HIV-1 Integrase Pharmacophore: Discovery of Inhibitors through Three-Dimensional Database Searching†

Marc C. Nicklaus,*,‡ Nouri Neamati,§ Huixiao Hong,‡ Abhijit Mazumder,§ Sanjay Sunder,§ Julie Chen,§ George W. A. Milne,‡ and Yves Pommier§

Laboratories of Medicinal Chemistry and Molecular Pharmacology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, 37 Convent Drive, Bethesda, Maryland 20892-4255

Received August 19, 1996⊗

Starting from a known inhibitor of human immunodeficiency virus type 1 (HIV-1) integrase (IN), caffeic acid phenethyl ester (CAPE), a putative three-point pharmacophore for binding of inhibitors to IN was derived. This pharmacophore was used to search the National Cancer Institute three-dimensional (3D) structural database. Out of the open, nonproprietary part of this database, comprising approximately 200 000 compounds, 267 structures were found to match the pharmacophore in at least one conformation, and 60 of those were tested in an *in vitro* assay against HIV-1 IN. Out of these, 19 were found to inhibit both the 3′-processing and strand transfer of IN at micromolar concentrations. In order to test the validity of this pharmacophore, a small 3D database of 152 published IN inhibitors was built. A search in this database yielded a statistically significant correlation of the presence of this pharmacophore and the potency of the compounds. An automated pharmacophore identification procedure performed on this set of compounds provided additional support for the importance of this pharmacophore for binding of inhibitors to IN and hinted at a possible second pharmacophore. The role of aromatic moieties in the binding of ligands to HIV-1 IN through interactions with divalent metal cations, which are known to be necessary for activity of the enzyme, was explored in *ab initio* calculations.

Introduction

During the life cycle of the human immunodeficiency virus (HIV), the insertion of the viral DNA into the host cell's chromosome is a step that is essential for the efficient replication of the virus. This "integration" of HIV-1 DNA, priorly synthesized by HIV reverse transcriptase, proceeds by a series of defined cutting and joining reactions performed on both the viral and host DNA. In an initial step, two nucleotides are excised from each 3′-end of the linear, blunt-ended, viral DNA. This 3′-processing reaction is a prerequisite for, and is followed by, the actual DNA integration into the host cell genome. These reactions are all catalyzed by the protein HIV integrase (IN), which is one of the products of the *gag-pol* fusion protein precursor and is contained in the virus particle. (See Rice *et al.*,[1] Katz and Skalka,[2] and Vink and Plasterk[3] for recent reviews on retroviral DNA integration.)

Because of its essential nature in the replicative cycle of HIV, IN is an attractive target for the development of anti-AIDS drugs. A number of compounds have been found to inhibit IN in *in vitro* assays,[4−13] but no successful drug based on IN inhibition has emerged so far. The crystal structure of the catalytic core domain of IN has been determined recently,[14] as well as the structure of a related integrase.[15,16] However, no X-ray or NMR structure of an inhibitor−IN complex has been published to date, and the details of the binding of the substrate or inhibitors to the active site of IN are not known. Important steps toward progress in the development of IN inhibitors are therefore identification of the possible pharmacophore(s) as well as discovery of new lead compounds.

Three-dimensional (3D) database searching allows one to achieve both goals simultaneously, thus playing an increasingly important role in this phase of drug development projects. It allows the rapid identification of compounds that possess the pattern(s) of atoms constituting the putative pharmacophore(s), with the most important characteristic of these atoms typically being their capability to form hydrogen bonds. If the 3D database is linked to a repository of samples of the compounds and an assay is available, then very rapid cycles of proposing, testing, and rejecting, refining, or confirming of pharmacophore hypotheses are possible, at the same time yielding useful lead compounds in an oftentimes rapid fashion.[17,18]

The National Cancer Institute (NCI) has conducted for over 4 decades extensive testing of both synthetic and natural compounds for possible activity against various forms of cancer. In the course of this endeavor, nearly one-half million compounds have been tested, and the two-dimensional structural data as well as test results and other information were concurrently collected in computer databases.[19] Recently, in a project that has been described elsewhere,[20] these structural data have been converted to 3D coordinates for more than 400 000 of the compounds. Although the inventory of the NCI repository changes over time, samples are usually available for more than one-half of the compounds (65% at the time of writing). This 3D database is increasingly, and successfully,[21] used in the laboratory's drug design efforts, and its application in the development of lead compounds for IN inhibition is described here.

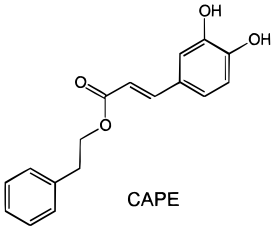
† This is the first in a series of papers describing pharmacophore-based development of HIV-1 integrase inhibitors.
‡ Laboratory of Medicinal Chemistry.
§ Laboratory of Molecular Pharmacology.
⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1997.

S0022-2623(96)00596-1

Computational Chemistry

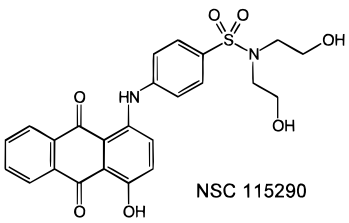
Initial Derivation of the Pharmacophore. Caffeic acid phenethyl ester (CAPE) (**1**), a principal component of propolis, the substance that bees use to reduce the size of the entrance and seal holes in their hives, is a compound that was found to be selectively cytotoxic to tumor cells[22] and has also been of interest in IN inhibitor development. CAPE exhibits moderate potency as an inhibitor of both the 3′-processing and strand transfer steps[7,9] (42 and 10 μM, respectively, measured in a preincubation assay, as described in the Experimental Section).

Among a series of compounds investigated by our laboratory,[9] which were evaluated for IN inhibitory activity, a bis(hydroxyethyl)sulfonamido-substituted 1-amino-4-hydroxyanthraquinone derivative (**2**) had the highest potency (5 μM). Compounds **1** and **2** were used



CAPE

1



NSC 115290

2

as starting points for the initial derivation of the pharmacophore because the limited number of atoms in **1** that are capable of forming hydrogen bonds restricts the choice of equivalent atoms in **2**, while this latter compound helps in narrowing down the choice of atom types (*i.e.*, hydrogen bond donor or acceptor) that are admitted in the subsequent 3D database searches.

Models of both molecules were built using the Molecular Modeling package QUANTA[23] and energy-minimized with its associated molecular mechanics program CHARMm.[24] Low-energy conformers of compounds **1** and **2** could be superimposed in such a way that three hydrogen bond-accepting atoms in each structure matched an equivalent set in the other compound. A reasonably good fit was obtained both in position of the atoms (rms = 0.6 Å) and in their orientation with respect to the formation of hydrogen bonds (Figure 1). The set thus selected suggests that this pharmacophore consists of three atoms that are all hydrogen bond acceptors and/or can form complexes with positively charged ions. The atom triad of **2**, plus a steric exclusion sphere which takes into account the orientation of the carbonyl oxygen in CAPE, was adopted as the putative pharmacophore used in the search of the 3D database (Figure 2).

