

Tautomers of styrylquinoline derivatives containing a methoxy substituent: Computation of their population in aqueous solution and their interaction with RSV integrase catalytic core[⊛]*

Mohammed Ouali¹, Cyril Laboulais¹, Hervé Leh², David Gill¹, Emil Xhuvani¹, Fatima Zouhiri³, Didier Desmaële³, Jean d'Angelo³, Christian Auclair², Jean-François Mouscadet² and Marc Le Bret¹✉

¹Laboratoire de Physicochimie et de Pharmacologie des Macromolécules Biologiques, CNRS UMR 8532, LBPA, Ecole Normale Supérieure de Cachan, 61 avenue du Président Wilson, 94235 Cachan, France; ²CNRS UMR 8532, Institut Gustave Roussy, PR11, 39 rue Camille Desmoulins, 94805 Villejuif, France; ³Unité de Chimie Organique, CNRS UPRES A 8076, Centre d'Etudes Pharmaceutiques, 5 rue J.B. Clément, Université de Paris Sud, 92296 Chatenay Malabry, France

Received: 25 October, 1999

Keywords: solvent effects, docking, integrase inhibitors, styrylquinoline

8-Hydroxy-2-[2-(3-hydroxy-4-methoxyphenyl)ethenyl]-7-quinoline carboxylic acid and 8-hydroxy-2-[2-(3-methoxy-4-hydroxyphenyl)ethenyl]-7-quinoline carboxylic acid inhibit the processing and strand transfer reactions catalyzed by HIV-1 integrase with an IC₅₀ of 2 μM. Some of their spectral properties are briefly reported. Their fluorescence is so weak that it is of no use in an experimental determination of the binding to the protein and we resorted to computer simulation. Both styrylquinoline derivatives, in their monoanionic form, have several dozens of tautomers and each of these forms has four planar rotamers. In this work computer simulations have been performed to determine which tautomer is the most abundant in aqueous solution and which binds

*Presented at the 7th International Symposium on Molecular Aspects of Chemotherapy, September 8-11, 1999, Gdańsk, Poland.

⊛This work has been made possible through the access to IDRIS computational facilities (project #980807) and to the SGI/O2000 multiprocessor of the Pôle Parallélisme IdF Sud, and through the financial support of ARC #2080, CNRS, ANRS and Sidaction. H.L., C.L. and E.X. are grateful for ANRS, MENRT and TEMPUS fellowships.

✉To whom correspondence should be addressed; phone: (33) 147 405 997, fax: (33) 147 402 479, e-mail: mlebret@lbpa.ens-cachan.fr

Abbreviations: ASV, avian sarcoma virus; B3LYP, Becke's three parameter hybrid functional using the Lee Yang & Parr correlation functional; CPCM, polarizable conductor calculation model; HF, Hartree-Fock level of theory; IC₅₀ drug concentration that reduces the enzymatic activity to 50%; IN, integrase; RSV, Rous sarcoma virus.

to the Rous sarcoma virus (RSV) integrase catalytic core. As the substituents on the quinoline moiety are the same as on salicylic acid, the energies of hydroxy benzoic acid tautomers were also computed both *in vacuo* and embedded in a continuous medium which had the dielectric constant of bulk water, using the recent CPCM technique. The CPCM method was then applied to the two integrase inhibitors to estimate the tautomer population in water. The binding site of the compounds on the RSV integrase catalytic core was determined through a docking protocol, consisting of coupling a grid search method with full energy minimization. The designed method is a way leading to identification of potent integrase inhibitors using *in silico* experiments.

AIDS is essentially a viral disease and should be treated by antiretroviral agents (Ho *et al.*, 1995). The advent of combination antiretroviral therapy has made it possible to suppress the replication of HIV-1 in infected persons to such an extent that the virus becomes undetectable in the plasma for more than two years, but seems to persist in peripheral-blood mononuclear cells. This means that HIV-1 infection can be controlled but not eradicated by current treatments (Zhang *et al.*, 1999; Furtado *et al.*, 1999). It is therefore important to search for agents that could block the virus at these steps of its replicative cycle, which are not affected by current treatments. From this standpoint, HIV DNA integration into the genomic DNA of host cells, a crucial step in the replicative life cycle of the virus, constitutes a particularly attractive target for AIDS chemotherapy drugs, including potential synergy with currently available HIV reverse transcriptase and protease inhibitors (De Clercq, 1995; Pommier & Neamati, 1999). In HIV-1 and other retroviruses such as RSV, integration is mediated by a viral enzyme, integrase, which is necessary for the production of progeny viruses (LaFemina *et al.*, 1992). The integration function is composed of two steps both consisting of the nucleophilic attack of a phosphoester bond by the lone pair of a hydroxyl group (Engelman *et al.*, 1991). In the first step, called 3' processing, the enzyme catalyzes the hydrolysis of a phosphoester bond, which removes two nucleotides from either of viral DNA strands and forms free 3'-OH groups at a conserved CA dinucleotide sequence. In the second step, called strand transfer, these groups are used

as nucleophilic agents and attack phosphoester bonds in opposite strands of the genomic DNA separated by five base pairs in HIV-1 and four base pairs in RSV (reviews: Brown, 1990; Katz & Skalka, 1994).

