Eur J Med Chem (1996) 31, 775–786

A model allowing the design of modified nucleosides as HIV-RT inhibitors

G Pèpe[1], M Meyer[1], P Faury[2,3], JC Graciet[2,3], JC Chermann[2], JL Kraus[2,3]

[1]*CRMC2-CNRS – Laboratoire associé aux Universités d'Aix-Marseille II et III, Campus universitaire de Luminy, Case 913, 13288 Marseille Cedex 9;*
[2]*INSERM unité 322, Unité des rétrovirus et maladies associées, Campus universitaire de Luminy, BP 33, 13273 Marseille Cedex 9;*
[3]*Laboratoire de chimie biomoléculaire, Faculté des sciences de Luminy, Case 901, 13288 Marseille Cedex 9, France*

(Received 22 January 1996; accepted 18 March 1996)

Summary — A chemical, structural, molecular electrostatic potential (MEP) analysis of modified nucleosides allows the understanding of how nucleosides interact with different receptors. The interaction with kinases is sensitive to base modifications, while the interaction with the reverse transcriptase receptor HIV active site is more affected by ribose modifications. The model herein indicates a geometrical lower limit in the width of the modified sugar that corresponds to the 3' *erythro* position. This characteristic allows one to predict a potential activity of the 3' substituted compounds. The 4'-hydroxymethyl group position with respect to the nucleic base is also important for antiviral activity. The model gives the geometric parameters of this position (related to kinetic effects) that corresponds to an increase in the activation energy required to fit the active site of the kinases and the RT. Our model is compatible with the 3D structure of the HIV RT active site. It allows the design of potent new active compounds where the sugar can be substituted by any group answering to the defined parameters.

HIV RT / inhibitor / nucleoside

Introduction

To achieve an optimal therapy of HIV infection, it is likely that one or several steps in the viral replication cycle will need to be inhibited. Therefore, combinations of antiviral agents, acting at different stages of the virus replication cycle, may be used to achieve an additive or synergistic effect.

Antiviral agents [1–9] include: inhibitors of viral binding and cell entry, inhibitors of reverse transcription, antiproteases, and agents which act at other sites of viral replication. However, the antiretroviral agents developed to date have been primarily the inhibitors of HIV reverse transcriptase (HIV RT). The HIV RT has been one of the most intensively studied viral targets in the development of anti-HIV drugs. A broad family of 2',3'-dideoxynucleoside analogues have been identified to have a potent antiretroviral activity. These analogues can be metabolized to form a potent RT chain terminator [10].

The dideoxynucleosides are of special interest because a simple chemical modification of their sugar moiety can convert a normal substrate used for nucleic acid synthesis into a potent inhibitor for the replication of HIV and its cytopathic effects [11]. The search for more effective anti-HIV therapies includes the discovery of new drugs with different reaction mechanisms and the improvement of drugs known to be efficacious, particularly nucleosides.

To assist in the identification and development of more effective chemotherapies for AIDS, we have developed computer-assisted structure–activity correlations of dideoxynucleoside analogues. Most of the compounds included in this study are in the public domain and have been tested against HIV cell culture assays targeting reverse transcriptase. Nucleosides are the focus of the present study because the literature on these compounds is extensive enough to permit further detailed analysis.

Subsequent to the synthesis [13, 14] and the discovery [12] of the anti-HIV activity of 3'-azido-3'-deoxythymidine (AZT), thousands of modified nucleosides have been tested. Some results suggest a link between nucleoside modifications and anti-HIV activity [15, 16]. Such nucleoside modifications can be achieved either on the nucleoside base or on the ribose.

Despite the 3D structure determinations of RT [17–19], no effective model asserting anti-HIV activity of modified nucleosides has been defined so far. Using geometry and molecular electrostatic potential (MEP) systematic analysis, we propose a simple efficient model capable of designing new modified nucleosides with potent anti-HIV activity. This model is consistent with the 3D structure of the HIV RT active site.

Criteria defining the model

It is well known that the modified nucleosides must be phosphorylated by various kinases before acting on the RT active site as chain terminators [20, 21]. The kinases are specific to the nucleosides, which means that their active sites are probably sensitive to nucleic base structure, rather than to sugar modifications. Base modifications can facilitate or hinder the enzymatic phosphorylation step [15, 16]. The RT active site is sensitive to ribose structural changes. The interaction of the triphosphate moiety with the RT allows the sugar to reach the active site. This can be deduced from the 3D RT structure. The base is then oriented by association with its complementary base in the RNA virus strand. In order to terminate a chain, the modified nucleoside must fit into the RT active site corresponding to the ribose ring and the base must link with its complementary base to be integrated in the DNA strand and to block efficiently the next polymerization step. Thus, base modifications must not affect the molecular part of nucleoside which participates in base pairing, and other substitutions must not disturb sterically the specific base association. Therefore it appears that the geometry of the sugar part of the molecule is essential for HIV activity. The pathway of the drug before reaching the active site of the HIV RT seems to be very complex, because the molecule has to penetrate through the cellular membrane. In this respect the hydrophilicity can be improved by adding to the molecule a lipophilic group which can be released inside the cell. The drug can also be phosphorylated in order to pass over the phosphorylation step mediated by kinases. However during the last step if the molecule does not fit the HIV RT active site it will be ineffective. For these reasons we defined a model characterizing compounds whose structure corresponds to the best fit with the HIV RT active site. In this last step if the molecule does not respond to specific parameters it will be inactive. This statement does not mean that a compound answering to these criteria will present an activity.

The metabolites of modified nucleosides differ from those of natural bases and as such can be toxic. We suggest the use preferentially of unmodified nucleic bases. Since cytosine nucleosides (Lamivudine (3TC), Zalcitabine (ddC)) belong to the most potent drugs, cytosine linked to modified ribose was selected to define the model.

Specificity of the sugar RT receptor

The HIV sugar RT receptor appears to act very specifically on nucleoside structure. For example, a drug like Ganciclovir [22] presents anti-hepatitis B virus (HBV) activity but no anti-HIV activity. This observation could imply that other parts of this molecule should interact with other sites of kinases and of the HBV RT.

