Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 199–215
Bioorganic & Medicinal Chemistry

2,6-Bis(3,4,5-trihydroxybenzylydene) derivatives of cyclohexanone: novel potent HIV-1 integrase inhibitors that prevent HIV-1 multiplication in cell-based assays

Roberta Costi,[a] Roberto Di Santo,[a] Marino Artico,[a,*] Silvio Massa,[b] Rino Ragno,[c] Roberta Loddo,[d] Massimiliano La Colla,[d] Enzo Tramontano,[d] Paolo La Colla[d,*] and Alessandra Pani[d]

[a]*Istituto Pasteur-Fondazione Cenci Bolognetti, Dipartimento di Studi Farmaceutici, Università degli Studi di Roma 'La Sapienza', P.le A. Moro 5, I-00185 Rome, Italy*
[b]*Dipartimento Farmaco Chimico Tecnologico, Università degli Studi di Siena, Via A. Moro 5, San Miniato, I-53100 Siena, Italy*
[c]*Dipartimento di Studi di Chimica e Tecnologia delle Sostanze Biologicamente Attive, Università degli Studi di Roma 'La Sapienza', P.le Aldo Moro 5, I-00185 Rome, Italy*
[d]*Dipartimento di Biologia Sperimentale, Sezione di Microbiologia, Università degli Studi di Cagliari, Cittadella Universitaria, I-09042 Monserrato, Cagliari, Italy*

Received 4 April 2003; accepted 7 October 2003

Abstract—A number of 2,6-bisbenzylidenecyclohexane-1-one derivatives have been synthesized and tested as HIV-1 integrase (IN) inhibitors with the aim of obtaining compounds capable to elicit antiviral activity at non-cytotoxic concentrations in cell-based assays. 3,5-Bis(3,4,5-trihydroxybenzylidene)-4-oxocyclohexaneacetic acid (**20d**) resulted one of the most potent and selective derivatives in acutely infected MT-4 cells (EC_{50} and CC_{50} values of 2 and 40 μM, respectively). In enzyme assays with recombinant HIV-1 integrase (rIN), this compound proved able to inhibit both 3′-processing and disintegration with IC_{50} values of 0.2 and 0.5 μM, respectively. In order to develop a model capable to predict the anti HIV-IN activity and useful to design novel derivatives, we performed a comparative molecular field analysis (CoMFA) like 3-D-QSAR. In our model the ligands were described quantitatively in the GRID program, and the model was optimized by selecting only the most informative variables in the GOLPE program. We found the predictive ability of the model to increase significantly when the number of variables was reduced from 20,925 to 1327. A Q^2 of 0.73 was obtained with the final model, confirming the predictive ability of the model. By studying the PLS coefficients in informative 3-D contour plots, ideas for the synthesis of new compounds could be generated.

1. Introduction

Among the virus-coded enzymes that are essential for HIV-1 replication, integrase (IN) plays a fundamental role by inserting the retro-transcribed viral DNA into the host chromosome. Initially, IN recognizes the LTR termini of the linear double-stranded viral DNA molecule, of which it removes two nucleotides from each 3′ end (3′ processing reaction), thus leaving recessed 3′-OH termini. Then, IN catalyzes joining of the latter to the 5′ ends of host DNA strand breaks. Removal of mispaired nucleotides and gap repair lead to provirus formation. Due to its peculiar properties and to the absence of cellular counterparts, IN is an attractive target for selective chemotherapeutic intervention.

Studies performed so far have led to the identification of a great number of HIV-1 IN inhibitors. Most of them have been described in enzyme assays with purified recombinant IN (rIN) and 21^{mer} duplex oligonucleotides reproducing the U5 end of HIV-1 LTRs. However, only a few compounds are able to inhibit rIN at concentrations of 1 μM or lower[1–5] and also to prevent the HIV-1 multiplication in cell-based assays at non-cytotoxic concentrations. Therefore, the search for novel

Keywords: Polyhydroxylated aromatics; Cyclohexanone derivatives; Anti-HIV-1-IN agents; QSAR studies.
* Corresponding authors. Tel./fax: +39-06-446-2731; e-mail: marino.artico@uniroma1.it (M. Artico); Tel.: +39-070-6754147; fax: +39-070-6754210; e-mail: placolla@unica.it (P. La Colla).

0968-0896/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2003.10.005

anti-IN agents active in cell-based assays, other than in enzyme assays, continues.

Natural products and related synthetic analogues[6] are among the compounds investigated as potential IN inhibitors. One of the most thoroughly investigated class[7,8] is represented by hydroxylated aromatics, such as CAPE (**1**),[9] flavones (**2**),[9] curcumin (**3**),[10–13] tyrphostins (**4**),[14] biscatechols (**5**),[15] dicaffeoylquinic acids (**6**),[16,17] L(−)chicoric acid (**7**),[17,18] and digalloyl-L-tartaric acid (**8**).[19] With the exception of **5** and **8**, the above compounds share a 3,4-dihydroxycinnamoyl moiety, sometimes incorporated into a ring structure, which is the likely pharmacophore since its integrity is crucial for maintaining the anti-IN activity. Nevertheless, the lack of this moiety in several potent IN inhibitors, such as styrylquinoline (**9**),[20] aryl dioxobutanoic acids (**10**)[21] and the flavone derivative baicalein (**11**)[9] suggests that additional pharmacophore groups (and, therefore, different modes of interaction with the target enzyme) may exist among hydroxylated aromatics (Fig. 1). This prevents definitive conclusions on their mode of action; in particular, it is still unclear whether the catechol hydroxyls act by chelating the divalent cations (Mg^{++}, Mn^{++}) required for IN catalysis[9] or by donating hydrogen-bonds to specific chemical functions of the enzyme's catalytic core domain.

Recently, we have been engaged in the design, synthesis and biological evaluation of novel curcumin-related derivatives such as **12**.[12] Although they lacked activity in cell-based assays and showed the characteristic cytotoxicity of the catechol system, many derivatives turned out to be potent inhibitors of the rIN in enzyme assays. Interestingly, potent anti-rIN activity correlated with the presence of two styryl moieties bearing an unsubstituted 3,4-dihydroxy group (catechol system) linked to: (i) a ketoalkane; (ii) a cycloalkanone, eventually containing an heteroatom; (iii) a benzene ring (Fig. 2). CoMFA and CoMSIA 3-D QSAR analyses and docking simulations have been recently performed on the above compounds[22] and the importance of hydrogen-bonding interactions in determining binding at the active site has been documented.

In this study, we present novel cinnamoyl derivatives synthesized to establish whether improvement in both cytotoxicity and anti-HIV-1 activity could be obtained maintaining the focus on the catechol system (Fig. 3).

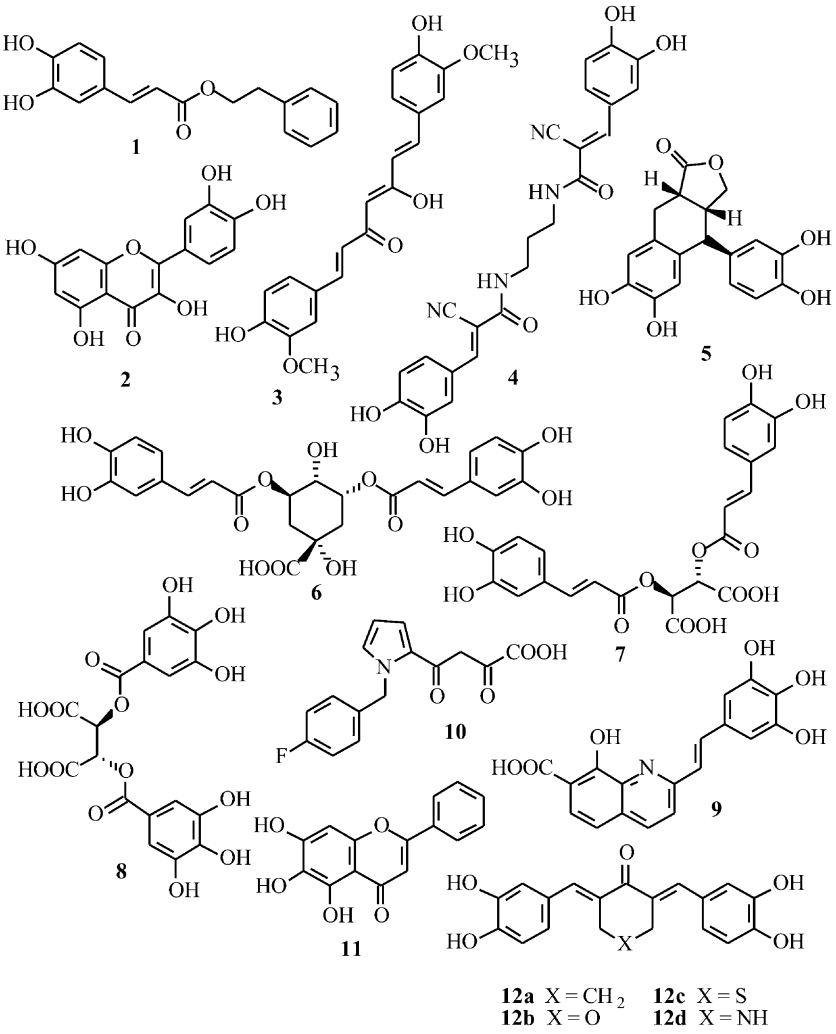


Figure 1. HIV-1 integrase inhibitors.

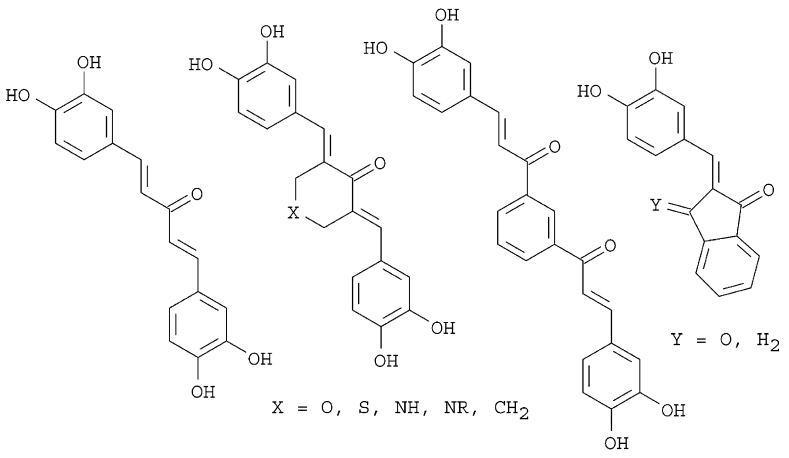


Figure 2. Previously synthesized curcumin-like derivatives.

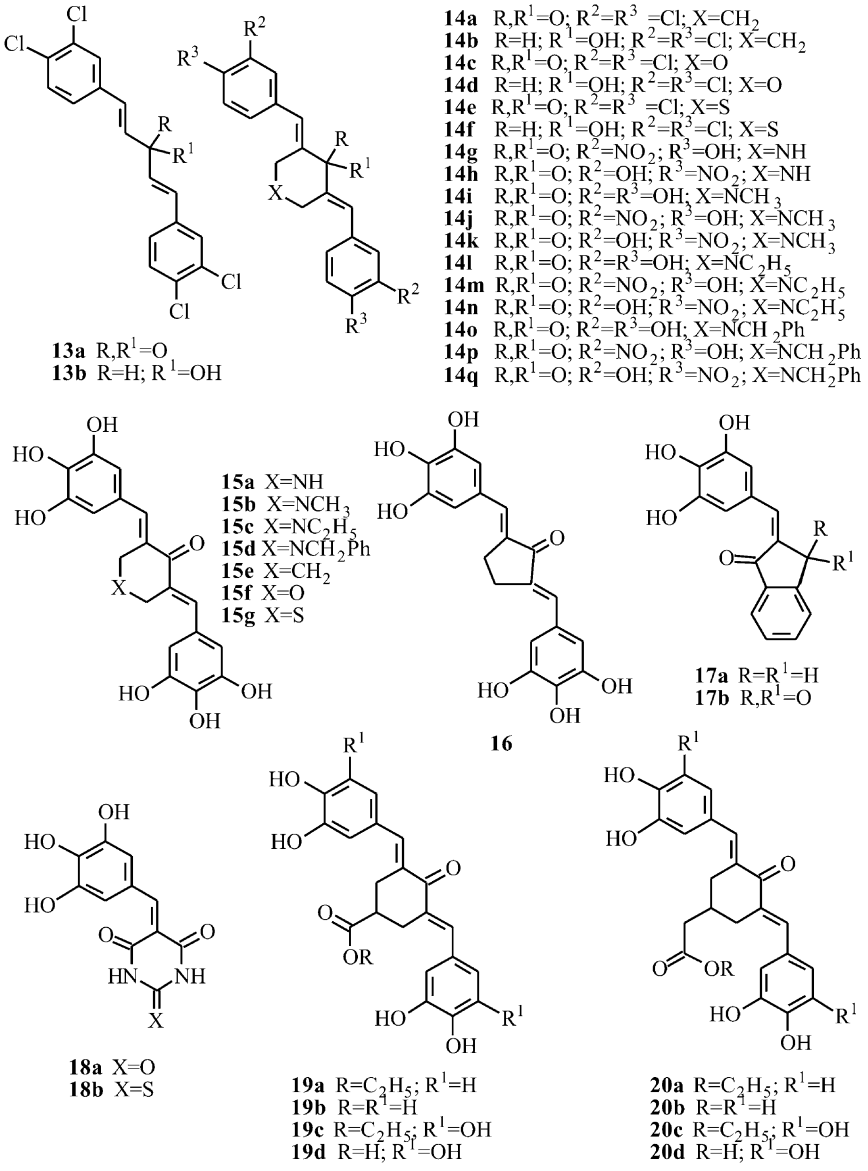


Figure 3. Newly synthesized curcumin-like derivatives.

