0026-895X/96/040846-10$3.00/0

MOLECULAR PHARMACOLOGY, 50:846–855 (1996).

Dicaffeoylquinic Acid Inhibitors of Human Immunodeficiency Virus Integrase: Inhibition of the Core Catalytic Domain of Human Immunodeficiency Virus Integrase

W. EDWARD ROBINSON, JR., MARA CORDEIRO, SAMIA ABDEL-MALEK, QI JIA, SAMSON A. CHOW, MANFRED G. REINECKE, and WILLIAM M. MITCHELL

Departments of Pathology and Microbiology and Molecular Genetics, University of California, Irvine, California 92697-4800 (W.E.R.), Department of Molecular and Medical Pharmacology, University of California, Los Angeles, California 90095 (M.C., S.A.C.), Department of Chemistry, Texas Christian University, Fort Worth, Texas 76129 (S.A.-M., Q.J., M.G.R.), and Department of Pathology, Vanderbilt University Medical Center, Nashville, Tennessee 37232 (W.M.M.)

Received April 30, 1996; Accepted June 27, 1996

SUMMARY

Integration of a cDNA copy of the human immunodeficiency virus (HIV) genome is mediated by an HIV-1-encoded enzyme, integrase (IN), and is required for productive infection of CD4+ lymphocytes. It had been shown that 3,5-dicaffeoylquinic acid and two analogues were potent and selective inhibitors of HIV-1 IN *in vitro*. To determine whether the inhibition of IN by dicaffeoylquinic acids was limited to the 3,5 substitution, 3,4-, 4,5-, and 1,5-dicaffeoylquinic acids were tested for inhibition of HIV-1 replication in tissue culture and inhibition of HIV-1 IN *in vitro*. All of the dicaffeoylquinic acids were found to inhibit HIV-1 replication at concentrations ranging from 1 to 6 μM in T cell lines, whereas their toxic concentrations in the same cell lines were >120 μM. In addition, the compounds inhibited HIV-1 IN *in vitro* at submicromolar concentrations. Molecular modeling of these ligands with the core catalytic domain of IN indicated an energetically favorable reaction, with the most potent inhibitors filling a groove within the predicted catalytic site of IN. The calculated change in internal free energy of the ligand/IN complex correlated with the ability of the compounds to inhibit HIV-1 IN *in vitro*. These results indicate that the dicaffeoylquinic acids as a class are potent and selective inhibitors of HIV-1 IN and form important lead compounds for HIV drug discovery.

HIV-1 has been demonstrated to be the causative agent in AIDS. Like other retroviruses, HIV-1 requires integration into a host cell chromosome for productive infection (1–6). The integration reaction is catalyzed by a preintegration complex that contains the viral DNA in a stable nucleoprotein complex that includes the retroviral protein IN (7–9).

The amino acid sequences of retroviral integrases share common structural motifs. In the central, protease-resistant core, there is a D,D(35)E domain consisting of the motif $DX_{39-58}DX_{35}E$, in which *X* is any amino acid. The domain is found in integrases of retroviruses and retrotransposons as well as in the transposases of some prokaryotic transposons (10, 11). The D,D(35)E domain alone can mediate disintegration (see below) (12, 13) and is believed to be the active site for all IN catalytic activities (14). Mutations of either the conserved aspartates or glutamate in the D,D(35)E domain abolish all catalytic activities of IN (15–17). In contrast to the D,D(35)E domain, the functional roles of the amino and carboxyl termini of IN are not well defined. The amino terminus has an HHCC domain, named for its zinc finger-like motif $HX_{3-7}HX_{23-32}CX_2C$ found in many DNA binding proteins (18). The HHCC domain seems to recognize features of the viral DNA (19) and may play a role in the formation of a stable complex between IN and viral DNA ends (20). The carboxyl terminus of IN, the least-conserved region among integrases, has DNA-binding activity and is presumed to be the target DNA-binding domain because the DNA binding does not show any specificity for viral LTR sequences (17, 21, 22).

Three separate IN activities can be measured *in vitro*, and the first two are likely to occur in a concerted reaction *in vivo* (for a review, see Ref. 23). HIV-1 IN first removes two nucleotides from the 3′ end of each LTR. This process, called 3′ end processing, results in a viral DNA that is reduced in size by two nucleotides at each 3′ end. In the second step, 3′ end joining or strand transfer, the processed 3′ end of one LTR is transferred to a 5′-phosphate via nucleophilic attack on the phosphodiester bond of the host chromosomal DNA. The final

This work supported in part by Grant AI31326 from the United States Public Health Service (W.E.R.).

ABBREVIATIONS: IN, integrase; HIV, human immunodeficiency virus; HIV-1, human immunodeficiency virus type 1; AIDS, acquired immune deficiency syndrome; LTR, long terminal repeat; DCQA, dicaffeoylquinic acid; 1-MO-3,5-DCQA, 1-methoxyoxalyl-3,5-DCQA; MCQA, monocaffeoylquinic acid; DCαRA, dicaffeoyl-α-resorcylic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

reaction that can be measured, disintegration, is the reverse of the strand transfer reaction. This latter reaction is unique in that it can be mediated by the core catalytic domain of HIV IN and does not require the full-length IN molecule, whereas the 3′ end processing and strand transfer reactions require the core catalytic domain and both the carboxyl and amino termini of the IN protein.

Structural analysis of retroviral IN has been hampered by its poor solubility and biophysical properties (24). Recently, the crystal structure of the central core domain (residues 50–212) of HIV-1 IN was determined at 2.5 Å resolution (13). To improve solubility, the core domain was modified through substitution of a lysine for a phenylalanine residue at position 185 (F185K); this substitution resulted in a protein with a markedly improved solubility. The structure of the F185K core domain is a 5-stranded β sheet flanked by helical regions. The overall topology shows a striking similarity to a family of polynucleotidyl transferases, including the Holliday junction resolving enzyme RuvC, the MuA transposase, and RNase H (13).

The crystallographic information on HIV-1 IN has stimulated intense interest in the synthesis of inhibitors of HIV-1 IN. To date, most of these inhibitors of IN have been highly effective in preventing the integration reaction in biochemical assays but have been ineffective at preventing virus replication in tissue culture (25–29), possibly due to an inability of the compounds to enter cells. Unfortunately, none of these compounds has yielded cocrystals with HIV-1 IN. Recently, however, a new class of IN inhibitors, the 3,5-DCQAs, has been reported that inhibits HIV replication in tissue culture (30).

We report the anti-HIV-1 activities of several structural isomers of 3,5-DCQA. These compounds, 1,5-DCQA, 3,4-DCQA, and 4,5-DCQA, are abundant (31). Like 3,5-DCQA and 1-MO-3,5-DCQA, these compounds inhibit both the full-length and the core catalytic domain of HIV-1 IN *in vitro* and HIV-1 replication *in vivo*. In addition, because cocrystals of drug and either full-length or truncated IN have been unsuccessful to date, molecular modeling of several IN inhibitors with the F185K point mutant of the core catalytic domain of HIV-1 IN was used to identify seven conserved amino acids likely to be important for inhibitor activity and HIV-1 integration.

Materials and Methods

Cells and Virus

All cell lines were cultivated in RPMI-1640 containing 25 mM HEPES (Mediatech, Herndon, VA) and supplemented with L-glutamine and 12% fetal bovine serum (Irvine Scientific, Santa Ana, CA) (growth medium). MT-2 cells are a T lymphoblastoid cell line that is highly susceptible for HIV infection and are lysed completely by HIV (32). HIV_{LAI} was obtained from the National Institutes of Health AIDS Reagent Repository and was propagated in H9 cells.

Cell Toxicity and Anti-HIV Assays

Cell toxicity and antiviral assays were performed as reported previously (32). Briefly, DCQAs were diluted 1:1 in growth medium, filter sterilized, then serially diluted 2-fold from 1:8 to 1:1280 in triplicate wells of a microtiter plate. To each 50 μl of diluted drug, 50 μl of growth medium was added, followed by 100 μl of MT-2 cell suspension (2×10^5 cells). Cells were incubated with drug for 48 hr at 37° and then harvested for cell viability in a neutral red dye assay as described previously (32). The lethal dose was defined as 50% inhibition of MT-2 cell growth in 48 hr (LD_{50}). Most of the compounds were not available in sufficient quantity or demonstrated a solubility profile that precluded determination of a true LD_{50}. For these compounds, cell toxicity is defined as greater than the maximum concentration of compound tested.

