ELSEVIER

Thermochimica Acta 291 (1997) 141–153

thermochimica acta

Prediction and structural analysis of the enthalpy of ionization of proteins

Boris P. Atanasov, Maria A. Miteva*

Biophysical Chemistry Laboratory, Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, BG-1113 Sofia, Bulgaria

Received 23 October 1995; accepted 24 June 1996

Abstract

A theoretical model is proposed for the prediction and analysis of the acid–base calorimetric titration of proteins. The model is based on the inclusion of the additive calorimetric contributions in a semi-empirical electrostatic method. Any electrostatic approach predicting pK_a values can be used for the analysis of calorimetric titration curves. The first step in the treatment is to find relationships between the ionization enthalpies of titratable amino acid residues and the relative solvent accessibilities of their ionizing atoms (AA_i). This is achieved on the basis of relations between the experimental values of enthalpies of the ionization of appropriate model compounds in aqueous organic solutions and their dielectric permeabilities. The predicted calorimetric titration curves of myoglobin, cytochrome *c*, ribonuclease *A*, lysozyme and α-chymotrypsin are compared with the available experimental data. Our results describe qualitatively the calorimetric titration of the first four proteins, while assuming a possible artifact in an experimental lysozyme calorimetric titration and predict the titration curve of α-chymotrypsin. This paper also presents the development of an analysis of the differential calorimetric titration curves, which can describe the contributions of individual ionizable groups. Such an analysis is demonstrated for the case of ferri- or ferro-cytochrome *c* as an example. © 1997 Elsevier Science B.V.

Keywords: Enthalpy of ionization; Calorimetric titration; Electrostatic interactions; Structure analysis

1. Introduction

Calorimetry is one of the most direct methods of biophysical molecular studies in the thermodynamics of biopolymers [1,2]. However, the integral character of calorimetric data restricts the amount of information that can be obtained at the atomic level. In order to overcome this difficulty, a number of models, based on the additivity of thermodynamic functions, have been developed [3–5]. However, most of these models describe mainly the temperature-dependent properties obtained by the method of adiabatic scanning calorimetry, i.e. measuring the temperature-dependence of heat capacity. This method usually evaluates the changes in protein stability with respect to pH [6], which reflect changes in the electrostatic terms of the free energies. A simplification is possible when measurements are carried out by this method in a narrow pH range. In this case, the heat capacity (C_p) is pH independent, which permits a unification of the C_p values. However, it is shown that the C_p values of a number of models of ionizable groups are pH-dependent in wide pH-intervals. Proteins, which

*Corresponding author. Fax: +359-2 700225; e-mail: mmiteva@bgcict.acad.bg.

PII S0040-6031(96)03075-4

include a number of ionizable groups, have also pH-dependent heat capacities as an intrinsic property.

Meanwhile, numerous calorimetric titrations of proteins and their complexes have been performed using batch and flow calorimetry [7–12]. A new sensitive instrument has been built for calorimetric measurements [13] specially adapted for any type of titration experiments, including acid–base titrations. In the case of acid–base calorimetric titration, an apparent quantity is the pH-dependent enthalpy (heat of ionization, ΔH_{ion}). The measured ΔH_{ion} values consist of enthalpy contributions of all ionizations in the given pH range. The structural dependence of the ionization enthalpy requires knowledge of the electrostatic interactions in the investigated system.

In order to explain the electrostatic interactions in proteins, a number of theoretical, microscopic and macroscopic approaches have been developed [14–20]. Most of them describe the pH-dependent states or processes and can predict the potentiometric titration curves or include them for parameterization. The ionization of each acid–base group is accompanied by enthalpy changes which can be experimentally measured by calorimetric titration. Consequently, both potentiometric and calorimetric titration curves are naturally related.

The purpose of the present work is to develop an appropriate model for the prediction and analysis of calorimetric titration curves of proteins. Different electrostatic methods determining pH-dependences can be used for such analysis. However, the models utilizing microscopic and finite-difference approaches are very time consuming when a number of pH-dependent calculations are made. A comparison between the semi-empirical and the conventional methods reveals approximately the same results [21,22]. A fast and adaptive semi-empirical approach has been chosen in this work [23]. Such an approach, and a model based on it, have been tested on a number of proteins. Good correlation between the theoretical and experimental potentiometric titration curves was achieved for lysozyme, myoglobin, cytochrome *c* and ribonuclease *A*. Experimental calorimetric titrations for these proteins have been reported elsewhere [7,8,10–12,24,25]. The pH-dependence of the heat of ionization of proteins can be reasonably predicted by the model presented in this paper. The differences between the measured and the calculated ΔH_{ion}(pH) of these proteins are analyzed.

2. Methods

2.1. Method of electrostatic calculations

The electrostatic calculations are based on a semi-empirical approach described earlier [23,26]. As input data, the method requires: (1) the atomic coordinates of protein molecules from the Brookhaven Protein Data Bank (PDB) [27]; (2) the list of the intrinsic pK_a (pK_{int}) of all titratable groups [17]; and (3) the parameters a_k of the empirical electrostatic potential function $W(r_{ij})$ of pair- interactions:

$$W(r_{ij}) = a_k/(r_{ij})^k \tag{1}$$

where r_{ij} is the distance between the charges i and j. The values of the parameters a_k are chosen in such a way, that the potential function obtained should correspond to the experimental titration curve of the protein [23].

The effect of a charged multipole on the dissociation of a given ionogenic group i, is described using:

$$pK_i = pK_{int,i} + \Delta pK_i \tag{2}$$

where

$$\Delta pK_i = \left(-\tfrac{1}{2.3RT}\right) \sum_{j \neq i} \sigma_j z_j (pH)[W(r_{ij}) - C] \times (1 - SA_{ij}). \tag{3}$$

In Eq. (3) $\sigma_j = 1$ for basic groups and $\sigma_j = -1$ for acid groups; z_j is the charge of the ionizable group; C is the Debye–Hückel correction accounting for the ionic strength and SA_{ij} is the static average accessibility to the solvent as described by Lee and Richards [28].