The rationale for the synthesis was based on the observation that compounds **6–10** are hydroxylated aromatics bearing one or two carboxyl groups and/or 3,4,5-trihydroxycinnamoyl moieties which are capable to inhibit rIN in enzyme assays as well as to prevent the HIV-1 multiplication in cell-based assays.

In particular, the activity shown by styrylquinolines (**9**) led us to hypothesize that compounds containing 3,4,5-trihydroxycinnamoyl moieties could retain the capability to chelate divalent metal ions, or to donate H-bonds, while losing the cytotoxicity peculiar of the catechol system. In addition, the correlation between the presence of carboxyl groups and the activity against both rIN in enzyme assays and the HIV-1 multiplication in acutely infected cells led us to introduce carboxyl groups into both 3,4-dihydroxycinnamoyl and 3,4,5-trihydroxycinnamoyl bearing compounds.

The new bis-2,6-benzylidene derivatives (**13–20**) (Fig. 3) were prepared as outlined in Schemes 1–4 and tested in enzyme and cell-based assays according to previously reported procedures.[12]

Based on previously reported 3-D QSARs[22,23] studies, our aim was to generate a model capable to help in the design of new active compounds starting from the analysis of cinnamoyl derivatives reported in a previous[12] and in this paper. Traditionally, SYBYL/CoMFA[24,25] (comparative molecular field analysis) is the method used to create this kind of models, but additional 3-D quantitative structure–activity relationship (QSAR) methods are available.[26,27] In order to generate molecular descriptors and the GOLPE program[26] for the multivariate regression analyses, the GRID program[27] was used.

2. Chemistry

Condensation of 3,4-dichlorobenzaldehyde with acetone or appropriate heterocycloalkanones in alkaline medium afforded the related bis-benzylidene derivatives **13a**, **14a**,**c**,**e**, which were then reduced to the corresponding alcohols **13b**, **14b**,**d**,**f** by treatment with sodium borohydride (Scheme 1).

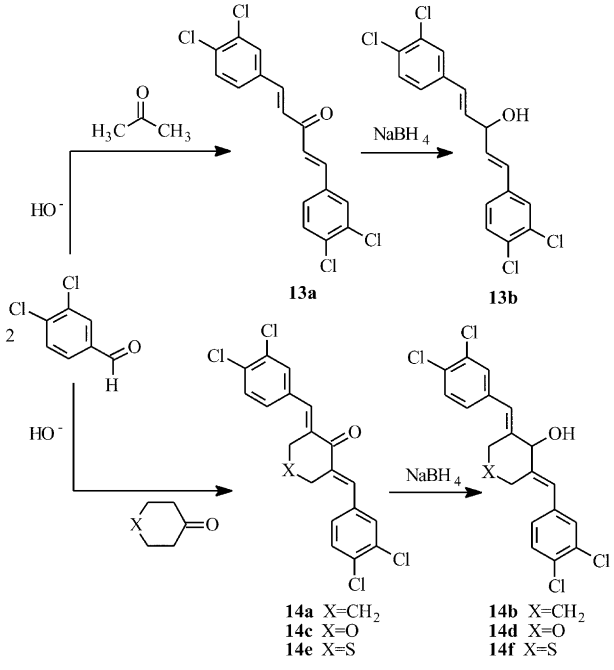
3,4,5-Trihydroxy, 3(4)-OH, 4(3)-NO_2 and 3,4-dihydroxybenzylidene derivatives **14g–q**, **15–17** and **19–20** were synthesized by reacting the corresponding arylaldehyde with the appropriate compound containing active methylene groups, in glacial acetic acid under a stream of gaseous hydrochloric acid, as depicted in Schemes 2–4. Barbituric and thiobarbituric derivatives **18a**,**b** were obtained by condensation of 3,4,5 - trihydroxybenzaldehyde with barbituric or thiobarbituric acid, respectively, in boiling water (Scheme 3).

Ethyl 4-oxocyclohexaneacetate and the corresponding acid, used for the synthesis of derivatives **20a–d**, were synthesized according to literature.[28,29]

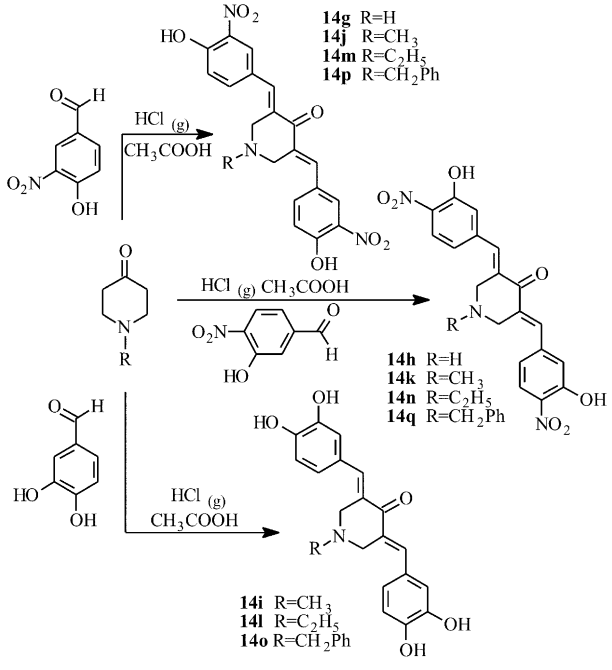
3. Results and discussion

3.1. Anti-HIV-1 cell-based assays and anti-rIN assays

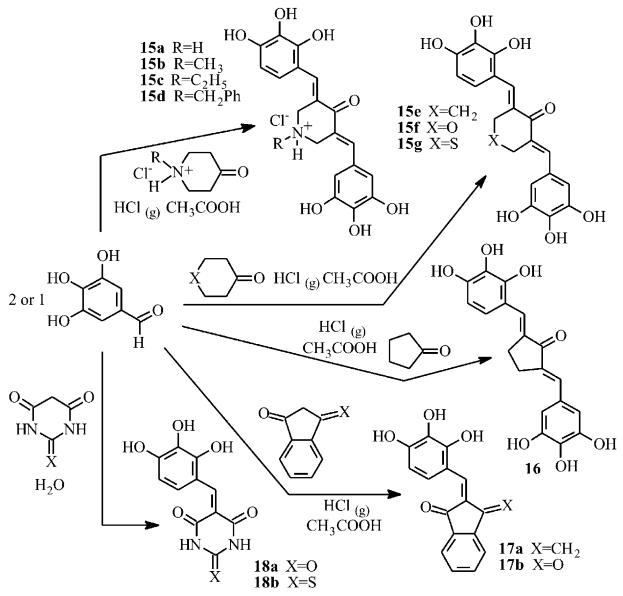
In a previous work, we described the synthesis and biological activity of a series of geometrically restrained cinnamoyl compounds including the 2,6-bis-(3,4-dihydroxybenzylidene)cyclohexanone (**12a**).[12] Although endowed with potent inhibitory activity against rIN in enzyme assays, these compounds proved highly cytotoxic and totally ineffective in preventing the HIV-1 multiplication in acutely infected cells.



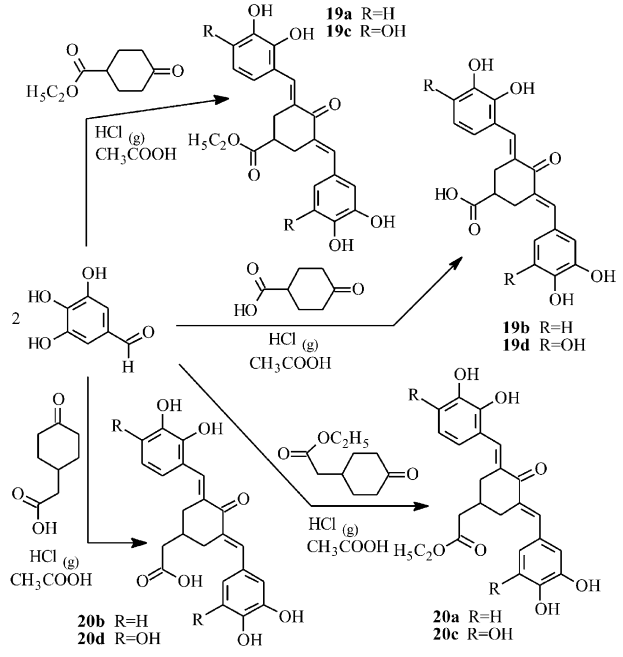
Scheme 1.



Scheme 2.



Scheme 3.



Scheme 4.

Therefore, to investigate whether the catechol moiety could be replaced by 3,4-disubstituted systems comprising or not OH groups, we first prepared and tested for biological activity a series of 3,4-dichlorophenyl, 4-hydroxy-3-nitrophenyl and 3-hydroxy-4-nitrophenyl derivatives together with **9**, **10** and **12a** used as reference drugs (Table 1). No matter whether ketones (**13a**, **14a**, **14c**, **14e**) or alcohols (**13b**, **14b**, **14d**, **14f**), all 3,4-dichlorobenzylidene derivatives were ineffective as inhibitors of the HIV-1 multiplication in cell-based assays as well as of the HIV-1 rIN in enzyme assays. Also inactive were all hydroxy-nitro derivatives (**14g**, **14h**, **14j**, **14k**, **14m**, **14n**, **14p**, **14q**). Partially positive results were obtained with N_1-alkyl (**14i**, **14l**) or N_1-benzyl (**14o**) derivatives of 3,5-bis(3,4-dihydroxybenzylidene)piperidin-4-one. In fact, although fairly cytotoxic and ineffective in preventing the HIV-1 multiplication in cell-based assays, they were found to inhibit rIN at submicromolar concentrations in 3′-processing and strand transfer assays. Nevertheless, **14i** and **14l**, which were substituted at position 1 of the 4-piperidinone moiety with methyl and ethyl groups, respectively, were 2.5-fold less potent than their parent compound **12d** reported previously[12] ($IC_{50} = 0.2$ μM), and **14o**, which was substituted with a benzyl group, was 8-fold less potent.

Differently from catechols (**14i**, **14l** and **14o**) which shared inactive in cell-based assays but active in enzyme assays, nitrophenols (**14g**, **14h**, **14j**, **14k**, **14m**, **14n**, **14p**, **14q**) were inactive either in cellular or in enzyme assays. Therefore, replacement of one of catechol hydroxyls with a nitro group led to compounds deprived of anti-IN activity.

Then, based on the observation that the addition of carboxyl groups and/or a third hydroxyl group to the catechol moiety of CAPE-like analogues favours the anti-IN activity,[17,18,20] with the aim of optimizing the antiviral activity/cytotoxicity profile of our cynnamoyl compounds, we designed novel analogues (Fig. 3) characterized by one or more of the following features: (i) the presence of a third hydroxyl group at the 5′-position of both phenyl rings connected to a piperidinone (**15a–d**), a cyclohexanone (**15e**), a pyranone (**15f**), a thiopyranone (**15g**) or a cyclopentanone (**16**) ring; (ii) alkylation or benzylation of the 4-piperidinone moiety (**15b–d**); (iii) the presence of carboxyl or acetic groups (free or esterified with ethanol) at position 4 of the cyclohexanone nucleus (**19a–d**; **20a–d**); (iv) monobenzylidene substitution on indanone (**17a**), indandione (**17b**), barbituric (**18a**) or thiobarbituric acid (**18b**).