Anti-HIV assays were performed as described previously (32). Based on cell toxicity data, DCQAs were diluted in growth medium so that a final 1:4 dilution of the sample would result in a concentration of sample that inhibited MT-2 cells by 5% (LD_5). The LD_5 falls within 1 SD of controls treated only with solvent and is therefore defined as a nontoxic concentration. The drugs were then serially diluted 2-fold in triplicate. To each 50 μl of diluted compound, 50 μl of HIV_{LAI} was added, and the virus/drug mixture was incubated for 1 hr at 37°. Next, 100 μl of MT-2 cell suspension (2×10^5 cells) was added to each well, and cells were incubated for 72 hr at 37°. The final multiplicity of infection was 1–5. Cells were harvested to quantify cytopathic effect using a neutral red dye assay as described previously (32). The antiviral concentration reported is the concentration of sample necessary to protect MT-2 cells from 50% viral-induced cell death; this is referred to as the 50% effective dose (ED_{50}).

Isolation and synthesis of DCQAs. Caffeic acid, quinic acid, and chlorogenic acid were purchased from Aldrich Chemical (Milwaukee, WI). The DCQAs were obtained as follows: 1,5-DCQA was isolated from *Achyrocline alata*, 4,5-DCQA was isolated from *Baccharis genistelloides*, 3,5-DCQA was obtained from *B. genistelloides* as described previously (33), and 3,4-DCQA was synthesized as described previously (34). Compounds were pure by ^{1}H NMR analysis. All compounds were layered with argon and stored at −80° as sterile-filtered aqueous solutions until tested for cell toxicity and anti-HIV activity.

IN reactions. The 3′ end processing, strand transfer, and disintegration activities of IN in the presence and absence of inhibitors was assayed *in vitro* as modified from Chow *et al.* (19) and Vincent *et al.* (35). The following oligonucleotides (Operon Technologies, Alameda, CA) were used as DNA substrates: T1 (16 mer), 5′-CAGCAACGCAAGCTTG-3′; T3 (30 mer), 5′-GTCGACCTGCAGCCCAAGCTTGCGTTGCTG-3′; V2 (21 mer), 5′-ACTGCTAGAGATTTTCCACAT-3′; V1/T2 (33 mer) 5′-ATGTGGAAAATCTCTAGCAGGCTGCAGGTCGAC-3′; C220 (21 mer), 5′-ATGTGGAAAATCTCTAGCAGT-3′; and B2-1 (19 mer), 5′-ATGTGGAAAATCTCTAGCA-3′. The oligonucleotides were purified by electrophoresis through a 15% denaturing polyacrylamide gel. Oligonucleotides T1, C220, and B2-1 were labeled at the 5′ end using T4 polynucleotide kinase and [γ-^{32}P]ATP (6000 Ci/mmol; Amersham). The substrate for 3′ end processing and strand transfer reactions, which corresponds to the terminal 21 nucleotides of the U5 end of viral DNA, was prepared by annealing the labeled C220 strand with its complementary oligonucleotide, V2. The preprocessed substrate, which resembles the viral U5 end after 3′ end processing and was used to assay only strand transfer activity, was prepared by annealing the labeled B2-1 strand with the V2 strand. The substrate for assaying disintegration activity, the Y oligomer, was prepared by annealing the labeled T1 strand with oligonucleotides T3, V2, and V1/T2 (35). In a 20-μl volume, the DNA substrate (0.1 pmol) was incubated with recombinant IN for 30 min at 37° in a buffer containing a final concentration of 20 mM HEPES, pH 7.5, 10 mM dithiothreitol, 0.05% Nonidet P-40, and 10 mM $MnCl_2$. For assays of inhibitor activity, 1 μl of inhibitor at various concentrations in solvent or solvent alone was added to 19 μl of each reaction mixture containing buffers, labeled-oligonucleotides, and enzyme. The reaction was stopped by the addition of EDTA to a final 18 mM concentration. Reaction products were heated at 90° for 3 min before analysis by electrophoresis on a 15% polyacrylamide gel with 7 M urea in Tris/borate/EDTA buffer (1× = 89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.0). All reactions using the full-length IN were performed at enzyme excess, and reactions were stopped within the linear range of the reaction (19). For experiments using the core

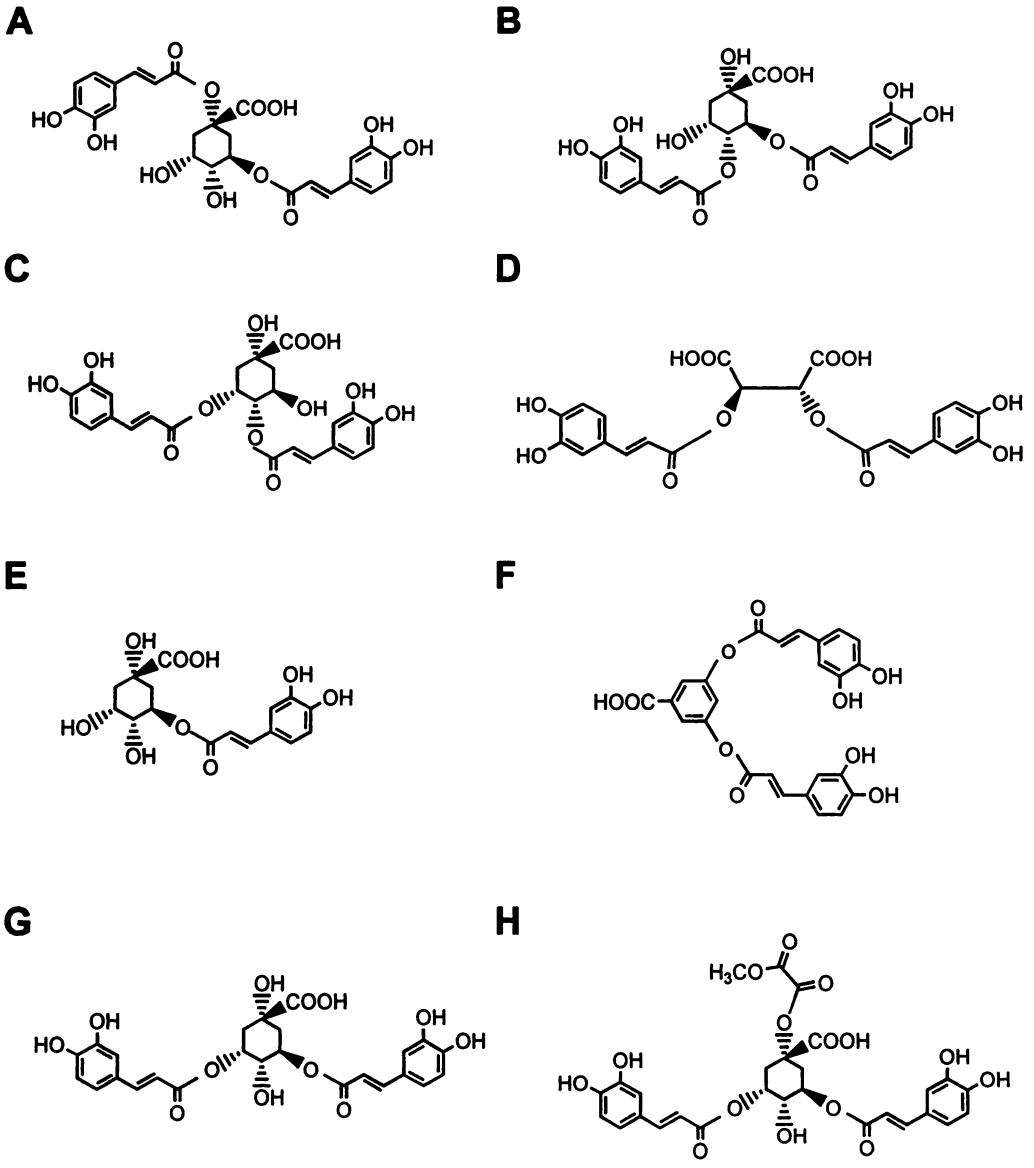

Fig. 1. The DCQAs and several analogues: 1,5-DCQA (A), molecular weight = 516; 4,5-DCQA (B), molecular weight = 516; 3,4-DCQA(C), molecular weight = 516; L-chicoric acid (D), molecular weight = 474; chlorogenic acid (E), molecular weight = 353; DCαRA (F), molecular weight = 478; 3,5-DCQA (G), molecular weight = 516; 1-MO-3,5-DCQA (H), molecular weight = 602.

domain, disintegration reactions were allowed to proceed to completion (60 min). IC_{50} analysis was performed on logarithmic dilutions of inhibitor, and the IC_{50} value was determined from a minimum of three experiments. The log IC_{50} for each reaction were obtained using the logistic curve model of nonlinear regression according to the following formula:

$$\% \text{ inhibition} = \frac{100}{1 + \left[\frac{\text{conc.}}{IC_{50}}\right]^n}$$

where conc. is the concentration of the compound and n is the Hill slope.

