Monatshefte für Chemie 130, 1347–1363 (1999)

Monatshefte für Chemie
Chemical Monthly

Printed in Austria

Molecular Calculations on the Conformation of the HIV-1 Reverse Transcriptase Inhibitor 1-((2-Hydroxyethoxy)methyl)-6-(phenylthio)thymine (*HEPT*)

Luckhana Lawtrakul[1], **Supa Hannongbua**[2], **Anton Beyer**[3], and **Peter Wolschann**[1,*]

[1] Institut für Theoretische Chemie und Strahlenchemie der Universität Wien, A-1090 Vienna, Austria
[2] Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
[3] Research Institute of Molecular Pathology, A-1030 Vienna, Austria

Summary. 1-((2-Hydroxyethoxy)methyl)-6-(phenylthio)thymine (*HEPT*) is an effective inhibitor of HIV-1 reverse transcriptase. Its conformations were analyzed by molecular calculations in order to gain some information about the energetical minima and the rotational barrier around the carbon-sulfur single bond. The calculated structures were compared to the results of X-ray investigations on *HEPT* associated with HIV-1 reverse transcriptase. The NMR spectra of *HEPT* were calculated to obtain information about its structure in solution. The conformation of the molecules in the complex was analyzed with respect to the intermolecular interactions between the inhibitor and the surrounding protein which determines the geometry of the inhibition complex. Docking simulations were performed to rationalize the experimentally estimated structure of *HEPT* in the complex.

Keywords. Non-nucleoside HIV-1 reverse transcriptase inhibitors; *HEPT*; Molecular calculations; *ab initio* Calculations; Density functional theory; Docking experiments.

Molekülrechnungen zur Konformation des HIV-1 Reverse Transcriptase-Inhibitors 1-((2-Hydroxyethoxy)-methyl)-6-(phenylthio)thymin *(HEPT)*

Zusammenfassung. Die Konformationen des HIV-1 Reverse Transcriptase-Inhibitors 1-((2-Hydroxyethoxy)-methyl)-6-(phenylthio)thymin (*HEPT*) wurden mit Hilfe von Molekülrechnungen bestimmt, um Informationen über Energieminima und die Rotationsbarriere der Kohlenstoff-Schwefel-Einfachbindung zu erhalten. Die Ergebnisse wurden mit der Kristallstruktur der Substanz im Komplex mit HIV-1 Reverse Transcriptase verglichen. Die NMR-Spektren von *HEPT* wurden berechnet, um Information über seine Konformation in Lösung zu erhalten. Die Struktur von *HEPT* im Komplex wurde in Hinblick auf die intermolekularen Kräfte, die die Geometrie des Komplexes bestimmen, analysiert. Zum Zweck der Erklärung der Struktur der Verbindung im Komplex wurden Docking-Experimente durchgeführt.

* Corresponding author

Introduction

1-((2-Hydroxyethoxy)methyl)-6-(phenylthio)thymine (*HEPT*) is *in vitro* active against HIV-1 reverse transcriptase (RT) [1–2]. It is a non-nucleoside inhibitor (NNI), interacting noncompetitively with an allosteric site. It was found that *HEPT* inhibits HIV-1 replication at nanomolar concentration [3] which is about 5 orders of magnitude below the cytotoxicity threshold. As the effectivity of *HEPT* and the other NNIs is diminished by the rapid drug-resistance mutations of HIV-1 [4–5], it is necessary to search for other active derivatives. To improve the biological activity of non-nucleoside inhibitors and to develop new drugs, a definite and detailed understanding of drug-protein interactions and the physicochemical mechanism of the inhibition process is a prerequisite.

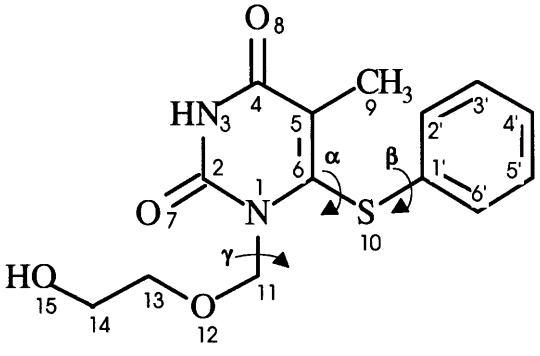
The structures of free and complexed forms of RT have been reported [6–7]. HIV-1 RT is a heterodimer that consists of a 66 kDa subunit (p66) and a 51 kDa subunit (p51). Both the p66 and p51 polymerase domains contain four subdomains (finger, palm, thumb and connection) which are arranged differently in the two subunits, and a ribonuclease H domain in the larger subunit only. It is reasonable to expect that comparing the structure of unliganded HIV-1 RT with the geometries of HIV-1 RT complexed with inhibitors should lead to valuable insights into drug-protein interactions and the mechanisms of NNI inhibition. In previous studies, conventional QSAR studies of *HEPT* analogues and a 3D-QSAR study using the comparative molecular field analysis (CoMFA) have been presented [8–10]. However, the explanation of structure-activity relationships still requires detailed information about the three-dimensional structures of both inhibitor and receptor site.

As found by a systematic search based on an empirical force field [10], *HEPT* is a rather flexible molecule. In the complex, however, restrictions on these conformational possibilities lead to only one conformation in the associate. Rotation around the carbon-sulfur single bonds between the aromatic ring systems and the carbon-carbon and carbon-oxygen single bonds of the side chain determine the conformational space of *HEPT* [11] and must be analyzed with respect to all energetically accessible conformations.

In the present paper, the results of such an analysis is given, performed by means of various *ab initio* and DFT methods.

Results and Discussion

Molecular calculation on the geometry of HEPT



The geometry of *HEPT* was fully optimized with the semiempirical methods MNDO, AM1, and PM3 and by an *ab initio* method (HF/3-21G basis set) starting from the same geometry. An *Ab initio* calculation using the HF/6-31G** basis set was then performed starting from the structure obtained from the HF/3-21G optimization. Optimization with the DFT B3LYP/6-31G** method used the optimized HF/6-31G** geometry as initial conformation. Some selected computationally determined bond lengths, bond angles, and torsion angles compared with the X-ray diffraction data of *HEPT* in the complex with RT are presented in Table 1.

The results indicate that all methods give sufficiently good results for the bond lengths except for the C-S bond distances which differ for about 0.1 Å calculated by the semiempirical methods AM1 and MNDO. As expected, the standard deviations are significantly smaller for results obtained by *ab initio* calculations than for those of the semiempirical methods (Table 1). As is generally known, by error cancellation the HF/3-21G basis set leads to a good geometry close to the results of the accurate B3LYP/6-31G** calculations. The differences between experimental and calculated bond angles are somewhat larger. Surprisingly, AM1 results show a good agreement with the crystal data. The DFT method leads to the most accurate results. Generally, bond lengths as well as bond angles are not drastically influenced by the environment in the crystal, and a comparison to the calculations in gas phase can be performed.

