Pure Appl. Chem., Vol. 71, No. 7, pp. 1285–1290, 1999.
Printed in Great Britain.

Aqueous solutions of pyrimidine nucleic acid bases. Solute–solvent interactions*

Wojciech Zielenkiewicz

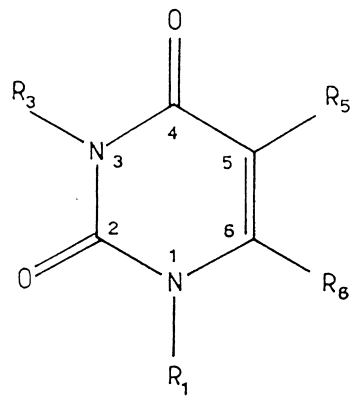
Institute of Physical Chemistry, Polish Academy of Sciences, 01–224 Warsaw, Kasprzaka 44152, Poland

Abstract: The data previously reported on the enthalpy of hydration, partial molar volumes and heat capacities of alkylated pyrimidine nucleic acid bases, e.g. uracil, thymine and cytosine, were correlated with the following structural parameters: molecular volume, volume of solvation shell, accessible molecular surface area and its atomic partition as well as and electrostatic potential charge. The simple additivity scheme of group contributions was also taken into consideration. The correlations obtained are presented and discussed. The results of the investigations clearly show that polar interactions are decisive in structuring the water around the solute molecules.

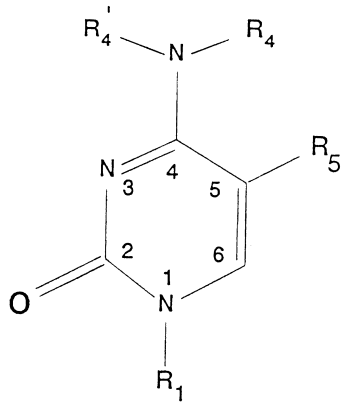
INTRODUCTION

The knowledge of the hydration scheme and hydration energy of individual nucleic acid bases is of fundamental importance for the explanation of the effect of the aqueous environment on bases pairing and stacking interactions and thus on the spatial organisation of polynucleotide chains in aqueous solutions.

Our interest is to explore to what extent the thermodynamic investigations of pyrimidine nucleic acid bases, e.g. uracil, thymine (5-methyluracil) and cytosine (Fig. 1) can be useful for the description of interactions of particular atoms of the skeleton of the base with its liquid environment, viz., water.



Uracil (R_1=R_3=R_5=R_6=H)



Cytosine (R_1=R_4=R'_4=R_5=H)

Fig. 1 Structural formulas of uracil and cytosine.

* Lecture presented at the 15th International Conference on Chemical Thermodynamics, Porto, Portugal, 26 July–1 August 1998, pp. 1167–1306.

MATERIALS AND METHODS

The method we have chosen is to screen a functional group on the skeleton of the base by methyl and others alkyl groups to see what happens when certain polar and apolar atoms are withdrawn from direct interaction with the hydration shell. This was the major reason to choose methylated and other alkylated derivatives of pyrimidine nucleic acid bases as objects of the study. An additional advantage of this approach is that methylation of nucleic acid bases plays an important role in the formation of biologically active conformations of nucleic acids leading in some cases to a modified specificity of nucleic acid molecules.

The experimental enthalpies of solution $\Delta_{sol}H_m$, enthalpies of sublimation $\Delta_{sub}H_m$, partial molar volumes V_2° and partial molar heat capacities $C^\circ_{p,2}$ reported earlier [1–19] are the basis of the present considerations. Investigations were made for the following compounds:

1. uracil ($C^\circ_{p,2}$ [1], V°_2 [2], ΔH°_{sol} [3]) **and its substituted derivatives: 2**. 1-methyl ($C^\circ_{p,2}$ [1], V°_2 [4], ΔH°_{sol} [5]); **3**. 3-methyl (V°_2 [6]; **4**. 5-methyl ($C^\circ_{p,2}$ [1], V°_2 [2], ΔH°_{sol} [12]); **5**. 6-methyl (V°_2 [6]); **6**. 1,5-dimethyl (V°_2 [6]); **7**. 1,6-dimethyl (V°_2 [6], ΔH°_{sol} [11]); **8**. 3,6-dimethyl (V°_2 [6]); **9**. 5,6-dimethyl (V°_2 [6]); **10**. 1,3-dimethyl ($C^\circ_{p,2}$ [1], V°_2 [4], ΔH°_{sol} [5]); **11**. 1,3,5-trimethyl ($C^\circ_{p,2}$ [1], V°_2 [6], ΔH°_{sol} [5]); **12**. 1,3,6-trimethyl ($C^\circ_{p,2}$ [1], V°_2 [6], ΔH°_{sol} [5]); **13**. 1,3,5,6-tetramethyl ($C^\circ_{p,2}$ [7], V°_2 [7], ΔH°_{sol} [8]); **14**. 5,6-trimethylene-1,3-dimethyl ($C^\circ_{p,2}$ [4], V°_2 [4], ΔH°_{sol} [9]); **15**. 5,6-tetramethylene-1,3-dimethyl ($C^\circ_{p,2}$ [4], V°_2 [4], ΔH°_{sol} [9]); **16**. 5,6-pentamethylene-1,3-dimethyl ($C^\circ_{p,2}$ [4], V°_2 [4]); **17**. 1,3-dimethyl-5-ethyl ($C^\circ_{p,2}$ [1],V°_2 [6], ΔH°_{sol} [9]); **18**. 1,3-diethyl-5-methyl ($C^\circ_{p,2}$ [4], V°_2 [4], ΔH°_{sol} [5]); **19**. 1,3-dimethyl-5-propyl ($C^\circ_{p,2}$ [7], V°_2 [6], ΔH°_{sol} [9]); **20**. 1,3-dimethyl-5-butyl ($C^\circ_{p,2}$ [7]), V°_2 [6], ΔH°_{sol} [9]); **21**. 1,3-dimethyl-6-ethyl ($C^\circ_{p,2}$ [10], V°_2 [10], ΔH°_{sol} [11]); **22**. 1,3-dimethyl-6-propyl ($C^\circ_{p,2}$ [10], V°_2 [10], ΔH°_{sol} [11]); **23**. 1,3-dimethyl-6-butyl ($C^\circ_{p,2}$ [10], V°_2 [10], ΔH°_{sol} [11]); **24**. 1,6-dimethyl-3-ethyl ($C^\circ_{p,2}$ [7], V°_2 [7], ΔH°_{sol} [8]); **25**. 1,6-dimethyl-3-propyl ($C^\circ_{p,2}$ [7], V°_2 [7], ΔH°_{sol} [8]); **26**. 1,6-dimethyl-3-butyl ($C^\circ_{p,2}$ [7], V°_2 [7], ΔH°_{sol} [8]); **27**. cytosine ($C^\circ_{p,2}$ [13], V°_2 [13], ΔH°_{sol} [14]) **and its substituted derivatives: 28**. 1-methyl ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]); **29**. 1,5-dimethyl ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]); **30**. 1,N^4-dimethyl ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]); **31**. 1,N^4,N^4-trimethyl ($C^\circ_{p,2}$ [15], V°_2 [15]); **32**. 1,5,N^4-trimethyl ($C^\circ_{p,2}$ [7]), V°_2 [7], ΔH°_{sol} [14]); **33**. 1,N^4-dimethyl-5-ethyl ($C^\circ_{p,2}$ [7], V°_2 [7], ΔH°_{sol} [14]); **34**. 1,4-dimethyl-5-propyl ($C^\circ_{p,2}$ [7], V°_2 [7]); **35**. 1,N^4-dimethyl-5-butyl ($C^\circ_{p,2}$ [7], V°_2 [7]); **36**. 1-methyl-N^4-hydroxy ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]); **37**. 1,5-dimethyl-N^4-hydroxy ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]); **38**. 1-methyl-N^4-methoxy ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]); **39**. 1,5-dimethyl-N^4-methoxy ($C^\circ_{p,2}$ [15], V°_2 [15], ΔH°_{sol} [14]);

In a search for a molecular model of interactions of pyrimidine bases with water, these thermodynamic data were correlated with the structural parameters [6,7,14,20,21] like molecular volume V^M, volume of solvation shell $V_{1,solv}$, relative density of solvation shell α, and accessible molecular surface areas S^M of the compounds studied. The atomic partition of the O, N and H atoms in S^M and the polarity P defined as the ratio of the molecular surface of the polar atoms and group exposed to the solvent to the total molecular surface area S^M of the molecule studied were evaluated. Electrostatic potential charges were calculated by a semi-empirical quantum-mechanical MNDO method [22].