Molecular modeling. *Atomic Coordinates of HIV IN.* The crystallographic structure for the F185K-mutated IN catalytic core at 2.5 Å resolution contains several amorphous individual amino acid R groups (D55, E138, E170, K188, and I191) and a 13-residue segment (I141-P-Y-N-P-Q-S-Q-G-V-I-E-S135) in which both the atomic coordinates for the R groups and the peptide backbone have not been resolved (13). This unresolved segment of IN occurs spatially between the functionally essential amino acids D65 and E159, making analysis of ligand docking to the catalytic site of IN impractical. To provide a suitable structure for ligand docking, the individual amorphous side chains were minimized by substituting the Cβ (alanine) of the amorphous amino acid with the correct R group and minimizing to convergence, allowing movement only of atoms distal to the Cα of the peptide backbone. The missing 13-residue segment was constructed with a commercial software peptide builder, minimized by molecular dynamics/conjugate gradients to convergence using the Dreiding II force field, and inserted into the missing region (BioGraf Software, Molecular Simulations, Waltham, MA). The intact IN core was then subjected to 449 psec of quenched dynamics (0.1 psec at 0.002-psec intervals with limited relief of strain using 200 deepest descent energy minimization steps) allowing movement of only the atoms of the inserted segment until energy convergence was achieved (Root-Mean-Square force = <0.01 kcal/mol/Å). The entire structure was then further minimized by the method of conjugate gradients (<200 steps) to final convergence. This structure, as well as a F185K-to-F back mutation, was used for all ligand docking, visualization, and internal free energy (E) calculations.

Ligand Docking. Each inhibitor was modeled according to the following protocol: 1) the potential energy of the ligand was minimized to convergence using the Dreiding II force field and the method of conjugate gradients (36). 2) The lowest free energy conformer was manually docked into the IN active site by approximat-

ing one of the phenyl rings of the ligand into the cavity in which D64 forms the base and E92 and K59 participate in the walls (13). 3) The energy of the resulting complex ($E_{complex}$) was then minimized to convergence, allowing movement of only the ligand atoms. 4) The interaction potential, ΔE_i, for the ligand was then calculated. This was accomplished by calculating the potential energy of the ligand in the absence of IN (E_{ligand}) and the energy of IN in the absence of ligand (E_{IN}). The interaction potential was calculated as $\Delta E_i = E_{complex} - (E_{IN} + E_{ligand})$.

The magnitude of ΔE_i was then correlated with the log IC_{50} as described previously (37). Deviations from linearity are believed to be entropic differences. The solvent-accessible surface and surface charge density were visualized with Grasp software (Anthony Nichols, Columbia University, New York, NY).

Results

Cell toxicity and anti-HIV activity of inhibitors. Five DCQAs, one MCQA, and two synthetic analogues, L-chicoric acid and DCαRA, were tested for cell toxicity and anti-HIV-1 activity in a tissue culture system. The eight compounds tested are illustrated in Fig. 1. The cell toxicity and anti-HIV-1 activity of 1-MO-3,5-DCQA, 3,5-DCQA, L-chicoric acid, and the MCQA, chlorogenic acid, have been described previously (30). The cell toxicity and antiviral results of one of the active compounds (L-chicoric acid, Fig. 2D) and one of the inactive compounds (chlorogenic acid, Fig. 2E) are illustrated in Fig. 2 for comparison with the activities of the newly described DCQAs. The LD_{50} values for all of the active compounds were >120 μg/ml against MT-2 cells (Table 1). The ED_{50} values for these compounds ranged from 1 μg/ml for the most active compound to 6 μg/ml for the least active compound (Table 1). The dose-response curves for each of the compounds are illustrated in Fig. 2.

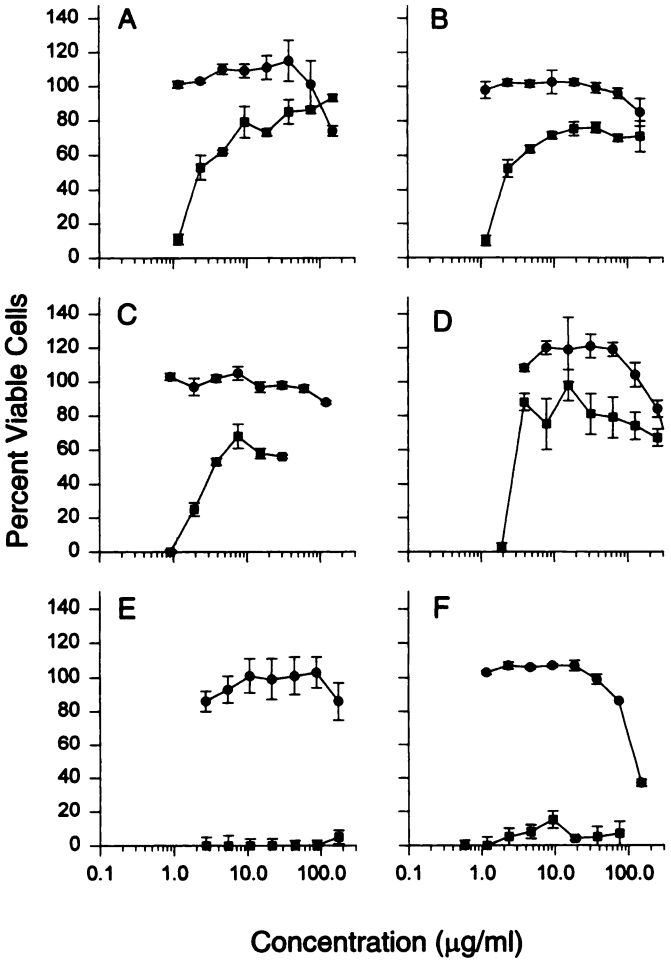

Fig. 2. Activity of DCQAs and their analogues *in vivo*. The cell toxicity of each compound ● in MT-2 cells and the anti-HIV ■ activity of each compound are illustrated. A, 1,5-DCQA. B, 4,5-DCQA. C, 3,4-DCQA. D, L-Chicoric acid. E, Chlorogenic acid. F, DCαRA.

DCQAs inhibit HIV-1 IN *in vitro*. All of the DCQAs inhibited HIV-1 replication at nontoxic concentrations. As the previous studies on 3,5-DCQA, 1-MO-3,5-DCQA, and L-chicoric acid demonstrated, these compounds inhibited HIV-1 IN *in vitro*. The activity of the other stereoisomers of DCQA were tested for inhibitory activity against HIV-1 IN. As illustrated in Fig. 3, three separate functions of HIV-1 IN can be measured *in vitro*. HIV-1 IN first removes two nucleotides from the 3′ end of each LTR. This process, called 3′ end processing, results in a DNA that is reduced in size by two nucleotides from each 3′ end. For *in vitro* reactions using oligonucleotides in place of the viral LTR, the resulting processed oligonucleotide can be visualized after separation on a denaturing polyacrylamide gel. In the second process, 3′ end joining or strand transfer, the processed 3′ end of one LTR is transferred to another oligonucleotide. The resulting integrated DNAs are all larger than the starting materials and can be visualized after separation of the oligonucleotides via denaturing polyacrylamide gel electrophoresis. The final reaction that can be measured, disintegration, is the reverse of the 3′ end joining step. This reaction is unique in that it can be mediated by the core catalytic domain of HIV IN and does not require the full-length IN molecule. Therefore, compounds that inhibit the disintegration reaction likely act at the catalytic core.