Our model takes into account this specificity implicitly. It should help to design drugs that will act only on the HIV RT receptor, avoiding side effects related to the drug action on other polymerases present in the human cells. All the parameters we will define, either geometric or electrostatic MEP, depend on the sugar ring conformation, related to the substituent nature. The same sugar conformation (C3' *exo*) is observed from X-ray analysis of ddC [23] and AZT [24]. As the base, the 3' substituent and the crystal packing of these compounds are different, van Roey and coworkers [25] assumed that the C3'-*exo* conformation was the only one active on the kinase and RT receptors. This last model, the Fisher's model [26] and others studies [27, 28] carried out on the nucleosides to find relationships between conformations and anti HIV activity, particularly at the sugar-ring level, require the knowledge of the active conformation in the cellular medium. This hypothesis expresses thermodynamically that the standard geometry of the natural substrate existing in the cellular medium corresponds to the minimum of the transition state energy, and any geometric modification of the drugs to put them in the geometry of the natural nucleoside will increase the activation energy [29]. The rate of the reaction corresponding to the integration of the molecule in the built DNA, will be lower than that of the standard substrate. To fit into the kinase or RT active sites, the drug must have an envelope (related to the conformation) equivalent to that of natural substrate in terms of both geometry and electrostatic potential. This feature is required in order to be complementary to the active site envelope [30]. Complementarity corresponds to a minimum in the coulombic contribution to the pertubation energy. The perturbation energy at the atomic level is analogous to the transition state energy at the macroscopic level.

The model

From previous considerations we can define a simple model eliminating non-potent active modified nucleosides. The molecules with modified or unmodified ribose were considered in the C3'-*exo* conformation as already mentioned by other authors [25]. If another conformation is selected as a reference molecule (based on the principle that only one molecular conformation of the natural substrate is active) corresponding to minimal transition state energy, other molecules will be strained to orient the 3'-CH_2OH group in the appropriate position with respect to the base. In such a case, different values for the parame-

ters will be obtained but the predicted activity order will remain the same. This assumption was verified through the results obtained with the sugar ring in the C4'-*exo* conformation.

All the calculations relative to this article were performed using GenMol software [31]; this was an original force field capable of treating atomic and molecular systems containing up to 10^5 atoms of 96 different types. Detailed information about the program has been described elsewhere [32]; in order to test the model on other molecules, the program, or any information about it, can be obtained from G Pèpe. The reliability of this software was measured through the root mean square value corresponding to the best molecular fit between the calculated and observed geometries of the two previous molecules: 0.03 Å for ddC and 0.01 Å for AZT (the molecule which corresponds to the C3'-*exo* conformation). As some clinical active anti-HIV drugs are compounds derived from cytidine. We selected the unmodified cytidine molecule ribose in C3'-*exo* conformation as the reference compound. All the molecules, optimized with the five-membered ring in C3'-*exo* conformation, were compared to cytidine so that the molecules can present the best envelope in order to fit into the different enzymatic active sites.

At the geometric level we then compared modified nucleosides to cytidine and calculated the energy necessary to orientate the molecules into cytidine-like geometry. The difference between the strain energy of the molecule in the C3'-*exo* conformation (with the imposed characteristic distances of the cytidine), and the minimum-strain energy after relaxation, corresponds to the increase of the transition-state energy with cytidine. The effect of this increase on the kinetic constants was calculated. This energetical increase implicitly means that the entropy contribution of the activation energy is equivalent for all the studied compounds.

This approximation can be justified by the fact that the parts of the molecule (nucleic base and triphosphate group) corresponding to the strongest interactions with the solvent and the protein are conserved.

However it is not possible to establish a classification in the affinity of these compounds for kinase and RT receptors. The increase of the kinetic constant values corresponding to the geometric fit are reported, only to indicate that small activation energy variations can have huge effects on the chemical reactivity.

Results

Hundreds of molecules were generated and analysed but we only report the results on a representative sample displayed in figure 1. In the compounds **1–10** the ribose ring structure is conserved. Compounds **11–19** include a heterocyclic five-membered ring. In the molecules **20–22**, the ring is replaced by a carbon chain: an allene in the compound **20**, a *cis*-butene in the compound **21**, and an imine for compound **22**, while in compounds **23** and **24** the sugar ring is replaced by a six-membered ring. We note that in compound **23** the base is not directly bonded to the ring.

Molecules with a 3'-erythro-substituted ribose

Since the most active compounds correspond to 3'-*erythro*-substituted molecules, we chose to analyse the corresponding effects on geometry and MEP parameters (compounds **1–10** in table I).

MEP effect

Since the MEP depends on the net atomic charge values, in order to validate those obtained from GenMol, we include in figure 2a the values given by AM1 [34] which are very similar. To show that the inductive effect of the base is negligible on the C3' neighbourhood and to support the MEP calculated in this region on modified sugar without the base, we report the net atomic charges on the modified sugar alone (fig 2b).

Figures 3a–d display respectively the MEP diagrams of compounds **1–4**. The corresponding projections are given above the MEP diagram (the MEP representation is calculated in a monopole approximation and net atomic charge evaluations are defined in previous papers [35, 36]).

A reference system placed on the atoms C4, O5 and C6 allows the orientation of the molecule (C4-O5 gives the *x*-axis, while C6 defines the *x,y* plane), which separates the regions of the MEP corresponding to the significantly active atoms. The molecule is cut in 1.0 Å equidistant parallel slices and the MEP is computed on the molecular edge corresponding to the van der Waals' solvent-accessible surface [37]. Once the centroid of the atoms belonging to the slice is determined, a vector centred on this point scans the slice from 0–360° counterclockwise, zero being given by *x*-axis direction. The MEP values are reported on the diagram. The numbers on the right-hand side of the diagram give the slice number and correspond to the zero line for the potential; the left-hand side of the diagram displays a scale in volts for the potential.

The region of the MEP corresponding to the 3'-substituent is located between 230° and 270° (fig 3a–d). The MEP minimum is displaced from the ribose to the 3'-azido substituted ring (270° (fig 3a), 250° (fig 3b), 230° (fig 3c)) and disappears in the 3'-deoxyribose (fig 3d). This result indicates a weak MEP selectivity for the corresponding part of the RT active site. This implies that a steric selectivity is more probable, as