Introduction of a third hydroxyl and carboxylic group in the structure of **12a** led to a progressive increase of potency in enzymatic tests (IC_{50}) and gave rise to anti-HIV-1 activity in cell-based assays (EC_{50}) (compare **12a** with **15e**, **19d**, **20b** and **20d**). As expected, unlike their catechol counterparts **19a**,**b** and **20a**,**b**, the trihydroxy-derivatives bearing carboxyl groups (**19d** and **20d**) in the cyclohehanone ring proved the most potent HIV-1 inhibitors in cell-based assays. Noteworthy, esterification with ethanol (**19c** and **20c**) diminished their potency.

Compounds **15a–g**, **16**, **17a–b**, **18a–b**, **19a–d**, **20a–d** turned out to be potent inhibitors of the rIN 3′-processing, strand transfer activities, showing IC_{50} values as low as 0.2 μM . In addition, they also inhibited the disintegration reaction (Table 1), which is the reversal of the strand transfer reaction.[30,31] Since the occurrence of disintegration requires only the IN core domain, it has been used to probe binding of the drugs to the enzyme.[6] Thus, since the above compounds inhibited the disintegration activity in the low micromolar range, they very likely bind to the rIN core region.

Table 1. Cytotoxicity, antiviral and anti-integrase activities of derivatives **13–20**

Compd	CC_{50}[a]	EC_{50}[b]	SI[c]	IC_{50}[d]		
				3′-Processing	Strand transfer	Disintegration
13a	>200	>200	—	>100	>100	nd
13b	121	>121	—	>100	>100	nd
14a	120	>120	—	>100	>100	nd
14b	102	>102	—	>100	>100	nd
14c	200	>200	—	9±2	9±2	nd
14d	165	>165	—	>100	>100	nd
14e	31	>31	—	>100	>100	nd
14f	119	>119	—	>100	>100	nd
14g	11	>11	—	>100	>100	nd
14h	1.2	>1.2	—	>100	>100	nd
14i	5	>5	—	0.5±0.2	0.9±0.4	nd
14j	25	>25	—	>100	>100	nd
14k	1.9	>1.9	—	>100	>100	nd
14l	5	>5	—	0.5±0.2	1.1±0.3	nd
14m	15	>15	—	>100	>100	nd
14n	4	>4	—	>100	>100	nd
14o	4.7	>4.7	—	1.7±0.3	2.2±0.3	nd
14p	19	>19	—	>100	>100	nd
14q	3.5	>3.5	—	>100	>100	nd
15a	35	22	1.6	0.3±0.1	0.5±0.2	0.5±0.2
15b	46	>46	—	0.4±0.1	0.6±0.2	1.0±0.3
15c	40	≥40	—	0.5±0.1	0.4±0.1	0.9±0.3
15d	117	30	3.9	0.7±0.2	1.1±0.2	2.0±0.4
15e	50	>50	—	1.4±0.4	2.3±0.6	1.6±0.3
15f	115	34	3.4	6.0±2.0	9.0±3.0	10.0±2.0
15g	100	20	5	0.7±0.2	0.5±0.1	1.9±1.4
16	180	60	3	1.6±0.3	0.9±0.3	0.2±0.1
17a	16	>16	—	0.3±0.1	0.7±0.2	nd
17b	19	>19	—	0.2±0.1	0.5±0.2	nd
18a	50	>50	—	3.0±0.5	4.0±1.0	5.3±1.0
18b	70	20	3.5	1.0±0.3	1.6±0.2	2.1±0.6
19a	5	>5	—	0.2±0.1	0.4±0.15	nd
19b	42	>42	—	1.2±0.3	1.0±0.7	0.1±0.05
19c	67	12.5	5.4	0.2±0.1	0.3±0.1	0.3±0.2
19d	70	4	17.5	0.2±0.1	0.3±0.1	1.2±0.5
20a	5	>5	—	2.8±0.4	4.3±1.0	nd
20b	41	>41	—	0.7±0.3	1.2±0.4	nd
20c	40	30	1.3	2.6±0.2	1.9±0.5	1.6±0.6
20d	40	2	20	0.2±0.1	0.2±0.1	0.5±0.2
9	>100	1.2	83.3	2.4	—	1.0
10	—	—	—	60	—	0.05
12a	7.4	>7.4	—	0.9±0.3	—	nd

nd, not determined.
[a] Compound concentration (μM) required to reduce the exponential growth of MT-4 cells by 50%, as determined by MTT method.
[b] Compound concentration (μM) required to achieve 50% protection of MT-4 cells from the HIV-1 induced cytopathicity, as determined by MTT method.
[c] Selectivity index: ratio CC_{50}/EC_{50}.
[d] Compound concentration (μM) required to reduce HIV-1 IN 3′-processing, strand transfer or disintegration activity by 50%. Data represent mean

3.2. Anti-rRT assays

Numerous IN inhibitors have been reported to affect additional viral enzymes in cell-free assays. Therefore, in order to exclude the possibility that title compounds act non-specifically as inhibitors of the various rIN activities, we tested selected representatives of the new cynnamoyl derivatives (**14g**, **14h**, **19d** and **20d**) against the HIV-1 rRT in enzyme assays (see Experimental). None of the compounds was active at 100 μM (data not shown).

3.3. Antiproliferative activity

The cytotoxicity of 3,4-dihydroxycinnamoyl compounds is known to be mainly related to: (i) oxidation to ortho-quinone reactive species, which leads to cross-linking of intracellular proteins;[32] (ii) the nucleophilic attack of the α,β-double bond by thiols (such as glutathion), which is promoted by electron-withdrawing substituents on the phenyl ring and by the overall lipophilicity. Although the above studies[32] have highlighted the physico-chemical features relevant to the cytotoxicity of cinnamoyl compounds, the biochemical complexity of the cell machinery is so huge that approaches to improve the selectivity index of catechol derivatives (by lowering the cytotoxicity) are still largely empirical. Nevertheless, in order to make a more realistic correlation between the structure and the cytotoxicity of our cynnamoyl derivatives, we tested representative compounds against several cell lines derived from haemathological and solid human tumors and against a reference 'normal' foreskin fibroblast cell line.

Among the catechol derivatives reported previously,[12] **12c** was the most potent, followed by **12a** and **12b**. The above compounds resulted equally potent against haemathological and solid tumor cell lines and against the reference 'normal' foreskin fibroblast cell line. Interestingly, the substitution of either of the OH groups of the catechol moiety with nitro groups yielded very cytotoxic derivatives. The 3-hydroxy-4-nitrophenyl (**14h**) emerged as the most cytotoxic compound, followed by the 3-nitrophenyl-4-hydroxy derivative. It is noteworthy that introduction of a third hydroxyl group at position 5 of the phenyl ring makes the catechol derivative less cytotoxic (compare **12d** with **15a**). Likewise, the introduction of carboxyl or carboxyethyl groups at position 4 of the cyclohexanone also leads to a significant decrease of cytotoxicity [more pronounced with the carboxyl (**19b**) than with the carboxyethyl group (**19a**)], which is more evident in the case of solid tumour-derived cell lines and the reference fibroblast cell line. As far as cytotoxicity is concerned, the addition to **19b** of a third hydroxyl group at position 5 of the phenyl ring (**19d**) is without further effects.

3.4. Protein-linked DNA breaks

Due to the fact that polyhydroxylated compounds have been reported to be topoisomerase inhibitors,[33] the most cytotoxic derivatives (**14g** and **14h**) were investigated in an intact cell assay (see Experimental) for anti-topoisomerase activity (Table 2).

This assay was preferred to an enzyme assay to take into account the capability of title compounds to penetrate the plasma membrane and to reach the nuclear compartment. Etoposide and camptothecin were used as reference drugs. None of the compounds significantly increased the amount of protein-linked DNA breaks (PLDB) in KB cells (data not shown) with respect to those of etoposide 10 and 100 μM (which gave a 12.5-fold and 21-fold increase in PLDB with respect to untreated controls, respectively) and camptothecin 1 and 10 μM (which gave a 4-fold and 7.4-fold increase in PLDB with respect to untreated controls, respectively).

3.5. Molecular modeling and 3-D QSAR studies

To gain insight into the main binding features of title compounds into the HIV-1 rIN, either derivatives described by Buolamwini et al.,[22] or the newly synthesized IN inhibitors were analyzed by means of 3-D QSAR studies. Due to the insufficient number of compounds, no effort was done to develop 3-D QSAR models for anti-HIV activity in cell-based assays. Due to the homogeneous results obtained in 3′-processing and strand transfer assays, and to the fact that the previously reported derivatives[12] were tested only in 3′-processing assays, only the IC_{50}s obtained in the latter assays were used for the 3-D QSAR studies.

3-D QSAR approaches rely on advanced statistical methods to correlate a dependent variable (usually the biological activity in enzyme assays) with 3-D chemical descriptors, such as molecular interaction potentials, which are interaction energies of the considered molecules with proper molecular probes. In the past, most of such studies have been carried out by means of the comparative molecular field analysis (CoMFA) developed by Cramer more than 10 years ago.[25] In the past

Table 2. Antiproliferative activity of compounds **12a–d**, **14g,h**, **15a**, **19a,b** and **19d**

Cell lines[b]	[a]IC_{50} (μM)									
	12a	**12b**	**12c**	**12d**	**14g**	**14h**	**15a**	**19a**	**19b**	**19d**
Leukemia/lymphoma										
Wil2-NS	25.7	30.7	5.5	50.9	2.7	0.97	38.4	19.0	37.4	54.2
CCRF-SB	6.7	12.8	2.6	43.6	2.5	0.8	30.6	13.7	55.3	62.5
Raji	8.6	15.3	2.5	62.7	3.5	0.8	60.2	15.2	55.8	43.7
CCRF-CEM	5.4	8.4	1.8	13.3	1.4	0.3	22.9	4.3	25.4	32.4
MOLT-4	9.3	16.2	1.9	84.2	2.3	0.3	29.5	3.7	23.2	38.6
MT-4	7.4	10.0	4.0	21.0	11.0	1.2	35.0	5.0	42.0	70.0
Carcinoma										
SK-MEL-28	8.2	12.2	3.2	8.9	3.9	1.9	>100	>100	>100	>100
MCF7	10.3	22.2	4.6	>100	16.4	2.4	61.1	>100	>100	>100
SKMES-1	5.9	12.3	3.7	63.1	4.6	1.4	75.6	>100	>100	>100
HepG2	90	21.3	6.8	66.5	7.5	1.9	>100	>100	>100	>100
DU145	7.0	12.4	3.3	10.9	2.4	1.3	>100	>100	>100	>100
HT-29	25.3	35.6	6.6	55.9	3.6	0.9	>100	>100	>100	>100
HeLa	20.0	27.5	7.6	46.2	4.2	1.0	87.3	90.5	>100	87.2
ACHN	24.0	47.2	6.0	78.4	3.7	1.7	>100	>100	>100	>100
5637	7.9	17.7	2.9	32.6	2.9	0.57	>100	>100	>100	>100
Normal cells										
CRL7065	85	16.9	5.7	72	8.8	0.9	30.2	>100	>100	>100

[a] Compound concentration required to reduce cell proliferation by 50%, as determined by the MTT method, under conditions allowing untreated controls to undergo at least three consecutive rounds of multiplication. Data represent mean values (±SD) for three independent determinations.

[b] Wil2-NS, human splenic B-lymphoblastoid cells; CCRF-SB, human acute B-lymphoblastic leukemia; Raji, human Burkitt lymphoma; CCRF-CEM and MOLT-4, human acute T-lymphoblastic leukemia; MT-4, human $CD4^+$ T-cells containing an integrated HTLV-1 genome; SK-MEL-28, human skin melanoma; MCF7, human breast adenocarcinoma; SKMES-1, human lung squamous carcinoma; HepG2, human epathocellular carcinoma; DU145, human prostate carcinoma; HT-29, human colon adenocarcinoma; HeLa, human cervix carcinoma; ACHN, human renal adenocarcinoma; 5637, human bladder carcinoma; CRL7065, normal foreskin fibroblast.

few years, the so-called GRID/GOLPE[34] approach has also been used. Although CoMFA and GRID/GOLPE rely on the same principles and similar statistical algorithms, GRID offers an excellent set of molecular probes for computation of interaction potentials, and GOLPE provides, besides a classical partial least-squares (PLS) method,[35] different useful tools for data pretreatment, selection of variables and interpretation of results. This often leads to more meaningful results and limits the risk of deriving over fitted models.[36]

In a recent study reported by Buolamwini[22] on curcumin-like derivatives[12] and other hydroxylated IN inhibitors,[8] a CoMFA was performed leading to the identification, at 3-D level, of the main molecular determinants which were likely responsible for inhibition of the 3′-processing activity of rIN. In that study, removal of some compounds from the training set was determinant to get statistically improved results for the 3-D QSAR.

