Acta Biochimica Polonica

Vol. 45 No. 4/1998

977–985

QUARTERLY

This paper is dedicated to Professor Maciej Wiewiórowski

Hydration of C–H groups in natural dithymidine nucleotide and its methylphosphonate analogues[o]

Katarzyna Kulińska[1✉], Tadeusz Kuliński[1] and Jacek Stawiński[2]

[1]*Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland,*
[2]*Stockholm University, Department of Organic Chemistry, Arrhenius Laboratory, S-106 91 Stockholm, Sweden*

Key words: oligonucleotide analogues, methylphosphonates, hydration pattern, molecular dynamics

In this paper we report our preliminary studies on the hydration pattern of selected C–H groups in natural thymidyl(3′-5)thymidine and its R_P and S_P-methylphosphonate analogues using Molecular Dynamic simulations in aqueous solutions. The methyl groups attached to the phosphorus center (P–Me) in methylphosphonate analogues are hydrated by water molecules as efficiently as the hydrophilic P=O group in the natural dithymidine nucleotide and better than the neutral P=O functions in these compounds, although the nature of the hydration is different. The C5–Me centers of the 3′-yl units seem to be better hydrated in the methylphosphonate analogues than in the natural dithymidine phosphate and than other centers of the thymine bases in methylphosphonate analogues. Due to chirality of the phosphorus center, the C5–Me group of the 5′-yl unit in the S_P diastereomer coordinates more water than that in the R_P diastereomer. The C6–H group in the 5′-yl unit of the S_P diastereomer exhibits a specific interaction with water.

It is generally accepted that hydrophobic effects play a crucial role in the organisation of biomolecules and cellular structures, and are very important factors governing molecular recognition and interaction processes. Although there is no comprehensive theory of hydrophobicity, studies on low molecular models indicate that in aqueous solutions, molecules of hydrophobic substances are surrounded by structurally distinct hydration shells, in which water molecules exhibit an increased number of solvent–solvent intermo-

[o]Financial support from the State Committee for Scientific Research, Poland, and the Swedish Council for Natural Sciences, is gratefully acknowledged.

[✉]Katarzyna Kulińska, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 14, 61-704 Poznań, Poland; tel. (48 61) 852 8503; fax: (48 61) 852 0532; e-mail kasiak@ibch.poznan.pl

Abbreviations: MD, molecular dynamic; FTIR, Fourier Transform Infrared Spectroscopy; RDF, radical distribution functions.

lecular hydrogen bonds [1]. These thermodynamically stable and more structured than bulk water, hydration shells, may have a leverage effect on other parts of a solute molecule and thus facilitate, e.g. in the instance of biomolecules, some biologically important conformational transitions.

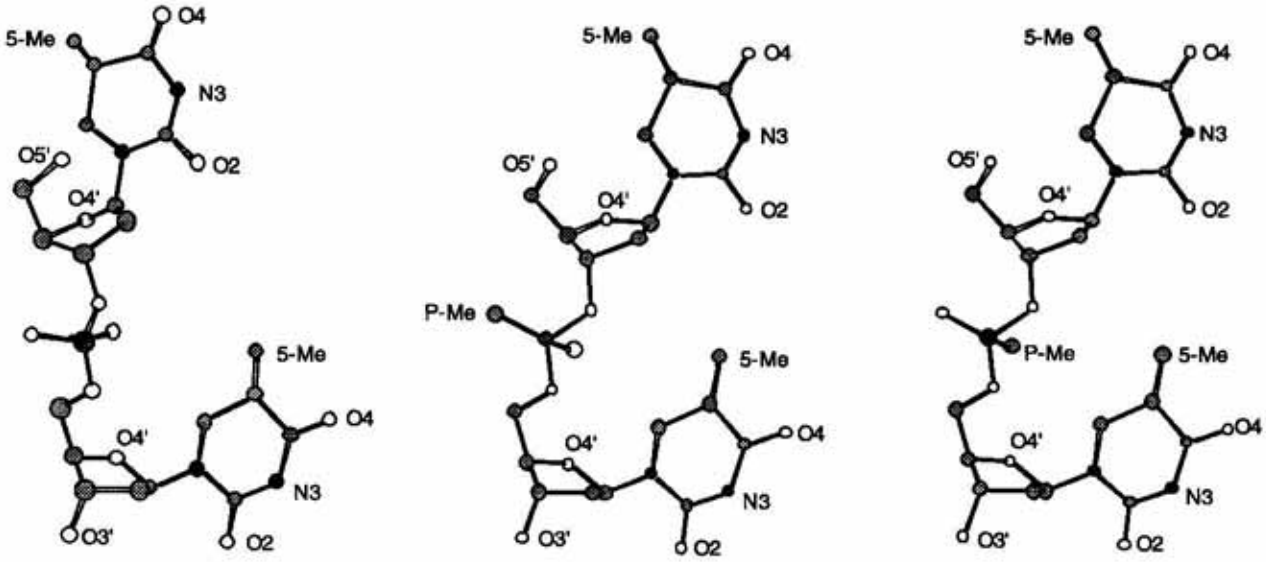
By the use of experimental and theoretical methods it has been proved that C–H groups may form stable hydrogen bonds with suitable acceptor sites [2]. C–H...Ow interactions have been found in various crystal structures of proteins, nucleic acids and in low-molecular hydrates [3]. In aqueous solutions, such interactions have been observed in short parallel-stranded duplexes formed by d(TCGA) oligonucleotides at acidic pH [4]. The C–H...Ow interactions have also been recognized as significant factors stabilizing specific conformations of nucleic acids [5].