As shown in Fig. 4A, all of the DCQAs as well as the L-chicoric acid inhibited both the 3′ end processing and 3′ end joining (strand transfer) reactions (Fig. 4A) at concentrations of 1 μg/ml. As illustrated in Fig. 4B, all of the DCQAs and L-chicoric acid also inhibited the disintegration reaction. The MCQA, chlorogenic acid, weakly inhibited these reactions but only at concentrations significantly higher than those of the DCQAs. DCαRA was unique in that it inhibited all three reactions at concentrations nearly 1 log greater than the weakest DCQA. It was much more effective at blocking the 3′ end processing reaction (Fig. 5A) than either the 3′ end joining reaction (Fig. 5B) or the disintegration reaction (Fig. 5C). The IC_{50} values for each of the compounds in each reaction are summarized in Table 1. Dose-response curves for the three reactions are illustrated in Fig. 5. All of the reactions were studied at enzyme excess in the linear portion of the curve. Results are mean ± 1 SD for three to five separate reactions. For all reactions, the mean percent conversion of substrate into products in solvent-treated control reactions was ~25% (range, 5.8% to 35.8%).

Because the highly conserved, core catalytic domain was used to determine the crystallographic coordinates of HIV-1 IN and because these compounds inhibited the disintegration reaction using the whole recombinant IN molecule, the activities of several of the DCQAs against the truncated polypeptide, amino acids 50–212 of HIV-1 IN, were studied. As illustrated in Fig. 6, the DCQAs inhibited the disintegration reaction catalyzed by HIV-1 IN at concentrations of 5 μg/ml. The IC_{50} values for these reactions were not deter-

TABLE 1

IC_{50} of IN inhibitors in biochemical assays

For specific assay conditions see Materials and Methods. A schematic representation of each assay is illustrated in Fig. 3.

Compound	HIV replication *in vivo*		Integration assays *in vitro*		
	LD_{50}	ED_{50}	3′ End processing	3′ End joining	Disintegration
			μg/ml		
Chlorogenic acid	>350	>176	>50	>50	>50
Caffeic acid	>250	>250	>50	>50	>50
L-Chicoric acid	333	2	0.07 ± 0.01	0.06 ± 0.04	0.15 ± 0.07
1,5-DCQA	>150	2	0.35 ± 0.04	0.56 ± 0.14	0.84 ± 0.08
4,5-DCQA	>150	2	0.13 ± 0.02	0.24 ± 0.08	0.30 ± 0.03
3,4-DCQA	>120	6	0.41 ± 0.09	0.28 ± 0.09	0.71 ± 0.31
3,5-DCQA	>150	1	0.33 ± 0.12	0.34 ± 0.22	0.66 ± 0.24
1-MO-3,5-DCQA	>350	4	0.27 ± 0.08	0.26 ± 0.08	0.47 ± 0.28
DCαRA	100	>75	0.66 ± 0.10	2.25 ± 0.46	6.92 ± 1.79

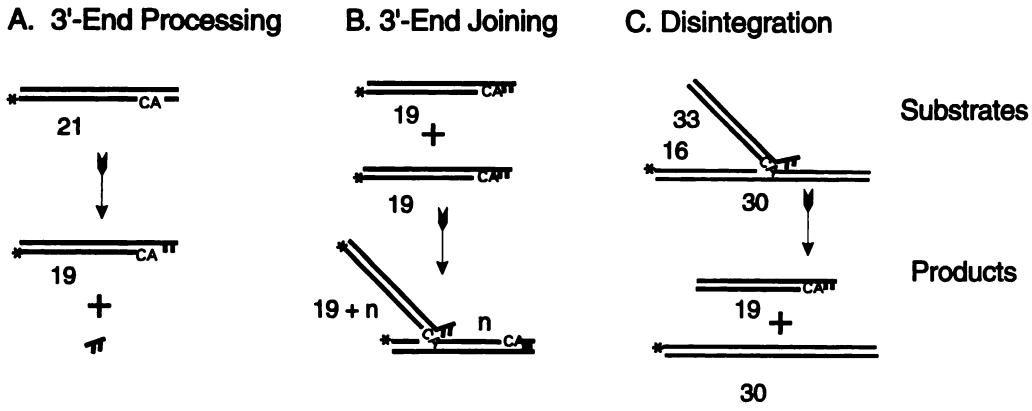

Fig. 3. *In vitro* measurements of IN activity. As shown, three different measures of HIV-1 IN can be measured *in vitro*. These include the removal of two nucleotides from the 3′ end of each oligonucleotide representing the HIV LTR (3′ end processing), the nucleophilic attack of processed oligonucleotides on the phosphodiester bonds of another oligonucleotide (3′ end joining or strand transfer), and the reversal of the strand transfer reaction measured using a Y substrate (disintegration).

mined. However, at that concentration, the activities of the compounds were similar to those obtained using the full-length IN protein (Figs. 4 and 5 and Table 1).

Molecular modeling of a truncated HIV-1 IN mutant and selected IN inhibitors. Because the X-ray crystallographic coordinates of the core catalytic domain containing the F185K mutation have been published (13), it was possible to accurately model the HIV-1 IN protein. Using BioGraf software, the HIV-1 IN protein was modeled, and the amorphous, nonvisualized amino acids within the core catalytic domain (amino acids 141–153) were reconstructed. Both the F185K mutant and a backmutant (F185K-to-F) were modeled. Once the coordinates were entered into the software, each IN inhibitor was modeled and allowed to achieve a minimum free energy state. The IN and ligand were then allowed to dock, and the change in internal free energy between the ligand and the ligand/IN complex was computed as described previously (38). As illustrated in Fig. 7, A–F, each of the compounds docked within the core domain of HIV-1 IN. Indeed, the compounds interacted with a region containing several amino acids described by others as important in the integration reaction. These amino acids and their predicted interatomic distances to key constituents of the L-chicoric acid are summarized in Table 2.

Once the modeling was completed, the change in internal free energy in kcal/mol was plotted versus the natural log of the IC_{50} value of each integration reaction. The plot for all three reactions demonstrated a relatively linear relationship, as illustrated in Fig. 8. Linear regression for each line using only the raw data resulted in an R^2 of 0.919 for the 3′ end processing reaction ($p < 0.001$), 0.66 for the strand transfer reaction ($p = 0.026$), and 0.675 for the disintegration reaction ($p = 0.023$).

Discussion

Integration of HIV-1 into the host chromosome is absolutely necessary for productive infection of the host cell. Since the X-ray crystallographic coordinates of the HIV-1 IN core domain were published, there has been a significant effort placed on synthesis of HIV-1 IN inhibitors. Several groups have described compounds that inhibit HIV-1 IN *in vitro*, yet none of these compounds have been reported to inhibit HIV-1 replication *in vivo*. Recently, 3,5-DCQA, 1-MO-3,5-DCQA, and L-chicoric acid were reported to inhibit full-length HIV-1 IN *in vitro* and HIV-1 replication *in vivo*. To determine whether such activity was reserved for the 3,5 substitution or whether DCQAs could inhibit HIV-1 replication *in vivo* regardless of the substitution pattern around the quinic acid ring, 1,5-DCQA, 3,4-DCQA, and 4,5-DCQA were isolated or synthesized. All of these stereoisomers were potent inhibitors of HIV-1 IN *in vitro*, and all inhibited HIV-1 replication at concentrations similar to the parent compounds described previously (Table 1 and Figs. 2, 4, and 5). Thus, activity against the full-length HIV-1 IN, like the anti-HIV activity of these compounds in tissue culture, was dependent not on the substitution pattern but, instead, on the presence of two caffeic groups rather than one. Each of the DCQAs was also active against the highly conserved core catalytic domain of HIV-1 (Fig. 6).