In the present 3-D QSAR studies, we selected the GRID/GOLPE approach because of the aforementioned advantages over the classical CoMFA approach. Indeed, the atom probes available in GRID[27] for the calculation of molecular interaction fields are much more numerous than those implemented in CoMFA, and most of them better resemble the functional amino acids constituting the putative binding sites of enzymes and receptors.

Another advantage of the GRID/GOLPE method is that it includes advanced mathematical tools such as variable pre-treatment, D-Optimal, and FFD variable selection methods as well as smart region definition (SRD).[37] This allows to obtain more robust and easily interpretable 3-D-QSAR models.

A GRID/GOPLE analysis was performed on three sets of derivatives selected as follows: (i) derivatives reported by Buolamwini paper[22] (Set A); (ii) newly synthesized derivatives described in the present work (Set B); (iii) both foregoing groups (Set C). The results of GRID/GOLPE analyses of the three molecular series depicted as Model A, Model B and Model C, respectively (carried out using the OH_2 probe) are summarized in Table 3 and Figure 4.

Models with high statistical coefficients were obtained for all the three training sets. It is noteworthy that the set B gave PLS models with better statistical figures (see Table 3), although the three models are comparable based on the fact that the A and C sets show a lower value of principal components. These findings support the hypothesis that A and B series can be considered as a whole set. In fact, the use of the B set as test set for the A set, and vice versa, showed low value of prediction standard error (SDEP) (Tables 4 and 5).

To assess the real predictive ability of the models obtained, the same external test set used by Buolamwini[22] was used. Our Model A displayed a SDEP of 0.60 on the external Test_Set, performing slightly better than the previous CoMFA model[22] (SDEP 0.64). The higher SDEP value of Model B against the external Test_Set is mainly due to large structural differences between the two sets. In fact, the joined Set C performed a satisfactory prediction with a SDEP value of only 0.69.

Table 3. Statistical results of the GRID/GOLPE analyses

Training_Set	PC[a]	Objects[b]	N[c]	R^2	Q^2	SDEP[d]
A	2	24	913	0.91	0.70	0.59
B	3	41	1293	0.97	0.90	0.33
C	2	64[e]	1327	0.84	0.73	0.55

[a] Number of optimal principal components from cross-validated analysis using five random groups.
[b] Number of compounds in the model.
[c] Number GRID variables selected from the fractional factorial selection.
[d] Standard error of prediction.
[e] Compound **7b** was in common for the A and B sets.

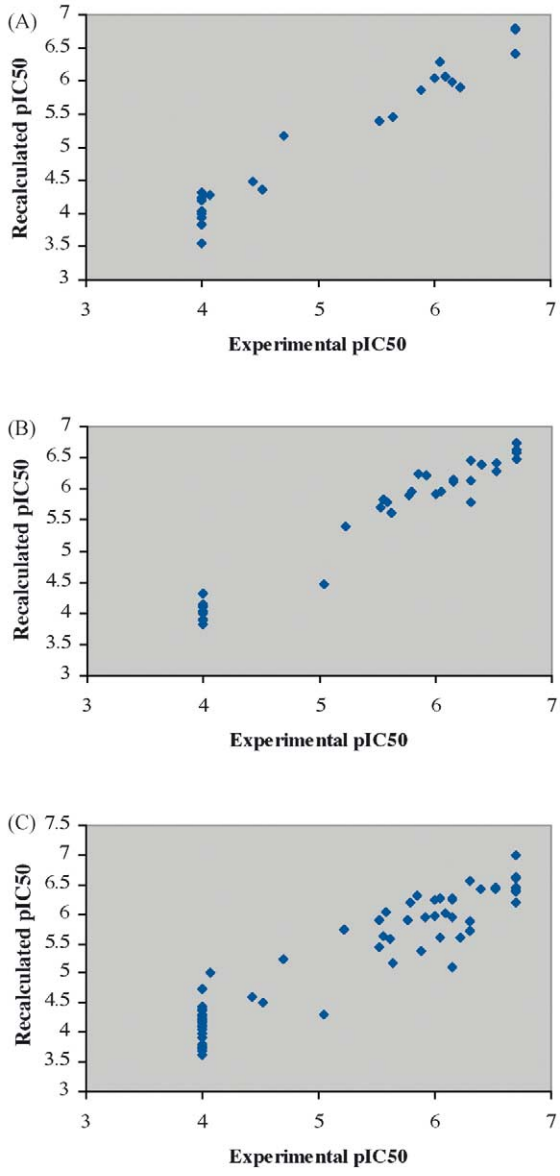


Figure 4. Experimental versus recalculated plots for the three 3-D QSAR models. Model A (A), Model B (B) and Model C (C).

Table 4. Predictivity of Models A, B and C

Model	SDEP[a]		
	Set A[b]	Set B[c]	External Test_Set[d]
A	—	0.92	0.60
B	0.95	—	1.03
C	—	—	0.69

[a] Standard deviation of errors of prediction.
[b] Derivatives reported in literature.[22]
[c] Derivatives reported in the present work.
[d] Test set as reported by Buolamwini.[22]

The GOLPE program allows also a graphical representation of the 3-D QSAR models by means of PLS coefficients. In order to identify which part of the molecules correlated with the variation of activity, we analyzed the PLS coefficient plots for the three models (Fig. 5). As can be seen, there are regions around the ligands where the different substituents of the derivatives could generate favourable/unfavourable interactions with the water probe. However in this plot the signs of the coefficients can induce to errors: in fact, coefficients have opposite meaning depending on the fact that the compound produces positive or negative field values in this areas. In Figure 5A–C, cyan polyhedra (negative PLS coefficients) delimitate regions where a favourable (negative) interaction leads to increased activity, whereas an unfavourable (positive) interaction leads to decreased activity. On the other hand, yellow polyhedra (positive PLS coefficients) represent regions where unfavourable (negative) interactions lead to decreased activity, whereas favourable (positive) interactions lead to increased activity.

In Figure 5A (contour plot for the Model A), cyan polyhedra appear in the region surrounding the meta hydroxyls groups of the highly active compound **7f**, suggesting a possible favourable hydrogen bond interaction for this derivative with those groups. However, the same polyhedra overlap the methoxyl groups of the less active compounds **8b** and **9b**, suggesting that either the loss of hydrogen bonding interactions, or the enhancement of steric hindrance exerted by the methyl groups, is detrimental for biological activity. Yellow polyhedra represent zones where the presence of even small groups is poorly tolerated. In fact, yellow polyhedra are found in the vicinity of the methoxyl group of **8b** and **9b** (right side of Fig. 5A) and in the region occupied by the methoxyl group of **11a** in the left side of Figure 5A.

In Figure 5B (contour plot for the Model B), yellow polyhedra have a more enlarged area than in Figure 5A. In fact, in the Set B the molecular diversity is mainly focused in the central ring bridging the two hydroxylated benzenes. A new yellow polyhedra can be observed near the region occupied by the carboxymethyl group in position 4 of the cyclohexanone moiety of the IN inhibitor **20d** (green structure in Fig. 5B). In fact, in this region the negative carboxyl group exerts a favourable interaction with the positive PLS coefficient. On the contrary, the inactive derivative **14p** (not shown) places in the same region a benzyl group, which interacts negatively, probably because of the enhanced steric hindrance. Interestingly, the negative contribution of the chlorine substituents should also be noted. Figure 5B reports compounds **13a** and **14c** in which chlorine atoms replace the hydroxyl groups of **20d**. In this case, the presence of either yellow or cyan polyhedra highlights the loss of hydrogen bond interactions. In fact, the favourable interaction mediated by the presence of OHs groups becomes unfavourable when chlorine atoms are introduced. In particular, negative cyan spots surround the positive hydrogens of the para hydroxyls of **20d** (favourable interactions) and, at the same time, they surround the partially negative charged and bigger chlorine atoms of **13a** and **14c** (unfavourable interactions).

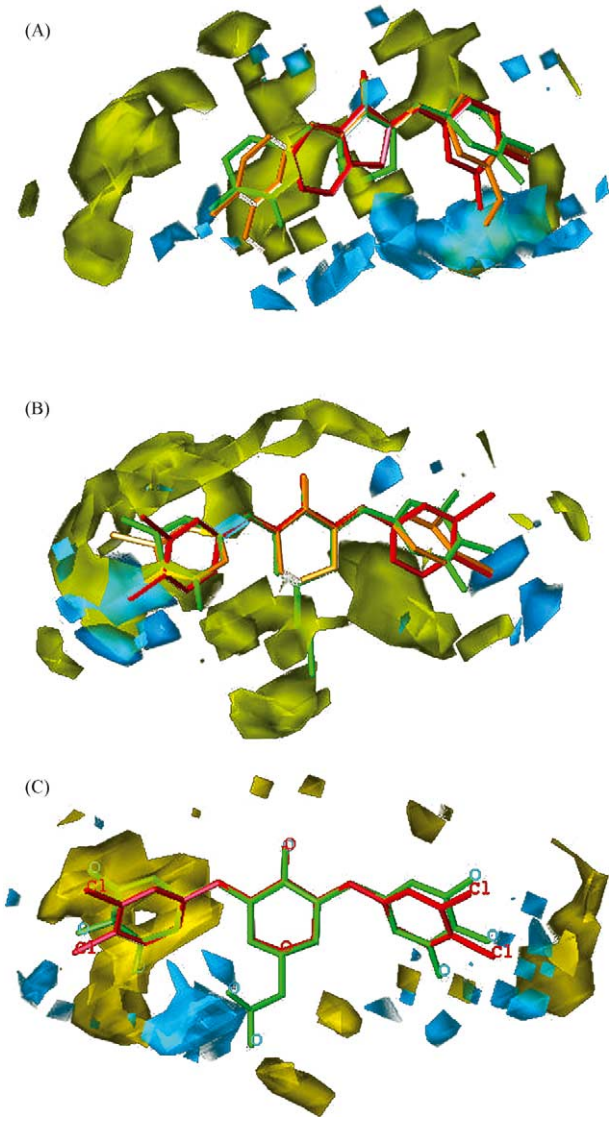


Figure 5. GRID/GOLPE PLS coefficients contour maps for the three 3-D QSAR models. Model A (A), Model B (B) and Model C (C) (Contour levels 0.0040 yellow, −0.0040 cyan, for colour code, see text). To aid interpretation, high (green), medium (orange) and low active derivative structures are displayed. In (A) derivatives **7f**, **8b** and **9b**, in (B) **13a**, **14c** and **20d** and in (C) **14c** and **20d**.