Recently, we have analysed, using FTIR spectroscopy [6] and MD simulation [7], the hydration pattern of dithymidine phosphate (**1**) and two dithymidine methylphosphonate diastereomers (**2a** and **2b**) to get a deeper insight how chirality and charge density at the phosphorus center may affect water distribution around various parts of these compounds.

In this paper we will discuss some preliminary data concerning the interaction of water molecules with three types of hydrophobic centers, namely, the phosphorus-bound methyl group (P–Me in **2a** and **2b**), the methyl group at the C5 position in the thymine residues (C5–Me in **1**, **2a** and **2b**), and the hydrogen atom attached to the carbon C6 (C6–H centers in **1**, **2a** and **2b**).

COMPUTATIONAL METHODS

The crystal structure of the 4-thio analogue of **2b** [8] was used as the starting coordinates for simulations. To model the nucleotide analogues, standard Amber force field [9] was applied. All the MD simulations were carried out using the general-purpose simulation package M. Dynamix [10], essentially as it was described previously [7]. The double-time



Scheme 1.

step method of Tuckerman was applied with 0.2 fs time step to describe the fast motional modes, such as bond stretching and angular bending. Torsional and long-distance non-bonded motions were integrated after each 1.0 fs. Simulations were carried out in NPT ensemble using periodic boundaries and minimum image convention.

Simulations were started from a cubic cell containing 256 water molecules, equilibrated prior to dissolving the solute. Overlapping water molecules were removed and the solution containing the single nucleotide residue and the remaining water molecules was equilibrated for 100 ps. The production was extended to 500 ps in all simulation runs. The atom-atom radial distribution function g(*ij*, *r*) values (RDF) were calculated during the simulation.

RESULTS AND DISCUSSION

Distribution of water molecules around the P–Me groups in methylphosphonates 2a and 2b

The arrangement of water molecules around the methyl groups P–Me in R_P and S_P methylphosphonate diastereomers **2a** and **2b** was