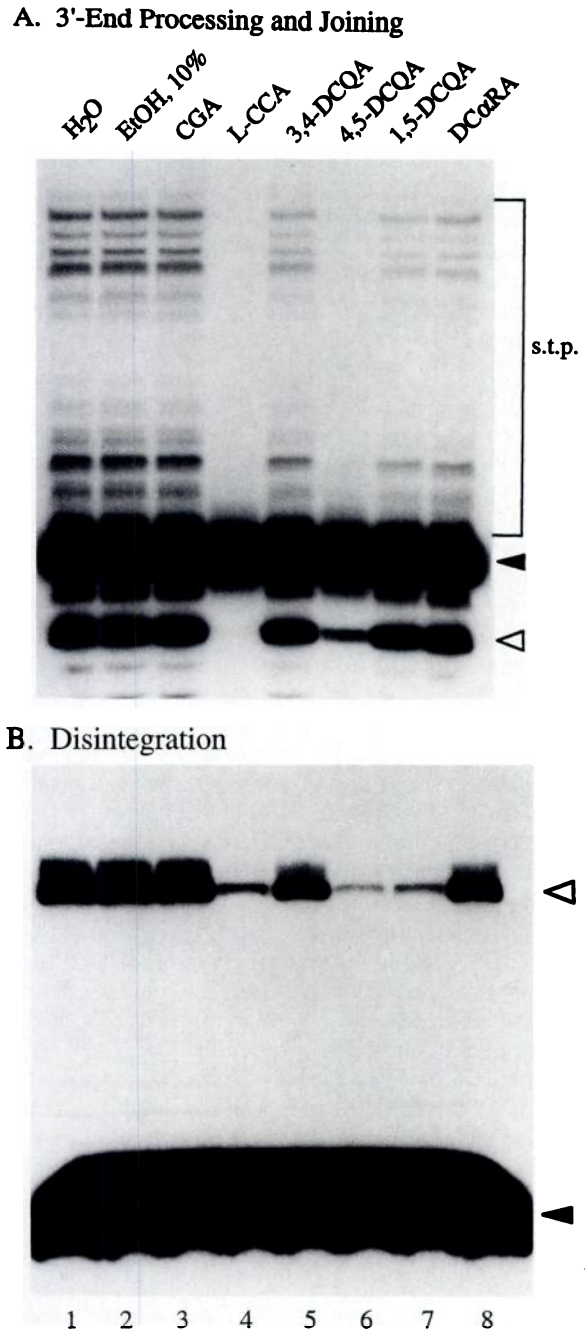

Fig. 4. Activity of the DCQAs against HIV-1 IN. Each compound was tested for the ability to block HIV-1 IN in three separate assays. A, The 3′ end processing and 3′ end joining reactions. *Open arrowhead*, products of the 3′ end processing reaction are two nucleotides smaller than the substrate; *closed arrowhead*, position of the labeled substrate; *s.t.p.*, strand transfer products (3′ end joining products), which are larger than the substrate. B, The disintegration reaction. *Open arrowhead*, disintegration reaction products are larger than the substrate; *closed arrowhead*, position of the labeled substrate. *Lane 1*, H_2O solvent; *lane 2*, 10% ethanol solvent (*EtOH*, 10%); *lane 3*, chlorogenic acid (*CGA*); *lane 4*, L-chicoric acid (L-*CCA*); *lane 5*, 3,4-DCQA; *lane 6*, 4,5-DCQA; *lane 7*, 1,5-DCQA; *lane 8*, DCαRA.

To date, no cocrystals between HIV IN and any HIV IN inhibitor have been described. Likewise, because of the difficulties associated with crystallization of IN, cocrystals may not be achieved. Thus, to determine whether molecular modeling of the HIV-1 IN and IN inhibitors could be used to

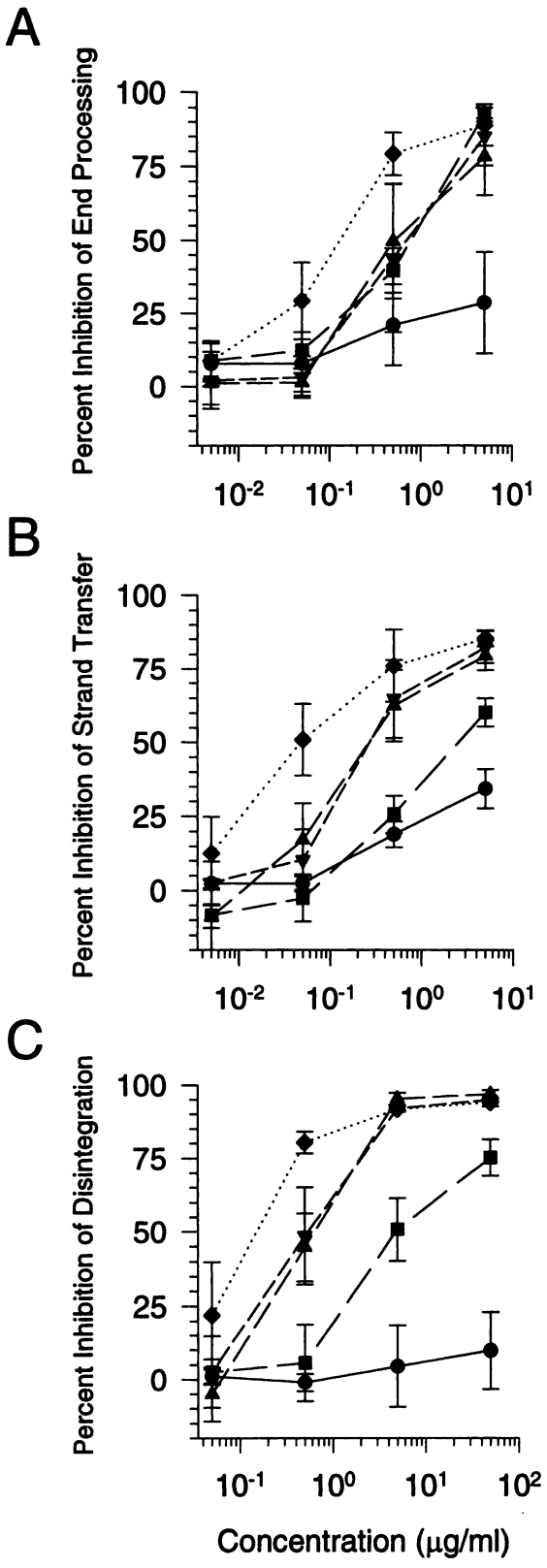

Fig. 5. Inhibition of full-length HIV-1 IN by DCQAs *in vitro*. Dose-response curves used to calculate IC_{50} values for five of the compounds. A, The 3′ end processing reaction. B, The 3′ end joining or strand transfer reaction. C, The disintegration reaction. Compounds tested in each reaction were chlorogenic acid (●), DCαRA (■), 3,5-DCQA (▲), 1-MO-3,5-DCQA (▼), and L-chicoric acid (◆). Note that the range for the disintegration reaction was 0.05–50 μg/ml (C) compared with 0.005–5.0 μg/ml for the other reactions (A and B).

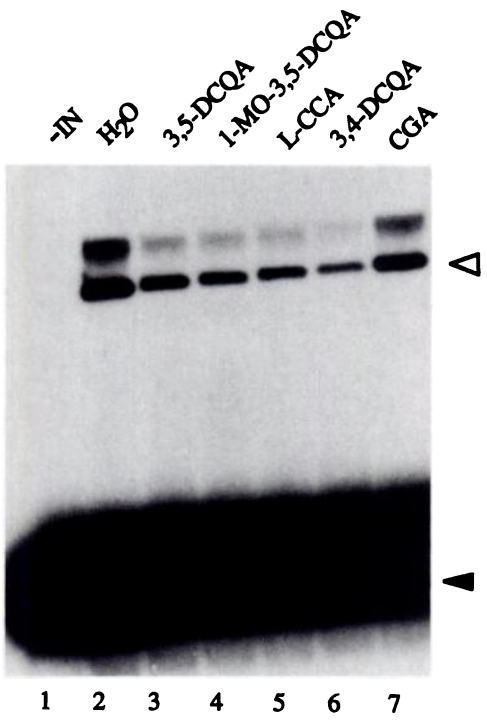

Fig. 6. Inhibition of HIV-1 IN core catalytic domain by DCQAs. The core catalytic domain of HIV-1 IN (amino acids 50–212) was expressed in *Escherichia coli* and purified. Disintegration catalyzed by this protein fragment was assessed in the presence of each DCQA or analog at a final concentration of 5 μg/ml. The compounds tested were *lane 1*, substrate without enzyme (−IN); *lane 2*, H_2O solvent; *lane 3*, 3,5-DCQA; *lane 4*, 1-MO-3,5-DCQA; *lane 5*, L-chicoric acid (*L-CCA*); *lane 6*, 3,4-DCQA; and *lane 7*, chlorogenic acid (*CGA*). ◁, products; ◀, labeled substrate.

