub>RF</sub> (MOI = 0.3). Serial dilutions of compounds were then immediately added to the cells in a final volume of 200  $\mu$ L. In each experiment, AZT and dextran sulfate were included as control compounds for anti-HIV activity. The cells were maintained at 37 °C with 5% CO<sub>2</sub>-containing humidified air for 6 days. Cell viability was quantified by absorbance at 450 nm after 4 h incubation with 2,3-bis[2-methoxy-4-nitro-5-sulphophenyl]-5-[(phenylamino)carbonyl]-2H-tetrazolium hydroxide (XTT) at 0.2 mg/mL. Antiviral activity was graded on the basis of the degree of anti-HIV protection as active (80–100% protection), moderate (50–79% protection), and inactive (0–49% protection). Toxicity of the compounds was determined simultaneously on the same plate in uninfected CEM-SS cells.

**Acknowledgment.** This work was supported by funds from the GlaxoSmithKline Drug Discovery Award to N.N.

## References

- Richman, D. D. HIV chemotherapy. *Nature* **2001**, *410*, 995–1001.
- De Clercq, E. Strategies in the design of antiviral drugs. *Nat. Rev. Drug Discovery* **2002**, *1*, 13–25.
- Brown, P. O. Integration. In *Retroviruses*; Coffin, J. C., Hughes, S. H., Varmus, H. E., Eds.; Cold Spring Harbor Press: Plainview, NY, 1999.
- Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J. A.; Espeseth, A.; Gabryelski, L.; Schleif, W.; Blau, C.; Miller, M. D. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science* **2000**, *287*, 646–650.
- Neamati, N. Structure-based HIV-1 integrase inhibitor design: A future perspective. *Expert Opin. Invest. Drugs* **2001**, *10*, 281–296.
- Neamati, N. Patented small molecule inhibitors of HIV-1 integrase: A ten-year saga. *Expert Opin. Ther. Pat.* **2002**, *12*, 709–724.
- Dayam, R.; Neamati, N. Small-Molecule HIV-1 Integrase Inhibitors: The 2001–2002 Update. *Curr. Pharm. Des.* **2003**, *9*, 1789–1802.
- Hazuda, D. J.; Young, S. D.; Guare, J. P.; Anthony, N. J.; Gomez, R. P.; Wai, J. S.; Vacca, J. P.; Handt, L.; Motzel, S. L.; Klein, H. J.; Dornadula, G.; Danovich, R. M.; Witmer, M. V.; Wilson, K. A.; Tussey, L.; Schleif, W. A.; Gabryelski, L. S.; Jin, L.; Miller, M. D.; Casimiro, D. R.; Emini, E. A.; Shiver, J. W. Integrase inhibitors and cellular immunity suppress retroviral replication in rhesus macaques. *Science* **2004**, *305*, 528–532.
- Yoshinaga, T.; Sato, A.; Fujishita, T.; Fujiwara, T. In Vitro Activity of a New HIV-1 Integrase Inhibitor in Clinical Development. 9th Conference on retroviruses and opportunistic infections, Seattle, WA, 2002.
- Billich, A. S-1360 Shionogi-GlaxoSmithKline. *Curr. Opin. Invest. Drugs* **2003**, *4*, 206–209.
- Guner, O. F. *Pharmacophore perception, development, and use in drug design*; International University Line: La Jolla, CA, 2000.
- Patel, Y.; Gillet, V. J.; Bravi, G.; Leach, A. R. A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. *J. Comput. Aided Mol. Des.* **2002**, *16*, 653–681.
- Kurogi, Y.; Guner, O. F. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. *Curr. Med. Chem.* **2001**, *8*, 1035–1055.
- Lyne, P. D.; Kenny, P. W.; Cosgrove, D. A.; Deng, C.; Zabludoff, S.; Wendoloski, J. J.; Ashwell, S. Identification of compounds with nanomolar binding affinity for checkpoint kinase-1 using knowledge-based virtual screening. *J. Med. Chem.* **2004**, *47*, 1962–1968.