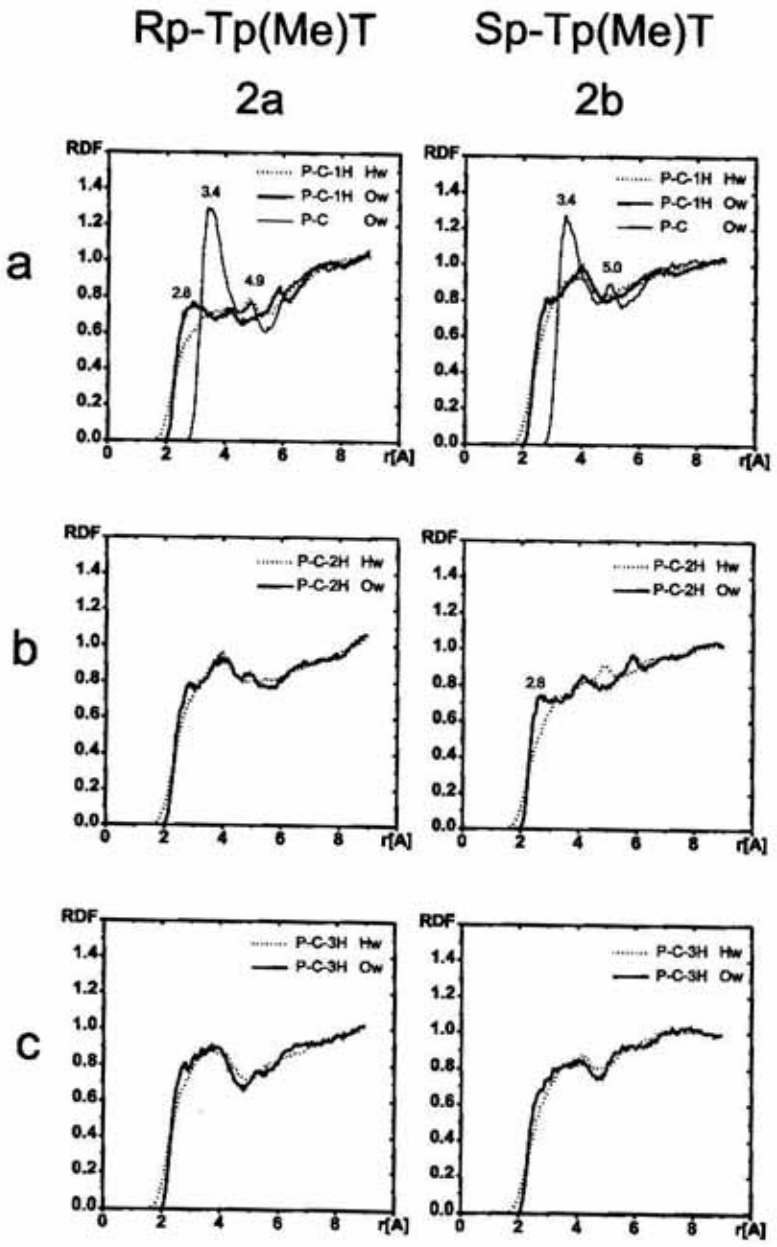


Figure 1. The radial distribution functions of water molecules around the P–Me centers in 2a and 2b.

The RDF describes the spatial organisation of atoms *j* around a central atom *i* in terms of the relative probability of finding a pair of atoms *ij* at the distance *r* apart and that of r_0, expected for a completely random distribution at the same density. Additional information can be gathered from an integral of this function (called the running integral of the RDF), which gives the relative number of atoms *j* (and hence the number of molecules they belong to) in a sphere of radius *r* around the atom *i*.

O_w and H_w refer to the oxygen and hydrogen atoms of water molecules, while other symbols stand for the appropriate functional groups in **1**, **2a** and **2b**, as indicated in Scheme 1.

The notation 3′-yl and 5′-yl units refers to the position of a deoxyribose ring to which the internucleotide bond is attached.

probed using the radial distribution functions (RDF) g(P-C-H Hw) and g(P-C-H Ow), separately for each proton of the methyl groups (Fig. 1, panels a, b, c) and g(P-C Ow) functions (Fig. 1, panel a). The broadness and diffuse profiles of g(P-C-H Hw) and g(P-C-H Ow) indicate that the hydrophobic type of interactions prevails. In the instance of H1 in **2a** and H2 in **2b**, the maxima of the functions g(P-C-H Ow) (at 2.8 Å) run about 1 Å ahead of that of g(P-C-H Hw) and may indicate that some percentage of the P-methyl protons can be involved in hydrogen bondings of type P-C-H...Ow-Hw. Differences in shapes of these functions observed for particular hydrogens may suggest that rotation of the P-Me group is partly restricted, in the time scale used in the analysis, due to proximity of the C5-Me and C6-H groups (the *S*p diastereomer) or the 5′-CH2 groups (the *R*p diastereomer) in the thymidin-5′-yl units.

Interesting information can be drawn from the comparison of g(P-C-H Ow) and g(P-C Ow) functions (Fig. 1, panel a). Both functions describe the distribution of water molecules (namely their oxygen atoms) around the same methyl centre, and a distinct difference in the shape of the functions is observed. The explanation of this discrepancy may be the much higher rate of the methyl group rotation than of the reconfiguration of the water molecules in the hydration shell. The jump of methyl hydrogen to an other rotamer configuration causes the change of the distance between particular atoms Me(H) and Ow by approximately 1.8 Å. The profiles of the g(P-C Ow) functions (Fig. 1, panel a), with sharp peaks of high intensity at 3.4 Å and distinct minima, indicate the presence of hydration shells around the P-Me groups, in which water molecules are more ordered than in the bulk of the solvent. This is consistent with theories of aqueous solutions which predict the increased solvent-solvent interaction as the hydrophobicity of the solute increases [1].