3D Database Search. The search was conducted in the nonproprietary, "open"[17,20] part (consisting of 206 876 compounds) of the 3D NCI database, which allows free usage of the retrieved structures in contrast to the "discreet" compounds which are bound by confidentiality agreements with their suppliers, and whose data may not be made public. The software used, for both initial building[20] and subsequent searching of the 3D database, was the program ChemDBS-3D, which is part of the chemical software suite Chem-X.[25] Conformational flexibility of compounds is explicitly taken into account by the software, which samples the conformational

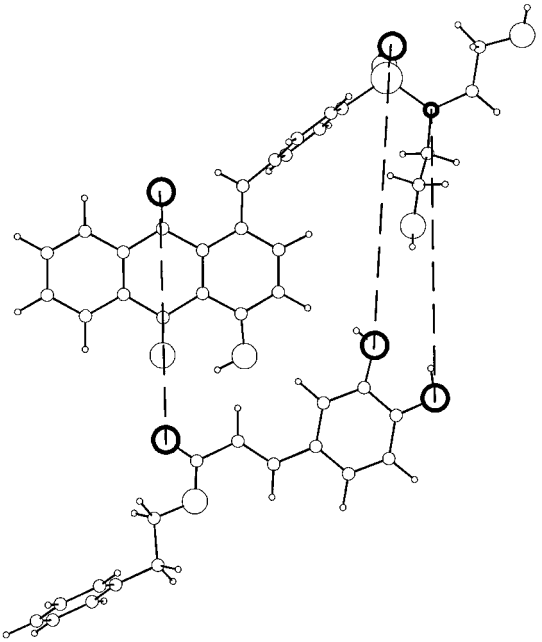


Figure 1. Correspondence of the atoms in low-energy conformers of **1** (bottom) and **2** (top) used in the putative pharmacophore. The tetrahedral nature of the sulfur atom in **2** makes the two sulfonyl oxygens and the sulfonamide nitrogen geometrically equivalent choices for the pharmacophore (after rotation of the phenyl−sulfur bond by 120°).

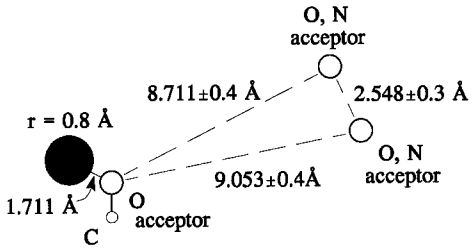


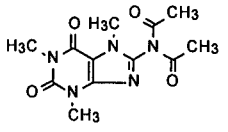
Figure 2. Pharmacophore used in the 3D search. Given are the dimensions and the admitted atom types and element equivalences used by ChemDBS-3D. The shaded sphere is an (steric) exclusion sphere (see text).

space of each molecule quite extensively during the database build phase (with up to millions of conformers per structure) and also during the search phase if a compound has passed initial, rapid screening tests.

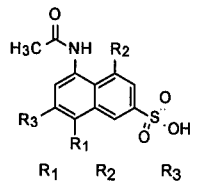
Since the distances used by the search algorithm are precise to the rounding errors of the hardware used, it is necessary to provide the program with upper and lower bounds for each of the distances in the pharmacophore in order to find any hits. For the searches performed, these tolerances were set to the values shown in Figure 2. These tolerances on the order of ±0.4 Å, plus a setting in ChemDBS-3D that steered the search algorithm away from very flexible structures (>15 rotatable bonds), resulted in 267 compounds matching the pharmacophore.

After review of this answer set, 60 compounds (Figure 3) were selected for testing on the basis of availability in the repository, estimated solubility, and heuristically assessed potential usefulness of the compound as a drug lead.

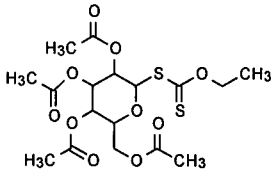
Validation Database. In order to investigate the general validity of the pharmacophore derived as described above, and used in the 3D searches, a validation database of known HIV-1 inhibitors was built and analyzed. The same software with which the large NCI



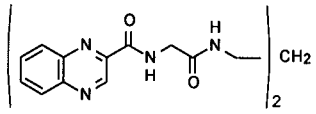
3



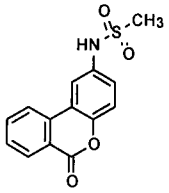
	R_1	R_2	R_3
4	H	H	H
5	NH_2	H	H
10	H	OH	SO_3H



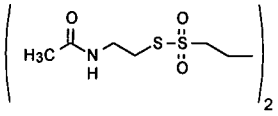
6



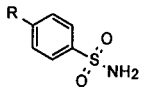
9

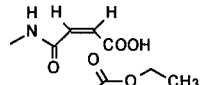


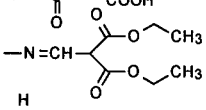
8

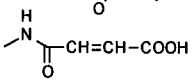


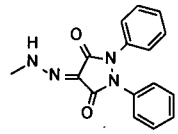
11

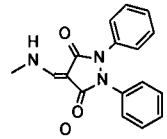


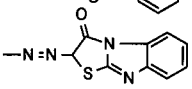
7 R = 

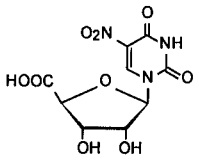
41 R = 

42 R = 

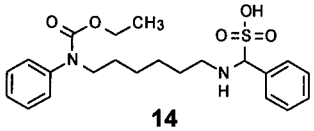
47 R = 

48 R = 

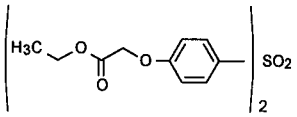
56 R = 



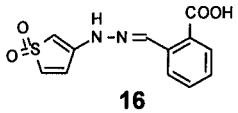
12



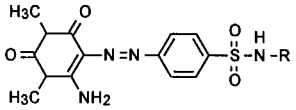
14



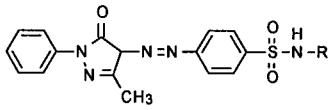
15



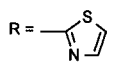
16

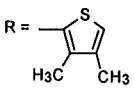


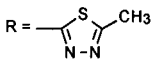
(17-23)

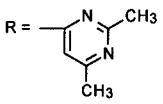


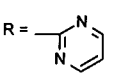
(24-30)

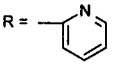
17 R = 

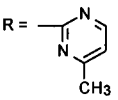
18/25 R = 

19/26 R = 

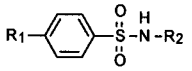
20/27 R = 

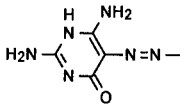
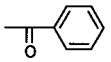
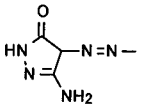
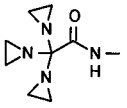
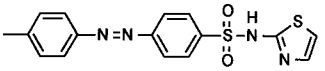
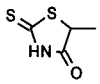
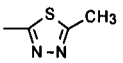
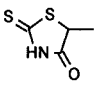
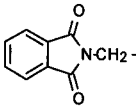
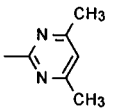
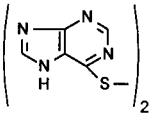
21/30 R = 