The HIV-1 integrase produced in an *Escherichia coli* expression system, can carry out both 3' processing and strand transfer *in vitro* in the presence of divalent cations such as Mg^{++} and Mn^{++} (Bushman & Craigie, 1991). It can also, in the same ionic environment, carry out the apparent reversal of the transfer step in the presence of a synthetic Y-shaped oligonucleotide (Chow *et al.*, 1992). Not the entire protein, but only its catalytic core, IN(50-212), containing the residues 50 to 212, is required for the disintegration step, showing that the core contains the enzyme active site. Site-directed mutagenesis experiments showed that catalytic activity is abolished by the substitution of any of the three absolutely conserved carboxylate residues, two aspartic residues D64, D116 and a glutamic residue E152 (the so-called D,D-35E motif). The same arrangement of catalytically essential carboxylates is conserved not only in all retroviral integrases but also in enzymes that catalyze cognate reactions such as prokaryotic transposases. This observation strongly suggested that their spatial conformation should be similar to the conformation of the corresponding amino acids D64, D121 and E157 in the RSV integrase catalytic core. Although this seemed to be disproved by the first crystal structures of the HIV-1 catalytic core (Dyda *et al.*, 1994), it was finally confirmed by two independent groups (Maignan *et al.*, 1998; Goldgur *et al.*, 1998). The dication is in the immediate vicin-

ity of the essential carboxylates, which is a common feature among enzymes in nucleic acid biochemistry (Cowan, 1998).

Some years ago, we designed molecules that could bridge two Mg^{++} cations, because, at the time, in the absence of crystallographic data, the integrase active site was thought to have a conformation similar to that of the β polymerase Klenow fragment (Beese & Steitz, 1991). These compounds had two moieties: the quinoline part could bind a cation through its 7 carboxyl or the region between N and O8 since 8-hydroxy quinoline is known to chelate bivalent and trivalent metal ions (Bardez *et al.*, 1997). The hydroxyl groups that reside on the phenyl ring could bind a second cation. Actually, styrylquinoline derivatives having a catechol or a resorcinol moiety have been shown to be potent HIV-1 integrase inhibitors *in vitro* with IC_{50} in the $1 \mu M$ range (Mekouar *et al.*, 1998). Moreover, these compounds displayed an antiviral activity in a *de novo* infection assay of CEM4fx cells and, therefore, they represent an exciting new lead for the design of new anti-HIV drugs. These styrylquinoline derivatives were likely to target the integrase catalytic core since they acted as inhibitors in the disintegration assay performed with the active site-containing deletion mutant (Mekouar *et al.*, 1998). However, a more accurate knowledge of molecular interactions between these compounds and their presumed target was needed to further improve their pharmacological properties. Consequently, Ouali *et al.* (2000) determined the thermodynamic and geometric parameters of their binding to the catalytic core of RSV integrase by intensive use of computer power. Each of the 14 styrylquinoline derivatives was systematically rotated and translated on a computer relative to the RSV catalytic core crystallographic coordinates. The solid displacements that gave a good overlap between the drug and protein surfaces were memorized. The 1100 best conformations were then minimized to partially take into account the flexibility of the ligand and of the

protein. This method allowed us to predict that the most likely binding sites of these 14 styrylquinoline derivatives were in the immediate vicinity of that Mg^{++} cation which is essential for the activity of the enzyme and that is reported in the crystallographic data (Bujacz *et al.*, 1995). The same could be achieved for various tautomeric forms. It was shown that the keto form (7-CO₂H quinoline 8-hydroxylate) had a more favorable interaction energy than the enol form (8-OH quinoline 7-carboxylate). Moreover, we showed that the interaction energy between drugs and RSV integrase catalytic core calculated for the best-fitted models perfectly correlates with biological data. Finally and surprisingly, the catechol moiety was much less important than the quinoline part for the binding. This suggested that substituting one of the hydroxy group by a methoxy should have only a minor effect on the binding. This work was undertaken to check this hypothesis and we present here the results about two styrylquinoline derivatives containing methoxy substitutions. The inhibitory power of a tautomer depends on two conditions: the active form should be available in water and have the greatest affinity for the protein. How much of the active form is available can be determined if solvent effects are taken into account in the computation of the electronic structure. Usually this is done by placing the solute molecule in a cavity surrounded by a polarizable continuum whose reaction field modifies the energy and the properties of the solute. We used here the conductor-like solvation model (COSMO) first proposed by Klamt & Schürmann (1993) and now available as the conductor type polarisable continuum model (CPCM) option in the GAUSSIAN98 package (Frisch *et al.*, 1998; Barone & Cossi, 1998; Rega *et al.*, 1999). In this implementation, each atom of the solute is surrounded by a spherical cavity with a conductive boundary. Water is assumed to be a conductor. Hydrogens are not treated as independent atoms, but are united with adjacent heavier atoms. The radius of

each atom is fitted to reproduce the solvation energies at a computational level (Hartree-Fock with a medium size basis set) allowing the study of relatively large systems (Barone *et al.*, 1997). This approach should yield an estimate of the availability of the various tautomers in water. Moreover, it is possible to classify the inhibitory power by docking the drugs to the RSV integrase catalytic core.