The first hydration shells in **2a** and **2b**, in which hydrophobic interactions prevail, extend from 2.7 to 4.4 Å and statistically contain about 1.5 molecules of water (Table 1). These centers are thus as efficiently hydrated as the most active hydrophilic P=O centers in natural dithymidine nucleotide **1** (Table 1).

Table 1. The average distribution of water molecules in first hydration shells around the phosphorus center in 1, 2a and 2b

	TpT [1]			Rp-Tp(Me)T [2a]			Sp-Tp(Me)T [2b]		
	distance [Å]	max [Å]	H_2O number	distance [Å]	max [Å]	H_2O number	distance [Å]	max [Å]	H_2O number
P-O′	2.3-3.2	2.6	1.5	2.3-3.1	2.6	0.5	2.3-3.1	2.6	0.6
P-O″	2.3-3.5	2.6	1.6	-	-	-	-	-	-
P-Me	-	-	-	2.7-4.4	3.4	1.4	2.8-4.5	3.4	1.5
P-O5′	2.4-3.5	3.0	0.6	*	*	*	*	*	*
P-O3′	2.5-3.4	2.9	0.5	*	*	*	*	*	*

*No clearly resolved hydration shells.

Distribution of water molecules around the C5-Me centers in 1, 2a and 2b

The distribution of water molecules around the C5-Me groups in **1**, **2a** and **2b** are shown in Fig. 2 (for the 3′-yl units) and Fig. 3 (for the 5′-yl units). The expected differences in the hydration of the 3′-yl and 5′-yl units should be related to the position of both parts in dinucleotides with respect to the phosphate group, reflecting intrinsic hydration properties of