In contrast, the dihedral angles of *HEPT* are influenced substantially by the intermolecular forces in the complex and cannot be compared directly. Nevertheless, the dihedral angles which determine the position of the phenyl ring (C5 = C6-S10-C1′ = α, C6-S10-C1′-C2′ = β) are equal for all calculation methods used and differ significantly from those obtained by the X-ray investigation [7]. There are larger deviations of the dihedral angle γ(C6-N1-C11-O12) and also for the other dihedral angles determining the conformation of the side chain. Due to the number of rotatable single bonds and the resulting flexibility, a complete conformational analysis of the side chain is not possible; the small energy differences obtained for various conformations of the side chain indicate the high mobility of this chain. Additionally, in the crystal structure of a derivative of *HEPT* the flexibility of the chain has been experimentally verified [16]. Figure 1 shows a superposition of the lowest energy conformation calculated by the *ab initio* approach (3-21G basis set on HF level) and the crystal structure of *HEPT* in the complex.

The comparison between the two structures shows no agreement for the side chain, but only relatively small deviations of the positions of the aromatic ring. For the determination of the other conformational minima of *HEPT*, the rotational potentials of the dihedral angles α and β were analyzed first with the AM1 method and, additionally, with the aid of *ab initio* calculations at the HF/3-21G level. Both rotational potentials for α are given in Fig. 2.

AM1 calculations show two energy minima, one close to $\alpha = 70°$, a second at $\alpha = 245°$. The energy difference between these two minima is 3.4 kJ/mol, the activation barriers amount to 21.2 and 14.4 kJ/mol. The rotational potential calculated by the *ab initio* method shows qualitatively a similar picture: the first minimum is slightly shifted to 68°, the second one appears at a dihedral angle of

Table 1. Structural parameters of *HEPT* calculated by different methods compared with X-ray diffraction data [7] (dipole moments in *Debye*, bond lengths in Å, bond angles and torsional angles in degrees)

	AM1	PM3	MNDO	HF/3-21G	HF/6-31G**	B3LYP/6-31G**	X-Ray
Dipole moment	4.100	4.300	3.881	4.587	5.142	4.490	
Bond length							
N1-C2	1.419	1.433	1.432	1.380	1.378	1.399	1.394
C2-N3	1.398	1.416	1.406	1.364	1.364	1.377	1.387
N3-C4	1.403	1.424	1.420	1.383	1.375	1.397	1.391
C4-C5	1.479	1.479	1.497	1.468	1.474	1.471	1.457
C5-C6	1.368	1.360	1.370	1.328	1.339	1.363	1.354
C6-N1	1.398	1.435	1.423	1.401	1.401	1.407	1.374
C2=O7	1.254	1.228	1.225	1.219	1.199	1.224	1.232
C4=O8	1.242	1.220	1.225	1.212	1.195	1.222	1.230
C5-C9	1.479	1.485	1.508	1.506	1.505	1.502	1.531
N1-C11	1.459	1.498	1.494	1.484	1.476	1.488	1.475
C11-O12	1.419	1.405	1.394	1.409	1.374	1.394	1.431
O12-C13	1.425	1.417	1.402	1.448	1.408	1.433	1.433
C13-C14	1.528	1.542	1.560	1.520	1.515	1.523	1.526
C14-O15	1.411	1.400	1.394	1.440	1.398	1.419	1.418
C6-S10	1.723	1.797	1.711	1.837	1.790	1.797	1.822
S10-C1′	1.696	1.771	1.699	1.850	1.792	1.803	1.822
C1′-C2′	1.399	1.397	1.415	1.380	1.387	1.400	1.390
C2′-C3′	1.394	1.390	1.405	1.384	1.386	1.395	1.390
C3′-C4′	1.395	1.391	1.406	1.383	1.384	1.395	1.392
SD	0.041	0.028	0.046	0.018	0.025	0.016	
Bond angle							
C2-N1-C6	119.6	119.9	118.8	120.0	120.0	120.9	121.8
N1-C2-N3	119.3	119.2	116.1	115.4	115.9	115.1	115.7
C2-N3-C4	122.7	121.2	124.3	127.2	127.2	128.0	126.0
N3-C4-C5	116.5	117.1	114.3	114.2	114.7	114.3	114.4
C4-C5-C6	120.1	121.2	119.1	118.8	118.6	119.0	120.5
N1-C6-C5	121.8	120.4	121.3	122.8	122.8	122.5	121.5
C2-N1-C11	117.5	118.0	117.4	115.3	114.1	114.3	118.0
O7=C2-N3	120.0	118.9	121.4	122.3	121.2	122.3	121.7
N3-C4=O8	118.5	115.9	117.8	120.9	120.7	120.4	120.1
C4-C5-C9	116.0	115.0	116.6	115.4	115.2	115.8	119.5
S10-C6=C5	120.2	122.4	123.9	120.2	120.7	120.3	120.4
N1-C11-O12	113.4	110.7	109.0	115.6	114.2	115.0	111.4
C11-O12-C13	116.0	117.8	122.4	120.1	119.4	118.0	111.5
O12-C13-C14	111.8	114.0	109.7	105.7	107.3	107.2	111.7
C13-C14-O15	112.1	113.4	110.2	109.3	111.0	111.2	111.7
C1′-S10-C6	105.0	104.7	108.6	99.1	102.5	102.0	106.0
C2′-C1′-S10	118.3	117.5	117.8	117.1	117.0	116.6	119.8
C6′-C1′-C2′	120.1	119.7	118.8	121.2	121.1	120.4	119.8
C1′-C2′-C3′	119.8	120.0	120.6	119.5	119.9	119.7	120.2
C2′-C3′-C4′	120.1	120.1	120.1	120.0	120.2	120.3	120.0
C3′-C4′-C5′	120.0	120.0	119.7	119.8	119.9	119.6	119.9
SD	1.9	2.7	3.0	3.3	2.7	2.6	

Table 1 (*continued*)

	AM1	PM3	MNDO	HF/3-21G	HF/6-31G**	B3LYP/6-31G**	X-Ray
Torsional angle							
N1-C2-N3-C4	359.3	356.2	359.0	357.8	358.0	359.9	0.0
C4-C5=C6-N1	1.9	2.0	2.1	4.5	4.6	3.6	0.1
O7=C2-N1-C6	179.7	174.5	203.5	191.0	191.2	184.7	180.4
C2-N3-C4=O8	181.2	188.7	164.6	176.0	175.8	178.1	179.8
N3-C4-C5-C9	177.1	169.8	199.1	188.3	188.5	183.0	178.5
C5=C6-N1-C11	182.4	195.6	138.8	160.6	160.1	167.1	178.3
C6-N1-C11-O12	84.8	82.0	264.9	116.5	116.7	117.1	252.1
N1-C11-O12-C13	81.8	91.5	104.0	288.8	288.6	289.1	139.7
C11-O12-C13-C14	72.2	67.3	201.5	124.6	124.7	128.7	168.2
O12-C13-C14-O15	233.3	238.7	69.8	304.0	303.8	301.2	43.3
C5=C6-S10-C1′	76.1	67.5	82.1	68.0	67.7	64.6	75.0
C6-S10-C1′-C2′	240.5	235.7	234.8	220.0	221.3	212.3	183.8
S10-C1′-C2′-C3′	183.8	184.9	184.1	182.3	182.3	183.8	177.8
C1′-C2′-C3′-C4′	359.5	359.1	359.8	359.3	359.2	359.0	0.1
C2′-C3′-C4′-C5′	0.4	0.4	0.4	0.4	0.4	0.4	359.9