An iterative procedure was applied to calculate pK_i values for the titratable groups according to Refs. [23,26]. Following Ref. [15], the influence of the solvation effects could be expressed as:

$$pK_{int}^{p} = pK_{int}^{w} + [\Delta G_s^{x-p}(X^-) - \Delta G_s^{w-p}(XH)]/(2.3RT) \tag{4}$$

$\Delta G_s^{w-p}(X-)$ and $\Delta G_s^{w-p}(XH)$ are the differences between the solvation energies in water and protein

mediums, of unprotonated and protonated forms, respectively. A linear relationship between ΔG_s^{w-p}, pK_a, and normalized atomic accessibility (AA_i) is proposed [23]:

$$pK_{int}^{p} = pK_{int}^{w} - \sigma_i k(1 - AA_i) \quad (5)$$

where k is a coefficient which is liable for parameterization and it is assumed to be 1 in this work.

2.2. The calorimetric model

The main assumption in the present model is the proportionality between the degree of ionization and the heat of ionization ($\Delta h_{ion,i}$) of each titratable group i. A linear correlation has been found between the enthalpy and free energy of ionization for some carboxylic acids [29]. It is assumed that such a linear relationship could be applied for each other type j of ionizable groups in proteins in different solvents:

$$\Delta H_{ion,j} = A \times pK_j - B, \quad (6)$$

where A and B are different constants for the different types of ionizable groups. On the basis of the proportionality between $\Delta H_{ion,j}$ and pK_j, Eq. (5) can be multiplied by some constant k' giving:

$$\Delta H_{ion,j} = \Delta H^0_{ion,j} - k'(1 - AA_j) \quad (7)$$

where $\Delta H_{ion,j}$ is the enthalpy of ionization of the jth type titratable group in a protein, $\Delta H^0_{ion,j}$ is the enthalpy of ionization of the same group in aqueous solution. Therefore:

$$-\Delta H_{ion,j} = a.AA_j + b \quad (8)$$

where $a = -k'$ and $b = -\Delta H^0_{ion,j} + k'$. The linear relationship obtained between $\Delta H_{ion,j}$ and AA_j, is a reasonable assumption. Such linear relationships were experimentally demonstrated for some atomic surfaces and their hydration energies [30].

In order to obtain the values of the slopes in the Eq. (8) for the different types of titratable amino acid residues, we have used the experimental data from the literature (see references in Table 1) for the heat of ionization of a number of model compounds in water organic solutions at different concentrations. The values of $\Delta H_{ion,i}$ of the model compounds depend

Table 1
Experimental data of ionization enthalpies of model compounds of amino acid residues

Residue	Model compound	Organic solvent	Molar part of the organic solvent (x_1)						Reference	R
			0.0	0.2	0.4	0.6	0.8	1.0		
1	2	3	4	5	6	7	8	9	10	11
Asp Glu CTR	Acetic acid	CH_3OH	(7.5) [a]	1.7	−1.2	—	—	—	[41]	0.981
His	Imidazole	CH_3OH	(38.9)	(35.9)	35.1	31.4	30.9	29.7	[42]	0.974
Cys	Glycodimercaptoacetate	CH_3OH	(28.9)	(23.8)	(18.0)	(15.5)	(13.0)	(10.9)	[43]	0.977
Tyr	Phenol	CH_3OH	23.0	22.6	20.9	19.7	18.8	18.0	[44,45]	0.990
Lys NTR	n-Butyl amine	CH_3COCH_3	58.6	36.4	25.1	19.7	12.6	10.5	[46]	0.942
Dielectric constant ε of aqueous-organic solutions at corresponding molar parts of the organic solvent										
		CH_3OH	80	63	52	43	37	33	[47]	
		CH_3COCH_3	80	55	40	34	25	22		

Columns 4–9 show values with opposite sign of the ionization enthalpies (kJ mol $^{-1}$) of the model compounds at given molar parts (0.0–1.0) of the organic solvent in aqueous-organic solution.
R – correlation coefficient obtained by linear regression analysis on the shown experimental values.
[a] – the values in parenthesis are averaged on the basis of the precise experimental values by linear regression analysis.

on the effective dielectric permeabilities (ε) of the local surroundings. The value of ε is frequently presented as a function $\varepsilon(R)$, where R is the radius-vector of the given charge to the centre of the molecule. Another approach, based on static accessibility (SA_i), can be applied to improve this approximate evaluation. Such an approach describes the relative exposure of the ionizable group to the solvent. The specificity of acid–base equilibria requires the use of more precise values, i.e. the relative atomic accessibilities (AA_i) of the ith proton binding site of each ionizable group. The contact of water molecules with the proton binding sites, rather than the entire amino acid residue, is important for the ionization process. The empirical correlations between the $\Delta H_{ion,j}$ of the jth type of model compounds of ionizable groups in proteins and AA_j are computed in two consecutive steps:

1. The experimental values of $\Delta H_{ion,j}$ for a model compound are plotted against the values of molar part (x_1) of organic solvents (mainly methanol) in aqueous-organic solutions (Table 1). The values of ε corresponding to the molar parts x_1 for the solutions are also taken from the literature (see values in Table 1). The correlation between $\Delta H_{ion,j}$ and ε was obtained by excluding x_1 values.
2. A relationship between AA_j and of ε the surroundings of the jth group, assumed on the basis of the reverse Bjerum plot [15], was used in the range AA_j=0.3–1.0 which corresponds to $\varepsilon = 35 - 80$.

The values of the parameter a in Eq. (8) were obtained by linear regression analysis of the experimental calorimetric data (shown in Table 1) as a function of AA_j. The correlation coefficients obtained for the linear dependencies corresponding to Eq. (8) are also presented in Table 1. The values of the parameter b in Eq. (8) were calculated on the basis of parameters a and the experimental values of ΔH^0_{ion} of amino acids in water (Table 2). The values of the numerical parameters a and b are presented in Table 2. The pK_a values of all arginine residues are very high and they have pH-dependent contributions in the highly alkaline pH range (usually > 12). In this pH range, the described model is not applicable. In such a case, a constant value of the ionization enthalpy of arginine ($\Delta H_{ion} = -56.9$ kJ mol^{-1}) was used.