MATERIALS AND METHODS

Chemicals. The chemical formulae are shown in Fig. 1. Their syntheses have been described by Zouhiri *et al.* (2000). Drug 1 was an orange-red powder of sodium carboxylate: R7 = CO₂Na, R8 = OH, R3' = OCH₃ and R4' = OH. Drug 2 was a yellowish powder: R7 = CO₂H, R8 = OH, R3' = OH and

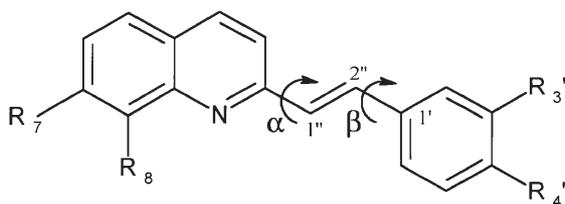


Figure 1. Formulae of the compounds studied. Drug 1: R3' = OCH₃, R4' = OH, R7 = CO₂H and R8 = OH. Drug 2: R3' = OH, R4' = OCH₃, R7 = CO₂H and R8 = OH. The dihedral angles α and β are equal to zero on this drawing.

R4' = OCH₃ that, in contrast to the former compound, was difficult to dissolve in Tris buffer. Spectra were measured on an Uvikon 941.

Modelling of protein structure. Before making any docking attempt, it was necessary to know the conformation of the catalytic core. First, the structure of the catalytic core of RSV integrase published by Bujacz *et al.* (1995) was minimized with our minimizer (Le Bret *et al.*, 1991) and the AMBER force field (Cornell *et al.*, 1995) to remove any close contact. The missing atoms in the crystallo-

graphic coordinate file were set using the XRAY option in the EDIT module of AMBER. The total coulombic charge of the catalytic core (with the divalent bication) was +10 proton charges. The effect of the minimization on the RSV integrase catalytic core is very small.

Docking algorithm. The docking method that was used in this study has been incorporated into our program MORCAD (Le Bret *et al.*, 1991). It consists of several steps: a grid search fit by rotating (15° increments) and translating (0.916 Å increment) the drug while holding the protein rigid as described by Katchalski-Katzir *et al.* (1992). A score function that takes into account the surface overlap and the electrostatic interaction (Rogers & Sternberg, 1984; Gabb *et al.*, 1997) is computed. The 1100 best conformations are minimized using a quasi newtonian algorithm (Le Bret *et al.*, 1991) to allow some flexibility. The force field is the AMBER forcefield (Cornell *et al.*, 1995) for amino acids, the Aqvist (1990) parameters for the Mg⁺⁺ divalent cation and a sigmoidal dielectric constant to mimic the aqueous medium (Lavery *et al.*, 1986). The parameters for the drugs were taken from those of the unmethylated compounds (Ouali *et al.*, 2000). A 10 Å cut-off was used in the nonbonded interactions during minimization. Once minimization of the whole system (protein + drug) was completed, the interaction energy was re-evaluated without the cut-off. The drug-catalytic core interaction energy is the sum of all interactions between any drug atom and any catalytic core atom. Such grid searches have the advantage of getting to the neighborhood of the correct solution (Shoichet, 1996). The full cycle requires 24 h of SGI R10000 195 Mhz processor time per tautomer.

Quantum mechanical calculations. The structure of the drugs was optimized through the *ab initio* program GAUSSIAN98 (Frisch *et al.*, 1998) at the restricted Hartree Fock approximation, using the 6-31G(d) basis. The electrostatic charges were calculated by the Merz-Kollman (Besler *et al.*, 1990) procedure

in each tautomer and each rotamer for use in the docking procedure. The Gibbs free energy difference was computed at 298.27 K using the Freq option of GAUSSIAN98.

RESULTS

Spectroscopic properties

The absorption spectrum of drug **1** is shown in Fig. 2. Drug **1** is weakly fluorescent in 0.1 M Tris buffer. This is not a homogeneous solution since the emission spectrum depends on the excitation wavelength. For instance, there is no detectable fluorescence when drug **1** is excited at 550 nm but a very weak fluores-

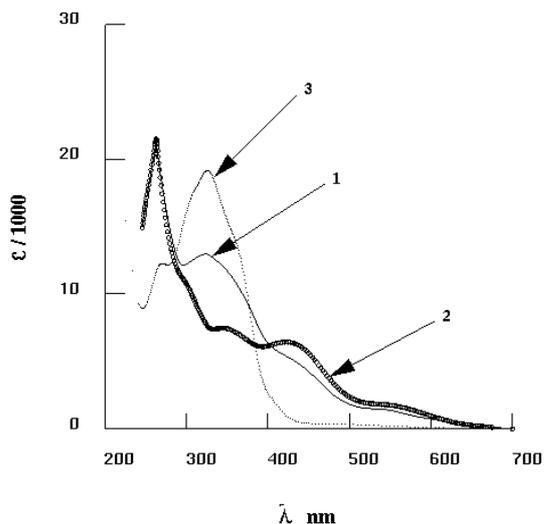


Figure 2. Absorption spectra in 0.1 M Tris, pH = 7.5.