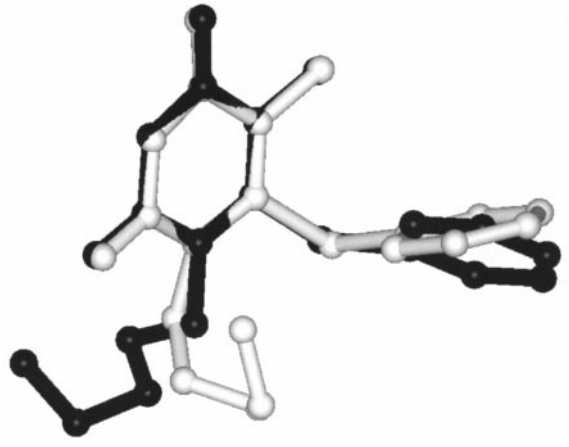


Fig. 1. Lowest energy conformation of *HEPT* calculated by *ab initio* method (HF/3-21G, grey) superimposed with the X-ray structure of *HEPT* in the association complex with HIV-1 RT (black)

$\alpha = 250°$. The rotational barrier is significantly higher (36 and 27.4 kJ/mol). Moreover, two local minima can be recognized along the rotational potential, one close to the absolute minimum (around $\alpha = 105°$), the second as a shoulder close to the second minimum at $\alpha = 280°$. The two rotational barriers are caused by the steric interactions between the phenyl ring and the methyl group of the thymine ring systems and the interaction between the phenyl ring and the side chain. Table 2 gives some structural data of both minima obtained by AM1 and *ab initio* calculations of different basis sets and methods.

Thus, for *HEPT* two conformational minima are found. AM1 favours the minimum around $\alpha = 76°$ as absolute minimum and the second minimum at $\alpha = 250°$, 3.35 kJ/mol higher in energy. *Ab initio* calculations with the small HF/3-

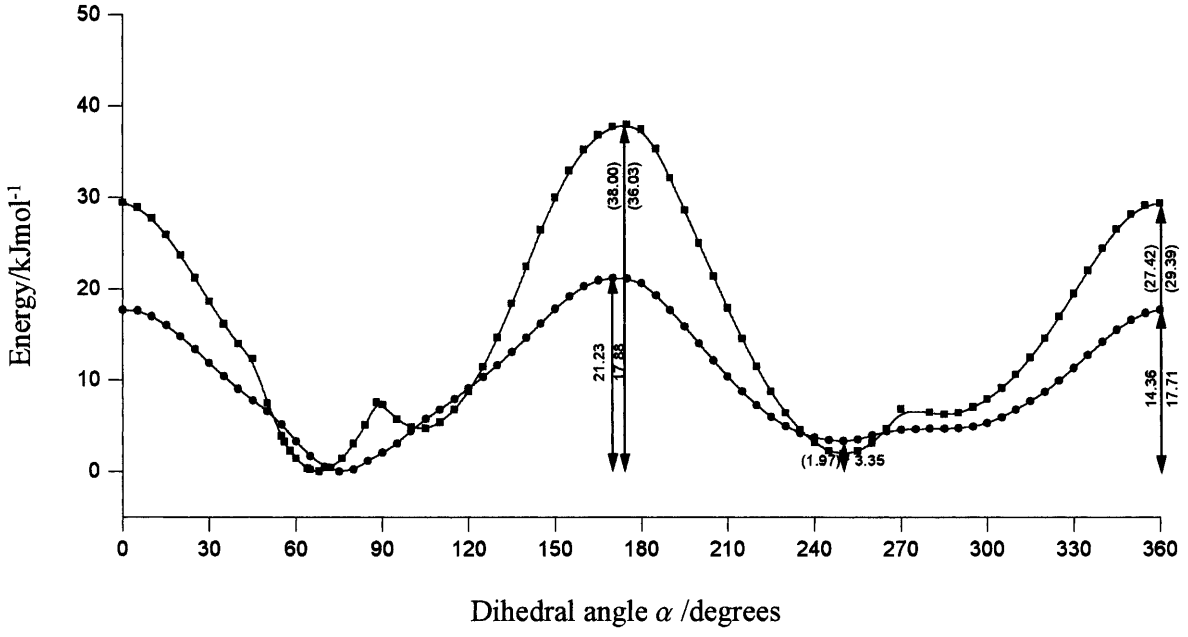


Fig. 2. Calculated AM1 (–●–) and HF/3-21G (–■–) rotational potentials for the dihedral angle α

21G basis set support these results in principle, but with a slightly smaller energy difference (1.97 kJ/mol). The more accurate higher basis set (HF/6-31G**) as well as the DFT method (B3LYP/6-31G**) indicate, however, that rather the second conformational minimum might be the global minimum. Nevertheless, all methods confirm that two conformational minima exist with a rather small energy difference in between. In Fig. 3, both conformational minima are shown, again with a comparison to the X-ray structure.

As pointed out before, the conformation of *HEPT* in its complex with HIV-1 reverse transcriptase correlates with the calculated conformational minimum at a dihedral angle α of 65°. Taking into account the rotation of the phenyl ring (dihedral angle β), a two-dimensional energy surface has to be considered. This surface, *i.e.* its energy contour plot as obtained by AM1 calculations, is shown in Fig. 4.

The contour plot should be symmetric due to the rotation of the phenyl ring. Small deviations from this symmetry are caused by a slightly different conformational behaviour of the side chain. The valley in Fig. 4 for α-values around 80° correlate with the minimum of the rotational potential in Fig. 2. The rotation of the phenyl ring is less sterically hindered for values of the dihedral angle α around the energy minimum. The reaction path for the rotation around α, shown in Fig. 2, is included in the contour plot and shows the synchronous change of the dihedral angle β during the variation of α. The calculation of the complete energy hypersurface by *ab initio* methods would need too much calculation time. Therefore, only small parts of the total hypersurface were calculated using the HF/3-21G basis set. The energetically favourable regions of the hypersurface are presented in Fig. 5.

Table 2. Selected structural parameters of the two conformational minima of *HEPT* (energies in kJ/mol, dipole moments in *Debye*, bond lengths in Å, bond angles and torsion angles in degrees)

	1st minimum				2nd minimum			
	AM1	HF/ 3-21G	HF/ 6-31G**	B3LYP/ 6-31G**	AM1	HF/ 3-21G	HF/ 6-31G**	B3LYP/ 6-31G**
ΔE	0	0	8.62	2.72	3.35	1.97	0	0
Dipole moment	1.614	4.587	5.142	4.490	1.847	6.778	6.668	5.866
Bond length								
C6-S10	1.723	1.837	1.790	1.797	1.720	1.839	1.840	1.794
S10-C1′	1.696	1.850	1.792	1.803	1.695	1.847	1.845	1.803
Bond angle								
S10-C6=C5	120.2	120.2	120.7	120.3	118.5	119.5	119.5	119.2
C1′-S10-C6	105.0	99.1	102.5	102.0	104.5	100.8	100.6	103.0
C2′-C1′-S10	118.3	117.1	117.0	116.6	118.1	117.7	118.0	117.3
Torsional angle								
C4-C5=C6-N1	1.9	4.5	4.6	3.6	1.6	0.8	0.6	0.2
O7=C2-N1-C6	179.7	191.0	191.2	184.7	181.2	180.3	180.2	178.4
C2-N3-C4=O8	181.2	176.0	175.8	178.1	181.1	180.4	180.5	182.3
N3-C4-C5-C9	177.1	188.3	188.5	183.0	178.9	180.0	180.0	176.8
C5=C6-N1-C11	182.4	160.6	160.1	167.1	181.2	182.6	183.2	186.6
C6-N1-C11-O12	84.8	116.5	116.7	117.1	78.6	110.7	110.7	113.8
N1-C11-O12-C13	81.8	288.8	288.6	289.1	80.9	303.4	303.4	297.7
C11-O12-C13-C14	72.2	124.6	124.7	128.7	72.9	139.1	139.1	144.2
O12-C13-C14-O15	233.3	304.0	303.8	301.2	234.1	300.3	300.3	299.7
C5=C6-S10-C1′	76.1	68.0	67.7	64.6	250.0	250.0	250.1	243.9
C6-S10-C1′-C2′	240.5	220.0	221.3	212.3	227.9	220.0	224.3	213.2
S10-C1′-C2′-C3′	183.8	182.3	182.3	183.8	183.7	182.6	182.5	183.4