Table 2
Ionization enthalpies and parameters of the correlation lines $-\Delta H_{ion,j} = a\,AA_j + b$ for titratable amino acid residues

Residue	$-\Delta H^0_{ion}$ [c]	Parameters	
		a	b
Asp	7.5	21.75	−14.25
Glu	1.7	21.75	−20.05
CTR			
His	29.7	9.24	20.46
Cys	28.9	17.84	11.06
Tyr	25.1	5.37	19.73
Lys	53.6	45.33	8.27
NTR	43.9	45.33	−1.43
Arg	56.9	0.00	56.9

a and b – The parameters from Eqs. (8) and (9).
c – Ionization enthalpies (kJ mol^{-1}) of amino acids in aqueous solution [48–53].

The model combines the pH dependent degree of dissociation, α_i, with ΔH_{ion} values with a negative sign at deprotonation [8]. Thus, the ionization enthalpy which depends on the AA_j of each ionogenic amino acid residue ($\Delta h_{ion,i}$), is calculated as a function of pH using:

$$\Delta h_{ion,i} = -\alpha_i(\mathrm{pH})(a\mathrm{AA}_j + b). \quad (9)$$

The ΔH_{ion}(pH) values for the whole protein are obtained as a sum of the ionization enthalpies $\Delta h_{ion,i}(\alpha_i)$ of each titratable group i in the computing process.

The model described here consists of two parts: 'electrostatic' and 'calorimetric'. It should be noted that the first one can be derived from any appropriate method capable of producing correct pK_a values. The total accuracy is mainly determined by the precision of the α_i values, as well as the uncertainty in the experimental $\Delta H_{ion,j}$ values (see Table 1). The errors due to the uncertainty of the degree of dissociation are up to 5%. The errors arising from the transformation of the experimental data for $\Delta H_{ion,j}(\varepsilon)$ to $\Delta H_{ion,j}(AA_j)$, because of the uncertainty of the slopes of correlation lines, are in the range of 6–8%. Thus, the total

inaccuracy of the model does not exceed 11–13%. It should be noted that these errors lie in the pH range in which conformational changes of protein structure do not occur and the model is acceptable. The main contribution to the calorimetric titration curve is provided by the relation $\Delta h_{\mathrm{ion},i}(\alpha_i)$, (see Fig. 6A, curve 2). The correction from $\Delta H_{\mathrm{ion},j}(\mathrm{AA}_j)$ is relatively small but includes the specificity of the protein structure (Fig. 6B, curve 1). Different types of groups participate with different relative errors, depending on the absolute values of their $\Delta H_{\mathrm{ion},j}$. The relative uncertainties in the measurements for carboxyl-containing residues (Asp, Glu, α-COOH terminal group (CTR)) are high as their absolute calorimetric contributions are small. In the cases of His, Tyr, Lys, and α-NH_2 terminal group (NTR), the situation is reversed. Therefore, the percentage of errors are smaller in the neutral and alkaline regions.

The developed model predicts both potentiometric Z(pH) and calorimetric ΔH_{ion}(pH) titration curves. In the case of Z(pH), all types of ionizable groups are only distinguishable by the sign of the charge Z. However, they have different group-specific contributions to ΔH_{ion} (Table 2). Therefore, an independently obtained acid–base calorimetric titration curve should be considered more informative with respect to the contributions of the individual residues than the potentiometric titration curve.

2.3. Analysis of the differential titration curves

The ionization of each ionizable group is represented by the logarithmic curve $(\log(\alpha/1-\alpha) = f(\mathrm{pH}))$ based on the Henderson–Hasselbach equation. The position of the maximum of $\delta\Delta h_{\mathrm{ion},i}$ in the pH scale appears at the middle of $\Delta \mathrm{p}K_{\mathrm{a},j}$. The value of $\delta\Delta h_{\mathrm{ion},i}$ is proportional to $\Delta \mathrm{p}K_{\mathrm{a},j}$ provided that $-1 \leq \mathrm{p}K_{\mathrm{a},i} \leq 1$ (which is the most usual case). Beyond this interval, $\delta\Delta h_{\mathrm{ion},i}$ tends to reach a maximal, constant value of $\Delta h_{\mathrm{ion},i}$. The pH-dependent contributions, $\delta\Delta h_{\mathrm{ion},i}$, of the individual ionizable groups with maximal $\Delta \mathrm{p}K_{\mathrm{a},i}$ can be evaluated using this model and then incorporated on the graph together with the total $\delta\Delta H_{\mathrm{ion}}$(pH)-curve. The comparison between $\delta\Delta h_{\mathrm{ion}}$(pH) of the individual group and $\delta\Delta H_{\mathrm{ion}}$(pH) of whole protein gives information about the residues, which are most influenced by change in pH.

3. Results

The experimental acid–base calorimetric titration curves of myoglobin, cytochrome c, ribonuclease A, lysozyme and α-chymotrypsin have been published elsewhere. These data were obtained at different starting points (pH) of the titration procedure. In order to compare all the theoretical and experimental curves, the experimental data measured in the direction from basic to acidic pH were considered to have an opposite sign compared to those obtained in the direction from acidic to basic pH. So, the enthalpy contributions of the deprotonation processes will be negative. The experimental calorimetric curves were fitted to the calculated curves with optimal overlapping in the acidic pH range. The acidic range was chosen because the absolute error values were lower, since the enthalpy contributions of deprotonation of the acidic groups were smaller than those of the basic groups. All the experimental data are shown in the figures as patterns of discrete circles. The predicted calorimetric titrations are presented as solid lines.

3.1. Myoglobin

The theoretical curve of the ionization enthalpy of sperm whale myoglobin was calculated with the help of its atomic coordinates (file "pdb4mbn.ent" [31] from PDB). The results are shown in Fig. 1 (curve 1). The experimental data for the ionization enthalpy of myoglobin [7] are also shown in Fig. 1 (curve 2). The theoretical and experimental curves show excellent agreement in a wide pH range, viz. 2.0–10.5. In the acid range, the enthalpy contributions are -145 kJ mol^{-1}. The ΔH_{ion} values in the neutral pH range vary from -190 to -750 kJ mol^{-1}, because of the large number of His. residues in myoglobin. The magnitude of the enthalpy contribution is -125 kJ mol^{-1} in the 10.0–11.5 pH range, respectively. The alkaline denaturation is associated with positive enthalpy contributions above pH 11.5 (curve 2).