- Lengauer, T.; Lemmen, C.; Rarey, M.; Zimmermann, M. Novel technologies for virtual screening. *Drug Discovery Today* **2004**, *9*, 27–34.
- Dror, O.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J. Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. *Curr. Med. Chem.* **2004**, *11*, 71–90.
- Daeyaert, F.; de Jonge, M.; Heeres, J.; Koymans, L.; Lewi, P.; Vinkers, M. H.; Janssen, P. A. A pharmacophore docking algorithm and its application to the cross-docking of 18 HIV-1 NRTI's in their binding pockets. *Proteins* **2004**, *54*, 526–533.
- van Drie, J. H. Pharmacophore discovery—lessons learned. *Curr. Pharm. Des.* **2003**, *9*, 1649–1664.
- Langer, T.; Krovat, E. M. Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. *Curr. Opin. Drug Discovery Devel.* **2003**, *6*, 370–376.
- Nicklaus, M. C.; Neamati, N.; Hong, H.; Mazumder, A.; Sunder, S.; Chen, J.; Milne, G. W.; Pommier, Y. HIV-1 integrase pharmacophore: Discovery of inhibitors through three-dimensional database searching. *J. Med. Chem.* **1997**, *40*, 920–929.
- Neamati, N.; Hong, H.; Mazumder, A.; Wang, S.; Sunder, S.; Nicklaus, M. C.; Milne, G. W.; Prokha, B.; Pommier, Y. Depsides and depsidones as inhibitors of HIV-1 integrase: Discovery of novel inhibitors through 3D database searching. *J. Med. Chem.* **1997**, *40*, 942–951.
- Hong, H.; Neamati, N.; Wang, S.; Nicklaus, M. C.; Mazumder, A.; Zhao, H.; Burke, T. R., Jr.; Pommier, Y.; Milne, G. W. Discovery of HIV-1 integrase inhibitors by pharmacophore searching. *J. Med. Chem.* **1997**, *40*, 930–936.
- Hong, H.; Neamati, N.; Winslow, H. E.; Christensen, J. L.; Orr, A.; Pommier, Y.; Milne, G. W. Identification of HIV-1 integrase inhibitors based on a four-point pharmacophore. *Antivir. Chem. Chemother.* **1998**, *9*, 461–472.
- Neamati, N.; Hong, H.; Sunder, S.; Milne, G. W.; Pommier, Y. Potent inhibitors of human immunodeficiency virus type 1 integrase: Identification of a novel four-point pharmacophore and tetracyclines as novel inhibitors. *Mol. Pharmacol.* **1997**, *52*, 1041–1055.
- Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. Developing a dynamic pharmacophore model for HIV-1 integrase. *J. Med. Chem.* **2000**, *43*, 2100–2114.
- Mustata, G. I.; Brigo, A.; Briggs, J. M. HIV-1 integrase pharmacophore model derived from diverse classes of inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1447–1454.
- Herr, R. J. 5-Substituted-1H-tetrazoles as carboxylic acid isosteres: Medicinal chemistry and synthetic methods. *Bioorg. Med. Chem.* **2002**, *10*, 3379–3393.
- Shionogi & Co Ltd. Indole derivatives with antiviral activity. WO9950245-A, 1999.
- Shionogi & Co Ltd. Integrase inhibitors containing aromatic heterocycle derivatives. WO0117968-A, 2001.
- 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.
- Sechi, M.; Derudas, M.; Dallochio, R.; Dessi, A.; Bacchi, A.; Sannia, L.; Carta, F.; Palomba, M.; Ragab, O.; Chan, C.; Shoemaker, R.; Sei, S.; Neamati, N. Design and synthesis of novel indole  $\beta$ -diketo acid derivatives as HIV-1 integrase inhibitors. *J. Med. Chem.* **2004**, *47*, 5298–5310.