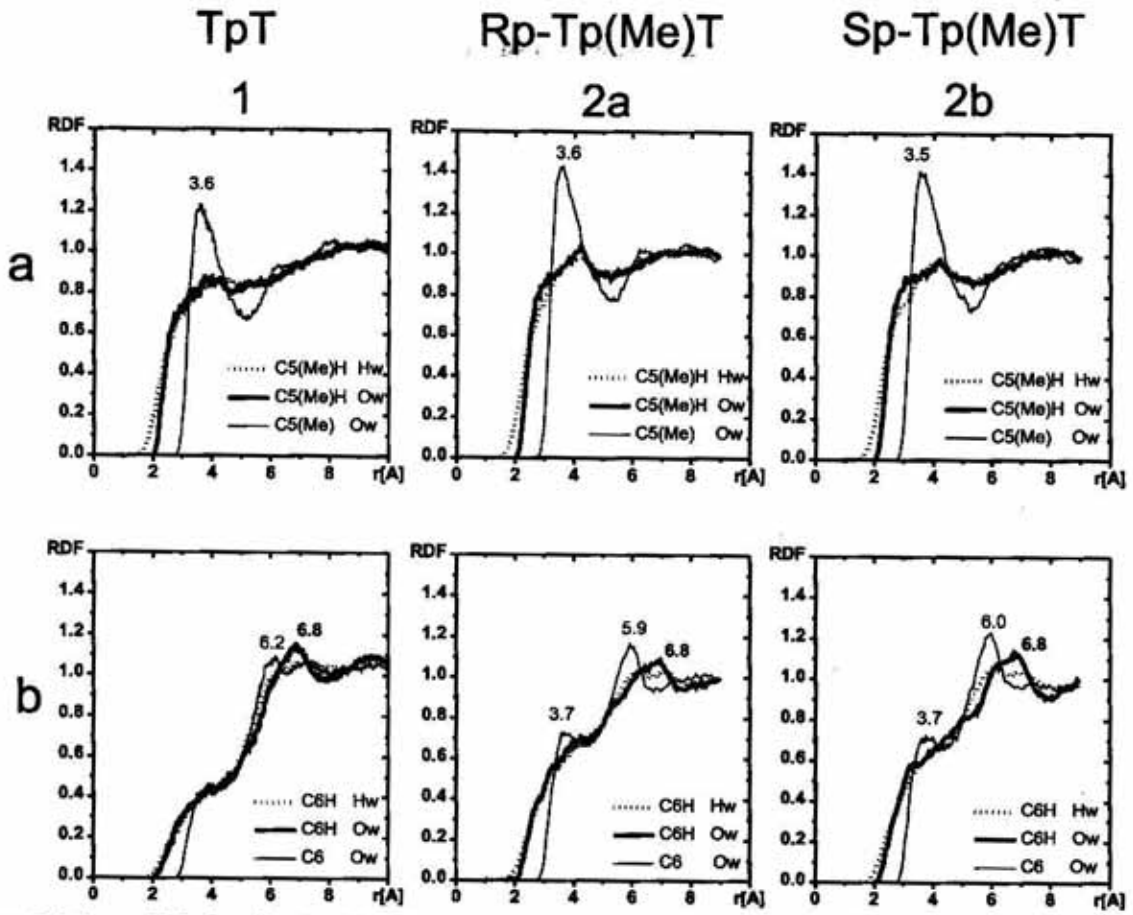


Figure 2. The radial distribution functions of water molecules around the C5–Me and C6–H groups of thymine bases in the 3′-yl units in 1, 2a and 2b.

In Figs. 2 and 3 the RDF for only one proton of each C5–Me group is shown, as shapes of these functions for the other protons appeared almost identical.

the base itself or properties influenced by the local arrangement due to the dinucleotide structure.

As inferred from the shapes of the functions g[C5–Me(H) Hw] and g[C5–Me(H) Ow][4] (Fig. 2 and 3, panels a) (overlapping of both functions and the lack of distinct minima) the interactions of the C5–Me groups with water molecules mainly have a hydrophobic character. Since the corresponding RDFs are almost identical for all protons attached to these methyl groups, it seems that these groups can freely rotate. Similarly as for the P–Me groups, the shapes of the functions g(C5–Me Ow) indicate that around C5–Me groups the hydrophobic hydration shells, centred at a distance of about 3.5 Å, are formed. For the natural dithymidine nucleotide **1**, these hydration shells (2.7–5.1 Å) contain statistically 1.9 molecules of water in the 3′-yl thymidine unit and 1.7 water molecules around the C5–Me group in the 5′-yl unit (Table 2). It is possible that the slightly less efficient hydration of the C5–Me groups in the 5′-yl thymidine unit can be due to proximity of the P=O group.

In both methylphosphonate analogues, **2a** and **2b**, the hydration shells around the C5–Me groups in the 3′-yl units are centred at 3.6 Å (similarly as for **1**) but they contain more water (statistically 2.2 molecules). Since the size of these hydration shells is almost identical in **1**, **2a** and **2b**, more water molecules accommodated around the C5–Me in **2a** and **2b** may indicate that, in the latter shells, water is more ordered. There are also noticeable differences between hydration of C5–Me groups in the diastereomers **2a** and **2b**. In the

Table 2. The average distribution of water molecules in first hydration shells around thymine centers in 1, 2a and 2b

		TpT [1]			Rp-Tp(Me)T [2a]			Sp-Tp(Me)T [2b]		
		distance [Å]	max [Å]	H_2O number	distance [Å]	max [Å]	H_2O number	distance [Å]	max [Å]	H_2O number
3'-yl unit	C2=O2	2.8–5.1	4.0	1.4	2.8–5.1	3.8	1.7	2.7–5.3	3.8	1.9
	N3H	2.7–5.2	3.5	1.7	2.7–4.4	3.7	1.3	2.8–4.3	3.7	1.1
	C4=O4	2.7–5.0	3.9	1.6	2.8–5.2	3.8	2.0	2.8–5.3	3.8	2.1
	C5Me	2.7–5.1	3.6	1.9	2.7–5.1	3.6	2.2	2.7–5.1	3.5	2.2
	C6	2.8–4.1	3.9	0.4	2.8–4.4	3.7	0.8	2.8–4.4	3.8	0.9
5'-yl unit	C2=O2	2.8–4.8	3.9	1.0	2.8–5.2	3.8	1.8	2.7–4.7	3.9	1.4
	N3H	2.9–4.2	3.6	0.7	2.7–4.3	3.4	1.1	2.7–4.2	3.6	1.1
	C4=O4	2.8–4.7	3.9	0.9	2.8–5.5	3.8	2.2	2.8–4.9	3.8	1.7
	C5Me	2.8–5.2	3.5	1.7	2.7–5.1	3.5	1.7	2.7–5.1	3.5	2.3
	C6	2.8–4.1	3.5	0.4	2.8–4.1	3.5	0.4	2.8–4.6	3.7	1.1