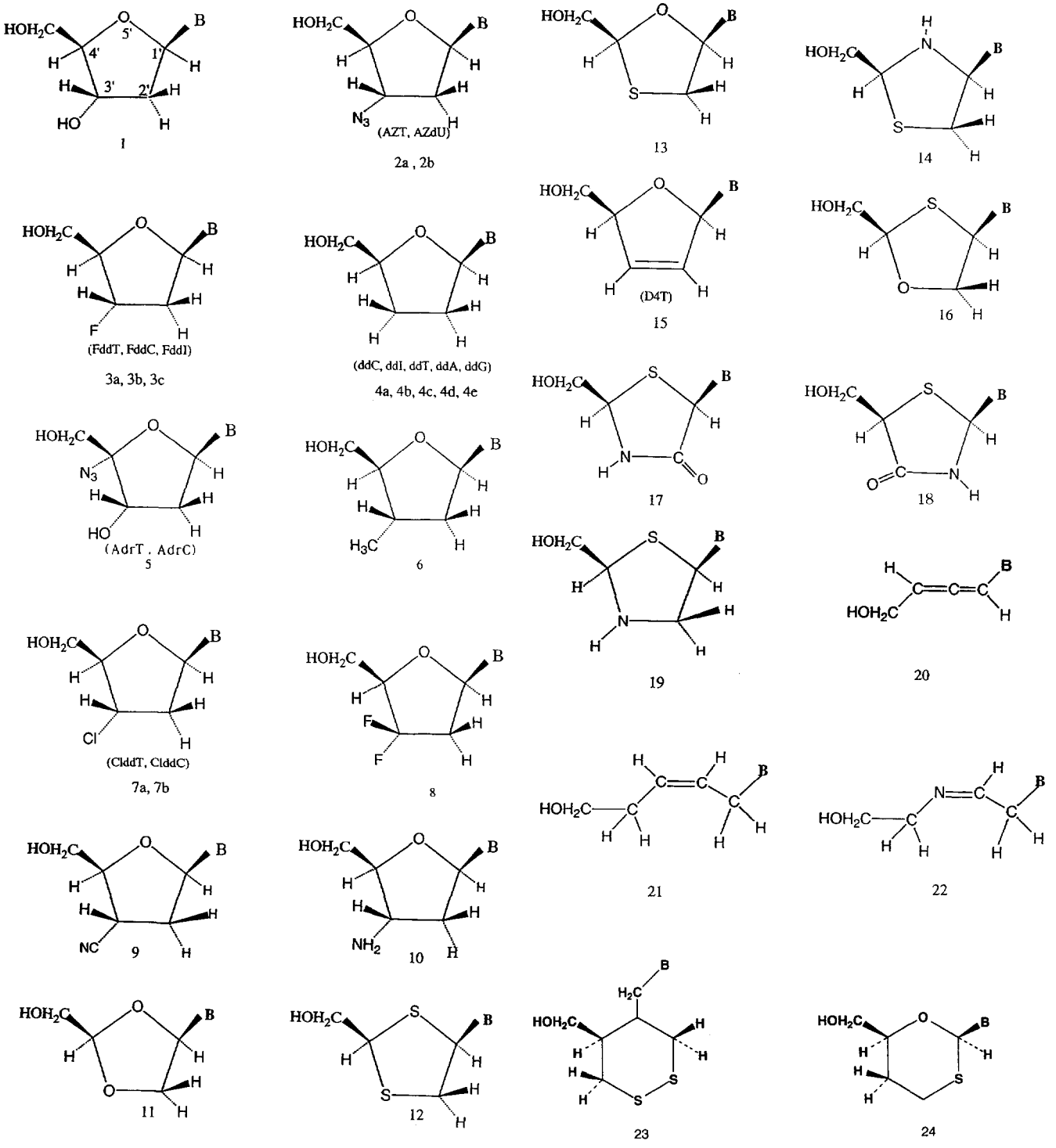

Fig 1. Sample of the modified riboses studied: **1–10**, 2'- and 3'-substituted ribose; **11–19,** ribose ring modified; **20–22**, ribose replaced by a chain; **23** and **24**, ribose replaced by a six-membered ring.

Table I. Principal parameters refering to the model and corresponding to the studied molecules.

Compound	*Width (Å)*	$\Sigma \lvert \Delta di \rvert$ *(Å)*	ΔE *kcal/mol*	*1/k*	R_{AB} *(%) on MEP*	*Receptor affinity predicted*	*In vitro anti-HIV activity*
1	3.6	0	0	1.00	100	+	–***
2	2.3	0.7	0.3	0.66	54**	+	+
3	3.4	0.5	0.7	0.50	67	+	+
4	1.9	0.2	0.2	0.76	91	+	+
5	3.6	0.4	1.0	0.24	48	+	+
6	3.8*	0.2	0.4	0.57	85	–	–
7	4.1*	0.3	0.3	0.66	90	–	–
8	3.4	0.4	0.2	0.76	69	+	–
9	4.7*	0.4	0.2	0.76	77	–	–***
10	3.6	0.1	0.4	0.57	97	+	–***
11	< 2.3	0.2	0.8	0.32	83	+	+
12	–	1.4	2.3*	0.04	86	–	?
13	–	0.3	0.8	0.32	89	+	+
14	–	0.7	0.9	0.28	78	+	?
15	–	0.3	0.3	0.66	89	+	+
16	–	0.8	1.0	0.20	87	+	+
17	–	2.8	8.0*	10^{-5}	75	–	?
18	–	2.6	8.0*	10^{-5}	65	–	?
19	–	0.9	1.4*	0.14	83	–	?
20	–	0.9	0.2	0.76	57	+	+
21	–	0.6	1.0	0.24	63	+	?
22	–	0.5	0.7	0.50	47	+	?
23	–	0.2	0.1	0.87	86	+	?
24	–	0.4	0.3	0.66	63	+	?

Width = steric limit of the 3'-substituent under the sugar plane. $\Sigma |\Delta di|$ = sum of distance differences between the standard compound and the other molecules, characterizing the position of 4'-CH_2OH group with respect to the base. ΔE = strain energy necessary to put the 4'-CH_2OH in the standard position with respect to the base. $1/k$ = ratio of the fit kinetic constant of the compound to the standard substrate constant taken as 1. R_{AB}% = measure of the MEP analogy between the compound and the standard substrate. * Width greater than 3.7 Å. **Value obtained with the cytidine base to be homogeneous with the other values. *** Molecules responding to the defined parameters. They perfectly fit the active site and do not stop the polymerization. The molecules of which the width is greater than 3.7 Å (see fig 6) present no anti-HIV activity. Those with a fit energy greater than 1 kcal/mol are also non-active or predicted to be so. Small values of the fit energy have important kinetic effects. (+) Positive receptor affinity means that the calculated parameters are in agreement with parameters defined to allow a good fit with the enzyme active site. (–) Negative receptor affinity.

attested by the R_{AB} values [38, 39] corresponding to MEP comparison [40] between the reference compound and the compounds studied. There is no relation between the R_{AB} value and the compound activity. However, as compounds **2,3** bearing 3' electron-withdrawing substituents F or N_3 are more active than the deoxyribose derivatives, the MEP is probably slightly positive in the corresponding region of the RT receptor. This agrees with the fact that the ribose displays a negative MEP in this region. Most of the substituted ribose compounds have been tested through our model. Due to the size limit for a substituent in that position, the possibility of finding new active compounds belonging to this family is very weak.