Table 5. Residual of the predictions by the GRID/GOLPE models

Present work Set B			JMC1998[12] Set A			JMC2002[22] External Test Set				
Compd	Exp. pIC_{50}	A[a]	Compd	pIC_{50}	B	Compd	pIC_{50}	A	B	C
12	6.05	0.69	**10a**[b]	5.89	0.09	**26**[c]	5.15	−0.05	−0.24	−0.09
13a	4.00	−0.43	**10b**[b]	4.70	−1.30	**27**[c]	4.22	−0.28	−1.21	−0.42
13b	4.00	−0.74	**10c**[b]	4.00	−1.41	**28**[c]	4.00	−0.61	−1.12	−0.85
14a	4.00	−1.17	**10d**[b]	5.64	−0.12	**29**[c]	4.00	−0.70	−0.87	−0.45
14b	4.00	−1.35	**11a**[b]	4.00	−0.20	**30**[c]	4.00	−0.67	−1.34	−1.03
14c	5.05	−0.48	**11b**[b]	6.15	1.71	**31**[c]	4.00	−1.30	−1.33	−1.23
14d	4.00	0.34	**11c**[b]	4.00	−0.34	**32**[c]	5.70	1.13	0.28	0.55
14e	4.00	−1.33	**11d**[b]	4.00	−0.19	**33**[c]	5.70	0.48	0.21	0.40
14f	4.00	−0.86	**12a**[b]	6.10	1.12	**34**[c]	4.26	−0.30	−0.94	−0.87
14g	4.00	−1.43	**12b**[b]	4.52	−0.71	**35**[c]	4.00	−0.33	−1.56	−1.05
14h	4.00	−1.91	**13**[b]	4.00	−0.79	**36**[c]	4.00	−0.12	−1.52	−0.93
14i	6.30	0.85	**7a**[b]	4.00	−1.79	**37**[c]	5.05	1.22	−0.35	0.33
14j	4.00	−0.75	**7b**[b]	6.05	0.01	**38**[c]	5.10	0.94	−0.34	0.21
14k	4.00	−1.08	**7c**[b]	6.00	0.19	**39**[c]	4.36	−0.28	−0.95	−0.68
14l	6.30	1.52	**7d**[b]	6.70	0.70	**40**[c]	4.82	0.23	−0.53	−0.15
14m	4.00	−0.59	**7e**[b]	6.70	0.79	**41**[c]	4.17	−0.25	−1.19	−0.67
14n	4.00	−1.00	**7f**[b]	6.70	0.66	**42**[c]	4.74	−0.17	−0.67	−0.35
14o	5.77	0.71	**8a**[b]	5.52	−0.52	**43**[c]	5.10	0.62	−0.46	−0.06
14p	4.00	−0.74	**8b**[b]	4.07	−1.92	**44**[c]	4.00	−0.05	−1.49	−0.92
14q	4.00	−1.02	**9a**[b]	4.00	0.18	**45**[c]	4.85	0.26	−0.70	−0.34
15a	6.52	−0.23	**9b**[b]	4.43	−0.51	**46**[c]	4.00	0.00	−1.45	−0.77
15b	6.40	0.30	**9c**[b]	6.22	0.69	**47**[c]	4.12	−0.23	−1.55	−0.95
15c	6.30	0.24	**9d**[b]	4.00	−1.35	**48**[c]	4.40	−0.65	−0.95	−0.71
15d	6.15	0.27	**3**[b]	4.52	−0.43					
15e	5.85	−0.53								
15f	5.22	−1.31								
15g	6.15	−0.29								
16	5.80	−0.01								
17a	6.52	1.50								
17b	6.70	1.57								
18a	5.52	−0.10								
18b	6.00	0.38								
19a	6.70	1.63								
19b	5.92	0.67								
19c	6.70	0.46								
19d	6.70	0.44								
20a	5.55	0.44								
20b	6.15	0.99								
20c	5.59	0.03								
20d	6.70	0.85								
9	5.62	1.02								

[a] A = Model A, B = Model B, C = Model C.
[b] Numeration of compounds as in the original paper.[12]
[c] External test set compounds numbered as in the CoMFA work.[22]

Figure 5C shows the contour plot of the PLS coefficient for the Model C, which summarize the previous plots.

4. Conclusions

The overall evaluation of the biological activity revealed that, in order to show anti-HIV-1 activity in cell-based assays, title compounds require the presence of a trihydroxyl moiety coupled with a carboxyl group at position 4 of a cyclohexanone ring. The assemblage of the above chemical requisites on the curcuminoid cyclic structure of **12a** led to obtain compound **20d**, a potent anti-IN derivatives as active as **9** and **10** in infected cells.

This interesting result encourage further synthetic and biological studies on cyclohexanone analogues of **20d**, with the aim of increasing anti-HIV-activity in infected cells and of exploring whether additional mechanisms of action, for instance fusion, are involved in the inhibition activity of title compounds.

Using molecular dynamics generated conformations, the GRID interaction fields and the statistical tools of the program GOLPE, we developed significant and predictive models, indicated by the high cross-validation coefficients and the low SDEP values for the external test set.

From the PLS coefficient map the models could be used as potential generator of design ideas for new anti-IN compounds.

Moreover, it is interesting to observe that the present statistical results are highly comparable with those found in previous 3-D-QSAR study conducted with CoMFA and CoMSIA on the Set A of compounds.[22] As an important complement to previous CoMFA results, GRID/GOLPE allowed us to elucidate the role

of hydrogen bonding in the receptor binding of our ligands, by using the OH_2 probe provided by GRID.

5. Experimental

5.1. Chemistry

Melting points were determined on a Büchi 530 melting point apparatus and are uncorrected. Infrared (IR) spectra (Nujol mulls) were recorded on a Perkin-Elmer 297 spectrophotometer. ^{1}H NMR spectra were recorded at 200 MHz on a Bruker AC 200 spectrometer using tetramethylsilane (Me_4Si) as internal reference standard. All compounds were routinely checked by TLC and ^{1}H NMR. TLC was performed by using aluminum-baked silica gel plates (Fluka DC-Alufolien Kieselgel 60 F_{254}). Developed plates were visualized by UV light. Solvents were reagent grade and, when necessary, were purified and dried by standard methods. Concentration of solutions after reactions and extractions involved the use of a rotary evaporator (Büchi) operating at a reduced pressure (ca. 20 Torr). Organic solutions were dried over anhydrous sodium sulfate. Analytical results agreed to within ±0.40% of the theoretical values. All compounds were analyzed for C, H, N, and, when present, S and Cl.

5.2. Syntheses

5.2.1. General procedure for the preparation of compounds 13a, 14a, 14c, 14e. Example: 2,6-bis(3,4-dichlorobenzylidene)pyran-4-one (14c). A solution of 3,4-dichlorobenzaldehyde (1.75 g, 9.9 mmol) and pyran-4-one (0.5 g, 4.9 mmol) in ethanol (15 mL) was added onto a well stirred solution of sodium hydroxide (0.8 g, 20 mmol) in water (15 mL). The mixture was stirred at room temperature for 1 h then treated with water (50 mL). Crystalline precipitate was filtered and recrystallized from N,N-dimethylformamide to obtain 1.74 g of **14c** (51% yield); mp 225–227 °C; IR cm^{-1} 1610 (CO); ^{1}H NMR (DMSO-d_6) δ 4.90 (s, 4H, CH_2), 7.42–7.81 (s, 8H, benzene H and =CH–).

Yield (%), reaction time (h), mp (°C), recrystallization solvent are reported for each already known compound:

13a:[38] 100%, 1 h, 175–177 °C, ethanol;

14a:[39] 30%, 0.5 h, 147–149 °C, tetrahydrofuran/water;

14e:[40] 21%, 0.5 h, 145–147 °C, ethanol.

5.2.2. General procedure for the preparation of compounds 13b, 14b, 14d, 14f. Example: 2,6-bis(3,4-dichlorobenzylidene) - 4 - hydroxypyrane (14d). Sodium borohydride (0.05 g, 1.4 mmol) was added into a well stirred solution of **14c** (1.0 g, 2.9 mmol) in anhydrous tetrahydrofuran (100 mL) and water (0.1 mL). The mixture was stirred at room temperature for 0.5 h. The solvent was removed and the residue was dissolved in ethyl acetate. The organic solution was washed with brine and dried. Evaporation of the solvent gave crude a product, which was chromatographed on silica gel column (chloroform as eluent) to obtain 0.91 g of pure **14d** (91% yield): mp 147–150 °C (from ethanol).

Yield (%), reaction time (h), mp (°C), recrystallization solvent; chromatographic system are reported for each compound:

13b: 39%, 0.5 h, 71–73 °C, benzene, silica gel/chloroform;

14b: 32%, 1.5 h, oil, silica gel/chloroform;

14f: 100%, 0.5 h, 122–124 °C, benzene/cyclohexane.

5.2.3. General procedure for the preparation of compounds 14g–q, 15, 16, 19 and 20. Example: 3,5-bis(3,4,5-trihydroxybenzylidene)piperidin - 4 - one hydrochloride (15a). A solution of 4-piperidinone hydrate hydrochloride (0.5 g, 3.2 mmol) in glacial acetic acid (20 mL), previously saturated with anhydrous hydrogen chloride, was treated with 3,4,5-trihydroxybenzaldehyde (1.68 g, 9.75 mmol). After stirring at room temperature for 48 h the precipitate which formed was filtered, washed in turn with water, ethanol and light petroleum ether to obtain 0.57 g of **15a** (46% yield): mp >300 °C (from ethanol); IR cm^{-1} 3150, 3300 (OH and NH) and 1580 (CO); ^{1}H NMR (DMSO-d_6) δ 4.37 (s, 4H, CH_2), 6.49 (s, 4H, benzene H), 7.57 (s, 2H, =CH–), 9.36 (br s, 8H, OH and NH).

By this procedure were prepared compounds **14g–q**, **15**, **16**, **19** and **20** starting from the appropriate arylaldehyde. Yield (%), reaction time (h), mp (°C), recrystallization solvent; chromatographic system if necessary, IR (cm^{-1}); ^{1}H NMR data are reported for each compound:

14g: 13%, 120 h, >300 °C, dimethylsulfoxide; IR cm^{-1} 3100 (OH) and 1600 (CO); ^{1}H NMR (DMSO-d_6) δ 4.46 (s, 4H, CH_2), 7.27 (s, 2H, benzene C5-H), 7.71 (s, 2H, benzene C6-H), 7.78 (s, 2H, =CH–), 8.04 (s, 2H, benzene C2-H), 10.53 (br s, 4H, OH and NH).

14h: 21%, 72 h, >300 °C, dimethylsulfoxide/water; IR cm^{-1} 3300 (OH) and 1610 (CO); ^{1}H NMR (DMSO-d_6) δ 4.46 (s, 4H, CH_2), 7.11 (d, J_o = 8.4 Hz, 2H, benzene C6-H), 7.32 (s, 2H, benzene C2-H), 7.80 (s, 2H, =CH–), 7.99 (d, J_o = 8.4 Hz, 2H, benzene C5-H), 10.71 (br s, 4H, OH and NH).

14i: 61%, 72 h, >300 °C, dimethylsulfoxide; IR cm^{-1} 3500 (OH and NH) and 1570 (CO); ^{1}H NMR (DMSO-d_6) δ 3.00 (s, 3H, CH_3), 4.60 (s, 4H, CH_2), 6.89–6.96 (m, 6H, benzene H), 7.70 (s, 2H, =CH–), 9.44 and 9.91 (2 br s, 4H, OH), 11.18 (br s, 1H, NH).

14j: 76%, 120 h, >300 °C, dimethylsulfoxide; IR cm^{-1} 3220 (OH) and 1610 (CO); ^{1}H NMR (DMSO-d_6) δ 2.97 (s, 3H, CH_3), 4.64 (s, 4H, CH_2), 7.32 (d, J_o = 8.2 Hz, 2H, benzene C5-H), 7.72 (d, J_o = 8.2 Hz, 2H, benzene C6-H), 7.83 (s, 2H, =CH–), 8.04 (s, 2H, benzene C2-H), 11.68 (br s, 3H, OH and NH).

14k: 82%, 72 h, 248–250 °C, dimethylsulfoxide; IR cm^{-1} 3300 (OH) and 1610 (CO); ^1H NMR (DMSO-d_6) δ 2.94 (s, 3H, CH_3), 4.59 (s, 4H, CH_2), 7.11 (d, $J_o = 8.6$ Hz, 2H, benzene C6-H), 7.32 (s, 2H, benzene C2-H), 7.81 (s, 2H, =CH–), 7.97 (d, $J_o = 8.6$ Hz, 2H, benzene C5-H), 11.43 (br s, 3H, OH and NH).

14l: 32%, 72 h, 270–272 °C, dimethylsulfoxide/water; IR cm^{-1} 3300 (OH and NH) and 1560 (CO); ^1H NMR (DMSO-d_6) δ 1.26 (t, 3H, CH_3), 3.35 (m, 2H, CH_2CH_3), 4.56 (s, 4H, CH_2), 6.89–6.99 (m, 6H, benzene H), 7.71 (s, 2H, =CH–), 9.42 and 9.83 (2 br s, 4H, OH), 11.36 (br s, 1H, NH).

14m: 33%, 72 h, >300 °C, dimethylsulfoxide; IR cm^{-1} 3220 (OH) and 1610 (CO); ^1H NMR (DMSO-d_6) δ 1.26 (t, 3H, CH_2CH_3), 3.37 (q, 2H, CH_2CH_3), 4.60 (s, 4H, CH_2), 7.32 (d, $J_o = 8.7$ Hz, 2H, benzene C5-H), 7.74 (d, $J_o = 8.7$ Hz, 2H, benzene C6-H), 7.84 (s, 2H, =CH–), 8.07 (s, 2H, benzene C2-H), 11.80 (br s, 3H, OH and NH).