22/29 R = 

23/28 R = 

24 R = H



	R_1	R_2
13		
31		
32		
37		
38		
39		
40		
43		
44		
46		

	R_1	R_2
51		
59		
61		

33 R =

34 R =

35 R =

36 R =

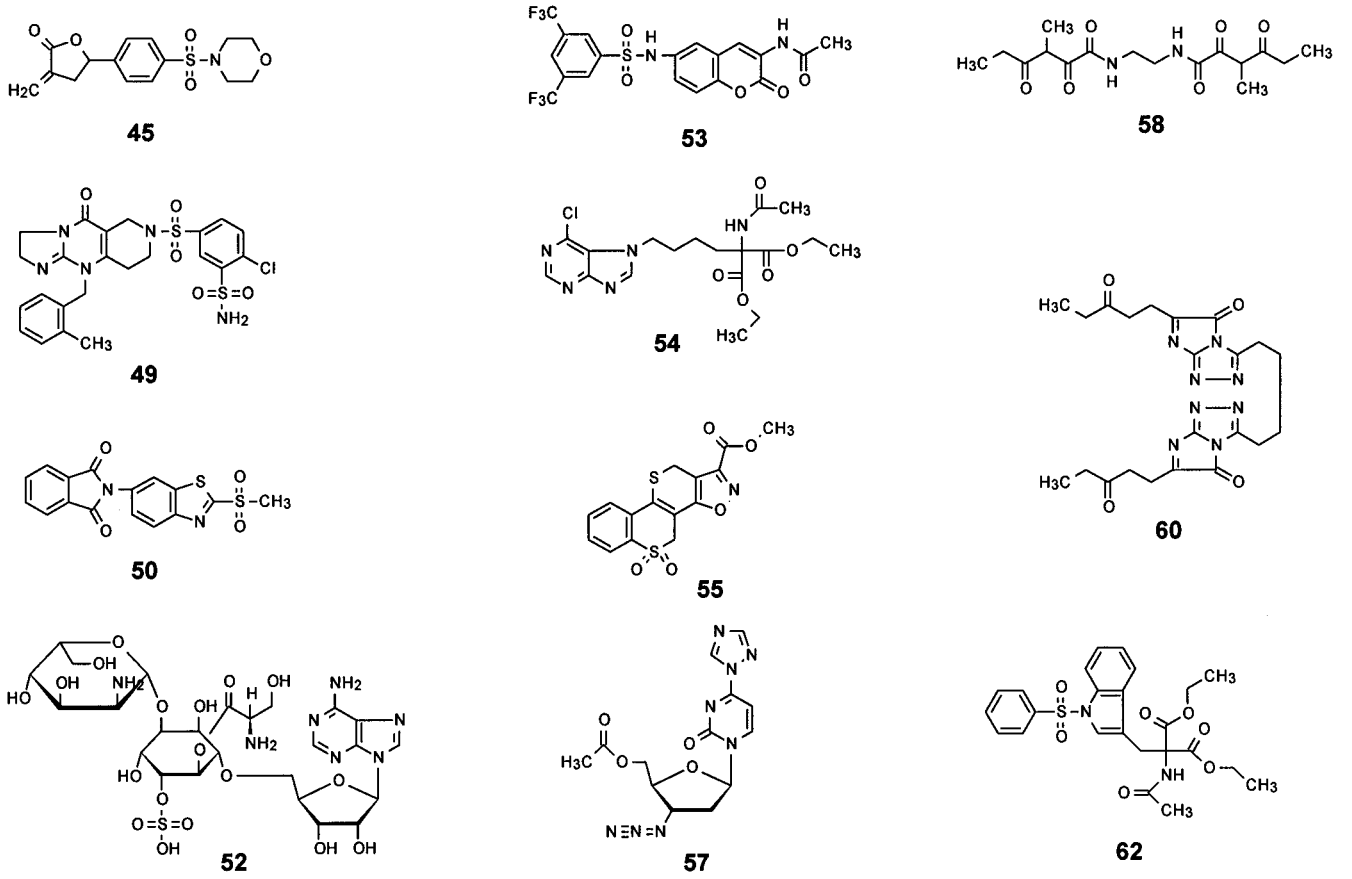


Figure 3. Structures of compounds assayed for IN inhibition.

database had been built, ChemDBS-3D, was used for the construction of, and subsequent searches in, this small 3D database. A total of 152 compounds (structural diagrams available as Supporting Information) were entered in it, all but three structures extracted from published literature.[4−9,11,12,26−33] The criterion for inclusion was that HIV-1 IN inhibitory test data had to be known. These were below 100 μM for slightly more than one-half of the compounds (see below). This set is a sizable portion of the entire body of compounds with IN inhibitory activity published to date. None of the structures in the database was among the 60 compounds presented in this paper. The compounds' conformational space was explored by employing flexible fitting to the pharmacophore ('FLEXIFIT') during the search.

Random Compound Selection. In order to determine the pharmacophore-independent background of anti-IN activity among the compounds in the NCI database, a random set of compounds was assayed. A set of numbers uniformly distributed across the entire range of compound numbers in the NCI database (ranging from NSC 1 to currently NSC 693 000) was produced using a computer random number generator. NSC numbers selected in this way that were either "discreet" compounds, compounds with no sample available according to the database records, or nonexisting entries (certain NSC number ranges are not in use) were discarded; all other compounds were retained. This procedure was repeated until a set of 61 compounds had been assembled. Samples of 59 of these were obtained from the repository and subsequently tested in the assay.

***Ab initio* Calculations.** The program system Gaussian 94[34] was used to calculate gas-phase interaction energies between divalent metal cations and aromatic structures. SCF calculations at the Hartree−Fock (HF) level were performed using the 6-311G** basis set for the complexes with Mg^{2+} and STO-3G for the Mn^{2+} complexes. After a calculation of the benzene−Mg^{2+} complex at the second-order Møller−Plesset perturbation (MP2) level of theory (MP2/6-311G**//MP2/6-311G**) had yielded only very small changes in both the optimized geometry and the interaction energy (less than 0.3 kcal/mol) when compared to the HF optimization, this latter level of theory was deemed sufficient for all subsequent calculations. No counterpoise calculations were performed to correct for the basis set superposition error (BSSE), since rather small BSSEs on the order of 1 kcal/mol or less have been found in previous studies of complexes involving small molecules and/or cations[35,36] and such small corrections would not materially affect our conclusions given the strength of the interactions found.

For both benzene and catechol, the metal ion was placed ca. 2 Å above the centroid of the aromatic ring, and the structure was then optimized without any degrees of freedom being frozen. In all cases, the ion stayed above the aromatic moiety in the optimized structure (local energy minimum), indicating a strong interaction between the π-cloud and the cation. From the total energy of this system was subtracted the sum of the individual energies of the metal ion and the aromatic molecule to yield the gas-phase interaction energy.

Results

The anti-HIV-1 IN activity was measured in an *in vitro* assay as described in the Experimental Section. All compounds were assayed at several concentrations

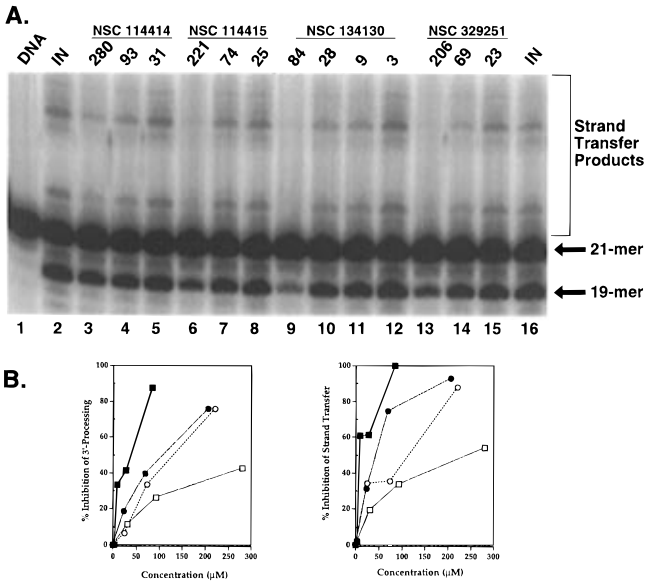
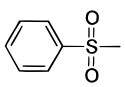


Figure 4. Inhibition of HIV-1 integrase-catalyzed 3′-processing and strand transfer using the blunt-ended DNA substrate by a series of inhibitors. (A) Lane 1 contains an enzyme-free control of just the substrate (DNA); lanes 2 and 16 contain inhibitor-free controls of the substrate with integrase (IN); and lanes 3−5, 6−8, 9−12, and 13−15 show the effects of adding, to substrate and integrase, different concentrations (as indicated, in μM) of the inhibitors NSC 114414 (**24**), NSC 114415 (**25**), NSC 134130 (**38**), and NSC 329251 (**51**), respectively. Normal enzymatic function of integrase produces strand transfer products which appear as bands above the 19-mer/21-mer region. Inhibition of integrase prevents these products from being formed and causes the bands in this region to diminish, as can be seen for all four test compounds, particularly at the higher concentrations. (B) Graphical representation of the inhibition of 3′-processing and strand transfer caused by the above four compounds, as a function of their concentration: (■) NSC 134130, (□) NSC 114414, (○) NSC 114415, and (●) NSC 329251.

to obtain IC_{50} values. Both the 3′-processing and DNA strand transfer reactions were measured in the same experiment and are shown in one representative gel (Figure 4).