Although the part of the surface considered is rather small, a similar shape of the energy hypersurface can be observed for the *ab initio* and AM1 results. As the rotational barrier between the two minima appears to be higher estimated by *ab initio* calculations, this energy hypersurface is not so smooth anymore. The two minima detected for the torsional potential can be also seen in the hypersurface at a dihedral angle β of about 140°.

From these results it can be concluded that the dihedral angle α is restricted to two pronounced minima with rather large activation barriers in between. On the other hand, the dihedral angle β has a very low rotational barrier if α relates to an energy minimum, but is much more restricted in non equilibrated geometries.

Calculated NMR spectra of HEPT

Molecular calculations lead to information about structures in the gas phase, whereas X-ray diffraction experiments describe the geometries in the solid state, in the case of *HEPT* the geometry of the complexed compound. To obtain some information about the geometry in solution, the NMR spectra of *HEPT* were

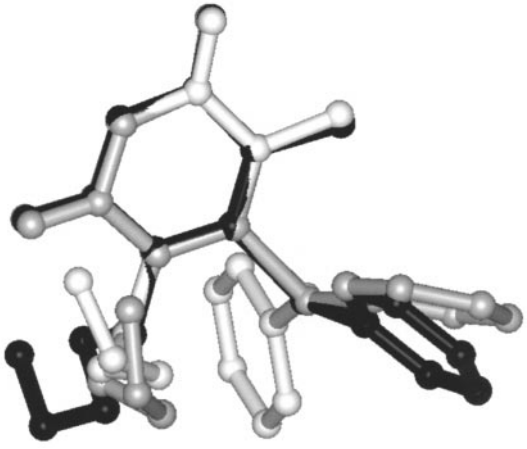


Fig. 3. Global minimum conformation of *HEPT* calculated by DFT method (B3LYP/6-31G**), ($\alpha = 244^{\circ}$, light grey) superimposed with the X-ray structure of *HEPT* in the association complex with HIV-1 RT (black) and the second minimum at $\alpha = 65^{\circ}$ (dark grey)

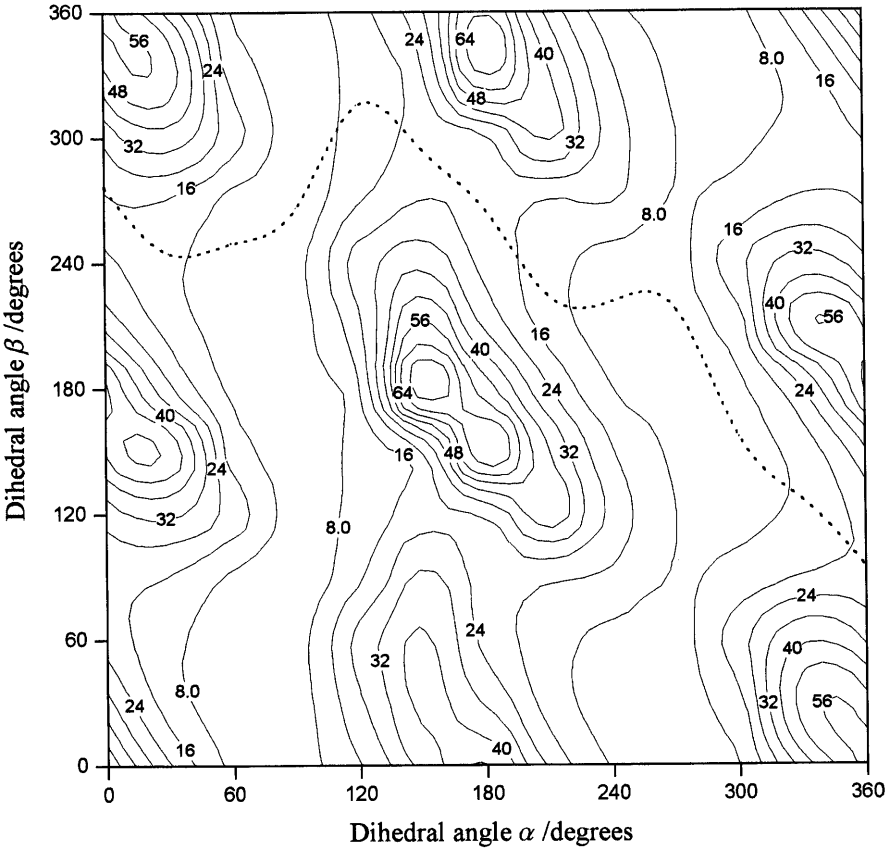


Fig. 4. Contour plot of the energy hypersurface of *HEPT* defined by the dihedral angles α and β; the absolute energy minimum is set to zero, the energy difference between the individual contour lines is 9 kJ/mol; the reaction path of the rotation of the dihedral angle α only (Fig. 2) is inserted in the drawing