14n: 26%, 48 h, 245–247 °C, dimethylsulfoxide/water; IR cm^{-1} 3300 (OH) and 1610 (CO); ^1H NMR (DMSO-d_6) δ 1.25 (t, 3H, CH_2CH_3), 3.32 (q, 2H, CH_2CH_3), 4.58 (s, 4H, CH_2), 7.14 (d, $J_o = 8.6$ Hz, 2H, benzene C6-H), 7.34 (s, 2H, benzene C2-H), 7.83 (s, 2H, =CH–), 8.02 (d, $J_o = 8.6$ Hz, 2H, benzene C5-H), 11.41 and 11.59 (2br s, 3H, OH and NH).

14o: 63%, 72 h, 180–182 °C, methanol; IR cm^{-1} 3200 (OH and NH) and 1570 (CO); ^1H NMR (DMSO-d_6) δ 4.51 (s, 6H, CH_2), 6.78–7.57 (m, 11H, benzene H), 7.73 (s, 2H, =CH–), 9.41 and 9.80 (2 br s, 4H, OH), 11.40 (br s, 1H, NH).**14p**: 60%, 120 h, 227–230 °C, *N*,*N*-dimethylformamide/water; IR cm^{-1} 3220 (OH) and 1610 (CO); ^1H NMR (DMSO-d_6) δ 4.46 (s, 2H, CH_2Ph), 4.53 (s, 4H, CH_2), 7.27–7.54 (m, 7H, benzene C5-H and benzyl H), 7.65 (d, $J_o = 7.8$ Hz, 2H, benzene C6-H), 7.84 (s, 2H, =CH–), 7.97 (s, 2H, benzene C2-H), 11.87 (br s, 2H, OH).

14q: 65%, 120 h, 211–213 °C, *N*,*N*-dimethylformamide/water; IR cm^{-1} 3300 (OH) and 1610 (CO); ^1H NMR (DMSO-d_6) δ 4.15 (br s, 6H, CH_2), 7.02 (d, $J_o = 8.5$ Hz, 2H, benzene C6-H), 7.20–7.36 (m, 7H, benzene C2-H and benzyl H), 7.71 (s, 2H, =CH–), 7.92 (d, $J_o = 8.5$ Hz, 2H, benzene C5-H), 11.26 (br s, 2H, OH).

15b: 30%, 72 h, >300 °C, ethanol; IR cm^{-1} 3300 (OH and NH) and 1580 (CO); ^1H NMR (DMSO-d_6) δ 3.00 (s, 3H, CH_3), 4.61 (s, 4H, CH_2), 6.52 (s, 4H, benzene H), 6.86 (s, 2H, =CH–), 9.61 (br s, 7H, OH and NH).

15c: 57%, 48 h, >300 °C, ethanol; IR cm^{-1} 3300 (OH and NH) and 1580 (CO); ^1H NMR (DMSO-d_6) δ 1.26 (t, 3H, CH_3), 3.33 (m, 2H, CH_2CH_3), 4.57 (br s, 4H, CH_2), 6.54 (s, 4H, benzene H), 6.86 (s, 2H, =CH–), 9.31 (br s, 7H, OH and NH).

15d: 68%, 120 h, >300 °C, dimethylsulfoxide; IR cm^{-1} 3300 (OH and NH) and 1580 (CO); ^1H NMR (DMSO-d_6) δ 4.41 (s, 6H, CH_2), 7.43–7.52 (s, 9H, benzene H), 7.66 (s, 2H, =CH–), 9.61 (br s, 7H, OH and NH).

15e: 30%, 1.5 h, >300 °C, dimethylsulfoxide/water, silica gel chloroform/methanol/formic acid 10:1:0.1; IR cm^{-1} 3300 (OH) and 1580 (CO); ^1H NMR (DMSO-d_6) δ 1.71 (m, 2H, $CH_2CH_2CH_2$), 2.83 (m, 4H, $CH_2CH_2CH_2$), 6.51 (s, 4H, benzene H), 7.35 (s, 2H, =CH–), 9.53 (br s, 6H, OH).**15f**: 55%, 48 h, >300 °C, dimethylsulfoxide; IR cm^{-1} 3300 (OH) and 1590 (CO); ^1H NMR (DMSO-d_6) δ 4.84 (s, 4H, CH_2), 6.38 (s, 4H, benzene H), 7.38 (s, 2H, =CH–), 8.21 (br s, 6H, OH).

15g: 60%, 2 h, >300 °C, dimethylsulfoxide/water; IR cm^{-1} 3300 (OH) and 1580 (CO); ^1H NMR (DMSO-d_6) δ 4.00 (m, 4H, CH_2), 6.40 (s, 4H, benzene H), 7.30 (s, 2H, =CH–), 9.38 (br s, 6H, OH).

16: 36%, 2 h, >300 °C, dimethylsulfoxide/water, silica gel chloroform/methanol/formic acid 10:1:0.1; IR cm^{-1} 3150, 3200 (OH) and 1580 (CO); ^1H NMR (DMSO-d_6) δ 4.75 (s, 4H, CH_2), 6.61 (s, 4H, benzene H), 7.12 (s, 2H, =CH–), 9.38 (br s, 6H, OH).

19a: 18%, 3 h, 102–104 °C, methanol, silica gel chloroform/methanol 20:1; IR cm^{-1} 3250 (OH), 1680 (CO ester) and 1580 (CO ketone); ^1H NMR (DMSO-d_6) δ 1.08 (t, 3H, CH_3), 2.82–3.16 (m, 5H, CH and CH_2), 4.03 (q, 2H, CH_2CH_3), 6.79–6.97 (s, 6H, benzene H), 7.49 (s, 2H, =CH–), 9.40 (br s, 4H, OH).

19b: 92%, 2 h, >270 °C, methanol/water, silica gel chloroform/methanol/formic acid 10:2:0.1; IR cm^{-1} 3260 (OH), 1680 (CO acid) and 1620 (CO ketone); ^1H NMR (DMSO-d_6) δ 2.69–3.17 (m, 5H, CH_2 and CH), 6.79–6.98 (m, 6H, benzene H), 7.49 (s, 2H, =CH–), 9.22 (br s, 4H, OH), 12.50 (br s, 1H, COOH).

19c: 37%, 2.5 h, >300 °C, methanol, silica gel chloroform/methanol/formic acid 10:1:0.1; IR cm^{-1} 3250 (OH), 1700 (CO ester) and 1580 (CO ketone); ^1H NMR (DMSO-d_6) δ 1.11 (t, 3H, CH_3), 2.51 (m, 1H, CH), 2.80–3.16 (m, 4H, CH_2), 4.05 (q, 2H, CH_2CH_3), 6.51 (s, 4H, benzene H), 7.39 (s, 2H, =CH–), 9.31 (br s, 6H, OH).

19d: 29%, 12 h, >300 °C, methanol, silica gel chloroform/methanol/formic acid 10:2:0.1; IR cm^{-1} 3320 (OH), 1600 (CO acid) and 1580 (CO ketone); ^1H NMR (DMSO-d_6) δ 2.73–3.21 (2 m, 5H, CH and CH_2), 6.53 (s, 4H, benzene H), 7.38 (s, 2H, =CH–), 9.33 (br s, 7H, OH).

20a: 10%, 1.5 h, 143–145 °C, methanol/water, silica gel chloroform/methanol/formic acid 10:1:0.1; IR cm^{-1} 3250 (OH), 1630 (CO ester) and 1580 (CO ketone); ^1H NMR (DMSO-d_6) δ 1.08 (t, 3H, CH_3), 2.60–3.06 (m, 7H, CH, CH_2 and OH), 3.98 (q, 2H, CH_2CH_3), 6.77–6.95 (s, 6H, benzene H), 7.48 (s, 2H, =CH–), 9.40 (br s, 4H, OH).

20b: 55%, 1.5 h, 230–232 °C, methanol, silica gel chloroform/methanol/formic acid 10:1:0.1; IR cm^{-1} 3250 (OH), 1680 (CO acid) and 1570 (CO ketone); ^1H NMR (DMSO-d_6) δ 2.07 (m, 1H, CH), 2.36 (d, 2H, CH_2COOH), 2.60–3.06 (m, 4H, CH_2), 6.78–6.97 (s, 6H,

benzene H), 7.48 (s, 2H, =CH–), 9.84 and 12.30 (2 br s, 5H, OH and COOH).

20c: 46%, 2.5 h, 135–137 °C, methanol, silica gel chloroform/methanol/formic acid 10:1:0.1; IR cm^{-1} 3320 (OH), 1680 (CO ester) and 1580 (CO ketone); ^1H NMR (DMSO-d_6) δ 1.13 (t, 3H, CH_3), 2.29–3.00 (2 m, 7H, CH and CH_2), 4.00 (q, 2H, CH_2CH_3), 6.49 (s, 4H, benzene H), 7.38 (s, 2H, =CH–), 9.60 (br s, 6H, OH).

20d: 8%, 2 h, >300 °C, methanol/water, silica gel chloroform/methanol/formic acid 10:2:0.1; IR cm^{-1} 3300 (OH), 1680 (CO acid) and 1560 (CO ketone); ^1H NMR (DMSO-d_6) δ 2.06 (m, 1H, CH), 2.32 (d, 2H, CH_2COOH), 2.60 and 3.06 (2m, 4H, CH_2), 6.51 (s, 4H, benzene H), 7.37 (s, 2H, =CH–), 8.30 and 15.30 (2 br s, 7H, OH).Derivatives **17a** and **17b** were obtained by reacting 1-indanone or 1,3-indandione, respectively, with trihydroxybenzaldehyde, in the 1:1 ratio.

17a: 41%, 2.5 h, 290–292 °C, dimethylformamide/water; IR cm^{-1} 3500, 3410 and 3120 (OH), 1600 (CO); ^1H NMR (DMSO-d_6) δ 4.00 (s, 2H, CH_2), 6.78 (s, 2H, benzene H), 7.30 (s, 1H, =CH–), 7.43–7.70 (m, 3H, indan C4-H, C5-H and C6-H), 7.76 (d, J_o = 7.4 Hz, 1H, indan C7-H), 9.05 (br s, 3H, OH).

17b: 27%, 2 h, >300 °C, dimethylformamide/water; IR cm^{-1} 3330 and 3100 (OH), 1640 (CO); ^1H NMR (DMSO-d_6) δ 7.57 (s, 1H, =CH–), 7.72 (s, 2H, benzene H), 7.92 (m, 4H, indan H), 9.50 and 9.73 (2br s, 3H, OH).

5.2.4. General procedure for the preparation of compounds 18. Example: 5-(3,4,5-trihydroxybenzylidene)-barbituric acid (18a). A mixture of 3,4,5 - trihydroxybenzaldehyde (0.67 g, 3.9 mmol) and barbituric acid (0.50 g, 3.9 mmol) in water (10 mL) was refluxed for 0.5 h. The precipitate which formed was filtered and washed with water, ethanol, and light petroleum ether, in turn, to obtain 0.91 g of pure **18a** (88% yield): mp >300 °C (from dimethylsulfoxide/water); IR cm^{-1} 3500, 3480, 3310 and 3210, (OH and NH), 1730, 1680 and 1650 (CO); ^1H NMR (DMSO-d_6) δ 7.54 (s, 2H, benzene H), 8.02 (s, 1H, =CH–), 9.50 (br s, 3H, OH), 11.10 (br s, 2H, NH).

Yield (%), reaction time (h), mp (°C), recrystallization solvent; IR (cm^{-1}); ^1H NMR data are reported for:

18b: 77%, 0.5 h, >300 °C, dimethylsulfoxide/water; IR cm^{-1} 3300 (OH and NH), 1640 (CO); ^1H NMR (DMSO-d_6) δ 7.60 (s, 2H, benzene H), 8.03 (s, 1H, =CH–), 9.62 (br s, 3H, OH), 12.22 (br s, 2H, NH).

5.3. Microbiology

5.3.1. Integrase and DNA substrates. Expression of the rIN protein with an amino-terminal polyhistidine tag was obtained by IPTG induction of the *Escherichia coli* strain BL21(DE3) containing the pINSD.His vector. Protein purification was carried out following the Novagen procedure, except for the presence of 5 mM CHAPS in Binding Buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris–HCl pH 7.9), Washing Buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris–HCl pH 7.9) and Elute Buffer (1 M imidazole, 0.5 M NaCl, 50 mM Tris–HCl pH 7.9, 5 mM β-mercaptoethanol).