DISCUSSION AND RESULTS

The enthalpies of hydration $\Delta_{hydr}H_m$, of methylated uracils **2–26** and methylated cytosine **28–32** are seen to vary not only with the number of CH_3 groups but also with the position of substitution on the ring. Substitution of the hydrogen atom by a methyl group on the C(5) or C(6) ring carbon atom of the skeleton, i.e. on the apolar side of the uracil ring, results in contribution to the enthalpy of hydration the average value of which is equal to 9.8 kJ/mol [5], and is practically equal to that obtained for aliphatic hydrocarbons [23]. The contribution of methyl group at an amide N-ring nitrogen N(1), N(3) on the uracil skeleton brings about reduction of $\Delta_{hydr}H_m$ with the mean increment close to -2 kJ/mol. This means, there exists evident difference of methyl substitution on polar and the apolar side of the uracil skeleton.

To explain the effects of substitution on the $\Delta_{hydr}H_m$ in terms of perturbations caused by the substituents according to the hydration scheme of uracil and cytosine, changes were analysed in accessible molecular surface areas of the molecules studied and their particular atoms. The existing differences in the series of methylated uracils can be seen in Fig. 2. For example, substitution of hydrogen by the CH_3 group on N(1) or N(3) polar atoms of the uracil skeleton reduces by half the molecular surface areas of the N atoms substituted per CH_3 groups (6.7 $Å^2 \rightarrow$ 3.6 $Å^2$) as well as reduces to some extent the surface area of the neighbouring atom O(2) (15.7 $\rightarrow$ 14.3 $Å^2$) in the case of **3**. In this case the electronic charge of N atoms drops from –0.31 to 0. However, substitution of the CH3 group on C(5) or C(6) atoms of the uracil skeleton causes rather inconsiderable changes in the surface accessible areas of the immediate neighbouring atoms: in the case of 4 the accessible surface area of O(4) changes from 16.0 $Å^2$ to 14.4 $Å^2$; in the case of **5** the surface of N(1) changes from 6.7 to 5.8 $Å^2$. The charges of the polar atoms remain unaffected.

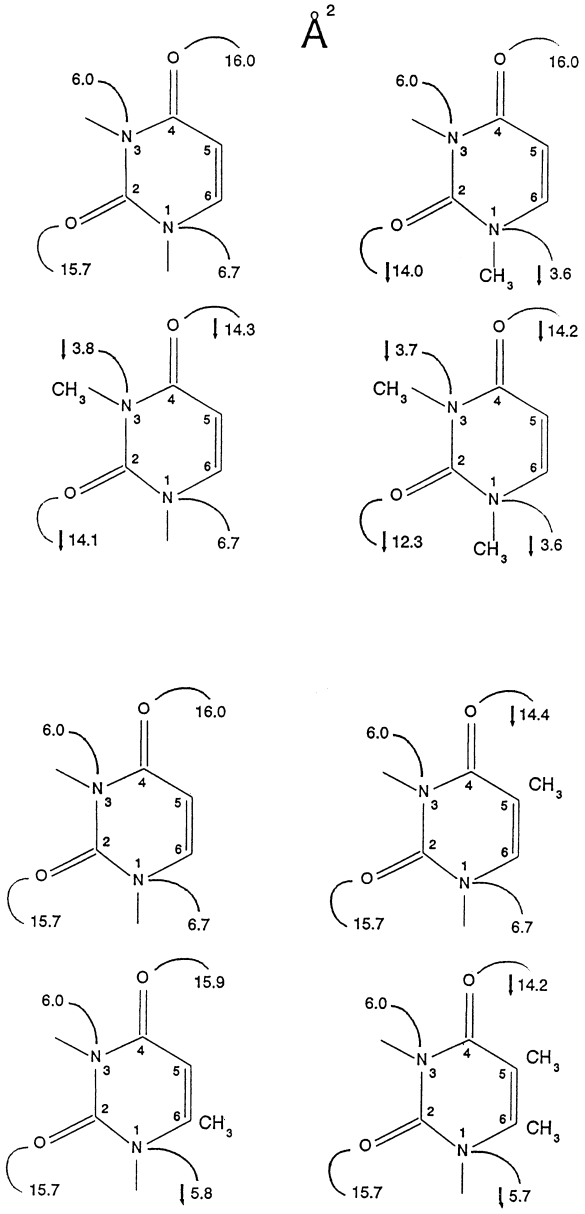


Fig. 2 Changes in the surface area and electronic charge caused by substitution of hydrogen by CH_3 groups on polar N(1), N(3) and apolar atoms of uracil skeleton.