3.2. Cytochrome c

The theoretical calorimetric curve of ferro-cytochrome c was computed using the atomic coordinates from X-ray structure of horse-heart cytochrome

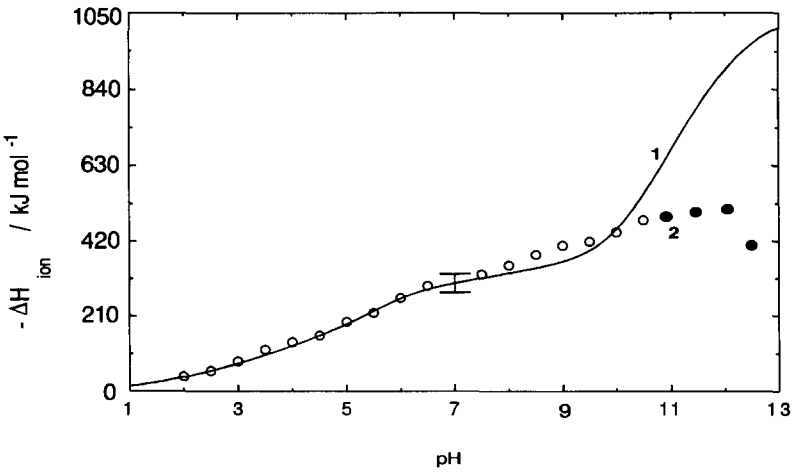

Fig. 1. The pH-dependence of the ionization enthalpy of myoglobin: (1) (solid line) calculated results; and (2) experimental results [7].

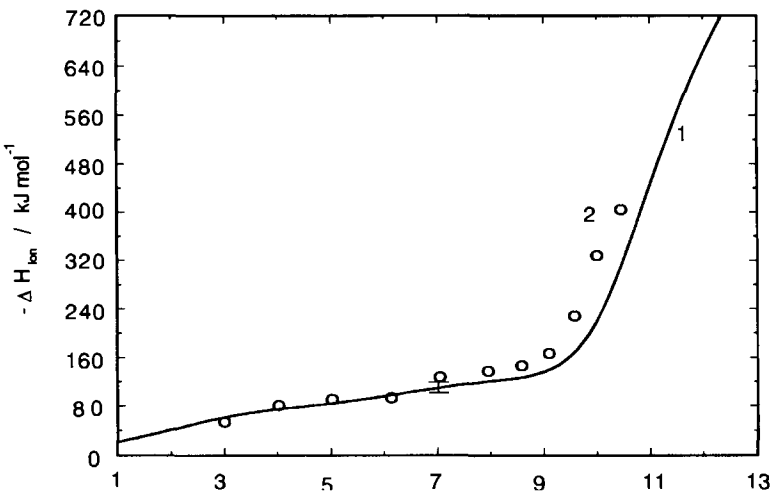

Fig. 2. The pH-dependence of the ionization enthalpy of horse heart ferro-cytochrome *c*: (1) (solid line) calculated results; and (2) experimental results [11].

c (files "pdb2pcb.ent" [32] and "pdblhrc.ent" [33] from PDB). Fig. 2 presents the obtained theoretical results (curve 1) compared with the experimental data [11,12] (curve 2). There is a good agreement between the theoretical and experimental results in the pH 3.0–8.5 range. The values of ΔH_{ion} change from −45 to −145 kJ mol^{-1} in the pH 2.0–8.5 range. The changes in the neutral range are minor, since the number of His. residues is two. There is an obvious difference between the experimental and theoretical curves above pH 8.5, which is significantly more than the uncertainties of the model. The values of ΔH_{ion} changed from −125 to −670 kJ mol^{-1} in the pH 9–12 range. The results obtained for ferri-cytochrome *c* were similar to that obtained for ferro-cytochrome *c*. The analysis of the ionization changes for the different

groups resulting from the appearance of a positive charge on the heme, Fe, in the ferro–ferri redox transition and the differential curves, calculated from the differences of calorimetric titrations of ferro- and ferri-cytochrome *c* (theoretical and experimental data), are presented in the discussion.

3.3. Ribonuclease A

Two sets of experimental data for calorimetric titration of ribonuclease *A* have been previously reported [8,10]. The coordinates of ribonuclease *A* (file "pdb7rsa.ent" [34] from PDB) were used in calculations of theoretical pH-dependent ΔH_{ion} values. The results are presented in Fig. 3, where curve 1 corresponds to an ionic strength I=0.1 and curve 3 corresponds to I=0. The curves of the ionization enthalpy of ribonuclease *A* can be divided into three parts. The first one is located in the acid pH 2–5 range and is characterized by relatively small changes in the ΔH_{ion} values. In the pH 5.5–8.0 range (where the titration of His. residues was carried out), the ΔH_{ion} values decreased from −146 to −230 kJ mol^{-1}. The values of ΔH_{ion} reached −545 kJ mol^{-1} in the 9.5–12.0 basic pH range. The corresponding experimental data are shown as curves 2 and 4, respectively [8,10]. The theoretical curves, 1 and 3, are in relatively good accordance with the experimental data obtained at an ionic strength I=0.15 (curve 2) [8], although they are not in agreement with the experimental data at a low ionic strength, I=0.05 (curve 4) [10]. The theoretical curves, 1 and 3, differ significantly for pH>10. The experimental data at the higher ionic strength (curve 2) are in relatively good accordance with curve 1. However, curve 4 significantly differs from curve 3 for pH>6. This difference is more than the expected uncertainty. Obviously, pH-dependent processes occur in this range.