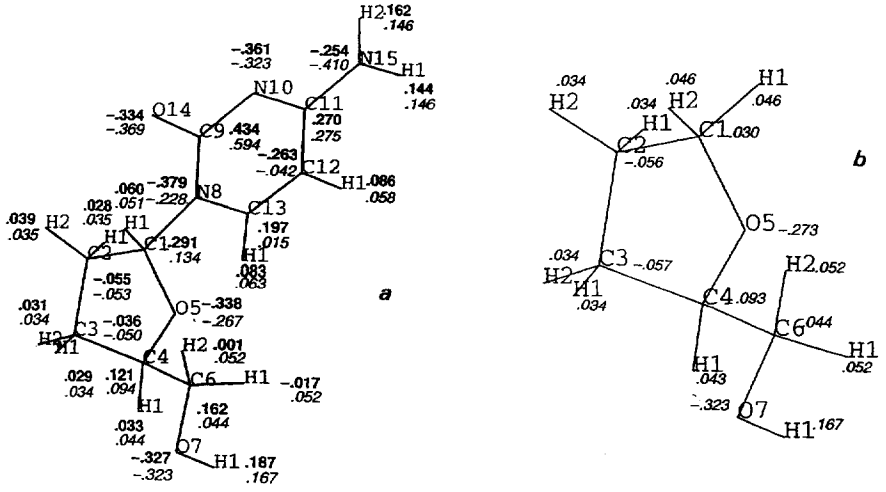

Fig 2. a. Similarities between the net atomic charges calculated by AM1 (bold) and GenMol (italic) on ddC accrediting the values given by GenMol. **b.** Net atomic charges on the sugar part of ddC indicating the weak inductive effect of the base on the C3' neighbourhood which validates the MEP calculated in this region on modified sugar without the base.

Geometry effect

Position of the hydroxymethyl with respect to the base. Three distances indicated in figure 4 were defined in order to locate the 4'-CH_2OH position with respect to the base. The differences between these distances and those of the corresponding distances in the standard position of cytidine (taken as the reference) were calculated. In all of these compounds the position of the hydroxymethyl group was almost the same $\Sigma \mid \Delta di \mid \leq$ 0.7 Å. The corresponding fit energies are $\leq$ 1 kcal/mol with a maximum (1 kcal/mol) for compound **5**.

Through the MEP calculations we showed that the electronic effects are not the decisive ones. The more important parameters for molecule activity were of geometric order.

The kinetic effect of this fit is also reported in table I (the cytidine kinetic fit constant value is arbitrarily chosen as 1). For the compounds **1–10** the effect of the 3'-substituent is negligible on the 4'-hydroxymethyl position with respect to the base. It cannot be responsible for the inactivity of some compounds belonging to this family.

Width of the substituted ribose. Compounds **2–4** have the most potent anti-HIV activity and should fit into the RT receptor with the greatest affinity like the ribose **1** or the 3'-amino derivative **10**. Ortep [33] projections has shown a limit width on the nucleic base side of the active molecules (fig 5). This probably means that the sugar fits the RT receptor on the opposite side to the one bearing the nucleic base. To have a measure of this steric limit, we calculated, as indicated in figure 6, the distance of the most remote atom of the 3'-substituent to the plane defined by the 1',2',4',5' atoms. We then added its van der Waals' radius. The results are shown in table I. The lower limit observed in the active compounds is about 3.6 Å. This indicates why the other molecules with larger 3'-substituents are predicted to be HIV RT inactive.

Compounds with modified sugar ring atoms

In molecules **11–19**, functions or heteroatoms have been introduced in the five-membered ring backbone. In compounds **9–14, 16** and **19** the 3' position is still *exo* as indicated in figure 4. Important displacements of the 4'-hydroxymethyl group are observed as compared with the standard compound **1** (table I). Compounds **11, 13, 15** and **16**, have been tested and present anti-HIV activity [16, 41]. In the active compounds the sum of the distance differences is generally lower than 1.0 Å, with a corresponding value for the fit energy lower than 1 kcal/mol. The kinetic effect of this fit is also reported in table I. The position of the hydroxymethyl group seems to be essential for the antiviral activity. From this test the model predicts a probable anti-HIV activity only for the compound **14** (3' = S, 5' = NH). The synthesis of molecule **14** is currently in progress in our laboratories.

Ribose replaced by an acyclic group or a six-membered ring

The fact that cytallene [42] analogue **20,** a rigid derivative, presents an anti-HIV activity and that our test