TABLE 2

Molecular interactions between amino acids 50–212 of the HIV-1 F185K mutant IN and L-chicoric acid

HIV-1 IN was modeled into BioGraf and Cerius II software based on the published X-ray crystallographic coordinates. The amorphous core domain was reconstructed, and the ligand was allowed to dock by computational docking as described in Materials and Methods. The predicted intermolecular distance between L-chicoric acid moieties and specific amino acids of the HIV-1 IN were determined.

L-Chicoric acid moiety	Amino acid	Predicted intermolecular distance
Carbonyl group	Asp64	7.8 Å
Phenol group	His67	3.0 Å
Phenol group	His67	8.0 Å
Phenol group	Glu92	6.22 Å
Carbonyl group	Asp116	6.9 Å
Phenol group	Asn117	4.6 Å
Carboxylic acid	Gln148	5.9 Å
Carboxylic acid	Lys159	7.15 Å
Phenol group	Lys159	5.26 Å

design better inhibitors of HIV-1 IN, the IN protein was modeled with several of the inhibitors. Docking of IN inhibitors at the active site of the enzyme was made possible by molecular modeling of the amorphous regions within the crystallographic structure. The lowest free energy conformer of each inhibitor was manually docked into the groove defined by the D,D(35)E domain of the active site. Because the X-ray structure is devoid of water at the catalytic site of IN, no solvation terms were used in these molecular modeling procedures. Using repetitive cycles of molecular dynamics and energy minimization, each inhibitor was allowed to conform to the fixed, desolvated macromolecule until energy convergence and lowest internal free energy of the ligand/protein complex were achieved. As illustrated in Fig. 7, one could predict which compounds best blocked HIV-1 IN based on the ability of the compounds to interact with a cleft region of high negative potential, identified in part by amino acids His67, Glu92, Asp116, Asn117, Gln148, Lys159, and Asp64 (at the base of the cleft). L-Chicoric acid, the most potent HIV-1 IN inhibitor, interacted with amino acids within this domain (Table 2). When L-chicoric acid, dicaffeoyl-α-resorcylic acid, and 1-MO-3,5-DCQA were modeled into the IN, only the compounds that filled a groove visualized along the IN protein potently inhibited HIV-1 IN (Fig. 7). The compounds that filled the groove are ranked from best to worst: L-chicoric acid > 1-MO-3,5-DCQA > DCαRA > chlorogenic acid.

In addition to direct visualization of the predicted binding of the DCQAs to HIV-1 IN, the change in internal free energy of the complex versus the sum of their component energies (i.e., ΔE_i) plotted against the natural log of the IC_{50} value for each integration reaction resulted in a linear relationship (Fig. 8). This same linear correlation has been reported for compounds that block gp120 binding to CD4 if the docking reaction measured is the interaction between each ligand and HIV-1 gp120(38a). The calculated correlation coefficient between IC_{50} and ΔE_i suggests that modeling of HIV-1 IN could be used to design better inhibitors of HIV-1 IN. It should be noted, however, that it is unknown whether the lack of activity of the DCαRA against HIV_{LAI} *in vivo* was due to an energetically unfavorable reaction (ΔEi of −13 kcal/mol) with IN or whether the compound failed to enter the cell; however, because the LD_{50} of the DCαRA was greater than that of the other compounds tested, this latter interpretation of the data is unlikely. Thus, modeling of new compounds that leads to more significant changes in internal free energy than that observed for the most active compound, L-chicoric acid, may lead to more active anti-HIV compounds.

The interactions of L-chicoric acid with Asp64, His67, Glu92, Asp116, Asn117, Gln148, and Lys159 predicted by molecular modeling are consistent with the functional roles of those amino acids. Previous work by others has indicated that mutation of several amino acids within the HIV-1 or HIV-2 IN affects the ability of HIV-1 to both productively infect CD4+ lymphocyte cell lines and function in *in vitro* integration assays (3, 15–17, 39). Mutations of Asp64 and Asp116, two components of the D,D(35)E domain, are associated with nonviable HIV-1 that cannot mediate *in vitro* integration. Mutation of Asn117 is associated with defective replication as indicated by delayed replication kinetics (3) and a partial loss of IN function *in vitro* (39). Mutation of Lys159 had no effect on viral replication, although van Gent *et al.* (39) described a partial loss of IN activity. Mutation of Gln148 has been shown to lead to a partial loss of IN activity *in vitro* (39). It should be stressed that mutations described by van Gent *et al.* were made in the HIV-2 IN rather than the HIV-1 IN described by others. Mutations of His67 and Glu92 have not been described to date. Thus, the amino acids identified as interacting with the L-chicoric acid by molecular modeling are also, for the most part, important in IN activity and viral replication. In addition, the use of HIV-1 IN inhibitors coupled with molecular modeling of IN/ligand interactions has indicated at least two additional amino acids that could be important in HIV-1 IN function that have not been