Molar extinction coefficients of a fresh solution (10 mg/l) of **1** (curve 1) and 24 h later (curve 2). The molar extinction coefficient of **2** in the same conditions did not vary (curve 3).

cence can be measured when it is excited at 350 nm. When the compound is left in the aqueous buffer, its absorption spectrum varies by 10% during the first hour and becomes stabilized after a day. The kinetics is rather slow with respect to the time required to measure the IC_{50} . The weak fluorescence vanishes with time. Obviously the drug becomes degraded. The degradation might be related to

the degradation of stilbene: the major products in the photolysis of *trans*-stilbene in the near ultra violet are *trans*- and *cis*-stilbene. In the case of *cis*-stilbene photolysis the major products are *trans*- and *cis*-stilbene but dihydrophenanthrene also can be formed (Waldeck, 1991). No fluorescence of drug **2** could be detected when it was directly dissolved in Tris buffer. Because of its low solubility, the drug was first dissolved in dimethylsulfoxide, then in Tris buffer. In contrast to drug **1**, no degradation of drug **2** was observed after it had stayed for 24 h in 0.1 M Tris buffer. Both drugs are stable when dissolved in Tris buffer, ethanol or dimethylsulfoxide. Both compounds were fluorescent in ethanol, a solvent in which the solute can be quantitatively measured at 10^{-5} M concentration. The emission maximum occurred at 545 nm for drug **1** and 520 nm for drug **2** when they were excited at 350 nm. These fluorescence yields are too small for determination of binding constants and we resort below to computer simulations. Divalent cations such as Ca^{++} or Mg^{++} at 10 mM concentration have no detectable effect. In contrast, Zn^{++} , Mn^{++} and Fe^{++} completely quench the fluorescence and modify the absorption spectrum in a way that is comparable to the effect of aging.

CPCM calculations on hydroxy benzoic acids

In our styrylquinoline derivatives, the carboxy and hydroxy substituents lie at adjacent positions R7 and R8, exactly as in salicylic acid. It has long been known that in salicylic acid, the neighborhood of hydroxyl and carboxyl substituents favors the formation of a chelate ring, in which a divalent hydrogen atom is a link (Branch & Yabroff, 1934). Hence, both pKs of salicylic acid significantly differ from those of *meta* and *para* hydroxy benzoic acids (Dunn & McDonald, 1969; Lange, 1967). It is generally accepted that the enol form of salicylic acid prevails both in ground and excited states (Lahmani &

Zehnacker-Rentien, 1997; Sobolewski & Domcke, 1998). It was interesting to estimate the validity of this model in our styrylquinoline derivatives (see Fig. 1) and compute the energy and the free energies both *in vacuo* and in water in a case where the *pK*s are experimentally known. If the computations were correct, the experimental *pK*s of the two compounds should follow the expression:

$$\Delta pK = \Delta G/RT \log 10$$

where *RT* is the Boltzmann factor, and ΔG is the free energy difference of the protonated and unprotonated forms of a compound. To predict a *pK* with an error of 1, we should have an error in the energy computations less than 1.4 kcal/M. Such an accuracy is still a challenge for the to-day state of art and we would be very satisfied if we could sort out the conformers and tautomers.

If we exclude the case in which both carboxylic oxygens are protonated, the *ortho* and *meta* hydroxy benzoic acids have 15 forms: 8 neutral forms, 6 with a charge of -1 and one with a charge of -2 (cf. Fig. 3). Because of its symmetry, the *para* compound has

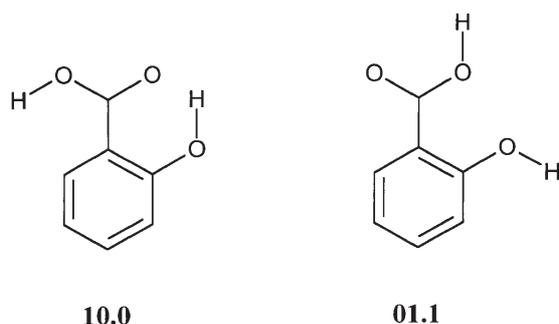


Figure 3. Rotamer coding.

The conformations are coded by two binary numbers separated by a dot. If the substituent is not protonated the number is replaced by a blank. The first number concerns CO_2H . The low bit of the first number is set to 1 if the oxygen that carries the proton is close to the hydroxyl group and is set to 0 otherwise. The high bit is set to 1 if the proton is close to the cycle. The second number concerns the hydroxyl group. A keto form is noted by a blank. If the proton points to the carboxyl group the bit is set to 0.

only 8 forms. Table 1 shows the results. From the *in vacuo* Hartree-Fock results, the *ortho* keto 00. form is predicted to be more favorable than the enol form .0 which is not realized experimentally in aqueous solution. When electron correlation is more accurately treated with B3LYP/6-311G(2d,p), the keto form is still more favorable than the enol form by 1.3 kcal/M (result not shown in Table 1). When solvent effects are included with CPCM, the enol form .0 gets more favorable than the keto form 00. (see Table 1). Moreover, at this level, the *meta* compound is more acidic than the *ortho* compound, in contrast to experimental data. We conclude that CPCM does not give good results when it is used with B3LYP. When CPCM is used at the HF/6-31+G(d) level, the ΔG correlates poorly with the experimental *pK*s, but the ordering is correct and this is a significant improvement. Finally, we note that the $\Delta G = 264.9$ kcal/M for water protonation is probably underevaluated. It should be greater than 277.2 kcal/M if all hydroxybenzoic acids have a charge of -1 when dissolved in water and less than half of 564.9 kcal/M (282.4 kcal/M) to prevent them from having a charge of -2.