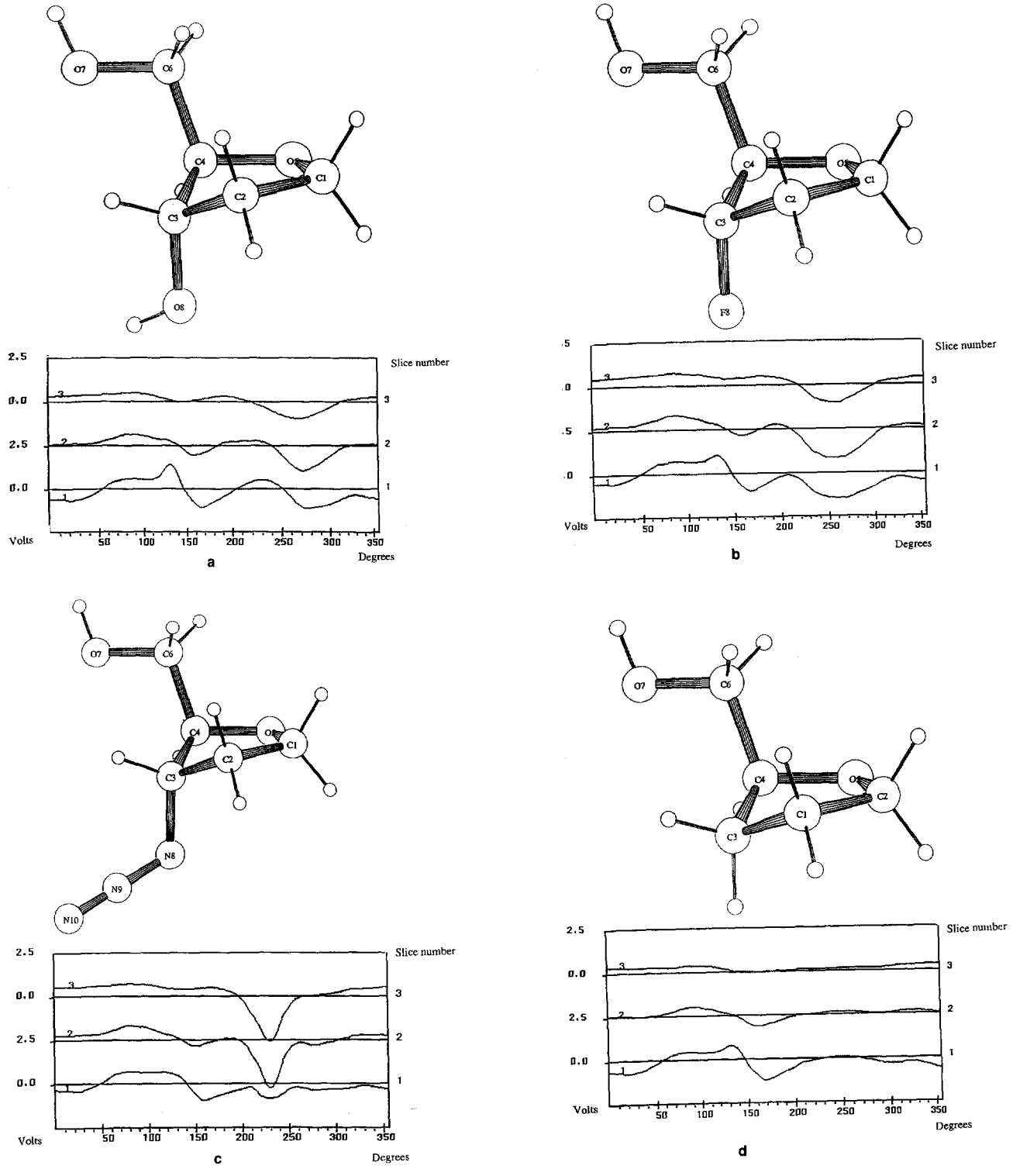

Fig 3. a. MEP diagram of ribose. **b.** MEP Diagram of 3'-*F-erythro*-dideoxyribose. **c.** MEP diagram of 3'-*N*3-*erythro* dideoxyribose. **d.** MEP diagram of dideoxyribose.

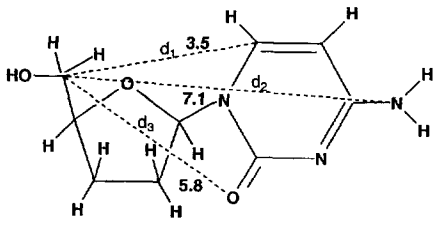

Fig 4. Distances chosen to locate the 4'-hydroxymethyl group with respect to the base (here cytosine).

gives a positive response for this molecule, led us to define molecules derived from *cis*-butene (compound **21**), with potential anti-HIV activity. Other compounds with carbon chains between the nucleic base and the hydroxymethyl group have been included in this computer analysis and are predicted non-RT inhibitors. For the imine derivative **22**, the hydroxymethyl group keeps the same position with respect to the nucleic base ($\Sigma[\Delta di]$ = 0.5 Å) compared to **21**. The negative MEP induced by the nitrogen atom corresponds to the negative MEP of the ribose 3'-hydroxyl group of compound **1** which is a favourable factor for antiviral activity, but this molecule can be predicted to be chemically unstable. In compounds **23** and **24** the sugar ring is replaced by a six-membered ring and in molecule **23** the base is not directly connected to the ribose ring, however for this compound as for the compound **24** the model predicts a positive anti-HIV activity. The calculations performed on equivalent compounds derived from thymidine and compared to thymidine as the reference molecule give the same results. This means that the difference of the observed activity, in the two families of compounds, is probably related to a phosphorylation problem or bioavaibility and seems to indicate a greater sensitivity of the thymine kinases to thymidine modifications.

Compatibility of the model and the active site structure of HIV RT

In the Protein Data Bank [43], partial or complete 3D structures of HIV RT alone and in molecular complexes are available.

Two structures, 1HAR [44] ↔ HIV-1 reverse transcriptase aminoterminal half (fingers and palm subdomains) at 2.2 Å resolution and 1HNI [45] ↔ reverse transcriptase in a complex with the non-nucleoside inhibitor ALPHA-APA R 95845 at 2.8 Å resolution, were used to analyse the active of HIV RT and its neighbourhood. Figure 7 gives the distances between the C_β of D110, D185 and D186 and between these atoms and the G112 C_α. The C_β positions were chosen because they are fixed with respect to the secondary structure; the geometry of the active site is highly conserved in the range limit of the structure resolution which allows us to model the molecular complex between the HIV RT active and the incoming triphosphorylated nucleoside.

Building of the molecular complex between the incoming molecule and the active site

In order to position one calcium cation with respect to the triphosphate group, we started from the X-ray structure of adenosine 5'-triphosphato manganese [46] (we note that in this structure the sugar ring is also in the C3'-*exo* conformation) where the maganese atom was replaced by a calcium cation, and we added a second calcium cation linked to three carboxylate groups borne by three methyl groups in the C_β position of 1HAR (this molecular complex was chosen because the HIV RT molecule is complete). The complex constructed in this manner was energetically optimized by GenMol and connected to a fragment of DNA molecule in A form. Then the three methyl groups were substituted by the C_β corresponding to residues D110, D185 and D186 as displayed in figure 8. Arrows indicate the three hydrogen bonds that must be created between the incoming molecule (triphosphorylated cytidine) and the template (that can be the RNA brought by the virus), the P···O bond with the primer strand, and the bond inside the triphosphate group that must be suppressed. Figure 9 displays the active site of HIV RT in 1HNI and the position of G112 that contacts the 3'-hydroxyl group of the incoming molecule. Figure 10 displays the active site position with respect to the structure of the A moiety of 1HNI. The residue D186 is buried under the residues V60 → V75 that hinder access to the active site, which indicates the softness of this protein region, because this region of the molecule must move to allow the building of the molecular complex around the active site.

Figure 11 shows in a solid representation the position of paired DNA in the active site neighbourhood, and the short contact corresponding to figure 5c.

The model and the 3D structure of the built complex

The 3'-substituent of the incoming molecule is in short contact with the glycine G112; the distance between the G112 C_α and the 3'-oxygen atom is lower than 5 Å. The structure resolution of 2.8 Å does not allow obtention of an accurate value for this contact, however this observation is consistent with the superior width limit ≈ 3.7 Å of the 3'-substituted sugar part in the active compounds.

Glycine is a neutral residue, non-specific to interactions with hydrophobic or hydrophilic atom groups; when the 3'-hydroxyl group is replaced by a hydrogen