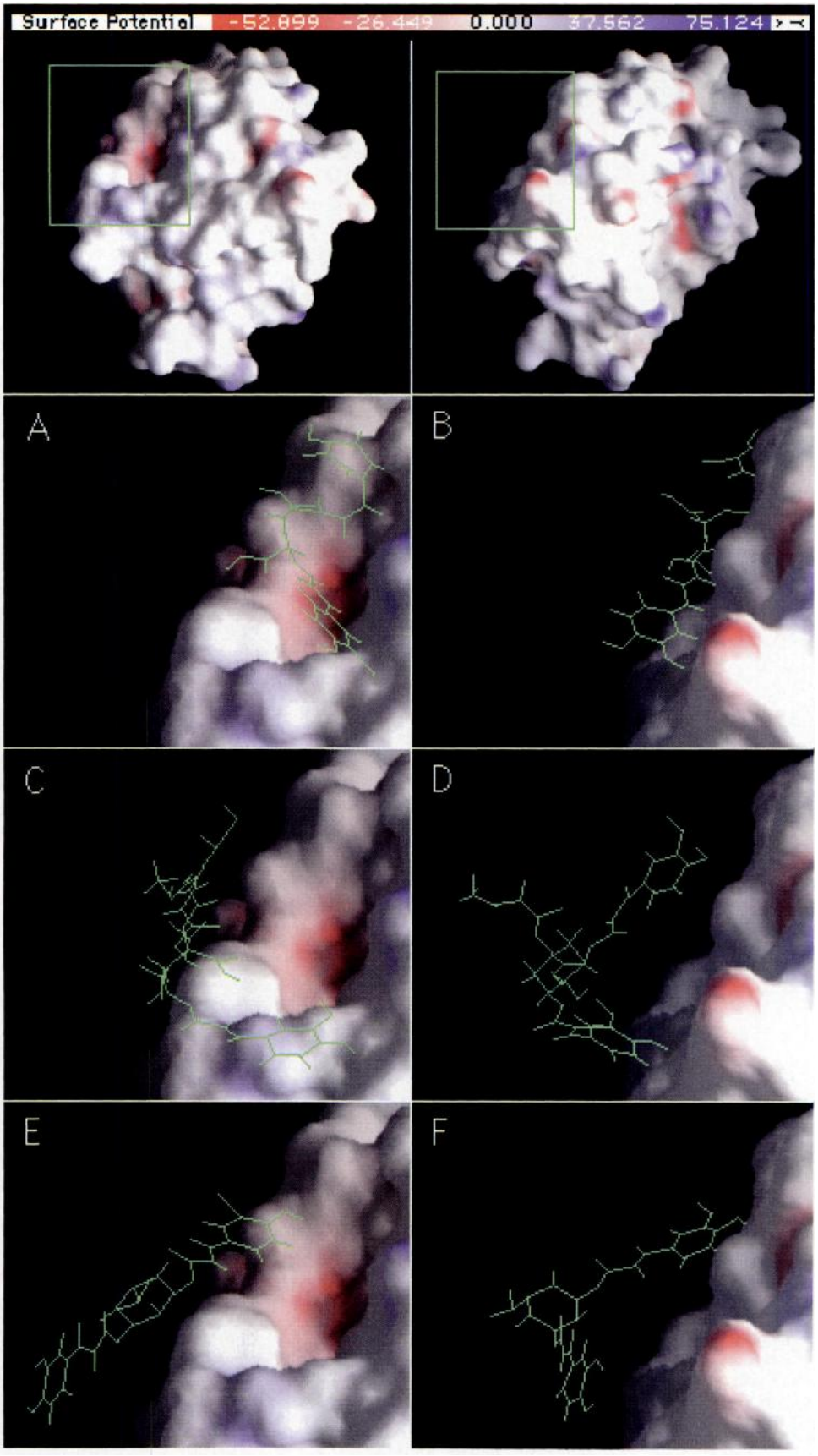

Fig. 7. Modeling of several DCQAs and their analogs with amino acids 50–212 of an F185K mutant of HIV-1 IN. Four ligands were modeled into the HIV-1 IN using computational docking (see Materials and Methods). *Top*, modeled F185K HIV-1 IN. *Green box*, area shown below each IN molecule. The compounds modeled included L-chicoric acid (A and B), 1-MO-3,5-DCQA (C and D), and DCαRA (E and F), The docking stations illustrated are station 0 (frontal view of compound docked into the cleft; A, C, and E) and docking station 1 (clockwise rotation of the IN/ligand complex 45° along the *y*-axis to illustrate ligands within a cleft in IN (B, D, and F). *Green*, IN was modeled using a solvent-accessible surface model with each compound. *Red*, regions with a calculated high net negative charge. *Blue*, regions with a calculated high net positive charge.

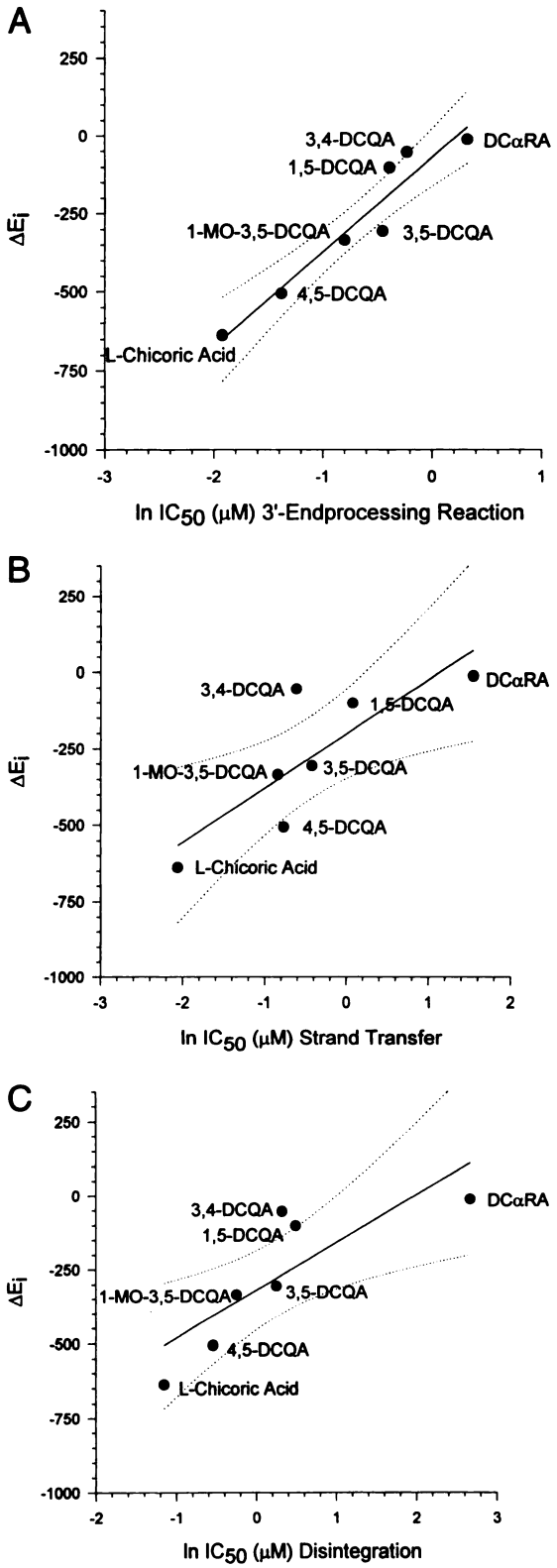

Fig. 8. Correlation between anti-HIV-1 IN activity and ΔE_i. *Top*, 3′ end processing reaction. *Middle*, 3′ end joining reaction. *Bottom*, disintegration reaction. The change in free energy (ΔE_i) for each ligand/IN product was plotted versus the natural log of the IC_{50} of each compound in each of the three separate IN reactions. *Solid lines*, linear regression analysis for each compound. ●, Observed IC_{50} values. *Curved, dashed lines*, 95% confidence intervals for each line. Data are for the F185K-to-F back-mutation; similar results were obtained using the F185K mutant.

identified through other methods. Mutation of these residues may lead to further information on amino acids that are important for HIV-1 IN activity. Although molecular modeling has limitations, the strong correlation between the modeling and the observed inhibition of IN (Fig. 8), combined with the significance of the amino acids identified to interact with the inhibitors (Table 2), makes the model, in the absence of cocrystals, a potentially important tool for designing better inhibitors of HIV IN.

Despite the critical role played by IN in the retroviral life cycle, there is little information concerning chemical compounds that show selective inhibition against the catalytic activities of IN. Other than the DCQAs, the major classes of IN inhibitors that have been reported to date include aurintricarboxylic acid (27) and cosalene analogues (28), caffeic acid phenylethyl ester (25, 26), DNA binding agents (25, 26), topoisomerase inhibitors (25), and bis-catechols (29). A majority of the compounds reported thus far are not selective for IN. Aurintricarboxylic acid and related compounds also inhibit reverse transcriptase and other phosphoryltransferases (28). Inhibition of IN by DNA binding agents and topoisomerase inhibitors is relatively weak and nonselective. The DCQAs reported in the current study have been tested for activity against a variety of HIV enzymes, including gp120/CD4 binding and reverse transcriptase. To date, these compounds are inactive or weakly active against all viral enzymes tested, demonstrating a remarkable degree of selectivity for HIV-1 IN.[1] Perhaps more importantly, most of the information on previously reported IN inhibitors is derived from *in vitro* experiments using purified IN, and a protective effect of IN inhibitors against HIV infection in tissue culture is either undetectable (29) or has not been examined. Furthermore, the mechanisms of action for any of these compounds in inhibiting IN have not been studied. The DCQAs, as described previously and demonstrated in the current study, are active not only against HIV-1 IN in biochemical reactions but also against HIV in tissue culture. These data suggest that IN can serve as a target for anti-HIV therapeutics and that the development of more potent and selective anti-HIV IN inhibitors could add to our armamentarium of anti-HIV agents.

The current data support the interpretation that the DCQAs and L-chicoric acid block HIV-1 replication directly through inhibition of HIV-1 IN. The results for DCαRA suggest that inhibition of 3′ end processing does not correlate with an *in vivo* effect but that inhibition of disintegration and 3′ end joining reactions can predict *in vivo* activity against HIV-1. In addition, effective IN inhibitors had an IC_{50} value against HIV-1 IN of <500 nM. Most of the other IN inhibitors reported in the literature have been significantly less potent. These results suggest that in the absence of complete crystallographic information, a combination of molecular modeling and *in vitro* assays may be used to both identify amino acids critical for IN activity and design more potent anti-IN compounds, which may ultimately prove to be useful in the treatment of HIV infection and AIDS.

Acknowledgments

We are grateful to Brenda McDougall for excellent technical assistance and Dr. Jerome Gabriel (Temple University, Philadelphia,

[1] W. E. Robinson, Jr., unpublished observations.

PA) for his expert advice and counsel in molecular modeling. We are indebted to Dr. Suzanne Sandmeyer for thoughtful suggestions on the manuscript.