Out of the 60 compounds tested (Figure 3), 19 inhibited IN at concentrations below 200 μM (Table 1). Fifteen of those inhibited IN at less than 100 μM, and three of these (**5**, **10**, **43**) showed potency below 10 μM.

Two of these three compounds, **5** and **10**, are naphthalenesulfonic acid derivatives, bearing one or two sulfonate groups, respectively. The second sulfonate group in **10** appears to be important for the inhibitory action, since it confers to the compound about 1 order of magnitude more potency when compared with **5**, making it the compound with the highest activity (IC_{50} of 0.47 μM for 3′-processing and 0.29 μM for strand transfer) in the set. Compound **43** belongs to a group of structures with a phenylsulfonamide core, out of which a considerable percentage had shown detectable potency in the assay. Two of the most active compounds of this group, **37** and **38**, are both nonionic molecules with a 2-thio-4-ketothiazolidine ring attached to this core, whereas **43** possesses a carboxyl functionality. All active compounds in the set contain a phenylsulfone



substructure, but not all compounds with this substructure were active. Some of the active compounds are molecules that are charged under physiological conditions, but many others are not.

All of the compounds with HIV-1 IN inhibitory potency presented here were either inactive in a cell-protection assay using T4 lymphocyte CEM-SS cells or too cytotoxic to these cells to exhibit a potential antiviral activity.

In order to examine the validity for anti-IN activity of the pharmacophore presented here, an independent

Table 1. HIV Integrase Inhibition[a]

		IC_{50} (μM)	
compd	NSC no.	3′-processing	strand transfer
3	34463	>200	>200
4	37072	27.9 ± 8.6	14.3 ± 0.6
5	37078	3.4 ± 0.8	2.7 ± 0.4
6	66060	>200	>200
7	66206	>200	>200
8	67722	>200	>200
9	71310	>200	>200
10	74489	0.47 ± 0.02	0.29 ± 0.04
11	76302	>200	>200
12	79075	>200	>200
13	84122	120 ± 32	96 ± 34
14	84424	>200	>200
15	85753	>200	>200
16	91757	>200	>200
17	114402	>200	>200
18	114403	>200	>200
19	114404	>200	>200
20	114405	>200	>200
21	114406	>200	>200
22	114407	>200	>200
23	114408	>200	>200
24	114414	>200	>200
25	114415	124 (132.7)	97 (115)
26	114416	134 ± 36	77 ± 14
27	114417	>200	>200
28	114418	49.0 ± 9.5	23.6 ± 1.9
29	114419	>200	>200
30	114420	>200	>200
31	125231	71 ± 44	61 ± 30
32	127501	>200	>200
33	134123	>200	>200
34	134124	70 ± 33	65 ± 25
35	134125	193	142.9
36	134126	244	178
37	134129	24 ± 8.2	19 ± 8.3
38	134130	28.6 ± 11.6	14 ± 9.7
39	142429	>200	>200
40	145113	140 ± 10	112 ± 29
41	157505	>200	>200
42	163484	>200	>200
43	165838	8.2 ± 2.4	8.8 ± 2.8
44	166298	>200	>200
45	176514	>200	>200
46	276745	>200	>200
47	279281	>200	>200
48	279288	>200	>200
49	281199	>200	>200
50	327338	>200	>200
51	329251	122 ± 22	80 ± 14
52	336242	>200	>200
53	361668	>200	>200
54	364040	>200	>200
55	375959	>200	>200
56	380502	209	126
57	382316	>200	>200
58	619850	>200	>200
59	630709	75.5 ± 17.3	44.6 ± 20.1
60	636340	>200	>200
61	644037	48.3 ± 25.8	13.7 ± 8.9
62	653265	>200	>200

[a] Results in parentheses are from second independent experiments; results with standard deviations are from at least three independent experiments.

compound set was analyzed. A 'validation' database of 152 known HIV-1 IN inhibitors was searched (using ChemDBS-3D) with the same pharmacophore (Figure 2) that had yielded, from the 200 000-compound DIS database, the 60 compounds assayed. With a 0.7 Å tolerance, this search identified 85 of the 152 compounds (55.9%) as matching the pharmacophore in at least one conformation. If the set was subdivided into two classes of "active" and "inactive" molecules, using as the cutoff criterion an IC_{50} of 100 μM (yielding 77 "actives", 50.7%, and 75 "inactives"), then comparison with the search results yielded the distribution shown (percentages given relative to the size of the entire set, *i.e.*, 152):

	inactive (>100 μM)	active (<100 μM)
pharmacophore not found	55 (36.2%)	12 (7.9%)
pharmacophore found	20 (13.2%)	65 (42.8%)

The two cells in the diagonal of this arrangement represent the correct predictions of both activity (lower right) and inactivity (upper left), *i.e.*, the "true positives" and "true negatives", respectively, whereas the upper right-hand cell could be called "false negatives" and the lower left-hand cell "false positives". The correctly identified compounds thus add up to 55 + 65 = 120, which means that 78.9% (120/152) of the predictions were accurate. This is statistically highly relevant at the $p < 0.005$ level. (The "actives" were predicted correctly at a somewhat higher rate, 65/77 = 84.4%, than the "inactives", 55/75 = 73.3%. One possible reason for this is that the pharmacophore can be identified in a structure in a conformation that is too high in energy for this compound to be active, since only steric "bump" checks but no conformational energy calculations were performed during the search.) If only the most active compounds are taken into account, defined by having a potency of 1.0 μM or better, the prediction rate rises to 100%; *i.e.*, all seven such compounds among the 152 structures were found to contain the pharmacophore.

A second, complementary kind of pharmacophore analysis was then applied to a subset of the validation database. The 77 active compounds with IC_{50} values below 100 μM were processed with Chem-X's Pharmacophore Identification functionality. This algorithm does not start with any predefined pharmacophore but instead searches for all possible pharmacophores in the molecule set analyzed and attempts to find, and rank, the pharmacophores common to the largest possible subset of compounds. In other words, if it identifies a pattern of (typically) three atoms with the same geometry (within predefined tolerances) and the same atom types (such as hydrogen bond donor/acceptor, charged atom, etc.) in at least one conformer of, say, 80% of the compounds analyzed, then this pattern would be ranked as the most common pharmacophore, whereas a different pattern showing up in 70% of the structures would be the second most common pharmacophore, etc.

When the results of these searches were analyzed, two main classes of pharmacophores emerged (Figure 5). For 65 out of the 77 compounds (84.4%), the program had identified pharmacophores very similar to the one shown in Figure 2: *i.e.*, an elongated triangle with two sides on the order of 9 Å and a much shorter side on the order of 2.5 Å. The program did present, however, indication of a second pharmacophore that may exist in the known HIV-1 IN inhibitors. A pattern of three

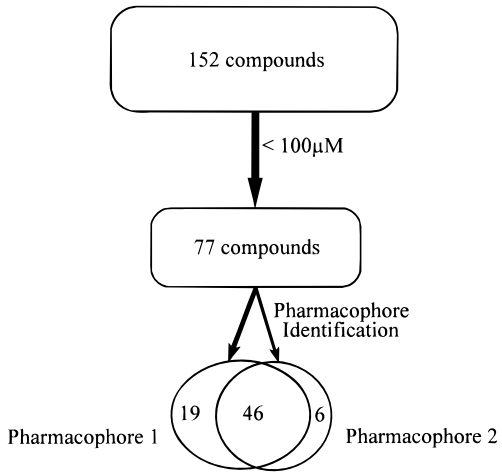


Figure 5. Results of Pharmacophore Identification analysis of 77 known HIV-1 IN inhibitors.