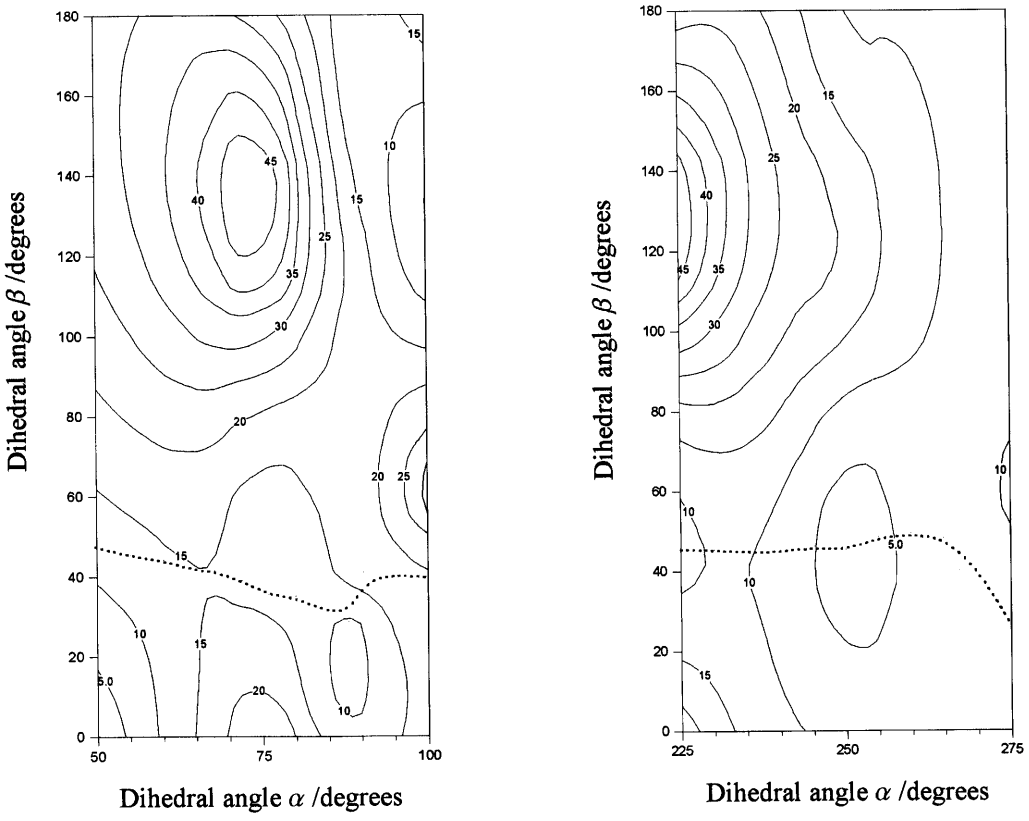


Fig. 5. Contour plot of two parts of the energy hypersurface of *HEPT* defined by the dihedral angles α and β; the absolute energy minimum is set to zero, the energy difference between the individual contour lines is 5 kJ/mol; the reaction path of the rotation of the dihedral angle α only (Fig. 2) is inserted in the drawing

calculated for both minima and compared to the experimental ^{1}H NMR spectra [17]. Table 3 gives the calculated NMR spectra for various *ab initio* basis sets and methods. The spectra were calculated using the respective optimized structures for each method except for the extended HF/6-311++G** and B3LYP/6-311++G** levels where the optimized structures of the corresponding B3LYP/6-31G** method were used.

As expected, the agreement between the chemical shifts as obtained with the small basis set HF/3-21G and enhanced *ab initio* calculations is quite poor, although the geometry is derived successfully from this basis. Therefore, only calculations on a rather high level deliver more proper results which can be compared with the experiment. The main difference between the spectra of both conformers occurs for the chemical shift of the methyl group at the heterocyclic ring and the protons of the methylene bridge bound to the nitrogen of the thymine ring, because these positions are predominantly influenced by the phenyl ring position. From the comparison of the experimental and the calculated spectra it seems that the calculated minimum at $\alpha = 65°$ gives a better correlation than the

Table 3. ^{1}H NMR chemical shifts (ppm) of *HEPT* relative to methane as obtained by different methods

	1st minimum					Expt.
	HF/ 3-21G	HF/ 6-31G**	B3LYP/ 6-31G**	HF/ 6-311++G**	B3LYP/ 6-311++G**	
$N(CH_2)O$	4.17	4.36	5.34	4.70	5.50	5.29
$N(CH_2)O$	4.81	4.47	5.08	4.53	5.15	5.29
5-Me	1.06	1.08	1.45	1.28	1.62	1.70
$OHCH_2(CH_2)O$	3.24	3.22	3.74	3.20	3.82	3.35
$OHCH_2(CH_2)O$	2.78	2.71	3.31	2.88	3.54	3.35
$OH(CH_2)CH_2O$	3.12	3.01	3.52	3.06	3.68	3.24–3.28
$OH(CH_2)CH_2O$	2.64	2.58	3.00	2.54	3.12	3.24–3.28
SPh	6.89	7.24	7.13	7.26	7.34	7.14–7.25
SPh	7.81	7.40	7.53	7.48	7.68	7.14–7.25
SPh	6.86	7.16	7.11	7.20	7.30	7.14–7.25
SPh	7.05	7.23	7.17	7.25	7.34	7.14-7.25
SPh	6.91	7.14	7.02	7.12	7.17	7.14–7.25
	2nd minimum					Expt.
	HF/ 3-21G	HF/ 6-31G**	B3LYP/ 6-31G**	HF/ 6-311++G**	B3LYP/ 6-311++G**	
$N(CH_2)O$	3.90	4.15	5.01	4.23	5.15	5.29
$N(CH_2)O$	4.61	4.23	4.79	4.16	4.86	5.29
5-Me	1.76	1.79	1.95	1.88	2.13	1.70
$OHCH_2(CH_2)O$	2.17	2.43	3.01	2.46	3.10	3.35
$OHCH_2(CH_2)O$	2.50	2.52	3.21	2.76	3.43	3.35
$OH(CH_2)CH_2O$	2.82	2.90	3.43	2.90	3.57	3.24–3.28
$OH(CH_2)CH_2O$	2.45	2.58	2.93	2.48	3.07	3.24–3.28
SPh	6.93	7.29	6.80	6.84	6.94	7.14–7.25
SPh	6.62	7.05	7.17	7.27	7.37	7.14–7.25
SPh	6.98	7.21	7.10	7.17	7.25	7.14–7.25
SPh	6.87	7.16	7.17	7.27	7.35	7.14–7.25
SPh	6.94	7.19	7.05	7.19	7.22	7.14–7.25

second (global) minimum. Figure 6 shows the experimental ^{1}H NMR spectrum of *HEPT* [17] together with the results of highest level DFT calculations.

The comparison of the experimental spectrum and the calculated one (B3LYP/ 6-311++G**) favours the presence of a conformation in solution which agrees with the first minimum at a dihedral angle α of about 65°.

From these calculations and from the experimental spectra it can be concluded that *HEPT* exists in two energetically similar conformations in the gas phase with respect to the position of the aromatic ring, whereas the side chain is rather flexible without pronounced energetic minimum geometry. In solution, one conformation seems to be predominant which corresponds principally to the conformation in which the molecules exists in the complex.