3.4. Lysozyme

The calorimetric titration curve of lysozyme is also theoretically predicted. The crystallographic coordinates of lysozyme (file "pdb71yz.ent" [35] from PDB) were used in the calculations. The results are shown in Fig. 4 (curve 1). The values of ΔH_{ion} change from −38 to −92 kJ mol^{-1} in the acid range, with pK_{app} about 4. The changes in the neutral pH range are minor because the lysozyme has only one His. residue. There is a −545 kJ mol^{-1} enthalpy contribution with pK_{app}: 11.8 in the basic range. The experimental data [24,25] are also shown in Fig. 5, curves 3 and 2, respectively.

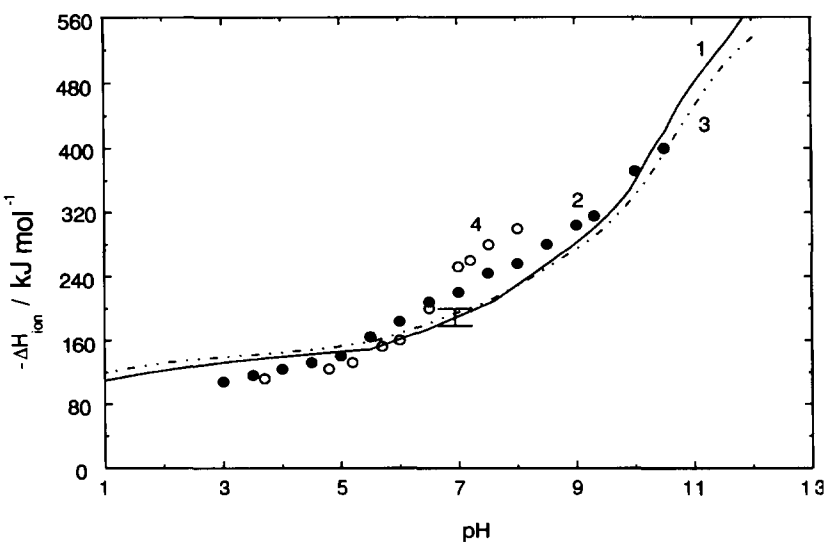

Fig. 3. The pH-dependence of the ionization enthalpy of ribonuclease *A*: (1) (solid line) calculated results, 0.15 M ionic strength; (2) experimental results [8], 0.15 M ionic strength; (3) (dotted line) calculated results, 0.05 M ionic strength; and (4) experimental results [10], 0.05 M ionic strength.

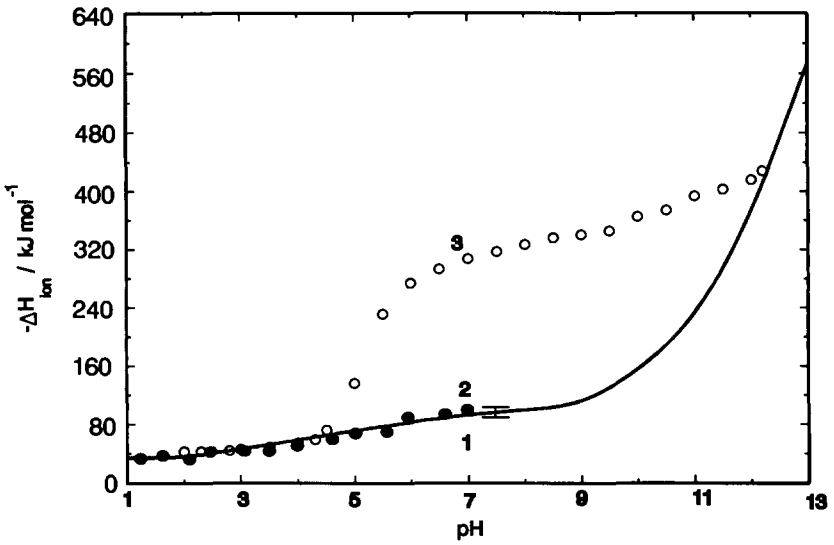

Fig. 4. The pH-dependence of the ionization enthalpy of lysozyme: (1) (solid line) calculated results; (2) experimental results [25]; and (3) experimental results [24].

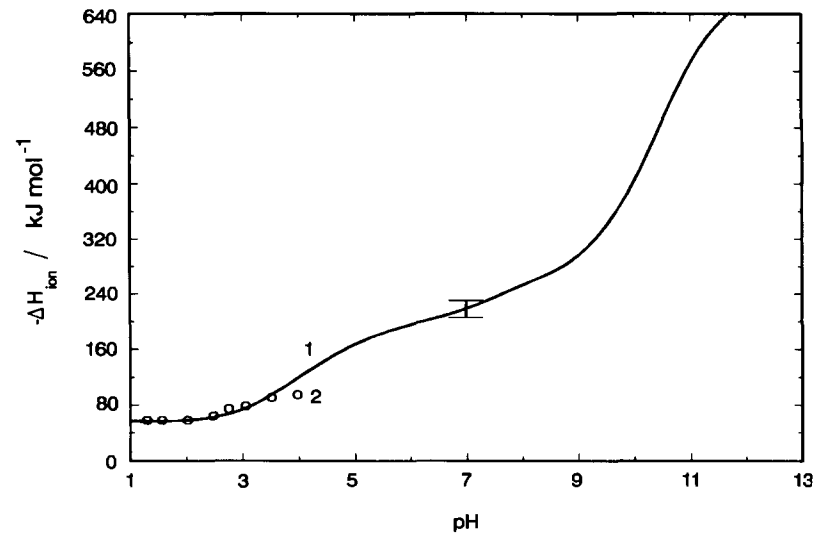

Fig. 5. The pH-dependence of the ionization enthalpy of α-chimotrypsin: (1) (solid line) calculated results; and (2) experimental results [9].

The theoretical curve (curve 1) is in good accordance with curve 2 and the part of curve 3 in the pH 2.0–4.4 range. However, there is a large difference between the curves 1 and 3 in the 4.5–12 pH range. The experimental curve (curve 3) reveals a complex sigmoidal transition with pK_{app}: 5.2 in the pH 4–7 range. The large discrepancy observed between the experimental data from elsewhere [24] (curve 3) and the theoretical curve cannot be explained by the dimerization of the molecules of lysozyme in the pH 4–7 range.