5'-yl units the R_P diastereomer (**2a**) contains only 1.7 molecules of water in the hydration shell around the C5–Me group, while the S_P isomer (**2b**), 2.3 molecules. These differences in the hydration of the C5–Me groups can be due to chirality of the phosphorus center. In the R_P diastereomer (**2a**) the hydration of the C5–Me group can be impaired due to proximity of the P=O group, while in the S_P isomer (**2b**), proximity of two hydrophobic groups, the P–Me and C5–Me, may result in an increased ordering of water in the corresponding hydration shells.

Interaction of water molecules with the C6–H centers in 1, 2a and 2b

Crystallographic data of nucleic acids fragments provide evidence that C6–H centers in pyrimidines as well as C8–H centers in purines can be involved in the intramolecular hydrogen bonds (e.g. with CO5' group, distance about 3.1–3.3 Å) stabilizing nucleotides conformation. The contribution of this C–H...O5' bonding was found to be greater for 5' nucleotides in which the phosphate group attached to O5' makes this atom more electronegative [11].

Analysis of the crystal structure of the 4-thio analogue of **2b** [8] revealed that the O5'H group may form an intermolecular hydrogen bond with the C2O carbonyl center of the adjacent dimer molecule (O5'H...O2T = 1.7 Å and O5'...O2T = 2.7 Å). In this conformation the O5'H function can also be involved in an intramolecular hydrogen bond with C6–H group (O5'...H–C6). The distances O5'...C6 (3.1 Å) and O5'...H–C6 (2.2 Å) strongly suggest that the O5' group can form intramolecular O5'...H–C6 hydrogen bond and that this interaction probably stabilizes the intermolecular hydrogen bond O5'H...O2T in the crystal structure.

The molecular dynamics simulation (Fig. 4) indicates that, in aqueous solution, there is an equilibrium between two rotamers of the O5'H group in the investigated compounds. For the S_P diastereomer **2b**, about 66% of molecules retain the conformation similar to that in a crystal state for the 4-thio analogue [8], i.e. the hydrogen atom of the O5'H group is pointing "out" and thus enables the formation of an intramolecular hydrogen bond O5'...H–C6. In the rest of population the hydrogen atom of the O5'H group is pointing "in" and, in principle, may be involved in