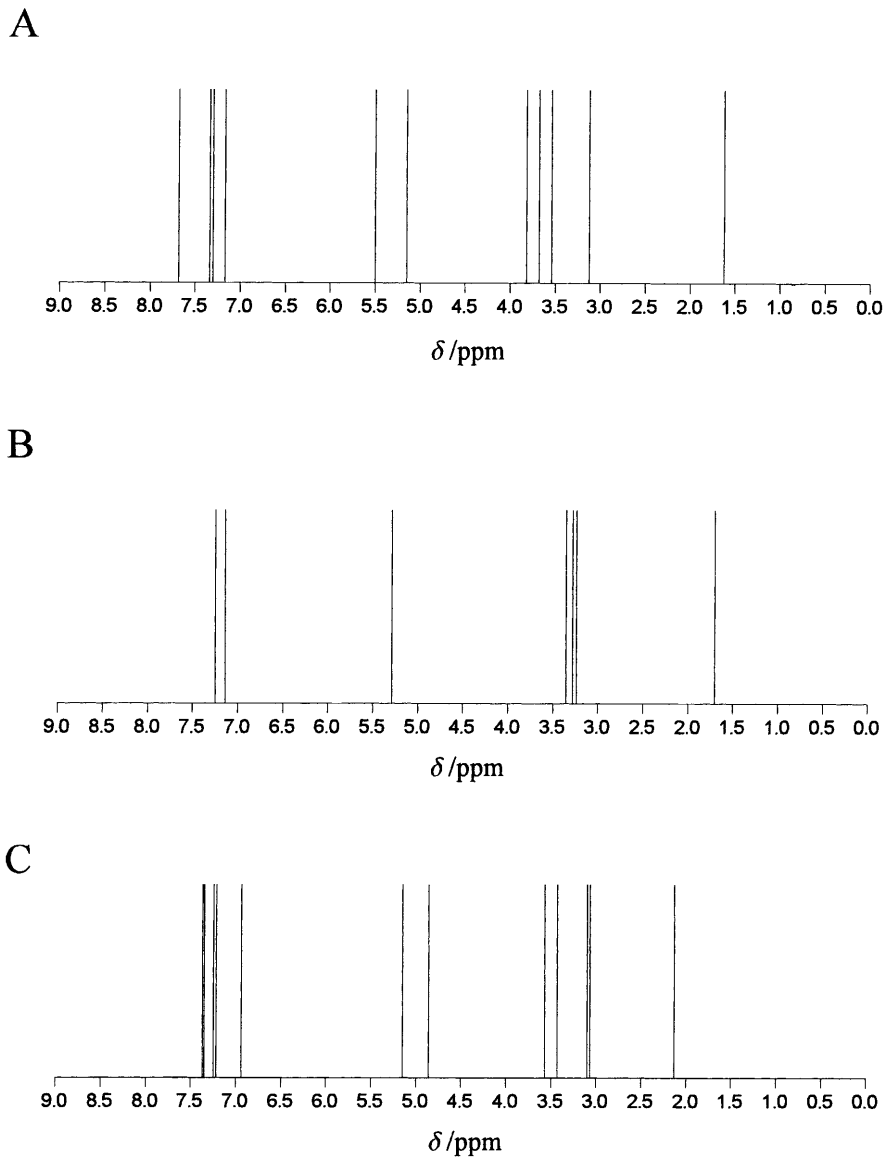


Fig. 6. Calculated and experimental ^{1}H NMR spectra of *HEPT*; *A*: energy minimum conformation with $\alpha = 65°$ (B3LYP/6-31++G**), *B*: experimental data [17], *C*: energy minimum conformation with $\alpha = 244°$ (B3LYP/6-31++G**)

Conformation of HEPT in the crystal of the complex with HIV-1 reverse transcriptase

The conformation of the association complex between *HEPT* and HIV-1 reverse transcriptase was obtained by X-ray diffraction [7]. Therefore, it was possible to investigate the interaction between the guest molecule and the protein in detail [18]. The amino acids which are close to the guest molecule could be detected, and the specific interactions between parts of the *HEPT* molecule and the environment were analyzed. The binding pocket of the HIV-1 RT with *HEPT* is shown in Fig. 7.

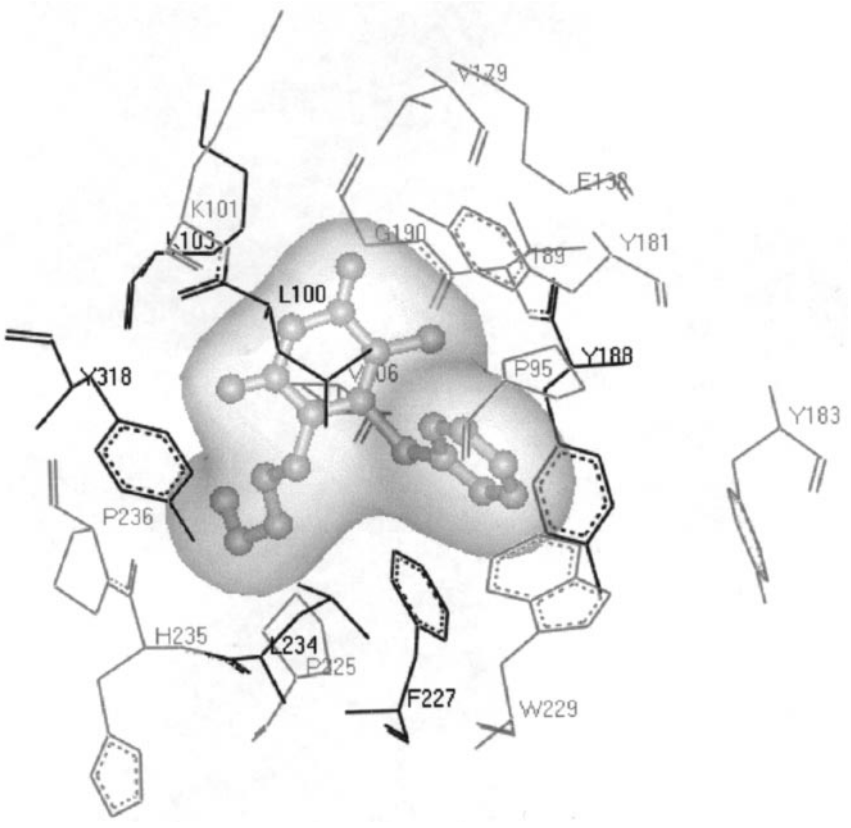


Fig. 7. Crystal structure of the binding pocket of *HEPT* with HIV-1 RT; only amino acids of the cavity with a distance of less than 6 Å are considered [7]; compared to the crystal structures of some other NNRTIs [19–21], only the amino acids which keep their position unchanged are drawn in black

Hydrogen bonding is established between the carbonyl oxygen of lysine K101 and the proton at the nitrogen of the thymine ring. Interactions (heavy atom distances lower than 3.6 Å) occur between the phenyl ring and tyrosine Y188 as well as with leucine L100. The heterocyclic ring interacts with the protein amino acids lysine K103, lysine K101, valine V179, and tyrosine Y188. The side chain is in contact with the amino acids histidine H235 and phenylalanine F227. Interestingly, the pocket is mainly hydrophobic [5, 7], and so *van der Waals* interactions play an important role for the association and, therefore, the inhibition.

Amino acids located at fixed positions are indicated in Fig. 7. These positions were estimated by a comparison between the crystal structures of various non-nucleoside HIV-1 RT inhibitors (NNRTIs) [5, 7, 19–21] and represent the molecular surface to which the complementary molecular surface of an inhibitor molecule has to be fitted exactly. The other amino acids are allowed to change their positions within some limits without preventing an association of a guest molecule.

The geometry of *HEPT* in the complex does not differ too much from one of the energy minimum conformations ($\alpha = 68°$; HF/3-21G basis set) of the free

molecule. In the complex, α is changed to 75°, and β from 220 to 184° (see Tables 1 and 2, Fig. 2) by the interaction with lysine L100 and tyrosine Y188 (heavy atom distances from 3.33 to 3.52 and 3.26 to 3.56 Å, respectively). The fact that the conformation of the molecule has not to be distorted to a large extent, which would be connected with a higher amount of activation energy, increases the binding ability of *HEPT* with the HIV-1RT and enhances the stability of the association complex. This is probably important for the stability of the complex and furthermore for its biological activity. In the conformation of the second minimum (α around 250°) the aromatic ring would experience a very strong repulsion with phenylalanine F227, and the side chain would have the same effect with leucine L100. This prevents the existence of that minimum conformation in the complex.