Calculations on the styrylquinoline derivatives

Tables 2 and 3 show the results concerning drugs **1** and **2**, respectively. The keto forms are favored *in vacuo*, but the enol forms prevail in the CPCM computations.

The docking results (Tables 2 and 3) may be summarized as follows :

- ◆ both drugs bind closely to the Mg^{++} divalent cation that is embedded in the crystallographic structure in the immediate vicinity of the D,D-35E motif. This remains valid whatever the tautomer or rotamer form. The atoms of the drug that are close to the cation are indicated in the Tables. In most cases, the drugs bind by the quinoline moiety. The keto form is favored.

Table 1. Energies of various tautomers and rotamers of *ortho*, *meta* and *para* hydroxy benzoic acids.

The rotamer conformations are encoded as described in Fig. 2. *In vacuo* ΔE s, are the optimised *ab initio* RHF /6-31G(d) energies. *In vacuo* ΔG s contain the thermal correction to the Gibbs energy multiplied by the *ad hoc* factor 0.9. The CPCM ΔG s contain the latter thermal correction and the optimised CPCM total free energy (with non electrostatic terms) computed at either HF/6-31+G(d) or B3LYP/6-311G(2d,p) level in a medium of dielectric constant equal to 78.39. The last line reports the deprotonation energies of H_3O^+ *in vacuo* and in water.

Conformation	CO ₂ dihedral (degrees)	<i>In vacuo</i>		In water (CPCM)		Experimental pKa
		ΔE (kcal/M) HF 6-31G(d)	ΔG (kcal/M) HF 6-31G(d)	ΔG (kcal/M) HF 6-31+G(d)	ΔG (kcal/M) B3LYP 6-311G(2d,p)	
<i>Ortho</i>						
00.0	0	0.0	0.0	0.0	0.0	
01.0	0	3.6	3.2			
10.0	16	8.7	8.3			
01.1	0	10.2	9.1			
00.1	0	11.0	9.6			
11.1	0	10.4	9.6			
10.1	46	18.5	17.1			
11.0	O ^a	27.2	25.8			
00.	0	348.2	340.1	282.8	288.1	
.0	0	348.9	340.8	275.7	286.4	2.98
10.	0	365.1	356.2			
01.	0	367.5	358.3			
.1	0	368.4	360.0			
11.	1	375.3	366.1			
.	42	823.7	806.1	570.8	589.7	12.38
<i>Meta</i>						
01.0	0	0.00	0.0	0.0		
.0	0	356.3	348.3	275.9	284.5	4.08
.1	0	358.9	350.6			
00.	0	363.0	353.9	286.4		
01.	0	364.8	356.9	286.5	292.3	
11.	27	363.0	355.1			
10.	32	363.0	355.9			
.	0	800.2	784.1	565.1	584.7	9.85
<i>Para</i>						
00.1	0	0.0	0.0	0.0		
00.	0	352.7	344.8	283.4	289.2	
10.	10	357.9	349.7			
.0	0	360.0	351.8	277.2	286.4	4.88
.	0	796.7	780.8	564.9	582.4	9.23
Water						
H ₃ O ⁺		174.8	167.9	264.9		

^aThis planar conformation is unstable. If it is distorted along one of its two imaginary normal modes it spontaneously goes to one of the other conformations.

Table 2. Energy and geometrical characteristics of various drug 1 tautomers, *in vacuo*, in water, and bound to the ASV integrase catalytic core.

Rotamer coding for R7 and R8 is the same as in Fig. 2. If R3' points towards R4' the bit is set to 0. If R4' points towards R3' the bit is set to 0.

Substituents					Drug alone		Drug bound to ASV IN catalytic core			
R3'	R4'	R7	R8	N1	$E(\textit{in vacuo})$	$E(\textit{CPCM})$	Energy	α	Orien-	Drug atoms
OCH ₃	OH	CO ₂ H	OH	H	kcal/M	kcal/M	kcal/M	deg	tation	near Mg ⁺⁺
1	1	11			1.0	3.1	-88.7	-1	<i>v</i>	8
1	1	11			0.0	3.4	-106.7	-161	<i>h</i>	8
1	1		0		5.5	0.0	-94.1	5	<i>v</i>	8
1			0	H	9.2	15.9	-88.4	38	<i>v</i>	7
0	1	11			3.1	6.6	-99.9	-5	<i>h</i>	7, 8
0	1		0		7.5	3.5	-99.3	-6	<i>v</i>	8
0			0	H	10.0	20.3	-91.9	3	<i>v</i>	8

◆ The drugs seem to bind either in a “horizontal” slit as in Fig. 4 or “vertically” as in Fig. 5. This is shown by the characters *h* or *v* in Tables 2 and 3. When it binds horizontally, the drug separates

the drug touches the following residues: D64, Q153, E157, N160, W176, R195 and V196.

◆ The drug that was planar in the rigid docking phase becomes slightly dis-

Table 3. Energy and geometrical characteristics of various drug 2 tautomers, *in vacuo*, in water, and bound to the ASV integrase catalytic core.

Rotamer coding for R7 and R8 is the same as in Fig. 2. If R3' points towards R4' the bit is set to 0. If R4' points towards R3' the bit is set to 0.