atoms, having one similarly short side but being considerably less elongated overall (side lengths of 5.62, 4.76, and 2.78 Å), was identified in 52 of the compounds (67.5%). These two subsets are not disjoint, however, since 46 of the compounds showed the presence of both pharmacophores. [Six compounds (7.8%) were found to possess (only) atom patterns that could not be categorized in either pharmacophore.] The pharmacophore presented here may be associated with higher potencies since five (out of 19) of the compounds matching only pharmacophore 1 (Figure 5) had activities below 2.5 μM, whereas only one (out of eight) of the compounds matching solely pharmacophore 2 was that active. The standard deviation of the activities in either group, however, was too large to allow more detailed conclusions to be drawn, and further work is required to elucidate the relationship between the two pharmacophores. This second pharmacophore is currently being investigated.[37]

A final control of the relevance of our pharmacophore was the analysis of 59 pharmacophore-independent compounds from the NCI database, selected through random number generation as described above. Out of these, six compounds inhibited IN at concentrations below 200 μM (*vs* 19 in the pharmacophore-selected group). Four of those showed potency below 100 μM (*vs* 15 found by the pharmacophore search). Two of these compounds, which showed an IC_{50} of 20−50 μM, contained either a reactive *N*-oxide group or a conjugated anthracycline moeity with known DNA-intercalating properties. The third compound, with an IC_{50} in the 60 μM range, did not contain the present nor the second pharmacophore. The most active of the "random" compounds, inhibiting HIV-1 IN in the 1 μM range, did indeed contain the pharmacophore presented here in a conformation 0.6 kcal/mol above the global energy minimum. It is, in fact, an analog of a series of HIV-1 IN inhibitors that are the focus of an ongoing study. Not excluding any of these compounds (which may be active in the assay through a mechanism different from blocking the active site), the NCI database's level of background activity (below 100 μM) against HIV-1 IN would then be four out of 59 compounds (6.8%). This contrasts with 15 out of 60 (25%) for the pharmacophore-containing compounds. This nearly 4-fold hit rate increase above the background level that was produced by the pharmacophore search is highly statistically relevant at the $p < 0.01$ level. (It is interesting to note

that the most active of the random compounds—although found to contain the pharmacophore in individual modeling—was not among the 267 compounds originally identified by the pharmacophore search performed using Chem-X. This shows that the 3D database build and search procedures still entail a certain degree of conformational 'coarseness' which is as likely to miss true inhibitors as it is to deliver false hits.)

Discussion

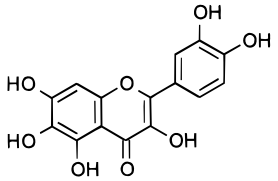
Starting from CAPE, a putative pharmacophore was derived in order to find new inhibitors of IN as well as to investigate the relevance of this pharmacophore for anti-IN activity in general. The application of 3D database techniques as described in this paper successfully achieved both goals. The search for, and subsequent assaying of, compounds from the NCI repository yielded 19 compounds with anti-IN activity. Similar 3D search techniques applied to an extensive collection of known HIV IN inhibitors originating from different sources provided independent support for the relevance of the proposed pharmacophore.

The inhibitors found are structurally quite diverse, although they show a number of common features. The distance (range) between the two oxygen atoms defining the corners of the short side of the pharmacophore that was used in the 3D search, 2.5 ± 0.3 Å (Figure 2), is the reason a considerable number of compounds were found in the NCI database that possessed a sulfone group, typically as part of a sulfonate or sulfonamide functionality. The distance between the two oxygens in a sulfone group is quite exactly 2.5 Å. This group, or more precisely, a phenylsulfone substructure, seems to play a role in the binding to IN since all active compounds contained the latter. However, it appears that it is more the *geometry* of the pattern of oxygen atoms that is decisive for activity than the exact chemical nature of these atoms. The starting point for the pharmacophore hypothesis, CAPE, is a case in point, and it was clearly confirmed by the fact that only one of the 16 most active compounds ($IC_{50} \leq 2.5\ \mu M$) in the *validation* database, which were all found to match the pharmacophore, contained a sulfone functionality.

Although the most active compound in the series, **10**, was a naphthalenesulfonic acid derivative, and other sulfonates[4,38] have been shown to possess strong HIV-1 IN inhibitory activity, the ionic nature of the sulfonated compound does not seem to be required since good inhibition was also achieved with at least three nonionic sulfonamides: **37**, **38**, and **61**.

It was recognized early on that integrases require a divalent metal cation as a cofactor for their enzymatic activities,[39,40] and several crystal structures of proteins belonging to the polynucleotidyl transferase superfamily, of which HIV-1 IN is a member, have been published in which one or two metal ions complexed with the (presumed) active site could be resolved.[16,41−43] The ion assumed to be the integration cofactor *in vivo* is Mg^{2+}, whereas in the *in vitro* assays the integrases have generally been found to work better with Mn^{2+},[9,44] the extent to which this is an artefact of the assay conditions currently being under discussion.[45−49] Assuming that the O=S=O substructure is crucially involved in the inhibitory activity of both sulfonate- and sulfone-containing compounds, then one might hypothesize that the amount of negative charge present at this group may have a bearing on the potency of the compound if the interaction causing the inhibition is mediated by one or two divalent cations. This could explain the trend toward somewhat higher activity of the ionic variant of the sulfone-containing compounds as compared to, *e.g.*, the sulfonamides. We will, however, show below that alternate binding modes involving the metal ion(s) are conceivable.

Structural features common to, and hence deduced to be important for the potency of, inhibitors of HIV-1 IN have been proposed previously.[9,50] The most important feature thus identified was the presence of one or more ortho-polyhydroxylated aromatic moieties in the compound. Our present pharmacophore is compatible with this structural feature and can, in fact, be seen as an extension of it. CAPE, which is a moderate inhibitor of HIV-1 IN, contains just a catechol moiety and an additional ester functionality. Quercategetin, being one of the most potent HIV-1 IN inhibitors known to date[9] (3′-processing, 0.8 μM; strand transfer, 0.1 μM), has



Quercategetin

been extensively analyzed as a prototype of the polyhydroxylated compounds.[50] It was included in our validation database and indeed found to contain the pharmacophore. Detailed analysis showed that a low-energy local minimum conformation (1.8 kcal/mol above the global minimum in vacuum) matched the pharmacophore within the 0.7 Å tolerance. Because of the numerous hydroxyl groups in quercategetin, however, several more possibilities of matching the pharmacophore with different triads of oxygen atoms become available if one allows for higher conformational energies and/or slightly larger distance tolerances, and this might confer to quercategetin an additional entropic advantage with regards to its binding affinity.

Either implicitly, such as in the ortho pair of *phenolic* hydroxyl groups found to be a minimum requirement for activity,[9] or explicitly in quantitative structure−activity analyses,[50] it has been repeatedly shown that an aromatic moiety next to (at least) two adjacent oxygens appears to be a structural element essential for activity. This structural feature was not explicitly used in our pharmacophore searches, but in fact, the phenylsulfone moiety found to be present in all active compounds in our set contains precisely this structural element. It appears possible that the aromatic moiety of these inhibitors interacts with the divalent cation (Mg^{2+} or Mn^{2+}, see above) in a "cation−π" type interaction. These interactions, which are gaining increasing interest in enzyme−ligand interactions,[51,52] are considered to be mostly electrostatic attraction between a positive charge and the quadrupole moment (*i.e.*, the negative partial charge of the π-face) of an aromatic system. They have been shown to be surprisingly strong for monovalent cations in both experimental[53−55] and computational[56,57] studies, with gas-phase binding energies between alkali metal ions and benzene in the range of 15−40 kcal/mol, the highest values being found for the smallest ions [$E(Li^+) > E(Na^+)$, etc.]. Since we are

not aware of equivalent calculations for divalent cations, we undertook *ab initio* computations of the gas-phase interaction energies of Mg^{2+} and Mn^{2+} with benzene and catechol, which are listed below for each of the four combinations of an ion with an aromatic molecule. As expected, substantially higher energies were obtained for these doubly charged ions when compared with the singly charged species such as Li^+, Na^+, etc. The differences in the interaction energies found between Mn^{2+} and Mg^{2+}, if scaled appropriately to account for the higher charges, agree very well with the equivalent differences observed for the alkali ions,[56] for which a roughly hyperbolic dependence of the energy on the ion radius was seen. The ionic radii for Mn^{2+} and Mg^{2+} are 0.67 and 0.72 Å, respectively.[58]

Gas-Phase Binding Energies (kcal/mol) between Metal Ions and Aromatic Structures

	Mg^{2+}	Mn^{2+}
benzene	−114.6[a]	−132.6[b]
catechol	−119.6[a]	−137.9[b]

[a] HF/6-311G**//HF/6-311G** energies (HF/STO-3G//HF/STO-3G energies were −119.8 and −124.6 kcal/mol for benzene and catechol, respectively). [b] HF/STO-3G//HF/STO-3G energies.