Taking into calculations three parameters: the molecular surface of accessible area, the electron charge and the number of CH_3, the correlation with the corresponding values of the enthalpy of hydration was obtained (with a regression coefficient of 0.844). That means that polar interactions are mainly responsible for changes the in $\Delta_{hydr}H_m$ upon substitution.

In the case of methylated cytosines [14], changes in the increments of $\Delta_{hydr}H_m$ are also correlated a with the corresponding changes in the molecular surface area. Substitution of the hydrogen atom by a CH_3 group on C(5) of the skeleton causes a change in the accessible molecular surface area by about 1.4 $Å^2$, whereas substitution of the hydrogen on N^1 by CH_3 diminishes the surface by 4.5 $Å^2$. The changes in the surface area following the substitution on N^4 atom are considerably larger: substitution of a hydrogen atom by OH increases the surface area by 6.7 $Å^2$, whereas replacement of the hydrogen on OH by CH_3 diminishes the surface by 9.6 $Å^2$. The $\Delta_{hydr}H_m$ data are too limited to permit general conclusions about all the differently substituted cytosine derivatives.

An increase in the number of uracil and cytosine derivatives by elongation of alkyl side chains or the investigations of cyclooligomethylenouracils **14–16** did not provide direct information about changes that accompanied screening of certain atoms in the diketopyrimidine skeleton. In the case of the series of 1,3-dimethyl-5-alkyluracils (**6,17,19,20**) [5,6,19], 1,6-dimethyl-3-alkyluracils (**7,24–26**) [8,11] and 1,3-dimethyl-6-alkyluracils (**10,21–23**) [5,11] erratic changes were observed to occur in the values of the enthalpy of hydration. The higher values of the enthalpy increment correspond to the odd number of CH_3 groups added while the lower values correspond to those with the even $^nCH_2$ number. We were seeking for the reason of this phenomenon in the structure of the solid compounds because the value of the enthalpy of hydration is influenced, first of all, by the value of the enthalpy of sublimation. X-Ray studies for one series of alkyluracils (1,6-dimethyl-3-alkyluracils) have shown [24] the calculated crystal packing energies to reproduce quite well the experimental values of the enthalpy of sublimation.

The experimental enthalpies of hydration have been applied [9] to describe the hydration process by evaluation of the enthalpy of interaction determined as the enthalpy of solvation (hydration) corrected by the term related to the energy required to make a cavity in liquid water that would accommodate a solute molecule. However, the numerical values of the enthalpy of interaction depend [25] on the choice of the method of evaluation of the energy required to make a cavity in liquid water.

Analysis of the data of partial molar volumes V°_2 and heat capacities $C^\circ_{p,2}$ can enrich our previous considerations. Among other things the general additivity scheme [26] was used and contributions to the partial molar values examined were calculated according to the general formula:

$$X = X_0 + \Sigma_i n_i Z_i$$

where X_0 is a constant; Z_i is the additive value for group i and n_i is the number of type i groups. For cytosine derivatives four types of contributions were distinguished: $Z_{CH2(C)}$, $Z_{CH2(N)}$, $Z_{CH2(O)}$, and Z_O which correspond to the substitution of hydrogen on C, N and O (in OH on N4) atoms by CH_3 group and to the replacement of hydrogen on the N4 atom by an OH group. for uracil and thymine derivatives $Z_{CH2(N)}$ and $Z_{CH2(O)}$ were evaluated. Occasionally, the increments $Z'_{CH2(C)}$ and $Z''_{CH2(C)}$ were distinguished in $Z_{CH2(C)}$, corresponding, respectively, to the substitution of the hydrogen atom by CH_3 on the skeleton and to the extension of an alkyl side chain by the methylene group. Values of X_0 and Z_i were estimated by using standard multiple linear regression routine based on least squares. Results of calculations of X_0 and Z_i and regression factor r (Table 1) allow us to formulate the following observations and conclusions: (a) similarly as in the case of enthalpy of hydration, the increment in partial molar values V°_2 and $C^\circ_{p,2}$ depend strongly on the position of substitution of CH_3 groups; (b) the regression coefficient demostrates the reliability of correlation; (c) the constrains of the CH_2-motion due to cyclization in **14–16** influence the results of correlations, similarly as uracil derivatives of elongated alkyl side chains (compounds **17–26**).