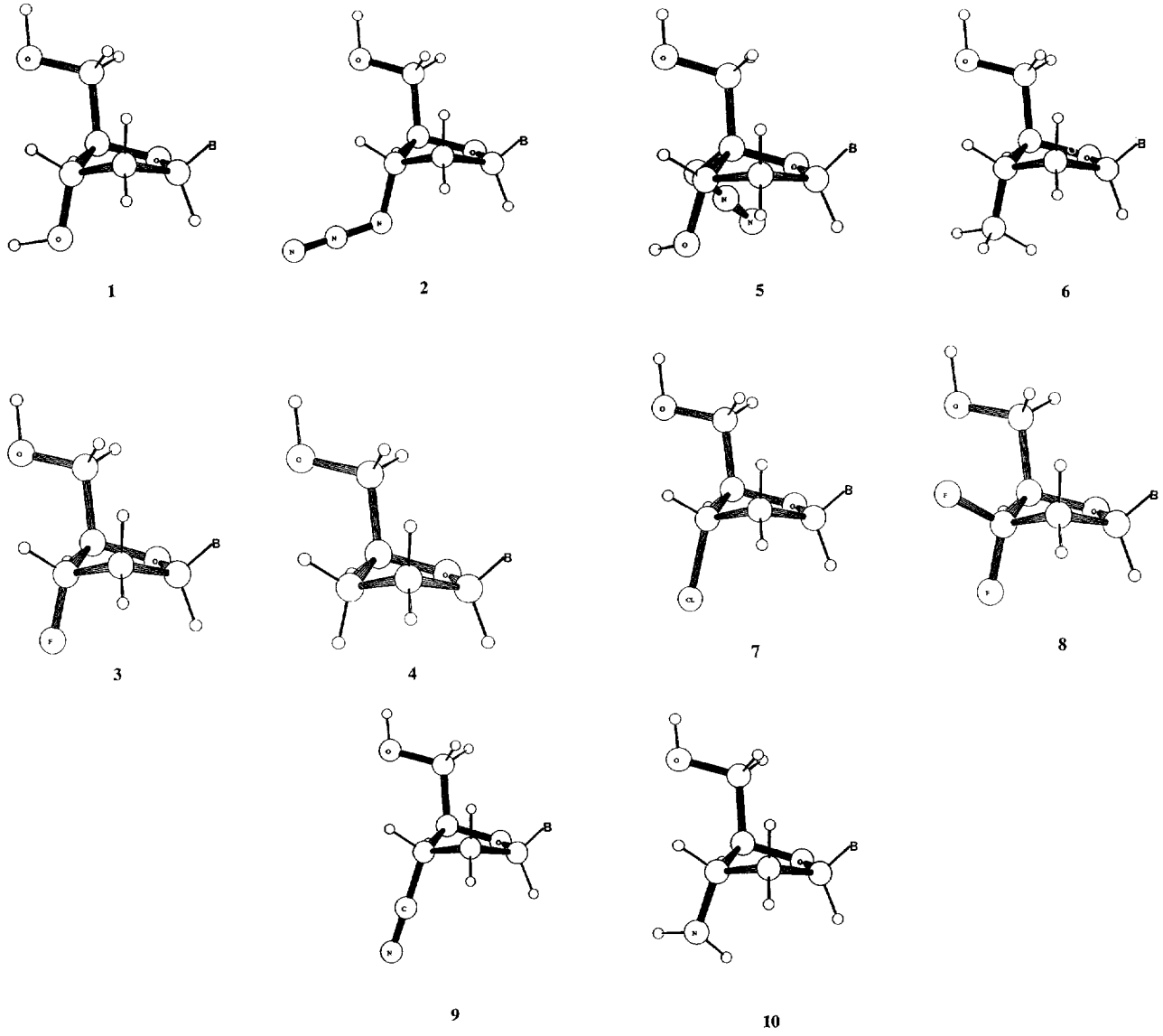

Fig 5. Ortep projections of ribose in the same orientation and in the same conformation C3'-*exo*, displaying the spatial position of the substituents.

atom the compound is still active, which agrees with the MEP analysis.

The template position, like the active site position and the built primer strand position imposes a geometry on the incoming molecule in order to correctly form all the required bonds. In order to block the polymerization of DNA, due to the absence of the 3'-hydroxyl group, the nucleotide must come in the position II of figure 8c, but before reaching this position it must necessarily pass by position I, and consequently responds to the defined parameters.

Conclusion

In order to avoid synthesizing and maybe testing inactive compounds, we define a simple model able to define new potent forms of anti-HIV active nucleoside derivatives. This model highlights the fact that the RT receptor receiving the ribose is very sensitive to the geometry of the 3'-*erythro* group and that a steric lower limit is found on this side of the molecule (van der Waals' limits of the substituent must be < 3.6 Å under the 1',2',4',5' ribose plane),

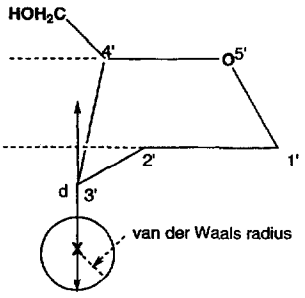

Fig 6. *d* is the van der Waals limit of the most remote atom of the 3'-*erythro* substituent to the plane defined by the atoms (1', 2', 4' and 5').

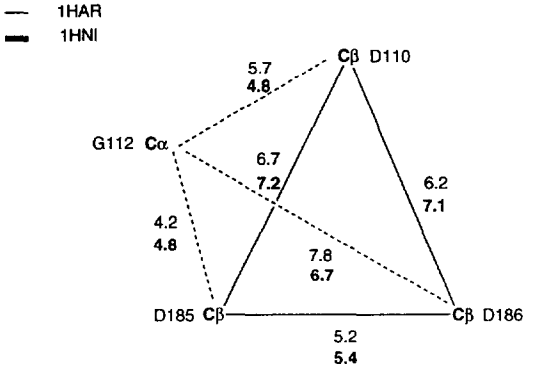

Fig 7. Region of the active site of the HIV RT highly conserved in two very different molecular assemblages 1HAR (bold) and 1HNI (standard).