Although these results represent gas-phase interactions, which have to compete, in an aqueous environment, with the solvation of the ions, they allow the conclusion that a participation in the ligand binding of the aromatic ring adjacent to the pharmacophore oxygens cannot be excluded. This possible occurrence of cation−π interactions in the HIV-1 IN active site does not rule out the presence of 'normal' charge−charge interactions of the metal ion(s) with ionic or partial charges of the ligand, since it has been shown that both types of interactions can coexist in the same binding site.[52] Also, different compounds could simply bind in different binding modes, "employing" the ion(s) in these various ways or, possibly, not at all. Additional, or alternative, hydrogen bond interactions are possible for the catechol hydroxyl oxygens both as donors and acceptors and as acceptors for the phenylsulfone oxygens.

Recently, the catalytically active core domain of HIV-1 IN was solved crystallographically.[14] This structure will hopefully allow us to tie in the pharmacophore(s) presented in this paper with the details of the active site and to refine the pharmacophore model. However, the protein was crystallized without metal ion or ligand, and a loop, carrying one of the crucial active site residues, was found to be disordered and thus could not be resolved crystallographically. Therefore, considerable modeling will be required to obtain a structure that will allow study of the binding of inhibitors to the active site of IN. Such modeling studies are currently underway.

The analyses performed on the validation database of 152 known HIV-1 IN inhibitors strongly supported the notion that, for many ligands binding to IN, the pharmacophore presented here and used in the 3D searches is indeed relevant. A particularly independent support was provided by the Pharmacophore Identification analysis, which starts from just the structures without predefined knowledge of any specific pharmacophore.

The results of this analysis indicated that (at least) one additional pharmacophore might be important in binding to HIV-1 IN. This is not surprising given the active site's rather open, and probably quite shallow, shape which can be perceived from its resolved part in the crystal structure.[14] The existence of multiple binding modes, and thus multiple pharmacophores (albeit possibly partially overlapping) for inhibitors that bind at the same active site of an enzyme, is increasingly being recognized[59,60] and has also been recently observed in a drug development project performed in this laboratory.[61]

Conclusion

Starting from a pharmacophore hypothesis derived from a known inhibitor of HIV-1 IN, a search of the NCI 3D structural database for compounds that matched this pharmacophore successfully identified a group of novel compounds that showed inhibitory activity against IN. Subsequent 3D searches in a validation database of known HIV-1 IN inhibitors, which had no overlap with the group of compounds found in the initial search, lent strong support for the existence of the postulated pharmacophore in many ligands binding to HIV-1 IN. In addition, it hinted at the existence of a possible second pharmacophore relevant in the binding to IN.

These analyses show that the 3D search technology, coupled with a repository of structures that can be rapidly obtained for testing, is a powerful tool both in the development of new drug lead compounds as well as in the elucidation of the mechanism of binding to, and inhibition of, enzymes.

Experimental Section

Preparation of Oligonucleotide Substrates. The HPLC-purified oligonucleotides AE117, 5′-ACTGCTAGAGATTTTCCACAC-3′, and AE118, 5′-GTGTGGAAAATCTCTAGCAGT-3′, were purchased from Midland Certified Reagent Co. (Midland, TX). Purified recombinant wild-type HIV-1 integrase was prepared as described previously.[62] Dr. R. Craigie (Laboratory of Molecular Biology, NIDDK, NIH) generously provided us with the expression system for the wild-type HIV-1 integrase. To analyze the extents of 3′-processing and strand transfer using 5′-end-labeled substrates, AE118 was 5′-end labeled with ^{32}P (asterisk) using T4 polynucleotide kinase (Gibco BRL) and [γ-^{32}P]ATP (DuPont-NEN):

$$\overset{*}{5'}\text{-GTGTGGAAAATCTCTAGCA}\overset{\downarrow}{\text{G}}\text{T-3'}$$

$$3'\text{-CACACCTTTTAGAGATCGTCA-5'}$$

The kinase was heat-inactivated, and AE117 was added to the same final concentration. The mixture was heated at 95 °C, allowed to cool slowly to room temperature, and run on a G-25 Sephadex quick spin column (Boehringer Mannheim) to separate annealed double-stranded oligonucleotide from unincorporated label.

***In Vitro* HIV-1 Integrase Assay.** Integrase was preincubated at a final concentration of 200 nM with inhibitor in reaction buffer (50 mM NaCl, 1 mM HEPES, pH 7.5, 50 μM EDTA, 50 μM dithiothreitol, 10% glycerol (w/v), 7.5 mM $MnCl_2$, 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 ^{32}P-labeled linear oligonucleotide substrate was added, and incubation was continued for an additional 60 min. Reactions were quenched by the addition of an equal volume (16 μL) of loading dye (98%

deionized formamide, 10 mM EDTA, 0.025% xylene cyanol, and 0.025% bromophenol blue). An aliquot (5 μL) was electrophoresed on a denaturing 20 polyacrylamide gel (0.09 M Tris-borate, pH 8.3, 2 mM EDTA, 20% acrylamide, and 8 M urea). Gels were dried, exposed in a molecular Dynamics Phosphorimager cassette, and analyzed using a Molecular Dynamics Phosphorimager (Sunnyvale, CA). Percent inhibition was calculated using the following equation:

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

where C, N, and D are the fractions of 21-mer substrate converted to 19-mer [3′-processing product generated after dinucleotide cleavage (arrow above)] or strand transfer products for DNA alone, DNA plus integrase, and integrase plus drug, respectively. The IC_{50} was determined by plotting the drug concentration versus percent inhibition and determining the concentration which produced 50% inhibition. The compounds were assayed up to a concentration of 200 μM.

Supporting Information Available: Structural diagrams of the compounds in the validation database (14 pages). Ordering information is given on any current masthead page.

References

(1) Rice, P.; Craigie, R.; Davies, D. R. Retroviral integrases and their cousins. *Curr. Opin. Struct. Biol.* **1996**, *6*, 76−83.

(2) Katz, R. A.; Skalka, A. M. The retroviral enzymes. *Annu. Rev. Biochem.* **1994**, *63*, 133−173.

(3) Vink, C.; Plasterk, R. H. A. The human immunodeficiency virus integrase protein. *Trends Genet.* **1993**, *9*, 433−438.

(4) Carteau, S.; Mouscadet, J. F.; Goulaouic, H.; Subra, F.; Auclair, C. Inhibitory effect of the polyanionic drug suramin on the in vitro HIV DNA integration reaction. *Arch. Biochem. Biophys.* **1993**, *305*, 606−610.

(5) Carteau, S.; Mouscadet, J. F.; Goulaouic, H.; Subra, F.; Auclair, C. Effect of topoisomerase inhibitors on the in vitro HIV DNA integration reaction. *Biochem. Biophys. Res. Commun.* **1993**, *192*, 1409−1414.