Substituents					Drug alone		Drug bound to ASV IN catalytic core			
R3'	R4'	R7	R8	N1	$E(\textit{in vacuo})$	$E(\textit{CPCM})$	Energy	α	Orien-	Drug atoms
OH	OCH ₃	CO ₂ H	OH	H	kcal/M	kcal/M	kcal/M	deg	tation	near Mg ⁺⁺
1	1	11			4.6	4.4	-93.9	-20	<i>v</i>	8
1	1	11			3.5	4.9	-84.9	-155	<i>h</i>	8, N
1	1		0		9.1	1.3	-87.0	20	<i>v</i>	1'
0	1	11			0.0	3.2	-83.4	3	<i>v</i>	8, N
0	1		0		4.5	0.0	-85.8	-3	<i>v</i>	8, N
	1		0	H	34.1	25.3	-93.2	7	<i>v</i>	7

the two aspartic acids from the glutamic residue on the catalytic core of the D,D-35E motif and touches the following residues: D64, N122, N149, Q153, N160. In the vertical binding site

torted after minimization. The distortion is always small so that only the final values of the angle α are indicated because they are always very close to their starting values (0° or 180°).

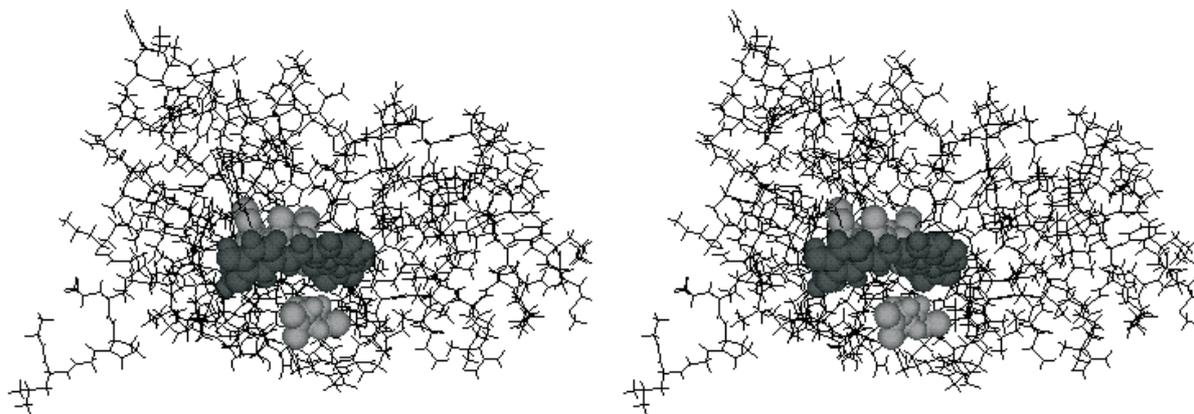


Figure 4. Stereoview of drug 1 in its most favorable binding to RSV integrase catalytic core.

The catalytic core is shown as a wire model. The three carboxylic acids D64, D121 and E157 are in light gray. The drug is darker and hides the Mg^{++} divalent cation.

DISCUSSION – CONCLUSIONS

Integrase inhibitors of possible medical use should have IC_{50} in a less than 10^{-9} M range. The best drugs in the styrylquinoline series have IC_{50} slightly smaller than 10^{-6} M and the IC_{50} of the inhibitors we have studied here is 2×10^{-6} M. It is also necessary to get more insight into the conformation of the drug in

solution and bound to what can be imagined as a target. This is rather simple in the case of many compounds. In the styrylquinoline series it is far from trivial because of the many possible tautomeric forms. The study of the tautomers *in vacuo* is only the first approach which should to be supported by computations such as CPCM where solvent effects are taken into account. Even if the method is not

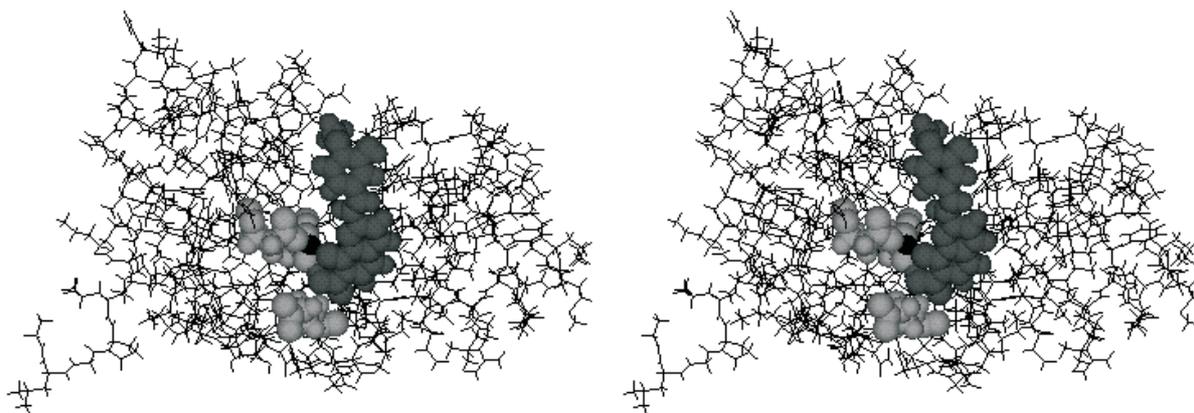


Figure 5. Stereoview of drug 2 in its most favorable binding to RSV integrase catalytic core.