Partial molar volume data were interpreted in terms of the new model of interaction of the solute with the solvent molecules [20,21] based on the assumption of the existence of a relationship between the volumetric properties and the solute structure. The α-parameter was calculated, corresponding to the ratio of V^M-V°_2 to the $V_{1,solv}$, defined as the relative density of the solvation shell, to compare the uracil and cytosine derivatives of different structure and polarity P. The dependencies were found to be linear (Table 2).

Table 1 Contributions of Zi for CH2(N), CH2(C), CH2(O) and O substituents to V°_2 and $C^{\circ}_{p,2}$

Compounds	X_0	$Z_{CH2(N)}$	$Z_{CH2(C)}$		$Z_{CH2(O)}$	Z_O	r
			$Z'_{CH2(C)}$	$Z''_{CH2(C)}$			
Partial molar volumes V°_2 (cm^3/mol)							
1–39	74.7 ± 1.8	18.0 ± 0.9	13.0 ± 1.1	15.7 ± 0.7	21.8 ± 3.7	6.1 ± 2.9	0.987
1–26	75.1 ± 4.2	18.4 ± 1.4	11.9 ± 1.5	15.6 ± 0.9			0.995
1–13,17–26	71.5 ± 1.1	18.5 ± 0.4	16.1 ± 1.5	16.4 ± 0.3			0.999
1–13,17–26	71.3 ± 1.1	18.5 ± 0.4	16.3 ± 0.2				0.999
27–39	74.0 ± 0.8	17.6 ± 0.3	16.5 ± 0.5	15.5 ± 0.3	21.8 ± 0.8	5.9 ± 0.7	0.999
27–39	72.4 ± 0.9	17.5 ± 0.4	15.9 ± 0.2		21.8 ± 0.9	6.0 ± 0.7	0.999
Partial molar heat capacities $C^{\circ}_{p,2}$ (J/K/mol)							
1,2,4,10–26	140 ± 47	88 ± 20	65 ± 8				0.910
1,2,4,10–13, **17–26**	137 ± 36	84 ± 15	76 ± 7				0.956
1,2,4,10–13	135 ± 9	86.2 ± 4.0	74.6 ± 4.6				0.996
27–39	165 ± 10	73 ± 5	91.6 ± 6.8	95.5 ± 4.0	101 ± 11	51 ± 9	0.998
27–39	164 ± 10	73 ± 4	94.2 ± 2.5		101 ± 10	50 ± 9	0.998

Table 2 The parameters α_0 and a of the relation $\alpha = \alpha_0 + aP$

Compounds	α_0	a	r	RSMD
1−39	−0.121 ± 0.002	0.104 ± 0.003	0.965	0.003
1−26	−0.121 ± 0.002	0.105 ± 0.004	0.965	0.002
27−39	−0.121 ± 0.002	0.018 ± 0.006	0.965	0.004

The resulting correlations bear out previous conclusion that, among the possible solute–solvent interactions, polar interactions are dominant. Structural parametrization was also used to interpret $C^{\circ}_{p,2}$ data. According to the concept of Spolar *et al*. [27] and the other by Freire *et al*. [28–30] the relation between $C^{\circ}_{p,2}$ data and the molecular surface areas of polar S_p and apolar S_n atoms of uracil derivatives was expressed by:

$$C^{\circ}_{p,2} = -604(\pm 25) + 8.1(\pm 1.8)Sp + 6.9(\pm 0.5)S_n, \qquad r = 0.974$$

The relationship between $C^{\circ}_{p,2}$ and S^M the molecular surface areas is expressed as:

$$C^{\circ}_{p,2} = -517(\pm 25) + 6.6(\pm 0.3)S^M, \qquad r = 0.912$$

Most interestingly, the volumetric properties V^M, V°_2 and $C^{\circ}_{p,2}$ were found to be interrelated.

$$C^{\circ}_{p,2} = -211(\pm 2.4) + 11.5(\pm 0.4)\beta, \qquad r = 0.975$$

The product $\alpha \cdot V_{1,solv} = \beta$ was applied, which in fact is equal to $V^{\circ}_2 - V^M$, represents a measure of the overall solute–solvent interactions [20] and is not related to the parametrization of the solvation shell.

ACKNOWLEDGEMENTS

This work was supported by the State Committee for Scientific Research under Project KBN 3T09A 05611.