(6) Carteau, S.; Mouscadet, J. F.; Goulaouic, H.; Subra, F.; Auclair, C. Inhibition of the in vitro integration of Moloney murine leukemia virus DNA by the DNA minor groove binder netropsin. *Biochem. Pharmacol.* **1994**, *47*, 1821−1826.

(7) Fesen, M. R.; Kohn, K. W.; Leteurtre, F.; Pommier, Y. Inhibitors of human immunodeficiency virus integrase. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2399−2403.

(8) Cushman, M.; Sherman, P. Inhibition of HIV-1 integration protein by aurintricarboxylic acid monomers, monomer analogs, and polymer fractions. *Biochem. Biophys. Res. Commun.* **1992**, *185*, 85−90.

(9) Fesen, M. R.; Pommier, Y.; Leteurtre, F.; Hiroguchi, S.; Yung, J.; Kohn, K. W. Inhibition of HIV-1 integrase by flavones, caffeic acid phenethyl ester (CAPE) and related compounds. *Biochem. Pharmacol.* **1994**, *48*, 595−608.

(10) Mazumder, A.; Cooney, D.; Agbaria, R.; Gupta, M.; Pommier, Y. Inhibition of human immunodeficiency virus type 1 integrase by 3′-azido-3′-deoxythymidylate. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5771−5775.

(11) Mazumder, A.; Raghavan, K.; Weinstein, J.; Kohn, K. W.; Pommier, Y. Inhibition of human immunodeficiency virus type-1 integrase by curcumin. *Biochem. Pharmacol.* **1995**, *49*, 1165−1170.

(12) Mazumder, A.; Gupta, M.; Perrin, D. M.; Sigman, D. S.; Rabinovitz, M.; Pommier, Y. Inhibition of Human Immunodeficiency Virus Type 1 Integrase by a Hydrophobic Cation: The Phenanthroline-Cuprous Complex. *AIDS Res. Hum. Retrovir.* **1995**, *11*, 115−125.

(13) Mazumder, A.; Neamati, N.; Sommadossi, J. P.; Gosselin, G.; Schinazi, R. F.; Imbach, J. L.; Pommier, Y. Effects of nucleotide analogues on human immunodeficiency virus type 1 integrase. *Mol. Pharmacol.* **1996**, *49*, 621−628.

(14) Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Engelman, A.; Craigie, R.; Davies, D. R. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **1994**, *266*, 1981−1986.

(15) Bujacz, G.; Jaskólski, M.; Alexandratos, J.; Wlodawer, A.; Merkel, G.; Katz, R. A.; Skalka, A. M. High-resolution structure of the catalytic domain of avian sarcoma virus integrase. *J. Mol. Biol.* **1995**, *253*, 333−346.

(16) Bujacz, G.; Jaskólski, M.; Alexandratos, J.; Wlodawer, A.; Merkel, G.; Katz, R. A.; Skalka, A. M. The catalytic domain of avian sarcoma virus integrase: conformation of the active-site residues in the presence of divalent cations. *Structure* **1996**, *4*, 89−96.

(17) Milne, G. W. A.; Wang, S.; Nicklaus, M. C. Molecular Modeling in the Discovery of Drug Leads. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 726−730.

(18) Wang, S.; Milne, G. W. A.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. Discovery of Novel, Nonpeptide HIV-1 Protease Inhibitors by Pharmacophore Searching. *J. Med. Chem.* **1996**, *39*, 2047−2054.

(19) Milne, G. W. A.; Miller, J. A. The NCI Drug Information System. 1. System Overview. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 154−159. Milne, G. W. A.; Feldman, A.; Miller, J. A.; Daly, G. P.; Hammel, M. J. The NCI Drug Information System. 2. DIS Pre-Registry. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 159−168. Milne, G. W. A.; Feldman, A.; Miller, J. A.; Daly, G. P. The NCI Drug Information System. 3. The DIS Chemistry Module. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 168−179. Milne, G. W. A.; Miller, J. A.; Hoover, J. R. The NCI Drug Information System. 4. Inventory and Shipping Modules. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 179−185. Zehnacker, M. T.; Brennan, R. H.; Milne, G. W. A.; Miller, J. A. The NCI Drug Information System. 5. DIS Biology Module. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 186−193. Zehnacker, M. T.; Brennan, R. H.; Milne, G. W. A.; Miller, J. A.; Hammel, M. J. The NCI Drug Information System. 6. System Maintenance. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 193−197.

(20) Milne, G. W. A.; Nicklaus, M. C.; Driscoll, J. S.; Wang, S.; Zaharevitz, D. National Cancer Institute Drug Information System 3D Database. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1219−1224.

(21) Wang, S.; Zaharevitz, D.; Sharma, R.; Marquez, V. E.; Lewin, N. E.; Du, L.; Blumberg, P. M.; Milne, G. W. A. The Discovery of Novel, Structurally Diverse Protein Kinase C Agonists through Computer 3-D Database Search. Molecular Modeling Studies. *J. Med. Chem.* **1994**, *37*, 4479−4489. (The pharmacophore in **2** that is relevant in the PK-C assay is different from the one assumed to be responsible for potency of the same compound in the IN assay.)

(22) Grunberger, D.; Banerjee, R.; Eisinger, K.; Oltz, E. M.; Efros, L.; Caldwell, M.; Estevez, V.; Nakanishi, K. Preferential cytotoxicity on tumor cells by caffeic acid phenethyl ester isolated from propolis. *Experientia* **1988**, *44*, 230−232.

(23) QUANTA is provided by Molecular Simulations, Inc., 9685 Scranton Rd, San Diego, CA 92121-3752.

(24) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMm: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(25) Murrall, N. W.; Davies, E. K. Conformational Freedom in 3-D Databases. 1. Techniques. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 312−316.

(26) Burke, T. R., Jr.; Fesen, M. R.; Mazumder, A.; Wang, J.; Carothers, A. M.; Grunberger, D.; Driscoll, J.; Kohn, K.; Pommier, Y. Hydroxylated aromatic inhibitors of HIV-1 integrase. *J. Med. Chem.* **1995**, *38*, 4171−4178.

(27) Cushman, M.; Golebiewski, W. M.; Pommier, Y.; Mazumder, A.; Reymen, D.; De Clercq, E.; Graham, L.; Rice, W. G. Cosalane Analogues with Enhanced Potencies as Inhibitors of HIV-1 Protease and Integrase. *J. Med. Chem.* **1995**, *38*, 443−452.

(28) LaFemina, R. L.; Graham, P. L.; LeGrow, K.; Hastings, J. C.; Wolfe, A.; Young, S. D.; Emini, E. A.; Hazuda, D. J. Inhibition of human immunodeficiency virus integrase by bis-catechols. *Antimicrob. Agents Chemother.* **1995**, *39*, 320−324.

(29) Eich, E.; Pertz, H.; Kaloga, M.; Schulz, J.; Fesen, M. R.; Mazumder, A.; Pommier, Y. (−)-Arctigenin as a lead structure for inhibitors of human immunodeficiency virus type-1 integrase. *J. Med. Chem.* **1996**, *39*, 86−95.

(30) Neamati, N.; Mazumder, A.; Sunder, S.; Pommier, Y. Diarylsulfones, a Novel Class of HIV-1 Integrase Inhibitors. Proceedings of the 96th General Meeting of the American Society of Microbiology, May 19−23, 1996, New Orleans, LA; Abstract No. T-21, p 584.

(31) Mazumder, A.; Wang, S.; Neamati, N.; Nicklaus, M. C.; Sunder, S.; Chen, J.; Milne, G. W. A.; Rice, W. G.; Burke, T. R., Jr.; Pommier, Y. Antiretroviral Agents as Inhibitors of both Human Immunodeficiency Virus Type 1 Integrase and Protease. *J. Med. Chem.* **1996**, *39*, 2472−2481.