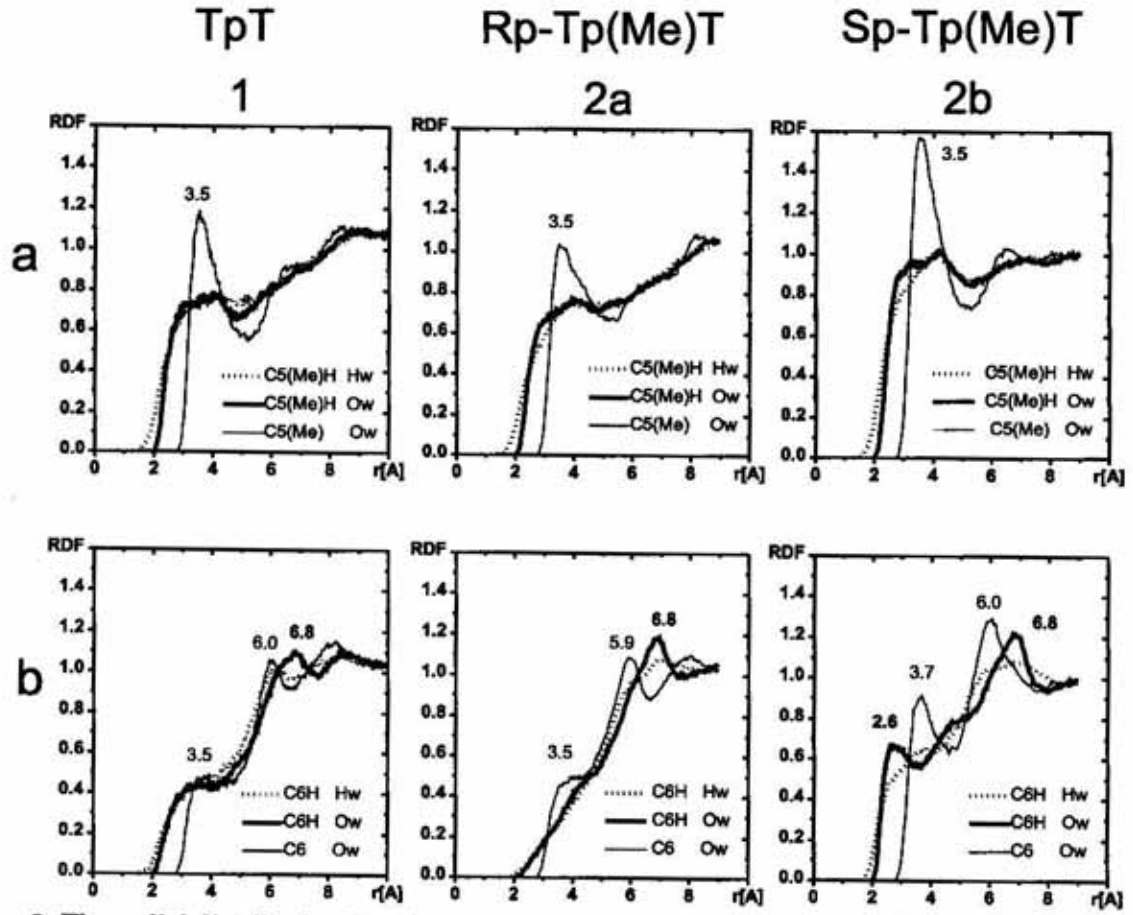


Figure 3. The radial distribution functions of water molecules around the C5–Me and C6–H groups of thymine bases in the 5′-yl units in 1, 2a and 2b.

intramolecular interaction with the donor oxygen O4′ (O5′H...O4′, 2.4 Å). Since the distance O5′...H–C6 remains unchanged during rotation around the C5′–O5′ bond, the C5′O group in this instance also can act as an acceptor for hydrogen bonding with the C6–H center. In the R_P diastereomer **2a** the distribution of these two conformers is more uniform (54% *vs* 46%).

The participation of C6–H groups in intramolecular hydrogen bonding is also apparent from the corresponding RDFs that probe distribution of water molecules around these centers (Figs. 2 and 3, panels b). In the natu-

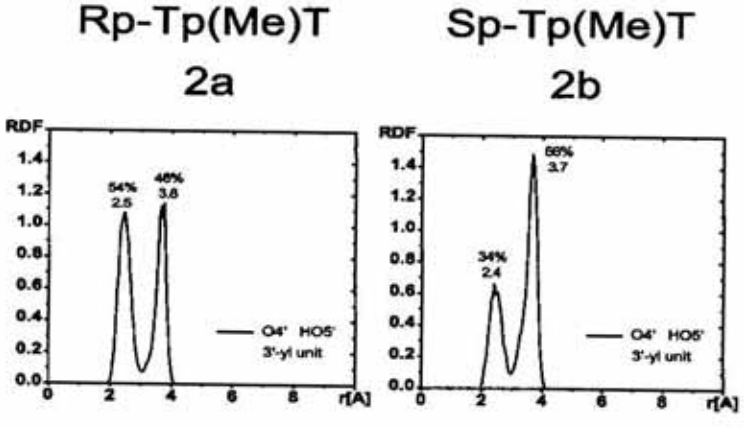


Figure 4. The radial distribution functions describing the intramolecular spatial organisation of the O4′ HO5′ atoms in the 3′-yl units in 2a and 2b.