The following oligonucleotides representing the terminal 21 nucleotides of the HIV-1 U5 LTR were used: **B**: 5′-ACT GCT AGA GAT TTT CCA CAC-3′ (minus strand); **C**: 5′ GTG TGG AAA ATC TCT AGC A-3′ (plus strand); **D** (dumbbell): 5′-TGC TAG TTC TAG CAG GCC CTT GGG CCG GCG CTT GCG CC-3′. For standard 3′-processing assays, **B** was annealed with **C** in 0.1 M NaCl by heating at 80 °C and slowly cooling to room temperature overnight. This double stranded substrate was labeled by introducing at the 3′ end of **C** the two missing nucleotides using α-[^{32}P]-dGTP, cold dTTP and Klenow polymerase. Unincorporated α-[^{32}P]-dGTP was separated from the duplex substrate by two consecutive runs through G-25 Sephadex quick spin columns. Dumbbell oligonucleotide was labelled at the 5′ end using γ-[^{32}P]ATP and T4 polynucleotide kinase, purified through a G-25 Sephadex quick spin column and self annealed.

5.3.2. rIN assays. Standard 3′-processing and strand transfer reaction conditions were: 10 mM Hepes pH 7.5, 10 mM $MnCl_2$, 1 mM DTT, 40 mM NaCl, 5 nM of 40mer oligo **B**/**C** and 100 nM rIN (considered as monomer). Incubation was carried out at 37 °C for 1 h in a volume of 15 μL. Reactions were stopped by adding 7.5 μL of sample buffer (96% formamide, 20 mM EDTA, 0.08% bromophenol blue and 0.25% xylene cyanol), samples were layered into a denaturing 15% polyacrylamide gel (7 M urea, 0.09 M Tris borate, pH 8.3, EDTA 2 mM, 15% acrylamide) and run for 1 h at 80 Watt. Reaction products were monitored and quantified with a Bio-Rad Personal FX Phosphoimager. For disintegration activity, the standard reaction conditions were the same as described for the 3′-processing assay with the exception of the presence of 3 nM oligo **D** (dumbbell substrate). Reactions were incubated at 37 °C for 1 h in a volume of 15 μL, reactions were stopped with sample buffer and analyzed by denaturing PAGE as described above.

5.3.3. Biological assays. Compounds were solubilized in DMSO at 200 mM and then further diluted. MT-4 cells were grown at 37 °C in a 5% CO_2 atmosphere in RPMI 1640 medium supplemented with 10% fetal calf serum (FCS), 100 IU/mL penicillin G, and 100 μg/mL streptomycin. Cell cultures were checked periodically for the absence of mycoplasma contamination with a MycoTect Kit (Gibco). HIV-1 stock solutions had titers of 4.5×10^6 50% cell culture infectious dose ($CCID_{50}$)/mL. Cytotoxicity evaluation was based on the viability of mock-infected cells, as monitored by the MTT method. Activity of the compounds against the HIV-1 multiplication in acutely infected cells was based on inhibition of virus-induced cytopathic effect in MT-4 cells and was determined by the 3-(4,5-dimethylthiazol-1-yl)-2,5-diphenyltetrazolium bromide (MTT) method.[41]

5.3.4. rRT assays. Purified rRT was assayed for its RNA-dependent DNA polymerase associated activity in

a volume of 50 μL containing 50 mM Tris–HCl pH 7.8, 80 mM KCl, 6 mM $MgCl_2$, 1 mM DTT, 0.1 mg/mL BSA, 0.5 OD_{260} unit/mL template:primer [poly(rC)-oligo$(dG)_{12-18}$] and 10 μM [^{3}H]dGTP (1 Ci/mmole). After a 30 min incubation at 37 °C samples were spotted on glass fiber filter (whatman GF/A) and the acid-insoluble radioactivity was determined.

5.3.5. Antitumor activities. Compounds were dissolved in DMSO at 100 mM and diluted in culture medium. Leukemia cells were grown in RPMI-1640 medium supplemented with 10% FCS, 100 units/mL penicillin G and 100 μg/mL streptomycin. Solid tumor-derived cells were grown in their specific media supplemented with 10% FCS and antibiotics. Cell cultures were incubated at 37 °C in a humidified, 5% CO_2 atmosphere. The absence of mycoplasma contamination was checked periodically by the Hoechst staining method. Exponentially growing human leukemia and solid tumor-derived cells were resuspended at a density of 1×10^5 cells/mL in growth medium containing serial dilutions of the drugs. Cell viability was determined after 96 h at 37 °C by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) method. Cell growth at each drug concentration was expressed as percentage of untreated controls. The concentration resulting in 50% growth inhibition (IC_{50}) was determined by linear regression analysis.

5.3.6. Quantitation of protein-linked DNA breaks. A modified in vivo K-SDS co-precipitation assay[33] was used to quantify protein-linked DNA breaks (PLDB). Briefly, KB cells were seeded at a density of 2.5×10^5/mL and labeled with [^{14}C]thymidine (0.75 μCi/mL) for 48 h. Monolayers were washed twice and, after trypsinization, cells were divided into aliquots of 3×10^5 cells/mL and incubated for 1 h at 37 °C. Duplicate samples were treated with test drugs (at concentrations 10-fold higher than respective CC_{50} values) and further incubated for 1 h. Then, cells were collected by centrifugation at 2000*g* for 10 min and resuspended in 1 mL of warm lysis buffer (1.5% SDS, 5 mM EDTA, 0.4 mg/mL salmon sperm DNA), and the viscous cell lysates were sheared by passage through a 22-gauge needle five times. After a 10 min incubation at 65 °C, KCl was added to a final concentration of 100 mM, samples were chilled in ice bath for 10 min and centrifuged at 3500*g* for 10 min at 4 °C. Pellets were resuspended in 1 mL of warm washing buffer (10 mM Tris–HCl, pH 8.0, 100 mM KCl, 1 mM EDTA, 0.1 mg/mL salmon sperm DNA) and incubated for 10 min at 65 °C, chilled in ice bath and centrifuged as above. After a second washing step, pellets were solubilized at 65 °C in 400 μL of water and the radioactivity was determined in a scintillation counter.

5.4. Molecular modeling and 3-D QSAR

5.4.1. Molecular modeling and molecular alignment. The compounds were built in the most extended conformation using standard bond distances and angles as defined in the Macromodel[41] graphical interface Maestro. Selection of the compound conformations was performed by using the AMBER force field and using molecular dynamics with simulated annealing as implemented in MacroModel version 7.1.[42] Each molecule has been energy minimized to a low gradient. The nonbonded cutoff distances were set to 20 Å for both van der Waals and electrostatic interactions. An initial random velocity to all atoms corresponding to 300 K was applied. Three subsequent molecular dynamics run were then performed. The first was carried out for 10 ps with a 1.5 fs time-step at a constant temperature of 300 K for equilibration purposes. The next molecular dynamics was carried out for 20 ps, during which the system is coupled to a 150 K thermal bath with a time constant of 5 ps. The time constant represents approximately the half-life for equilibration with the bath; consequently the second molecular dynamics command caused the molecule to slowly cool to approximately 150 K. The third and last dynamics cooled the molecule to 50 K over 20 ps. A final energy minimization was then carried out for 250 iterations using conjugate gradient. The minimizations and the molecular dynamics were in all cases performed for all the compounds in aqueous solution. The obtained conformer from the final minimization of each compound was used for the molecular alignment and used to calculate the 3-D-QSAR model. For the molecular alignment the atom by atom procedure was used using as fitting points a common substructure (Fig. 6).

5.4.2. GRID Calculations. The interaction energies were calculated by using GRID (version 20)[43] with a grid spacing of 1 Å and the grid dimensions (Å): X_{min}/X_{max}, −24.0/6.0; Y_{min}/Y_{max}, −2.0/24.0; and Z_{min}/Z_{max}, 20.0/44.0.

5.4.3. GOLPE analyses. PLS models were calculated with GOLPE 4.5.12.[44]

5.4.4. Probe selection. In this study we started with 10 probes (BOTH, C_{sp3}, DRY, OH_2, Mg^{2+}, NH_{amide}, N_{sp2}, N^+_{sp3}, O_{sp2}, O_{CO_2}) according to the nature of the rIN active site. For 3-D QSAR, water probes were selected on the basis of the PCA and PLS tools available in GOLPE. Plotting the PCA and PLS scores, we evinced that the water, NH_{amide}, N_{sp2}, N^+_{sp3}, O_{sp2} and O_{CO_2}, probes were those better describing the molecular diversity of the Set A (Figs. 7A and B). Selection of the water probe (OH_2) was mainly based on the experimental/recalculated plot (Fig. 7C) of a preliminary PLS analysis where the OH_2 probe visually performed a better correlation other than a good discrimination between active and inactive/less active derivatives (Fig. 7D).

5.4.5. Variable preselection. The resulting probe-target interaction energies for each compound were unfolded to produce one-dimensional vector variables for each compound, which were assembled in the so-called X matrix. This matrix was pretreated by first using a cutoff

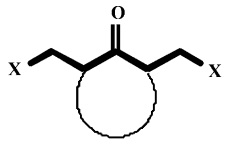


Figure 6. Common scaffold used for molecular alignment.

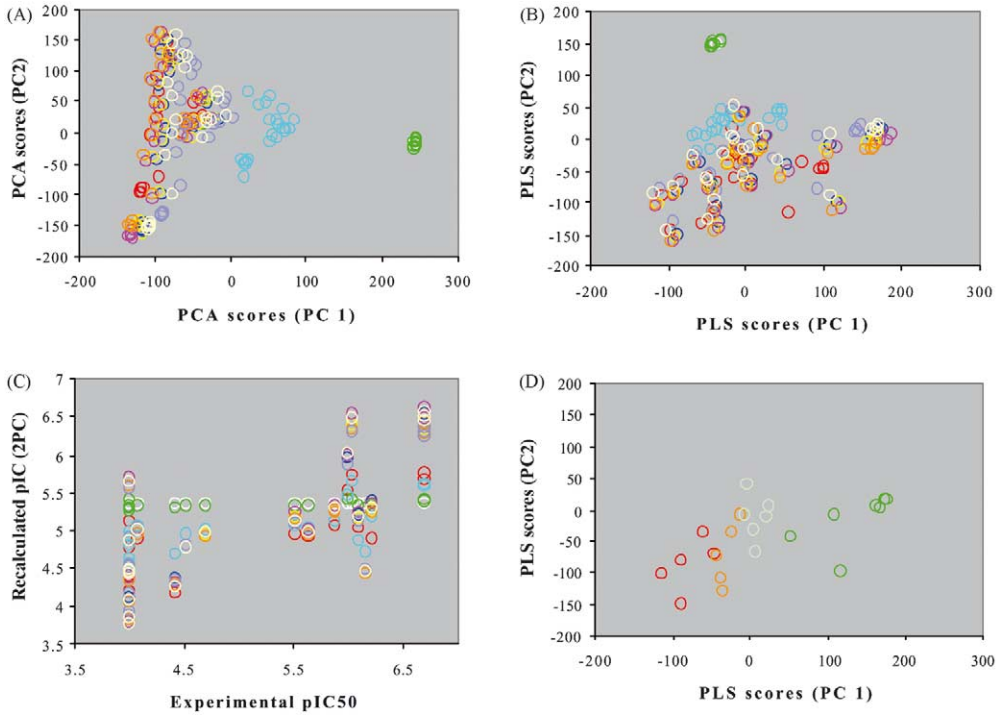


Figure 7. PCA and PLS score plots on the Set A using 10 probes (A and B). Experimental versus recalculated pIC_{50} from the preliminary PLS (C). PLS score plot on only the OH2 probe (D), red means less active and green more active. GRID probes are color coded as following: BOTH white, C_{sp3} red, DRY green, OH2 blu, Mg^{2+} cyan, NH_{amide} yellow, N_{sp2} magenta, N^{+}_{sp3} orange, O_{sp2} ivory, O_{CO_2} violet. For the plot the first two principal components were use.

of 5 kcal/mol to produce a more symmetrical distribution of energy values and then zeroing small variable values and removing variables with small standard deviation, using appropriate cutoffs. In addition, variables taking only two and three distribution were also removed. From an initial number of 20,925 the pretreatment selected 3051 active variable.