Figure 8 shows the superposition of the binding pocket of HIV-1 RT with *HEPT* [7] and of HIV-1 reverse transcriptase alone [19]. It can clearly be seen that most of the amino acids are not affected by the inhibitor molecule drastically. The association of the inhibitor into the protein's cavity does not take place with a synchronous reorientation of the amino acids in the pocket; the association

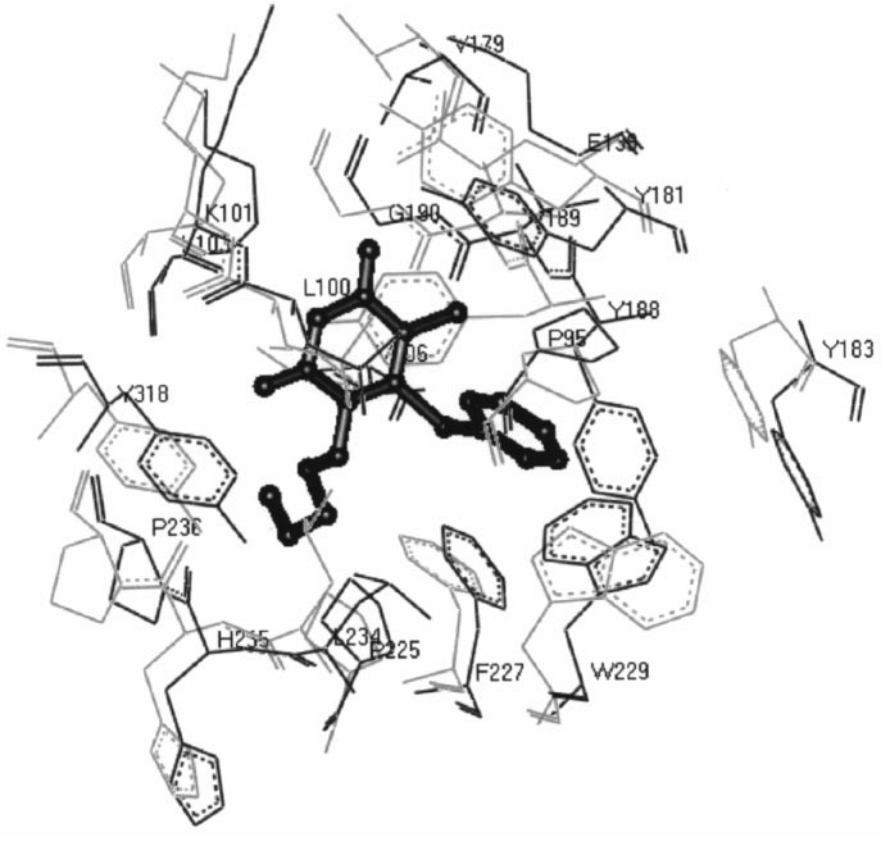


Fig. 8. Crystal structure of the binding pocket of *HEPT* with HIV-1 RT (black) [7] and of the unliganded HIV-1 RT cavity (light grey) [19]; only amino acids of the cavity with a distance of less than 6 Å are considered

Table 4. Interaction energies of docking simulations between *HEPT* and the amino acids of the inhibitor binding pocket

Starting geometry	Simulated geometry			Energy (kJ/mol)
	C5=C6-S10-C1′ α/°	C6-S10-C1′-C2′ β/°	C6-N1-C11-C12 γ/°	
1st minimum (*A*)	81.32	326.83	259.69	−292.8
Crystal structure (*B*)	82.16	347.08	269.94	−270.0
2nd minimum (*C*)	253.08	88.01	110.98	−207.0

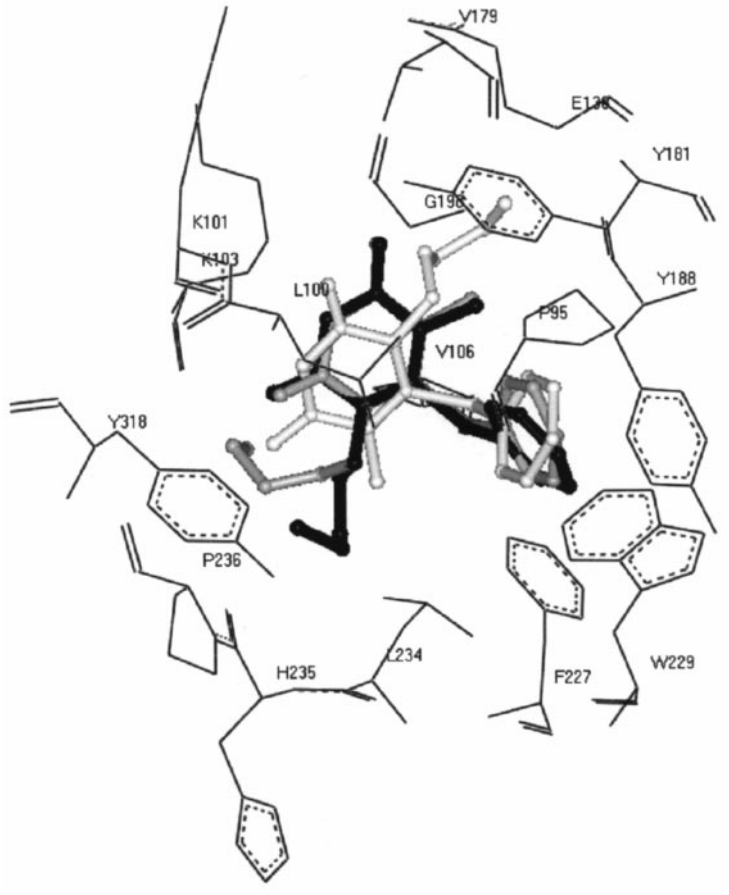


Fig. 9. The docking *HEPT* geometries analysis in the complex with HIV-1 RT binding pocket; the geometry which starts from the X-ray structure is drawn in black, the geometries starting from both minima (1st minimum: $\alpha = 65°$, 2nd minimum: $\alpha = 244°$) are drawn in dark grey and light grey, respectively

complex seems to be rather stable, and therefore *HEPT* prevents the biological activity of the protein.

Docking

To support the above mentioned hypothesis, docking experiments were performed to predict the interaction of *HEPT* with the inhibition binding pocket of the HIV-1 reverse transcriptase. A Monte Carlo simulated annealing technique was applied to find the interaction energy minimum of the complexation. The geometry of *HEPT* obtained from X-ray analysis of the complex with HIV-1 RT and both minima found in the molecular calculations (B3LYP/6-31G**) were taken as starting geometries. The program AutoDock successfully reproduced the crystallographically determined position of *HEPT* with some small deviations of the geometry only. Starting from the geometry of the first conformational minimum (*A*) a similar interaction energy is obtained, again with a geometry similar to the conformation found in the crystal structure (B).

A comparison of the result of a typical simulation run and the crystal structure is given in Fig. 9. Larger deviations of both structures are observed for the flexible side chain. In the docking simulation experiment, a hydrogen bond is established between the hydroxyl group of the side chain and amino acid lysine K103 which is not experimentally found in the crystal structure. The reason for that might be that no water molecules have been included in the simulation. Nevertheless, the interaction energy is somewhat higher than in the docking simulation starting from the crystal structure geometry. In Table 4 the lowest obtained interaction energies are given for the different starting geometries. Starting from the second minimum geometry (*C*), in no simulation run the crystal structure geometry could be obtained. A complete different orientation of the *HEPT* molecule in the binding pocket leads to much lower interaction energies. Moreover, the program fails in adjusting the dihedral angles α and β in such a way that a geometry similar to the crystal structure conformation of *HEPT* is obtained.