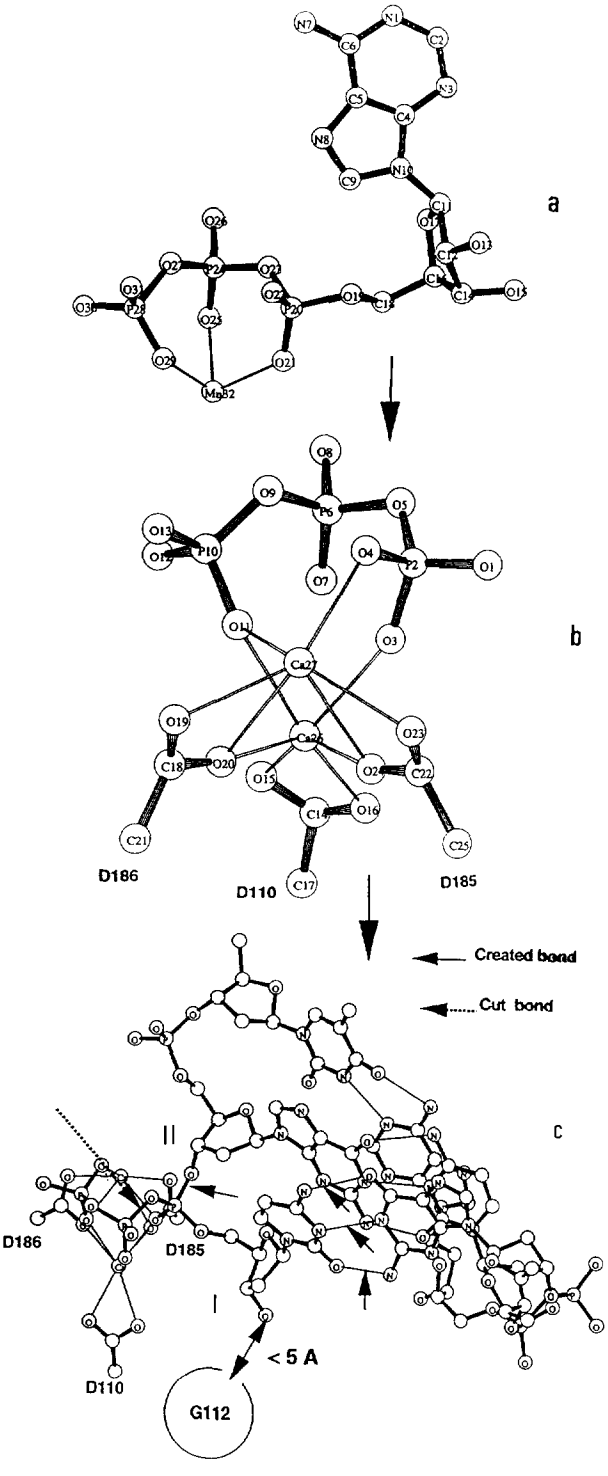

Fig 8. Building of the molecular complex of the HIV RT active site and the incoming triphosphorylated cytosine. **a.** The starting fragment structure of the triphosphate group linked to a bivalent cation observed in the X-ray structure of adenosine 5'-triphosphatomanganese. **b.** The active site: a molecular complex between the triphosphate group, the two calcium cations and the three carboxylate groups borne by methyl groups in the position corresponding to the C_β of aspartates **D110, D185** and **D186** in 1HNI [41]. **c.** Position of the DNA with respect to the active site. I. Incoming triphosphorylated molecule with a scheme of the transition state indicating the created and the cut bonds. II. Position of the nucleotide after integrating the primer strand of DNA, the other strand (template) can be DNA or RNA. For clarity we only represent three base pairs. The 3'-hydroxyl group is in close contact with the glycine **G112.**

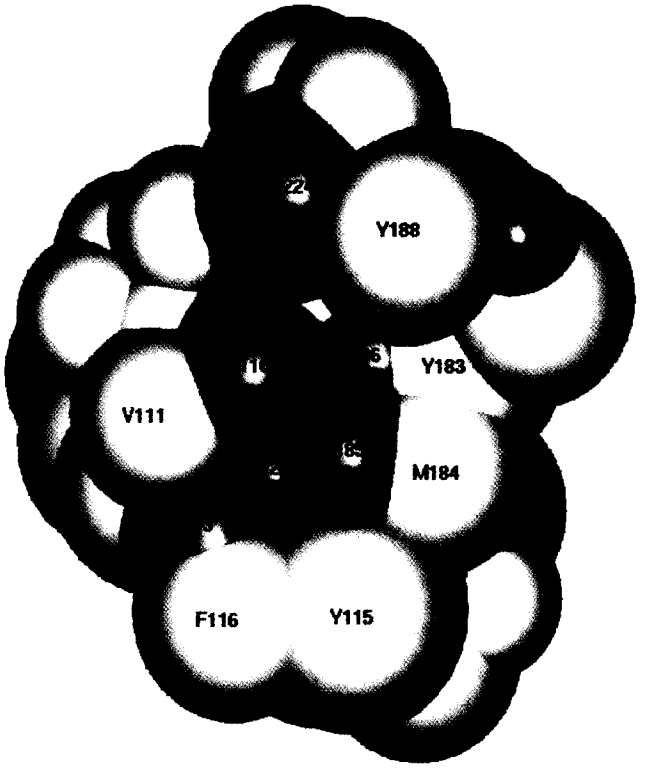

Fig 9. GenMol display of the active site of HIV RT in 1HNI in a Levitt's [47] representation showing the position of **G112** with respect to **D110**, **D185** and **D186**. Residue codes: **red**: charged + K and R, **blue**: charged - D and E, **yellow**: hydrophobic A C T V M P Y L I F and W, **green**: neutral H G N S and Q.

related to a short contact between this substituent and the residue G112. Hence it is probably useless to synthesize new molecules with other groups in this position. The position of the 4'-hydroxymethyl group with respect to the nucleic base for the anti-HIV activity, corresponding to the relative position of the template and the active site, is an important parameter. Introduction of heteroatoms or chemical functions in the ribose backbone ring suggests that a compound like **14** could be a potent anti-HIV drug. Moreover, the replacement of the ribose ring by an acyclic unsaturated chain or an acyclic imine (allenic compound **20** having already been studied) or a six-membered ring (compounds **23** and **24**) should lead to better RT inhibitor design. Syntheses of the suggested modified nucleosides are currently in progress in order to test the model.

Acknowledgments

We thank P Roycroft and P Winblattf for the English improvement.

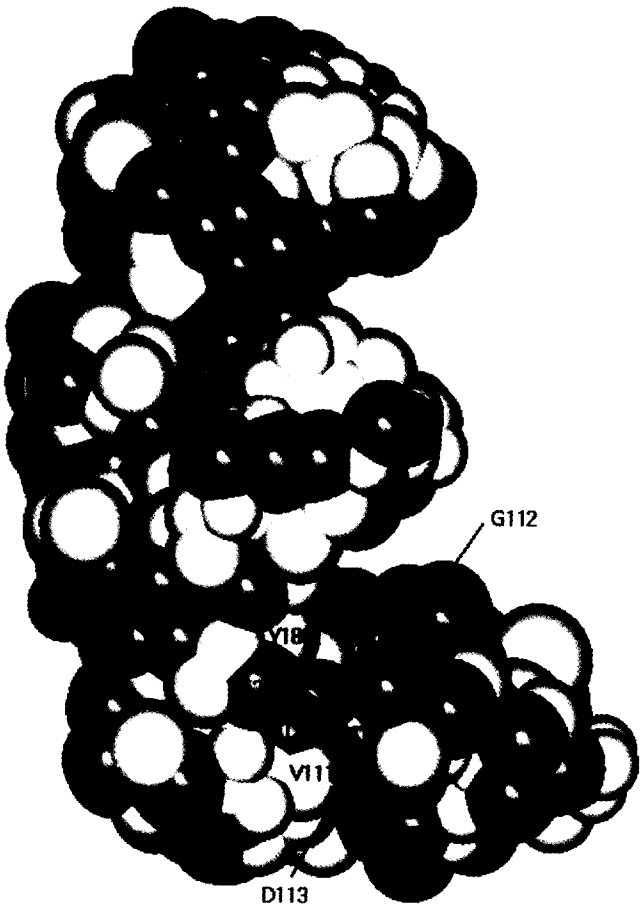

Fig 10. Position of the HIV RT active site in the A moiety of 1HNI in a Levitt's representation. The residue **D186** is buried. It indicates that the strand **V60–V75** located on the right of the active site must necessarily move to allow the DNA entry.