Here the divalent cation (black) can be seen.

reliable enough for pK_s to be predicted, the results reported here show that the enol form is favored in aqueous solution.

Docking computations have been performed to get insights into the binding. The computed interaction energies are very large. They are not related in a simple way to experimental dissociation constants because entropy and solvent effects have not been fully taken into account in our computations. The obtained values, in any case, do not correspond to experimental dissociation constants around 10^{-6} M, that are expected if the IC_{50} are straightly translated into binding constants. Although the absolute values are probably incorrect, the calculated energies may be interpreted as a score. Ouali *et al.* (2000) report scores in the -117 to -113 kcal/M range for styrylquinoline derivatives that have an IC_{50} less than 10^{-6} M and in the -80 to -50 kcal/M range for drugs that have an IC_{50} greater than 10^{-4} M. The presented interaction energies are perfectly compatible with drugs that have an IC_{50} of about 2×10^{-6} M. Therefore docking of molecules seems to be, in this case, a valid tool. The remarkable fitness of *ex nihilo* calculation to experimental *in vitro* data will allow us to make predictions and eventually lead to design of more active drugs.

These results show that the drug inhibits the integrase by complexing the protein close to the active site. This binary complex is of course much simpler to model than a ternary complex in which DNA is somehow involved. Consequently, drug design should not be as difficult as could have been imagined. While this manuscript was being revised, another inhibitor 5CITEP, was found by X-ray crystallography to be bound centrally in the active site of the HIV-1 integrase catalytic core (Goldgur *et al.*, 1999). As this compound has two cycles connected by a linker that is similar to ours, the localization found here *in silico* for styrylquinolines is probably right.

We found that the keto tautomer showed a better interaction energy with RSV integrase than the end tautomer. Unfortunately, it is

not abundant in solution. In the near future we should be able to design drugs that have the right tautomeric form.

The authors thank Miss Ange-Patricia Tchamedeu-Kameni for technical assistance.

REFERENCES

- Aqvist, J. (1990) Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* **94**, 8021–8024.
- Bardez, E., Devol, I., Larrey, B. & Valeur, B. (1997) Excited-state processes in 8-hydroxyquinoline: photoinduced tautomerization and solvation effects. *J. Phys. Chem. B* **101**, 7786–7793.
- Barone, V., Cossi, M. & Tomasi, J. (1997) A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* **107**, 3210–3221.
- Barone, V. & Cossi, M. (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* **102**, 1995–2001.
- Beese, L.S. & Steitz, T.A. (1991) Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: A two metal ion mechanism. *EMBO J.* **10**, 25–33.
- Besler, B.H., Merz, K.M. & Kollman, P.A. (1990) Atomic charges derived from semiempirical methods. *J. Comp. Chem.* **11**, 431–439.
- Branch, G.E.K. & Yabroff, D.L. (1934) The anomalous strength of salicylic acid. *J. Am. Chem. Soc.* **56**, 2568–2570.
- Brown, P.O. (1990) Integration of retroviral DNA. *Curr. Top. Microbiol. Immunol.* **157**, 19–48.
- Bujacz, G., Jaskólski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R.A. & Skalka, A.M. (1995) High-resolution structure of the catalytic domain of avian sarcoma virus integrase. *J. Mol. Biol.* **253**, 333–346.
- Bushman, F.D. & Craigie, R. (1991) Activities of human immunodeficiency virus (HIV) integra-

- tion protein *in vitro*: Specific cleavage and integration of HIV DNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1339–1343.
- Chow, S.A., Vincent, K.A., Ellison, V. & Brown, P.O. (1992) Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* **255**, 723–726.
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gougl, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. & Kollman, P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
- Cowan, J.A. (1998) Metal activation of enzymes in nucleic acid biochemistry. *Chem. Rev.* **98**, 1067–1087.
- De Clercq, E. (1995) Toward improved anti-HIV chemotherapy: Therapeutic strategies for intervention with HIV infections. *J. Med. Chem.* **38**, 2491–2517.
- Dunn, G.E. & McDonald, R.S. (1969) Infrared spectra of aqueous sodium benzoates and salicylates in the carboxyl-stretching region: Chelation in aqueous sodium salicylates. *Can. J. Chem.* **47**, 4577–4588.
- Dyda, F., Hickman, A.B., Jenkins, T., Engelman, A., Craigie, R. & Davies, D. (1994) Crystal structure of the catalytic domain of HIV-1 integrase: Similarity to other polynucleotidyl transferases. *Science* **266**, 1981–1986.
- Engelman, A., Mizuuchi, K. & Craigie, R. (1991) HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell* **67**, 1211–1221.
- Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G., Montgomery, J.A., Stratmann, R.E., Burant, J.C., Dapprich, S., Millam, J.M., Daniels, A.D., Kudin, K.N., Strain, M.C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G.A., Ayala, P.Y., Cui, Q., Morokuma, K., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Cioslowski, J., Ortiz, J.V., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P.M. W., Johnson, B.G., Chen, W., Wong, M.W., Andres, J.L., Gonzalez, C., Head-Gordon, M., Replogle, E.S. & Pople, J.A. (1998) Gaussian 98, Revision A.6, Gaussian, Inc., Pittsburgh PA.
- Furtado, M., Callaway, D., Phair, J., Kunstman, K., Stanton, J., Macken, C., Perelson, A. & Wolinsky, S. (1999) Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy. *N. Engl. J. Med.* **340**, 1614–1622.
- Gabb, H., Jackson, R. & Sternberg, M. (1997) Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol.* **272**, 106–120.
- Goldgur, Y., Dyda, F., Hickman, A., Jenkins, T., Craigie, R. & Davies, D. (1998) Three new structures of the core domain of HIV-1 integrase: An active site that binds magnesium. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9150–9154.
- Goldgur, Y., Craigie, R., Cohen, G., Fujiwara, T., Yoshinaga, T., Fujishita, T., Sugimoto, H., Endo, T., Murai, H. & Davies, D. (1999) Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: A platform for antiviral drug design. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13040–13043.
- Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M. & Markowitz, M. (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123–126.
- Katz, R.A. & Skalka, A.M. (1994) The retroviral enzymes. *Annu. Rev. Biochem.* **63**, 133–173.
- Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A.A., Aflalo, C. & Vakser, I.A. (1992) Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2195–2199.
- Klamt, A. & Schürmann, G. (1993) COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screen-