3.5. α-Chymotrypsin

The theoretical curve $\Delta H_{ion}=f(pH)$ of α-chymotrypsin was computed using the structure presented by atomic coordinates (file "pdb2cha.ent" [36] from PDB). The theoretical curve (curve 1) and experimental calorimetric data (curve 2) [9] are shown in Fig. 5. The calculated curve 1 consists of three regions. In the first, pH 1–4 range, the enthalpy contribution is 65 kJ mol^{-1}. In the second, pH 4.5–8 range, ΔH_{ion} values change from −145 to −250 kJ mol^{-1}. In the basic, pH 8.5–12.0 range, ΔH_{ion} values reach below −670 kJ mol^{-1}. Thus, the results obtained can be used for predicting experimental data.

4. Discussion

The potentiometric titration curve $Z(pH)$ is widely used as an integral characteristic of the electrostatic interactions in proteins. It can be predicted by a number of approaches based on finite difference techniques [19,20] or by utilizing semi-empirical methods [16]. However, the use of $Z(pH)$ provides an insufficient basis of data for construction of reasonable models for large complex systems such as protein molecules with multiple charges. Therefore, other integral characteristics are required for such purposes. In the present study, the acid–base calorimetric titration curve, $\Delta H_{ion}(pH)$ is analyzed in the same manner as the potentiometric titration curve $Z(pH)$. The values of $\Delta H_{ion,j}$ are specific for each type of groups j and they have different sensitivities to the solvent exposure of the ionizable sites. The introduction of a phase diagram, such as $Z/\Delta H_{ion}$ (excluding pH dependence from both the foregoing functions) in the semi-empirical calculations, will increase the accuracy of the developed electrostatic model.

The proportionality between the ΔH_{ion} and pK_a (both determined independently) [29] for the process of ionization in a given solvent reflects the enthalpy–entropy compensation phenomenon [37]. It is typical for processes in condensed phases (solute–solvent interaction) and is manifested very well in the case of proteins. Most probably, this relationship reflects the correlation between the ionization state of a given group and its hydration. Obviously all changes in the ionization state of acid–basic groups will alter the number and orientation of water dipoles, especially in the first hydration shells. Thus, a detailed analysis of the $\Delta H_{ion}(pH)$ curve gives an additional tool to understand the hydration reactions in proteins. Finally, it is obviously related to other more sensitive characteristics of the pH-dependent pre-hydration of charged protein surfaces, e.g. pH-dependent adiabatic compressibility.

The exploitation of the simplified relation $pK = pK^0_{int} - \sigma k(1 - AA_j)$ (see Eq. (5)) and its transformation into $\Delta H_{ion,j} = \Delta H^0_{ion,j} - k'(1 - AA_j)$ (see Eq. (7)) can be taken as an appropriate approximation. The results obtained from the described model and their comparison with the experimental data justifies this suggestion. If the hydration of an ionizable group is important for ΔH_{ion} production, there is a physical reason for the validity of such a simple relationship.

Using the relationship between $\Delta H_{ion,j}$ and AA_j introduces some details of the molecular structure into the model. Accounting for the protein structure by use of atomic accessibilities is very important, especially in the pH 2–6 and 9–11 ranges, where a number of groups with very close values of pK_a are being titrated. This was shown for ribonuclease *A* using two different theoretical models. Fig. 6A shows the theoretical calorimetric titration curve obtained from the proposed model (curve 1) and curve 2 calculated by the simple theoretical model, without the corrections of the ΔH_{ion} values which depend on the AA_j and the individual pK_a values. The differential curve obtained from the differences of the curves 1 and 2 is shown in the insertion (B). The nonlinearity of this differential curve shows that the corrections for AA_j values (based on the 3D structure) are not linear and are significant for the accuracy of the model. The greatest deviations occur in the pH 2–6 and 9–11 ranges (Fig. 6B). These deviations are apparently related to the ionization of the carboxylic and amino groups.

4.1. Parameters and absolute scale of calorimetric titration

A number of studies have provided values of the heats of ionization of amino acids in aqueous solution (Table 2). The ionization enthalpies of amino acid residues were obtained from experimental data for appropriate amino acid derivatives and other model

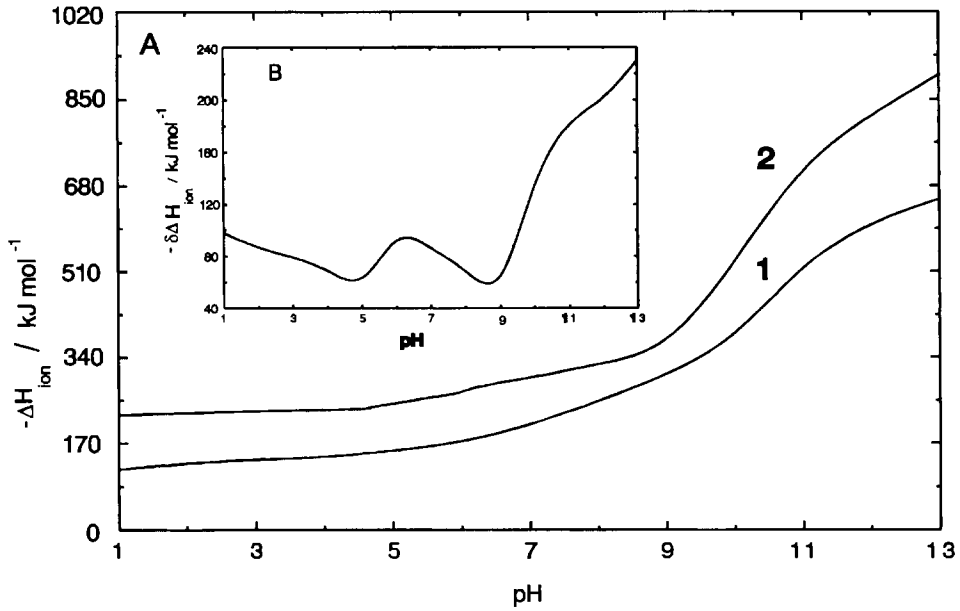

Fig. 6. Comparison of the calculated calorimetric titration curves of ribonuclease A by two theoretical models. Model A, curve 1: theoretical results of the model accounting for pK_a values and ΔH_{ion} with corrections for atomic accessibilities AA_j of ionizable amino acid residues; and curve 2: theoretical results of the model accounting for $pK_{int,j}$ and ΔH_{ion} values of ionizable amino acid residues. Model B: differential curve obtained from the differences of the curves 1 and 2 in Fig. 6A.