References

1 Posky B (1989) *Rev Infect Dis* 11, S1648–S1663
2 Sandström E (1989) *Drugs* 38, 417–450
3 Yarchoan R, Mitsuya H, Broder S (1989a) *Am J Med*, 191–200
4 Yarchoan R, Mitsuya H, Myers CE, Broder S (1989) *New Engl J Med* 321, 726–738
5 Broder S (1990) *Am J Med* 88, 1S–33S
6 Broder S (1990) *Ann Intern Med* 113, 604–618
7 De Clercq ET (1990) *Trends Pharmacol* 198–205
8 De Clercq E (1991) *J Acquired Immune Defic Syndr* 4, 207–218
9 Hirsch MS (1990) *J Infect Dis* 161, 845–857
10 Wagar MA, Evans MJ, Manly KF, Hughes RG, Huberman JA (1984) *J Cell Physiol* 402–408
11 Chien YW, Wearley LL (1989) *Drugs Today* 25, 19–25
12 Ostertag W, Roesler G, Grieg CJ et al (1974) *Proc Natl Acad Sci USA* 71, 4980–4985
13 Dube SK, Pragnell IB, Kluge N, Gaedicke G, Steinheider G, Ostertag W (1975) *Proc Natl Acad Sci USA* 72, 1863–1867

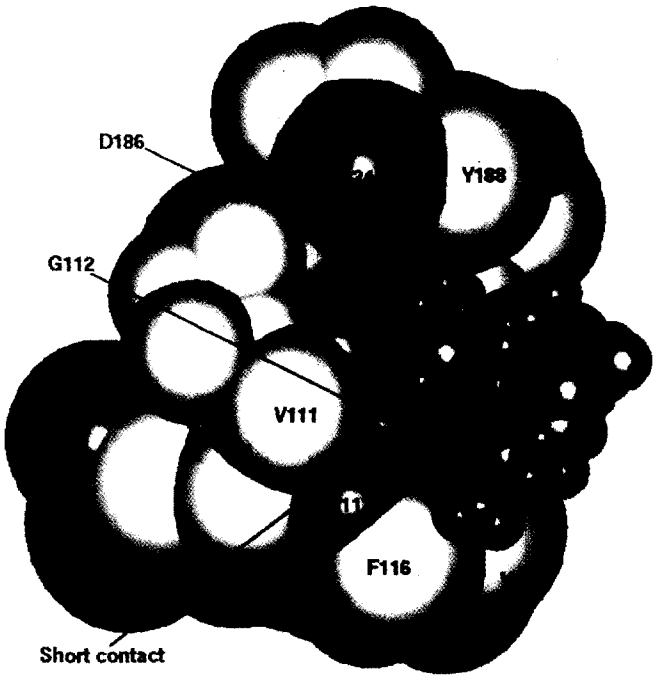

Fig 11. The active site of HIV RT in a Levitt's representation, the non-hydrogen atoms of the DNA fragment are displayed in van der Waals' representation. The short contact between the C3'-hydroxyl group and **G112** corresponding to figure 9 is indicated.

14 Mitsuya H, Broder S (1986) *Proc Natl Acad Sci USA*, 1911–1915
15 Nasr M, Litterst C, Gowan Mc (1990) *Antiv Res* 125–148
16 Nasr M, Cradock J, Johnston MI (1992) *Aids Res Hum Retroviruses* 8, 135–144
17 Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA (1992) *Science* 256, 1783–1790
18 Jacobo-Molina A, Ding J, Nanni RG et al (1993) *Proc Natl Acad Sci USA* 90, 6320–6324
19 Nanni RG, Ding J, Jacobo-Molina A, Hughes SH, Arnold E (1993) *Perspect Drug Discovery Design* 1, 129–150
20 Furman P, Fyfe J, St Clair et al (1986) *Proc Natl Acad Sci USA* 83, 8333–8337
21 St Clair M, Richards CA, Spector T et al (1987) *Antimicrob Agents Chemother* 31, 1972–1977
22 Civitico G, Wang YY, Luscombe C et al (1990) *J Med Virol* 31, 90–97
23 Birnbaum GI, Lin TS, Prusoff WH (1988) *Biochem Biophys Res Commun* 151, 608–614
24 Birnbaum GI, Giziewicz J, Gabe EJ, Lin TS, Prusoff WH (1987) *Can J Chem* 65, 2135–2139
25 Van Roey P, Taylor EW, Chu CK, Schinazi RF (1991) *Ann NY Acad Sci* 29–30
26 Fisher MA, Yadav PNS, Yadav J, Kristol D, Arnold E, Modak MJ (1994) *J Mol Recognition* 7, 211–214
27 Sundaralingam M (1975) In: *Structure and Conformation of Nucleic Acid and Protein–Nucleic Acid Interactions* (Sundaralingam M, Rao ST, eds) University Park Press, Baltimore, MD, 487–536
28 Pearlman DA, Kim SH (1985) *J Biomol Struct Dyn* 3, 99–125
29 Glasstone S, Laidler KJ, Eyring H (1941) *The Theory of Rate Process* McGraw Hill, New York
30 Scrocco E, Tomasi G (1978) *J Adv Quant Chem* 11, 115–193
31 Pèpe G, Siri D (1990) In: *Modeling of Molecular Structure and Properties* (Rivail JL, ed) Elsevier Science, Amsterdam, 93–101
32 Cavelier-Frontin F, Pèpe G, Verducci J, Siri D, Jacquier R (1992) *J Am Chem Soc* 114, 8885–8890
33 Johnson CK (1976) *Ortep II, Report ORNL- 5138*, Oak Ridge National Laboratory, Tennessee, USA
34 Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) *J Am Chem Soc* 107, 3902–3909
35 Pèpe G (1989) *J Mol Graphics* 7, 233–239
36 Pèpe G, Serres B, Laporte D, Del Ré G, Minichino C (1985) *J Theor Biol* 115, 571–575
37 Connolly MC (1983) *Science* 221, 709–712
38 Carbo R, Leyda L, Arnaud M (1980) *Int J Quantum Chem* 17, 1185–1189
39 Carbo R, Domingo L (1987) *Int J Quantum Chem* 32, 517–545
40 Good AC, Hodgkin EE, Richards WG (1992) *J Chem Inf Comput Sci* 32, 188–191
41 Belleau B, Brasili L, Chan L et al (1993) *Bioorg Med Chem Letters* 3, 1723–1728
42 Hayashi S, Phadstare S, Zemlicka J, Matsukura M, Mitsuya H, Broder S (1988) *Proc Natl Acad Sci USA* 85, 6127–6131
43 1HAR and 1HNI: entries of the Protein Data Bank Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973
44 1HAR: Unge T, Knight S, Bhikhabhai R et al (1994)
45 1HNI: Ding J, Das K, Tantillo C et al (1995)
46 Sabat M, Cini R, Haromy T, Sundaralingam M (1985) *Biochemistry* 24, 7827–7830
47 Levitt M (1976) *J Mol Biol* 104, 59–107