- ing energy and its gradient. *J. Chem. Soc. Perkin Trans.* **2**, 799–805.
- Lahmani, F. & Zehnacker-Rentien, A. (1997) Effect of substitution on the photoinduced intramolecular proton transfer in salicylic acid. *J. Phys. Chem. A* **101**, 6141–6147.
- LaFemina, R.L., Schneider, C.L., Robbins, H.L., Callahan, P.L., Legrow, K., Roth, E., Schleif, W.A. & Emini, E.A. (1992) Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J. Virol.* **66**, 7414–7419.
- Lange (1967) *Handbook of Chemistry* (Dean, J.A., ed.) 15th edition, Mc Graw-Hill Inc.)
- Lavery, R., Sklenar, H., Zakrzewska, K. & Pullman, B. (1986) The flexibility of the nucleic acids: (II). The calculation of internal energy and applications to mononucleotide repeat DNA. *J. Biomol. Struct. Dyn.* **3**, 989–1014.
- Le Bret, M., Gabarro-Arpa, J., Gilbert, J.C. & Lemaréchal, C. (1991) MORCAD, an object-oriented molecular modelling package running on IBM RS/6000 and SGI 4Dxxx workstations. *J. Chim. Phys.* **88**, 2489–2496.
- Maignan, S., Guilloteau, J.P., Zhou-Liu, Q., Clément-Mella, C. & Mikol, V. (1998) Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: High level of similarity of the active site with other viral integrases. *J. Mol. Biol.* **282**, 359–368.
- Mekouar, K., Mouscadet, J.F., Desmaële, D., Subra, F., Leh, H., Savouré, D., Auclair, C. & d'Angelo, J. (1998) Styrylquinoline derivatives: A new class of potent HIV-1 integrase inhibitors that block HIV-1 replication in CEM cells. *J. Med. Chem.* **41**, 2846–2857.
- Ouali, M., Laboulais, C., Leh, H., Gill, D., Desmaële, D., Mekouar, K., Zouhiri, F., d'Angelo, J., Auclair, C., Mouscadet, J.F. & Le Bret, M. (2000) Modeling of the inhibition of retroviral integrases by styrylquinoline derivatives. *J. Med. Chem.* (in press).
- Pommier, Y. & Neamati, N. (1999) Inhibitors of human immunodeficiency virus integrase. *Adv. Virus Res.* **52**, 427–458.
- Rega, N., Cossi, M. & Barone, V. (1999) Improving performance of polarizable continuum model for study of large molecules in solution. *J. Comput. Chem.* **20**, 1186–1198.
- Rogers, N. & Sternberg, M. (1984) Electrostatic interactions in globular proteins. Different dielectric models applied to the packing of alpha-helices. *J. Mol. Biol.* **174**, 527–542.
- Shoichet, B.K. (1996) Docking ligands to proteins; in *Protein Structure Prediction* (Sternberg, M.J., ed.) *The Practical Approach Series* (Rickwood, D. & Hames, B., eds.) chapter 11, pp. 263–290.
- Sobolewski, A. & Domcke, W. (1998) *Ab initio* study of excited state intramolecular proton dislocation in salicylic acid. *Chem. Physics* **232**, 257–265.
- Waldeck, D.H. (1991) Photoisomerisation dynamics of stilbene. *Chem. Rev.* **91**, 415–436.
- Zhang, L., Ramratnam, B., Tenner-Racz, K., He, Y., Vesanen, M., Lewin, S., Talal, A., Racz, P., Perelson, A., Korber, B., Markowitz, M. & Ho, D. (1999) Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. *N. Engl. J. Med.* **340**, 1605–1613.
- Zouhiri, F., Mouscadet, J.F., Mekouar, Kh., Desmaele, D., Savouré, D., Leh, H., Subra, F., Le Bret, M., Auclair, C. & d'Angelo, J. (2000) Structure-activity relationships and binding mode of styrylquinolines as potent inhibitors of HIV-1-integrase and replication of HIV-1 in cell culture. *J. Med. Chem.* (in press).