References

1. LaFemina, R. L., C. L. Schneider, H. L. Robbins, P. L. Callahan, K. LeGrow, E. Roth, W. A. Schleif, and E. A. Emini. Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J. Virol.* **66**:7414–7419 (1992).
2. Sakai, H., M. Kawamura, J. Sakuragi, S. Sakuragi, R. Shibata, A. Ishimoto, N. Ono, S. Ueda, and A. Adachi. Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *J. Virol.* **67**: 1169–1174 (1993).
3. Shin, C. G., B. Taddeo, W. A. Haseltine, and C. M. Farnet. Genetic analysis of the human immunodeficiency virus type 1 integrase protein. *J. Virol.* **68**:1633–1642 (1994).
4. Englund, G., T. S. Theodore, E. O. Freed, A. Engleman, and M. A. Martin. Integration is required for productive infection of monocyte-derived macrophages by human immunodeficiency virus type 1. *J. Virol.* **69**:3216–3219 (1995).
5. Engleman, A., G. Englund, J. M. Orenstein, M. A. Martin, and R. Craigie. Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *J. Virol.* **69**:2729–2736 (1995).
6. Roth, M. J., P. Schwartzberg, N. Tanese, and S. P. Goff. Analysis of mutations in the integration function of Moloney murine leukemia virus: effect on DNA binding and cutting. *J. Virol.* **64**:4709–4717 (1990).
7. Heinzinger, N. K., M. I. Bukrinsky, S. A. Haggerty, A. M. Ragland, V. Kewalramani, M.-A. Lee, H. Gendelman, L. Ratner, M. Stevenson, and M. Emerman. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc. Natl. Acad. Sci. USA* **91**:7311–7315 (1994).
8. Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G. Tarpley, and M. Stevenson. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. USA* **90**: 6125–6129 (1993).
9. Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* **3**:469–478 (1989).
10. Kahn, E., J. P. G. Mack, R. A. Katz, J. Kulkosky, and A. M. Skalka. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res.* **19**:851–860 (1991).
11. Johnson, M. S., M. A. McClure, D.-F. Feng, J. Gray, and R. F. Doolittle. Computer analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc. Natl. Acad. Sci. USA* **83**:7648–7652 (1986).
12. Bushman, F. D., A. Engelman, I. Palmer, P. Wingfield, and R. Craigie. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc. Natl. Acad. Sci. USA* **90**:3428–3432 (1993).
13. Dyda, F., A. B. Hickman, T. M. Jenkins, A. Engelman, R. Craigie, and D. R. Davies. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science (Washington D. C.)* **266**:1981–1986 (1994).
14. Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol. Cell. Biol.* **12**:2331–2338 (1992).
15. Engelman, A., and R. Craigie. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol.* **66**:6361–6369 (1992).
16. Leavitt, A. D., L. Shiue, and H. E. Varmus. Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. *J. Biol. Chem.* **268**:2113–2119 (1993).
17. Vink, C., A. M. Oude Groeneger, and R. H. Plasterk. Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type I integrase protein. *Nucleic Acids Res.* **21**:1419–1425 (1993).
18. Berg, J. M. Zinc fingers and other metal-binding domains. *J. Biol. Chem.* **265**:6513–6516 (1990).
19. Vincent, K. A., V. Ellison, S. A. Chow, and P. O. Brown. Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. *J. Virol.* **67**:425–437 (1993).
20. Ellison, V., and P. O. Brown. A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. *Proc. Natl. Acad. Sci. USA* **91**:7316–7320 (1994).
21. Woerner, A. M., M. Klutch, J. G. Levin, and C. J. Marcus-Sekura. Localization of DNA binding activity of HIV-1 integrase to the C-terminal half of the protein. *AIDS Res. Hum. Retroviruses* **8**:297–304 (1992).
22. Engelman, A., A. B. Hickman, and R. Craigie. The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *J. Virol.* **68**:5911–5917 (1994).
23. Goff, S. P. Genetics of retroviral integration. *Annu. Rev. Genet.* **26**:527–544 (1992).
24. Hickman, A. B., I. Palmer, A. Engelman, R. Craigie, and P. Wingfield. Biophysical and enzymatic properties of the catalytic domain of HIV-1 integrase. *J. Biol. Chem.* **269**:29279–29287 (1994).
25. Fesen, M. R., K. W. Kohn, F. Leteurtre, and Y. Pommier. Inhibitors of human immunodeficiency virus integrase. *Proc. Natl. Acad. Sci. USA* **90**:2399–2403 (1993).
26. Fesen, M. R., Y. Pommier, F. Leteurtre, S. Hiroguchi, J. Yung, and K. W. Kohn. Inhibition of HIV-1 integrase by flavones, caffeic acid phenethyl ester (CAPE) and related compounds. *Biochem. Pharmacol.* **48**:595–608 (1994).
27. Cushman, M., W. M. Golebiewski, Y. Pommier, A. Mazumder, D. Reymen, E. De Clercq, L. Graham, and W. G. Rice. Cosalane analogues with enhanced potencies as inhibitors of HIV-1 protease and integrase. *J. Med. Chem.* **38**:443–452 (1995).
28. Cushman, M., and P. Sherman. Inhibition of HIV-1 integration protein by aurintricarboxylic acid monomers, monomer analogs, and polymer fractions. *Biochem. Biophys. Res. Commun.* **185**:85–90 (1992).
29. LaFemina, R. L., P. L. Graham, K. LeGrow, J. C. Hastings, A. Wolfe, S. D. Young, E. A. Emini, and D. J. Hazuda. Inhibition of human immunodeficiency virus integrase by bis-catechols. *Antimicrob. Agents Chemother.* **39**:320–324 (1995).
30. Robinson, W. E., Jr., M. G. Reinecke, S. Abdel-Malek, Q. Jia, and S. A. Chow. Inhibitors of HIV-1 replication that inhibit HIV integrase. *Proc. Natl. Acad. Sci. USA* **93**:6326–6331 (1996).
31. Clifford, M. N. Coffee bean dicaffeoylquinic acids. *Phytochemistry* **25**: 1767–1769 (1986).
32. Montefiori, D. C., W. E. Robinson, Jr., S. S. Schuffman, and W. M. Mitchell. Evaluation of antiviral drugs and neutralizing antibodies against human immunodeficiency virus by a rapid and sensitive microtiter infection assay. *J. Clin. Microbiol.* **26**:231–235 (1988).
33. Veit, M., D. Strack, F.-C. Czygan, V. Wray, and L. Witte. Di-E-caffeoyl-meso-tartaric acid in the barren sprouts of *Equisetum arvense*. *Phytochemistry* **30**:527–529 (1991).
34. Wald, B., V. Wray, R. Galensa, and K. Hermann. Malonated flavonol glycosides and 3,5-dicaffeoylquinic acid from pears. *Phytochemistry* **28**: 663–665 (1989).
35. Chow, S. A., K. A. Vincent, V. Ellison, and P. O. Brown. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science (Washington D. C.)* **255**:723–726 (1992).
36. Fletcher, R., and C. M. Reeves. Function minimization by conjugate gradients. *Comput. J.* **7**:149–154 (1964).
37. Gabriel, J. L., and W. M. Mitchell. HIV gp120 docking interactions and inhibitor design based on an atomic structure derived by molecular modeling using the Dreiding II force field., in *Protein Folds: A Distance Based Approach* (H. Bohr and S. Brunak, eds.). CRC Press, Boca Raton, FL, 294–307 (1996).
38. Gabriel, J. L., and W. M. Mitchell. Proposed atomic structure of a truncated human immunodeficiency virus glycoprotein gp120 derived by molecular modeling: target CD4 recognition and docking mechanism. *Proc. Natl. Acad. Sci. USA* **90**:4186–4190 (1993).

38a. Gabriel, J. A., and W. M. Mitchell. Functional design of potential inhibitors of HIV binding to $CD4^+$ target cells: a molecular model of gp120 predicts ligand binding. *Drug Des. Disc.*, in press (1996).

39. van Gent, D. C., A. A. Oude Groeneger, and R. H. Plasterk. Identification of amino acids in HIV-2 integrase involved in site-specific hydrolysis and alcoholysis of viral DNA termini. *Nucleic Acids Res.* **21**:3373–3377 (1993).

Send reprint requests to: W. Edward Robinson, Jr., M.D., Ph.D., Department of Pathology, D440 Medical Science I-4800, University of California, Irvine, CA 92697-4800. E-mail: ewrobins@uci.edu