REFERENCES

1 J. Szemińska, W. Zielenkiewicz, K. L. Wierzchowski. *Biophys. Chem.* **10**, 409 (1979).

2 N. Kishore, R. Bhat, J. C. Ahluwalia. *Biophys. Chem.* **33**, 227 (1989).

3 M. V. Kilday. *J. Res. Natl. Bur. Stand. (US)* **83**, 547 (1978).

4 W. Zielenkiewicz, A. Zielenkiewicz, J.-R. E. Grolier, A. H. Roux, G. Roux-Desgranges. *J. Solution Chem.* **21**, 1 (1992).

5 A. B. Teplitsky, I. K. Yanson, O. T. Glukhova, A. Zielenkiewicz, W. Zielenkiewicz. *Biophys. Chem.* **11**, 17 (1980).

6 W. Zielenkiewicz, J. Poznański, A. Zielenkiewicz. *J. Solution Chem.* **27**, 543 (1998).

7 A. Zielenkiewicz, G. Roux-Desgranges, A. H. Roux, J.-P. E. Grolier, K. L. Wierzchowski, W. Zielenkiewicz. *J. Solution Chem.* **22**, 907 (1993).

8 W. Zielenklewicz, A. Zielenkiewicz, K. L. Wierzchowski. *Pure Appl. Chem.* **66**, 503 (1994).

9 A. B. Teplitsky, O. T. Glukhova, L. F. Sukhodub, K. J. Yanson, A. Zielenkiewicz, W. Zielenkiewicz, J. Kosiński, K. L. Wierzchowski. *Biophys. Chem.* **15**, 139 (1982).

10 J.-P. E. Grolier, A. H. Roux, G. Roux-Desgranges. I. Tomaszkiewicz, W. Zielenkiewicz. *Thermochim. Acta* **176**, 141 (1991).

11 W. Zielenkiewicz, A. Zielenkiewicz, K. L. Wierzchowski. *J. Solution Chem.* **22**, 975 (1993).

12 M. V. Kilday, *J. Res. Natl. Bur. Stand. (US)* **83**, 529 (1978).

13 A. Zielenkiewicz, K. Busserolles, G. Roux-Desgranges, A. H. Roux, J.-R. E. Grolier, W. Zielenkiewicz. *J. Solution Chem.* **24**, 623 (1995).

14 A. Zielenkiewicz, M. Wszelaka-Rylik, J. Poznański, W. Zielenkiewicz. *J. Solution Chem.* **27**, 235 (1998).

15 A. Zielenkiewicz, K. Busserolles, G. Roux-Desgranges, A. H. Roux, J.-P. E. Grolier, M. Dramifiski, A. Zgit-Wróblewska, J. Poznański, W. Zielenkiewiez, *J. Solution Chem.* **25**, 529 (1996).

16 M. Kamiński, W. Zielenklewicz. *J. Chem. Thermodynamics* **28**, 153 (1996).

17 M. Kamiński, unpublished data.

18 W. Zielenkiewicz. *J. Thermal Anal.* **45**, 761 (1995).

19 B. Verkin. I. K. Yanson, L. F. Sukhodub, A. B. Teplitsky. *Vzalinnodeistvija Blomolekuł*. Naukova Durnka, Kiev (1995).

20 W. Zielenkiewicz, J. Poznański. *J. Solution Chem.* **27**, 245 (1998).

21 W. Zielenkiewicz, J. Poznański. *J. Mol. Liquids* in press.

22 M. Dewar, E. Haselbach, *J. Am. Chem. Soc.* **85**, 590 (1970).

23 S. J. Gill, N. F. Nichols, I. Wadsö, *J. Chem. Thermodyn.* **8**, 445 (1976).

24 K. Suwin'ska. *Acta Cryst. B* **51**, 248 (1995).

25 W. Zielenkiewicz, P. Zielenkiewicz, P. V. Lapshov. *J. Thermal Anal.* **54**, 775 (1995).

26 I. Cabani, P. Clianni, V. Mollica, L. Lepori. *J. Solution Chem.* **10**, 563 (1991).

27 R. S. Spolar, J. R. Livingstone, M. T. Record Jr. *Biochemistry* **31**, 3947 (1992).

28 K. P. Murphy, E. Freire, Y. Paterson. *Proteins* **21**, 83 (1995).

29 D. Xie, E. Freire. *Proteins* **19**, 291 (1994).

30 K. P. Murphy, D. Xie, K. C. Garcia, L. M. Amzel, E. Freire. *Proteins* **15**, 113 (1993).