- (32) Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.
- (33) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.
- (34) Habib, N. S.; Soliman, R.; Ashour, F. A.; el-Taiebi, M. Synthesis and antimicrobial testing of novel oxadiazolylbenzimidazole derivatives. *Pharmazie* **1997**, *52*, 746–749.
- (35) Takasu, K.; Inoue, H.; Kim, H. S.; Suzuki, M.; Shishido, T.; Wataya, Y.; Ihara, M. Rhodacyanine dyes as antimalarials. 1. Preliminary evaluation of their activity and toxicity. *J. Med. Chem.* **2002**, *45*, 995–998.
- (36) Sudo, K.; Matsumoto, Y.; Matsushima, M.; Fujiwara, M.; Konno, K.; Shimotohno, K.; Shigeta, S.; Yokota, T. Novel hepatitis C virus protease inhibitors: Thiazolidine derivatives. *Biochem. Biophys. Res. Commun.* **1997**, *238*, 643–647.
- (37) Terashima, H.; Hama, K.; Yamamoto, R.; Tsuboshima, M.; Kikkawa, R.; Hatanaka, I.; Shigeta, Y. Effects of a new aldose reductase inhibitor on various tissues in vitro. *J. Pharmacol. Exp. Ther.* **1984**, *229*, 226–230.
- (38) El-Kabbani, O.; Ruiz, F.; Darmanin, C.; Chung, R. P. Aldose reductase structures: Implications for mechanism and inhibition. *Cell Mol. Life Sci.* **2004**, *61*, 750–762.
- (39) Cheng, A.; Diller, D. J.; Dixon, S. L.; Egan, W. J.; Lauri, G.; Merz, K. M., Jr. Computation of the physico-chemical properties and data mining of large molecular collections. *J. Comput. Chem.* **2002**, *23*, 172–183.
- (40) Egan, W. J.; Walters, W. P.; Murcko, M. A. Guiding molecules towards drug-likeness. *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 540–549.
- (41) Egan, W. J.; Lauri, G. Prediction of intestinal permeability. *Adv. Drug Deliv. Rev.* **2002**, *54*, 273–289.
- (42) Long, Y. Q.; Jiang, X. H.; Dayam, R.; Sanchez, T.; Shoemaker, R.; Sei, S.; Neamati, N. Rational Design and Synthesis of Novel Dimeric Diketoacid-Containing Inhibitors of HIV-1 Integrase: Implication for Binding to Two Metal Ions on the Active Site of Integrase. *J. Med. Chem.* **2004**, *47*, 2561–2573.
- (43) Smellie, A.; Kahn, S. D.; Teig, S. L. Analysis of conformational coverage. 2. application of conformational models. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 295–304.
- (44) Smellie, A.; Kahn, S. D.; Teig, S. L. Analysis of conformational coverage 0.1. validation and estimation of coverage. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 285–294.
- (45) Smellie, A.; Teig, S. L.; Towbin, P. Poling—Promoting conformational variation. *J. Comput. Chem.* **1995**, *16*, 171–187.
- (46) Goldgur, Y.; Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Craigie, R.; Davies, D. R. Three new structures of the core domain of HIV-1 integrase: An active site that binds magnesium. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9150–9154.
- (47) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (48) CCDC, C. C. E., UK *GOLD*; 1.2 ed.
- (49) Nissink, J. W.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. A new test set for validating predictions of protein–ligand interaction. *Proteins* **2002**, *49*, 457–471.
- (50) Weislow, O. W.; Kiser, R.; Fine, D.; Bader, J.; Shoemaker, R. H.; Boyd, M. R. New soluble-formazan assay for HIV-1 cytopathic effects: Application to high-flux screening of synthetic and natural products for AIDS antiviral activity. *J. Natl. Cancer Inst.* **1989**, *81*, 577–586.

JM0496077