5.4.6. Smart region definition (SRD). A number of seeds (1000) were selected using a D-optimal design criterion in the *weight space*. Structural differences between different molecules in the series will be reflected in groups of variables, and therefore groups were generated around each seed in the 3-D-space. Variables with a distance of no more than 2 Å to the seeds were included in the groups. If two neighbouring groups (with a distance smaller than 10 Å) contained the same information the groups were collapsed. The groups were used in the variable selection procedure replacing the original variables. The effect of the groups on the predictivity was evaluated and groups instead of individual variables were removed from the data file.

5.4.7. Region selection. The effect of the grouped variables on the predictivity was evaluated using a fractional factorial design (FFD) procedure. A number of reduced models (twice the number of variables) were built removing some of the variables according to the FFD design. The effect of dummy variables (20%) on the predictivity was calculated and only if a variable had a positive effect on the predictivity larger than the effect of the average dummy variable was the variable included in the final model. The FFD selection was repeated until the R^2 and Q^2 value did not increase significantly. In the FFD selection the cross validation was conducted using five random groups for 20 times and a maximum of tree principal components.

5.4.8. Cross-validation. The models were validated using random groups. Molecules were assigned in a random way to five groups of equal size. Reduced models were built keeping out one group at a time. The formation of the groups was repeated 20 times and using a maximum model dimensionality of three components.

6. Analytical results

Compd	Elemental analysis calculated/found				
	C	H	N	S	Cl
13b	54.58	3.23	—	—	37.91
	54.63	3.41	—	—	37.73
14b	58.00	3.89	—	—	34.24
	58.10	3.93	—	—	34.14
14c	55.11	2.92	—	—	34.24
	55.21	2.91	—	—	34.28
14d	54.84	3.39	—	—	34.08
	54.71	3.45	—	—	34.21
14f	52.80	3.27	—	7.42	32.81
	52.71	3.35	—	7.54	32.63

(*continued on next page*)

Compd	Elemental analysis calculated/found				
	C	H	N	S	Cl
14g	52.61	3.72	9.69	—	8.17
	52.51	3.65	9.78	—	8.28
14h	52.61	3.72	9.69	—	8.17
	52.57	3.81	9.71	—	8.23
14i	61.62	5.17	3.59	—	9.09
	61.75	5.28	3.31	—	9.10
14j	53.64	4.05	9.38	—	7.92
	53.71	4.16	9.28	—	7.93
14k	53.64	4.05	9.38	—	7.92
	53.71	4.15	9.21	—	7.83
14l	62.45	5.49	3.47	—	8.78
	62.21	5.63	3.44	—	8.56
14m	54.61	4.36	9.10	—	7.68
	54.75	4.21	9.31	—	7.53
14n	54.61	4.36	9.10	—	7.68
	54.58	4.31	9.28	—	7.51
14o	67.02	5.19	3.01	—	7.61
	67.18	5.25	3.10	—	7.35
14p	59.60	4.23	8.02	—	6.77
	59.71	4.14	8.09	—	6.81
14q	50.60	4.23	8.02	—	6.77
	50.73	4.18	8.10	—	6.71
15a	55.96	4.45	3.43	—	8.69
	55.98	4.31	3.21	—	8.75
15b	55.95	4.78	3.32	—	8.40
	56.78	4.83	3.37	—	8.35
15c	57.87	5.09	3.21	—	8.13
	57.93	5.21	3.11	—	8.01
15d	62.72	4.86	2.81	—	7.12
	62.85	4.97	2.56	—	7.21
15e	64.86	4.90	—	—	—
	65.00	4.71	—	—	—
15f	61.29	4.33	—	—	—
	61.35	4.21	—	—	—
15g	58.76	4.15	—	8.25	—
	58.81	4.21	—	8.14	—
16	64.04	4.53	—	—	—
	64.14	4.35	—	—	—
17a	71.64	4.51	—	—	—
	71.77	4.50	—	—	—
17b	68.09	3.57	—	—	—
	68.21	3.41	—	—	—
18a	50.10	3.05	10.60	—	—
	50.21	3.15	10.43	—	—
18b	47.14	2.88	10.00	11.44	—
	47.16	2.93	1.05	11.47	—
19a	67.31	5.40	—	—	—
	67.28	5.53	—	—	—
19b	65.97	4.74	—	—	—
	65.91	4.83	—	—	—
19c	62.44	5.01	—	—	—
	62.31	5.07	—	—	—
19d	60.87	4.38	—	—	—
	60.91	4.40	—	—	—
20a	67.91	5.70	—	—	—
	67.98	5.64	—	—	—
20b	66.66	5.09	—	—	—
	66.71	5.38	—	—	—
20c	63.15	5.30	—	—	—
	63.21	5.43	—	—	—
20d	60.58	4.84	—	—	—
	60.49	4.91	—	—	—

Acknowledgements

Authors thank the Italian Ministero della Sanità, Istituto Superiore di Sanità- XI Progetto AIDS 1999 (grants no. 40C.8 and no. 40C.47) for financial support. Acknowledgements are also due to Italian MURST (40%).

References and notes

1. Hong, H.; Neamati, N.; Wang, S.; Nicklaus, M. C.; Mazumder, A.; Zhao, H.; Burke, T. R., Jr.; Pommier, Y.; Milne, G. W. A. *J. Med. Chem.* **1997**, *40*, 930.
2. Neamati, N.; Hong, H.; Mazumder, A.; Wang, S.; Sunder, S.; Nicklaus, M. C.; Milne, G. W. A.; Proksa, B.; Pommier, Y. *J. Med. Chem.* **1997**, *40*, 942.
3. Nicklaus, M. C.; Neamati, N.; Hong, H.; Mazumder, A.; Sunder, S.; Chen, J.; Milne, G. W. A.; Pommier, Y. *J. Med. Chem.* **1997**, *40*, 920.
4. Zhao, H.; Neamati, N.; Hong, H.; Mazumder, A.; Wang, S.; Sunder, S.; Milne, G. W. A.; Pommier, Y.; Burke, T. R., Jr. *J. Med. Chem.* **1997**, *40*, 242.
5. Zhao, H.; Neamati, N.; Mazumder, A.; Sunder, S.; Pommier, Y.; Burke, T. R., Jr. *J. Med. Chem.* **1997**, *40*, 1186.
6. Pommier, Y.; Pilon, A. A.; Bajaj, K.; Mazumder, A.; Neamati, N. *Antiviral Chem. Chemother.* **1997**, *8*, 463.
7. Fesen, M. K.; Kohn, K. W.; Leteurtre, F.; Pommier, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2399.
8. Burke, T. R., Jr.; Fesen, M. R.; Mazumder, A.; Wang, Y.; Carothers, A. M.; Grunberger, D.; Driscoll, J.; Khon, K.; Pommier, Y. *J. Med. Chem.* **1995**, *38*, 4171.
9. Fesen, M. K.; Pommier, Y.; Leteurtre, F.; Hiroguchi, S.; Young, J.; Kohn, K. W. *Biochem. Pharmacol.* **1994**, *48*, 595.
10. Mazumder, A.; Raghavan, K.; Weistein, J.; Kohn, K. W.; Pommier, Y. *Biochem. Pharmacol.* **1995**, *49*, 1165.
11. Mazumder, A.; Neamati, N.; Sunder, S.; Schultz, J.; Pertze, H.; Eich, E.; Pommier, Y. *J. Med. Chem.* **1997**, *40*, 3057.
12. Artico, M.; Di Santo, R.; Costi, R.; Novellino, E.; Greco, G.; Massa, S.; Tramontano, E.; Marongiu, M. E.; De Montis, A.; La Colla, P. *J. Med. Chem.* **1998**, *41*, 3948.
13. Li, C. J.; Zhang, L. J.; Dezube, B. J.; Crumpacker, C. S.; Pardee, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1839.
14. Mazumder, A.; Gazit, A.; Levitzki, A.; Nicklaus, M.; Yung, J.; Kohlhagen, G.; Pommier, Y. *Biochemistry* **1995**, *34*, 15111.
15. LaFemina, R. L.; Graham, P. L.; LeGrow, K.; Hastings, J. C.; Wolfe, A.; Young, S. D.; Emini, E. A.; Hazuda, D. J. *Antimicrob. Agents Chemother.* **1995**, *39*, 320.
16. McDougall, B. R.; King, P. J.; Wu, B. W.; Hostomsky, Z.; Reinacke, M. G.; Robinson, W. E., Jr. *Antimicrob. Agents. Chemother.* **1998**, *42*, 140.

17. Robinson, W. E., Jr.; Reinecke, M. G.; Abdel-Malek, S.; Jia, Q.; Chow, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6326.
18. Lin, Z.; Neamati, N.; Zhao, H.; Kiryu, Y.; Turpin, J. A.; Aberham, C.; Strebel, K.; Kohn, K.; Wiotvrouw, M.; Pannecouque, C.; Debyser, Z.; De Clercq, E.; Rice, W. G.; Pommier, Y.; Burke, T. R., Jr. *J. Med. Chem.* **1999**, *42*, 1401.
19. King, P. J.; Ma, G.; Miao, W.; Jia, Q.; McDougall, B. R.; Reinecke, M. G.; Cornell, C.; Kuan, J.; Kim, T. R.; Robinson, W. E., Jr. *J. Med. Chem.* **1999**, *42*, 497.
20. Mekouar, K.; Mouscardet, J.-F.; Desmaële, D.; Subra, F.; Leh, H.; Savouré, D.; Auclair, C.; d'Angelo, J. *J. Med. Chem.* **1998**, *41*, 2846, and references cited therein.
21. Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J. A.; Espeseth, A.; Gabryelski, L.; Schleif, W.; Blau, C.; Miller, M. D. *Science* **2000**, *287*, 646.
22. Buolamwini, J. K.; Assefa, H. *J. Med. Chem.* **2002**, *45*, 841.
23. Raghavan, K.; Buolamwini, J. K.; Fesen, M. R.; Pommier, Y.; Kohn, K. W.; Weinstein, J. N. *J. Med. Chem.* **1995**, *38*, 890.
24. *SYBYL, Molecular Modeling Software*, version 6.2; Tripos Inc.: St. Louis, MO, 1995.
25. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
26. Clementi S. *GOLPE 3.0*; Multivariate Infometric Analyses (MIA): Perugia, Italy, 1995; SGI.
27. Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849.
28. Alonso, F.; Micò, I.; Nàjera, C.; Sansano, J. M.; Yus, M.; Ezquerra J. Yruretagoyena, B.; Gracia, I. *Tetrahedron* **1995**, *51*, 10259.
29. Ungnade, H. E.; Morris, F. V. *J. Am. Chem. Soc.* **1948**, *70*, 1898.
30. Chow, S. A.; Vincent, K. A.; Ellison, V.; Brown, P. O. *Science* **1992**, *255*, 723.
31. Chow, A.; Brown, P. O. *J. Virol.* **1994**, *68*, 3896.
32. Stanwell, C.; Ye, B.; Yuspa, S. H.; Burke, T. R., Jr. *Biochem. Pharmacol.* **1996**, *52*, 475.
33. Lee, K. H.; Imakura, Y.; Haurna, M; Beers, S. A.; Thurston, L. S.; Dai, H. J.; Chen, C. H. *J. Nat. Prod.* **1989**, *52*, 606.
34. Baroni, M.; Constantino, G.; Cruciani, G.; Valigi, R.; Clementi, S. *Quant. Struct.-Act. Relats.* **1993**, *12*, 9.
35. Dunn, W. J., III; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. *Quant. Struct.-Act. Relats.* **1984**, *3*, 131.
36. Matter, H.; Schwab, W. *J. Med. Chem.* **1999**, *42*, 4506.
37. Pastor, M.; Cruciani, G.; Clementi, S. *J. Med. Chem.* **1997**, *40*, 1455.
38. Beaver, D. J.; Roman, D. P.; Stoffel, P. J. *J. Am. Chem. Soc.* **1957**, *79*, 1236.
39. Dibella E. P. US 3,389,986 (Cl. 71-123,) 25 June 1968, *Chem. Abstr.* 1968, *69*, 51812y.
40. Rovnyak, G. C.; Millonig, R. C.; Schwartz, J.; Shu, V. *J. Med. Chem.* **1982**, *25*, 1482.
41. Pauwels, R.; Balzarini, J.; Baba, M.; Snoeck, R.; Schols, D.; Herdewijn, P.; Desmyster, J.; De Clerq, E. *J. Virol. Methods* **1988**, *20*, 309.
42. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.
43. *GRID* 20; Molecular Discovery Ltd.: Oxford, UK, 2001.
44. *GOLPE 4.5*.12; Multivariate Infometric Analyses: Viale del Castagni, 16 Perugia, Italy, 1999.