ral dithymidine nucleotide (**1**), as inferred from the shapes of the functions g(C6-H Hw), g(C6-H Ow) and g(C6 Ow), the C6-H groups are poorly hydrated and their interactions with water molecules mainly have a hydrophobic character. The first hydration shells (2.8-4.1 Å) contain statistically 0.4 molecule of water both in the 3′- and 5′-thymidine units (Table 2). In both methylphosphonate analogues, **2a** and **2b**, the hydration shells around the C6-H groups in the 3′-yl units, similarly as for the C5-Me groups, contain more water (statistically 0.8 molecule) than those in **1**. Noticeable differences between hydration of C6-H groups in both methylphosphonate analogues, **2a** and **2b**, are apparent in the 5′-yl units. In the R_P diastereomer (**2a**) the hydration of C6-H group is similar to that in the natural dithymidine nucleotide (**1**) (probably due to proximity of the P=O functions; statistically 0.4 water molecule in the first hydration shell), whereas in the S_P diastereomer **2b** water molecules form a rather well defined hydration shell of a hydrophilic character (statistically 1.1 water molecules in the first hydration shell). The C6-H group in the 5′-yl unit of the S_P diastereomer most likely forms C6H...Ow hydrogen bonds with well ordered water molecules trapped in the hydrophobic hydration shells between two methyl groups, the P-Me and the C5-Me.

In conclusion, these studies indicate that the methyl groups attached to the phosphorus center (P-Me) in **2a** and **2b** are hydrated by water molecules as efficiently as the hydrophilic P=O group in **1** (about 1.6 water molecules) and better than the neutral P=O functions in these compounds (about 0.5 molecule of water), although the nature of the hydration is different. The C5-Me centers of the 3′-yl units seem to be better hydrated in the methylphosphonate analogues **2a** and **2b** than in the natural dithymidine phosphate **1** and than other centers of the thymine bases in methylphosphonate analogues. Due to chirality of the phosphorus center, the C5-Me group of the 5′-yl unit in the S_P diastereomer **2b** coordinates more water (2.3 molecules) than that in the R_P diastereomer **2a** (1.7 molecules). Finally, the C6-H group in the 5′-yl unit of the S_P diastereomer **2b** exhibits a specific interaction with water, while in the other instances the hydrophobic hydration seems to prevail. This „opportunistic" hydrogen bonding in which the C6-H group is probably involved requires a stable microenvironment as it is too weak to retain a long living, stable bonding in dynamic environment.

REFERENCES

1. Hecht, D., Tadesse, L. & Walters, L. (1992) Defining hydrophobicity: Probing the structure of solute-induced hydration shells by Fourier Transform Infrared Spectroscopy. *J. Am. Chem. Soc.* **114**, 4336-4339.
2. Desiraju, G.R. (1996) The C-H...O hydrogen bond. Structural implications and supramolecular design. *Acc. Chem. Res.* **29**, 441-449.
3. Derewenda, Z.S., Lee, L. & Derewenda, U. (1995) The occurrence of C-H...O hydrogen bonds in proteins. *J. Mol. Biol.* **252**, 248-262.
4. Wang, Y. & Patel, D.J. (1994) Solution structure of the d(T-C-G-A) duplex at acidic pH. A parallel-stranded helix containing $C^+ \cdot C$, G·C and A·A pairs. *J. Mol. Biol.* **242**, 508-526.
5. Auffinger, P., LouiseMay, S. & Westhof, E. (1996) Hydration of C-H groups in tRNA. *Faraday Discuss.* **103**, 151-173.
6. Kulinska, K., Sarzynska, J., Szabó, T. & Stawinski, J. (1997) FTIR Study on nucleotide analogues. 1. Spectral characterization of dinucleoside methylphosphonates and dinucleoside 5′-methylenephosphonates in solution and in solid phase. *J. Biomol. Struct. Dynam.* **15**, 119-128.
7. Kulinska, K., Kulinski, T., Stawinski, J. & Laaksonen, A. (1998) Molecular dynamics

computer simulation study of nucleotide analogues. Comparison of the hydration pattern of dithymidine phosphate with those of dithymidine methylphosphonate diastereomers. *J. Biomol. Struct. Dynam.* **15**, 987-998.

8. Szabó, T., Noréus, D., Norrestam, R. & Stawinski, J. (1993) Molecular and crystal structure of *S*p-thymidin-3'-yl 4-thiothymidin-5'-yl methylphosphonate. *Nucleic Acids Res.* **21**, 3921-3926.

9. Weiner, S.J., Kollman, P.A., Nguyen, D.T. & Case, D.A. (1986) An all atom force field for simulation of proteins and nucleic acids. *J. Comput. Chem.* **7**, 230-252.

10. Lyubartsev, A. & Laaksonen, A. (1997) M. Dynamix User Manual. Stockholm University.

11. Saenger, W. (1984) *Principles of Nucleic Acid Structure;* pp. 80-81, Springer-Verlag, New York, Berlin.