Conclusions

The conformational analysis of *HEPT* yields two energy minima with respect to the orientation of both aromatic ring systems. The pronounced difference between these two energetically similar minima concerns the dihedral angle α which determines the position of the phenyl ring. The flexible side chain shows considerable disorder without a distinct energetic minimum geometry. One conformational minimum appears for a dihedral angle α of *ca.* 65–76°, the second for an α value of about 244–250°. The rotational potentials of the carbon-sulfur single bonds were investigated by semiempirical AM1 and *ab initio* HF/3-21G methods. The energy barrier due to the steric interactions between the phenyl ring and the side chain as well as the methyl group of the thymine ring system amounts to 18–21 and 36–38 kJ/mol. The comparison of calculated ^{1}H NMR spectra with the experiment indicates that the first minimum is the preferred conformation in solution which corresponds also to the geometry of the molecule in the association complex as determined by X-ray diffraction. The pocket in which *HEPT* binds is

characterized by a mainly hydrophobic surface. The binding force depends on the complementary molecular shape of the inhibition molecules and results from a number of *van der Waals* and ring stacking interactions and contributes to the electrostatic interactions as well as to hydrogen bonding. The conformation of *HEPT* in the binding pocket of HIV-1 reverse transcriptase is very similar to one calculated gas phase conformational minimum, leading to the conclusion that the amount of energy necessary for conformational changes before association takes place is rather small and does not diminish the association energy. Docking simulation experiments support the good fit of *HEPT* with the inhibition pocket.

Methods

An initial conformation for *HEPT* was generated by means of the ALCHEMY III molecular modelling software package [12]. The geometry optimizations were performed using semiempirical methods (AM1, PM3, and MNDO) as well as by *ab initio* methods (3-21G and 6-31G** basis sets at the HF level) as included in the GAUSSIAN 94 [13] program package. The torsional potentials of the dihedral angles given by the carbon-sulfur single bonds were calculated using the option SCAN in the GAUSSIAN 94 program. The same program package was used to optimize the *HEPT* conformation using density functional theory (DFT) at the B3LYP/6-31G** level. Theoretical calculations are compared to the experimental geometry, in particular to X-ray diffraction data of the reverse transcriptase – *HEPT* complex [7], as available from the Brookhaven Protein Data Base (1RTI) [14]. The NMR spectra of *HEPT* were calculated by the GIAO method using the HF/6-31G** and B3LYP/6-31G** method. Additionally, the spectra were calculated with the HF/6-311++G** and B3LYP/6-311++G** methods on the B3LYP/6-31G** optimized structures.

Docking procedures were performed by the program AutoDock 2.4 [15]. Different starting geometries of *HEPT* were used, and flexibility for all rotatable bonds of the molecule (eight) was allowed. A Monte Carlo (MC) simulated annealing technique was applied to find the conformational minimum (from 250 K (starting temperature) down to 20 K in 50 cycles; 10 runs were performed for one starting geometry).

Acknowledgements

This investigation was supported by the Austrian *Fonds zur Förderung der wissenschaftlichen Forschung* (project 11349-CHE), the *National Research Council of Thailand* under the *Austrian-Thailand Cooperative Science Program* (NRCT-FWF), and the *Development and Promotion of Science and Technology Talents Project* (DPST) of Thailand. The quantum chemical calculations were performed on the Cluster of Digital Alpha Servers (2100 4/275) of the computer center of the University of Vienna; generous supply of computer time on this installation is grateful acknowledged. The authors wish to thank Prof. *Alfred Karpfen* and Dr. *Hellfried Schreiber* for helpful discussions and critical comments.

References

[1] Tanaka H, Takashima H, Ubasawa M, Sekiya K, Nitta I, Baba M, Shigeta S, Walker RT, De Clercq E, Miyasaka T (1992) J Med Chem **35**: 337
[2] Tanaka H, Takashima H, Ubasawa M, Sekiya K, Nitta I, Baba M, Shigeta S, Walker RT, De Clercq E, Miyasaka T (1992) J Med Chem **35**: 4713
[3] De Clercq E (1995) J Med Chem **38**: 2491
[4] Buckheit RW Jr, Fliakas-Boltz V, Yeagy-Bargo S, Weislow O, Mayers DL, Boyer PL, Hughes SH, Pan BC, Chu SH, Bader JP (1995) Virology **210**: 186

[5] Hopkins AL, Ren J, Esnouf RM, Willcox BE, Jones EY, Ross C, Miyasaka T, Walker RT, Tanaka H, Stammers DK, Stuart DI (1996) J Med Chem **39**: 1589
[6] Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolft DJ, Debouck C, Harrison SC (1995) Proc Natl Acad Sci USA **92**: 1222
[7] Ren J, Esnouf R, Garman E, Somers D, Kirby CRI, Keeling J, Darby G, Jones Y, Stuart DI, Stammers D (1995) Nat Struct Biol **2**: 293
[8] Hannongbua S, Lawtrakul L, Limtrakul J (1996) J Comp Aided Mol Des **10**: 145
[9] Hannongbua S, Lawtrakul L, Sotriffer CA, Rode BM (1996) Quant Struct Act Relat **15**: 389
[10] Kireev DB, Chretien JR, Grierson DS, Monneret C (1997) J Med Chem **40**: 4257
[11] Lawtrakul L, Hannongbua S, Beyer A, Wolschann P (1999) Biol Chem **380**: 265
[12] ALCHEMY III (1992) Tripos Associates Inc, St. Louis, MO
[13] Frisch MJ, Trucks GW, Schlegel HB, Gill PMW, Johnson BG, Robb MA, Cheeseman JR, Keith T, Petersson GA, Montgomery JA, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martin RL, Fox DJ, Binkley JS, Defrees DJ, Baker J, Stewart JP, Head-Gordon M, Gonzales C, Pople JA (1995) Gaussian 94 (Version B 3), Gaussian Inc, Pittsburgh, PA
[14] Bernstein FC, Koetzle TF, William GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) J Mol Biol **112**: 535
[15] Goodsell DS, Morris GM, Olson AJ (1996) J Mol Recogn **9**: 1
[16] Maruenda H, Johnson F (1995) J Med Chem **38**: 2145
[17] Miyasaka T, Tanaka H, Baba M, Hayakawa H, Walker RT, Balzarini J, De Clercq E (1989) J Med Chem **32**: 2507
[18] Yadav PNS, Yadav JS, Modak MJ, Das K, Ding J, Arnold E (1998) J Mol Struct (Theochem) **423**: 101
[19] Esnouf R, Ren J, Ross C, Jones Y, Stammers D, Stuart D (1995) Nat Struct Biol **2**: 303
[20] Ding J, Das K, Zhang W, Clark AD, Jessen S Jr, Lu X, Hsiou Y, Jacobo-Molina A, Andries K, Pauwels R, Moereels H, Koymans L, Jansses PAJ, Smith RH, Kroeger Koepke M Jr, Michejda CJ, Hughes SH, Arnold E (1995) Structure **3**: 365
[21] Ren J, Esnouf R, Hopkins A, Ross C, Jones Y, Stammers D, Stuart D (1995) Structure **3**: 915

Received March 1, 1999. Accepted March 23, 1999