(32) Zhao, H.; Neamati, N.; Hong, H.; Mazumder, A.; Wang, S.; Sunder, S.; Milne, G. W. A.; Pommier, Y.; Burke, T. R., Jr. Coumarin-Based Inhibitors of HIV Integrase. *J. Med. Chem.* **1997**, *40*, 242−249.

(33) Zhao, H.; Neamati, N.; Mazumder, A.; Sunder, S.; Pommier, Y.; Burke, T. R., Jr. Arylamide Inhibitors of HIV-1 Integrase. *J. Med. Chem.*, in press.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe,

M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 94, Rev. C.3; by Gaussian, Inc., Carnegie Office Park, Bldg 6, Pittsburgh, PA 15106.

(35) Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. Molecular Recognition in Aqueous Media. New Binding Studies Provide Further Insights into the Cation-π Interaction and Related Phenomena. *J. Am. Chem. Soc.* **1993**, *115*, 9907−9919.

(36) Yan, X.; Wang, S.; Hodošček, M.; Milne, G. W. A. Prediction of geometries and interaction energies of complexes formed by small molecules using semiempirical and ab initio methods. *J. Mol. Struct.* **1994**, *309*, 279−294.

(37) Hong, H.; Neamati, N.; Wang, S.; Nicklaus, M. C.; Mazumder, A.; Zhao, H.; Burke, T. R., Jr.; Pommier, Y.; Milne, G. W. A. Discovery of Human Immunodeficiency Virus Type 1 Integrase Inhibitors by Pharmacophore Searching. *J. Med. Chem.* **1997**, *40*, 930−936.

(38) Clanton, D. J.; Moran, R. A.; McMahon, J. B.; Weislow, O. S.; Buckheit, R. W., Jr.; Hollingshead, M. G.; Ciminale, V.; Felber, B. K.; Pavlakis, G. N.; Bader, J. P. Sulfonic acid dyes: inhibition of the human immunodeficiency virus and mechanism of action. *J. Acquired Immune Defic. Syndr.* **1992**, *5*, 771−781.

(39) Brown, P. O.; Bowerman, B.; Varmus, H. E.; Bishop, J. M. Correct integration of retroviral DNA in vitro. *Cell* **1987**, *49*, 347−356.

(40) Ellison, V.; Abrams, H.; Roe, T.; Lifson, J.; Brown, P. Human immunodeficiency virus integration in a cell-free system. *J. Virol.* **1990**, *64*, 2711−2715.

(41) Katayanagi, K.; Miyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Ikehara, M.; Matsuzaki, T.; Morikawa, K. Three-dimensional structure of ribonuclease H from E. coli. *Nature* **1990**, *347*, 306−309.

(42) Katayanagi, K.; Miyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Nakamura, H.; Ikehara, M.; Matsuzaki, T.; Morikawa, K. Structural details of ribonuclease H from Escherichia coli as refined to an atomic resolution. *J. Mol. Biol.* **1992**, *223*, 1029−1052.

(43) Davies, J. F.; Hostomska, Z.; Hostomsky, Z.; Jordan, S. R.; Matthews, D. A. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* **1991**, *252*, 88−95.

(44) Vink, C.; Lutzke, R. A.; Plasterk, R. H. A. Formation of a stable complex between the human immunodeficiency virus integrase protein and viral DNA. *Nucleic Acids Res.* **1994**, *22*, 4103−4110.

(45) Foltin, S. K.; Holler, T. P.; Hupe, D. J. The DNA processing activity of HIV-1 integrase with magnesium as metal cofactor. *FASEB J.* **1995**, *9*, A1398.

(46) Lee, S. P.; Kim, H. G.; Censullo, M. L.; Han, M. K. Characterization of $Mg^{(2+)}$-dependent 3′-processing activity for human immunodeficiency virus type 1 integrase in vitro: real-time kinetic studies using fluorescence resonance energy transfer. *Biochemistry* **1995**, *34*, 10205−10214.

(47) Lee, S. P.; Censullo, M. L.; Kim, H. G.; Han, M. K. Substrate-length-dependent activities of human immunodeficiency virus type 1 integrase in vitro: differential DNA binding affinities associated with different lengths of substrates. *Biochemistry* **1995**, *34*, 10215−10223.

(48) Engelman, A.; Craigie, R. Efficient magnesium-dependent human immunodeficiency virus type 1 integrase activity. *J. Virol.* **1995**, *69*, 5908−5911.

(49) Pemberton, I. K.; Buckle, M.; Buc, H. The metal ion-induced cooperative binding of HIV-1 integrase to DNA exhibits a marked preference for Mn(II) rather than Mg(II). *J. Biol. Chem.* **1996**, *271*, 1498−1506.

(50) Raghavan, K.; Buolamwini, J. K.; Fesen, M. R.; Pommier, Y.; Kohn, K. W.; Weinstein, J. N. Three-dimensional quantitative structure-activity relationship (QSAR) of HIV integrase inhibitors: a comparative molecular field analysis (CoMFA) study. *J. Med. Chem.* **1995**, *38*, 890−897.

(51) Perutz, M. F. The role of aromatic rings as hydrogen-bond acceptors in molecular recognition. *Phil. Trans. R. Soc. A* **1993**, *345*, 105−112.

(52) Dougherty, D. A. Cation-π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **1996**, *271*, 163−168.

(53) Sunner, J.; Nishizawa, K.; Kebarle, P. Ion-Solvent Molecule Interactions in the Gas Phase. The Potassium Ion and Benzene. *J. Phys. Chem.* **1981**, *85*, 1814−1820.

(54) Taft, R. W.; Anvia, F.; Gal, J.-F.; Walsh, S.; Capon, M.; Holmes, M. C.; Hosn, K.; Oloumi, G.; Vasanwala, R.; Yazdani, S. Free energies of cation-molecule complex formation and of cation-solvent transfers. *Pure Appl. Chem.* **1990**, *62*, 17−23.

(55) Guo, B. C.; Purnell, J. W.; Castleman, A. W., Jr. The Clustering Reactions of Benzene with Sodium and Lead Ions. *Chem. Phys. Lett.* **1990**, *168*, 155−160.

(56) Kumpf, R. A.; Dougherty, D. A. A mechanism for ion selectivity in potassium channels: computational studies of cation-p interactions. *Science* **1993**, *261*, 1708−1710.

(57) Caldwell, J. W.; Kollman, P. A. Cation-p Interactions: Nonadditive Effects Are Critical in Their Accurate Representation. *J. Am. Chem. Soc.* **1995**, *117*, 4177−4178.

(58) *CRC Handbook of Chemistry and Physics*, 74th ed.; CRC Press, Inc.: Boca Raton, Ann Arbor, London, Tokyo, 1993; pp 12-8−12-9.

(59) Cody, V. Mechanisms of Molecular Recognition: Crystallographic Evidence for Multiple Inhibitor Binding Modes in Two Protein Classes. Abstract Book of the 12th Annual Conference of the Molecular Graphics Society, Interlaken, Switzerland, June 7−11, 1993; Abstract P087.

(60) Schulze-Gahmen, U.; Brandsen, J.; Jones, H. D.; Morgan, D. O.; Meijer, L.; Vesely, J.; Kim, S. H. Multiple modes of ligand recognition: crystal structures of cyclin-dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopentenyladenine. *Proteins* **1995**, *22*, 378−391.

(61) Wang, S.; Kazanietz, M. G.; Blumberg, P. M.; Marquez, V. E.; Milne, G. W. A. Molecular Modeling and Site-Directed Mutagenesis Studies of a Phorbol Ester Binding Site in Protein Kinase C. *J. Med. Chem.* **1996**, *39*, 2541−2553.

(62) Jenkins, T. M.; Engelman, A.; Ghirlando, R.; Craigie, R. A Soluble Active Mutant of HIV-1 Integrase-Involvement of Both the Core and Carboxyl-Terminal Domains in Multimerization. *J. Biol. Chem.* **1996**, *271*, 7712−7718.

JM960596U