compounds (Table 1). The published literature does not offer a full set of data for ΔH_{ion} values measured in the same aqueous organic solvent. An optimal set characterized by a minimal number of organic solvents (mainly methanol), is shown in Table 1. The data can be divided into two groups, one with large and another with small ΔH_{ion} values. The first group (Lys, Arg, NTR, Tyr) is mainly titratable in the basic range. The second group (Asp, Glu, CTR) is titratable in the acid range. These differences are probably due to different degrees of hydration of the sites, which are the number of water molecules directly hydrogen-bonded to the proton-binding sites, and the changes in the degree of hydration during the process of ionization. In the case of carboxylic groups, when the changes of ionization do not lead to change in hydration, the contributions of enthalpy ionization are minor (6 ± 1.5 kJ mol^{-1}). In the case of amino groups, the ionization enthalpy is higher because of mutual changes of ionization and hydration. Obviously, the basic difference between the two types of group is due to a hydrogen-bonded capacity and to differences in the free-electron lone-pairs (two at the oxygen and one at the nitrogen atoms). Thus, the carboxylic group remains mainly as a proton-acceptor for four water molecules, before and after ionization. The amino group is a proton acceptor for one water molecule and proton-donor for two water molecules before ionization and donor for three water molecules after ionization.

In the case when the model compounds represent the amino acid residues in proteins, the structural surroundings can strongly influence the hydration of a given group, mainly by decreasing it. This process can be simulated using a simple relationship between the dielectric constant in the aqueous-organic solvent and ΔH_{ion} (Table 1) and, alternatively, between the dielectric constant and the atomic accessibility [17]. A simple model describing these relationships is associated with the contact area of ionizable groups and the polarizability of the surrounding atoms. The linear relationship between the atomic/molecular surfaces and free energy of transfer from water to media with

different dielectric constants is well known and widely used [38].

The experimental data available on calorimetric titration show that many investigators use their own starting point of titration depending on sample preparation. In order to standardize the process, the origin of the scale may be taken as zero at very low values of pH or extrapolated to zero pH. All the deprotonations in this case will be accompanied by negative ΔH_{ion} values and all the calorimetric titration curves will be comparable. Such a choice of the origin of the scale is a matter of convention. The origin of the scale could also be placed in the high pH range. This is not convenient, however, for experimental measurements and their comparison with theoretical calculations. Since the calorimetric contributions are small in acidic regions, the curves would be relatively pH-insensitive and the deviations would be low. This is the reason why the origin of the calorimetric titration curves should be positioned at the most acidic pH.

4.2. *Coincidence between the experimental and theoretical calorimetric titration data*

As shown in the results, the agreement between the experimental and theoretical data for myoglobin is best from pH 2.0 to 10.5 (Fig. 1). Some small differences in the pH 7.5–9.5 range may arise from the abnormal titration of the half of the buried histidyl residues (B5, EF4, Cl, etc.). Large differences appear above pH 10.5, probably due to structural changes in the molecule, which are not taken into consideration in the model.

In the case of cytochrome *c* (Fig. 2), good agreement between theory and experiment is found in the pH 3–9 range. However, in alkaline pH, the experimental curve starts to deviate from the theoretical one. In this pH range, the main contribution is due to lysine titrations (19 Lys residues in cytochrome *c*), which are the most flexible charged groups clustered in one moiety of the molecule. It can be hypothesized that the Lys residues change their positions in such a way that they decrease the electrostatic interactions among themselves. Their pK_a undergo greater shifts and the result is that the whole basic part of the curve shifts to lower pH values.

The calorimetric titration of ribonuclease *A* is in good accordance with the predicted curve in a wide pH 3.0–10.5 range (Fig. 3). The small difference in pH (5–9) can be explained by taking into consideration pH-dependent changes between several conformational forms (which differ little in structure due to the presence of some histidyl residues). The measurements at high ionic strength (curve 4) show a remarkable difference in the pH 7–8 range. This difference cannot be explained as a salt-effect of the ionization, since the theoretical curve (curve 3) is the same as curve 1 in this region. If the experimental results are correct, this is evidence for possible ionic-strength-dependent conformational changes in ribonuclease *A*.

In the case of lysozyme (Fig. 4), excellent agreement was found between the theoretical and experimental results [25] in the pH 1–7 range. This protein was primarily used for testing the electrostatic model applied here and the coincidence between theory and experiment was expected. However, a great discrepancy (Fig. 4, curve 3) with the experimental data from another source [24] was observed. The only titratable group in the pH 4.5–6.0 range with sufficient contribution to ΔH_{ion} is the single His 15. However its ionization cannot explain the difference of 250 kJ mol^{-1} (from curve 3). Nevertheless, it is well known that lysozyme undergoes dimerization in the pH 5–7 range [39]. The hiding of ionizable groups which could be buried at dimerization, will decrease the apparent enthalpy (see Fig. 1). In such a case, the final effect would be opposite to the results seen in Fig. 4. Therefore, we can only assume that the difference in the pH 4.5–7.5 range arises from the non-compensated neutralization and dilution of the titrant, as previously mentioned [24]. Curve 3 (Fig. 3) is practically parallel to the theoretical curve at higher pH, up to pH 9.

Finally, the limited experimental data for the calorimetric titration of α-chymotrypsin in Fig. 5 are in accordance with the theoretical curve 1.

4.3. *Differential calorimetric titration of cytochrome c*

In order to deduce the structure–function relationships of proteins (enzymes) from the calorimetric model presented, it is necessary to compare the pH-dependent heat of ionization in different states (bound/unbound, oxidized/reduced, holo/apo-enzyme, etc.). In fact, the $\delta\Delta H_{ion}$(pH) curves give the "pure spec-

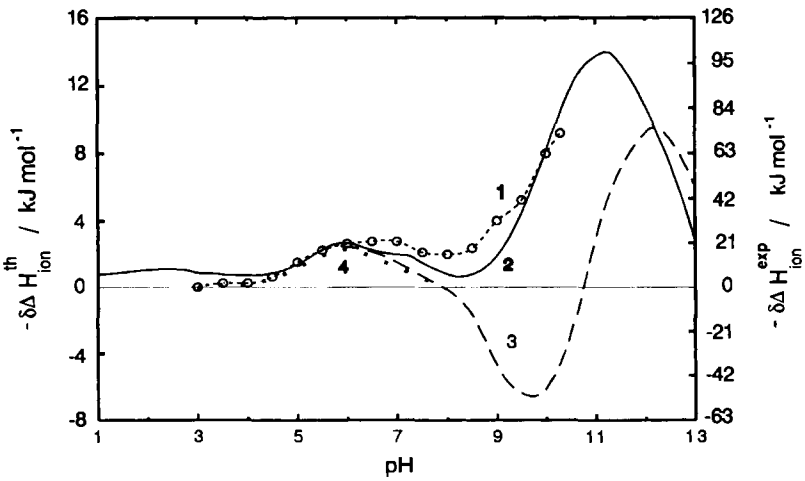

Fig. 7. Differential calorimetric titration curves of cytochrome *c*. (1) Differential curve obtained from the differences of experimental calorimetric titration curves of ferro- and ferri-cytochrome *c* [11] (circles); (2) differential curve obtained from the theoretical calorimetric titration curves of ferro- and ferri-cytochrome *c* without including water molecule in the vicinity of Met80 (solid line); (3) differential curve obtained from the theoretical calorimetric titration curves of ferro- and ferri-cytochrome *c* including the water molecule (dashed line); and (4) individual contribution of His 26 to the differential curve obtained from theoretical calorimetric titration curves of ferro- and ferri-cytochrome *c* (dots).

trum" of interactions of "state"-dependent changes of charged groups, widely distributed in the molecule. It is possible to make structural assignments of some differential peaks with the specific $\Delta h_{\mathrm{ion},i}$ contributions of corresponding ionizable groups.

In the present study, such an analysis for the ferro- or ferri-cytochrome *c* has been made. The results are demonstrated in Fig. 7. Both differential curves – experimental curve 1 and theoretical curve 2 – have the same qualitative shape, but they differ in their absolute scale. The similarity confirms the participation of the same groups in their interaction with the "excessive" electron (reduced Fe in the heme). It can be assumed that the difference between the curves 1 and 2 in the pH 7.5–9.5 range might be due to the influence of one "unusual" water molecule with $pK_{1/2}$ 9.8 in the ferri- and $pK_{1/2}$ 10.6 in the ferro- form. This water is buried in the vicinity of Met80 and Fe from the heme [40]. The enthalpy contribution due to this water molecule was not included in curve 2. However, the inclusion of $\Delta h_{\mathrm{ion},i}$(pH) for this water molecule in calorimetric calculations (shown as the differential theoretical curve 3 in Fig. 7) gives the opposite effect. Thus, the difference between curves 1 and 2 in the pH 7.5–9.5 range is not caused by the water molecule.

A detailed analysis can be accomplished by calculating the differential calorimetric titration curve of each individual titratable group of the protein. In the case of ferri- or ferro-cytochrome *c*, the contribution of His 26, with $pK_{1/2}$ 5.7 in ferri-cytochrome *c* and $pK_{1/2}$ 5.9 in ferro-cytochrome *c* forms is the only difference compared to the other ionizable groups (Fig. 7, curve 2). The pH-dependence of the differential calorimetric curve of His 26 is presented in Fig. 7 (curve 4).

It is expected that the titration calorimetry will undergo an extensive development in the next few years and we hope that this study will contribute to this process. Undoubtedly, the experimental results of acid–base calorimetric titration could be used in order to improve the semi-empirical models for electrostatic analysis in proteins. Thus, the models have to comply with both proton titration and calorimetric titration curves.

Acknowledgements

This work was supported by Grant No. X-429 from the National "Scientific Investigations" Fund, Sofia, Bulgaria.

References

[1] P. Privalov and S. Gill, Adv. Protein Chem. 39 (1979) 191.
[2] G. Makhatadze and P. Privalov, J. Mol. Biol. 232 (1993) 639.
[3] K. Murphy and S. Gill, Thermochim. Acta 172 (1990) 11.
[4] K. Murphy and S. Gill, J. Mol. Biol. 222 (1991) 699.
[5] D. Hayne and E. Freire, Anal. Biochem. 216 (1994) 33.
[6] P. Privalov, N. Khechinashvili and B. Atanasov, Biopolymers 10 (1971) 1865.
[7] J. Hermans and G. Rialdi, Biochemistry 4 (1965) 1277.
[8] G. Kresheck and H. Scheraga, J. Am. Chem. Soc. 88 (1966) 4588.
[9] R. Biltonen, A. Schwartz and I. Wadso, Biochemistry 10 (1971) 3417.
[10] M. Flogel and R. Biltonen, Biochemistry 14 (1975) 2603.
[11] M. Marini, C. Martin and R. Berger, Biopolymers 19 (1980) 899.
[12] M. Marini, C. Martin, R. Berger and L. Forlani, Biopolymers 20 (1981) 2253.
[13] T. Wiseman, S. Williston, J. Brandts and L. Lin, Analyt. Biochem. 179 (1989) 131.
[14] S. Shire, G. Hanania and F. Gurd, Biochemistry 13 (1974) 2967.
[15] A. Warshel and S. Russell, Quart. Rev. Biophys. 17 (1984) 283.
[16] B. Atanasov and A. Karshikov, Studia Biophysica 105 (1985) 11.
[17] J. Matthew, Ann. Rev. Biophys. Chem. 14 (1985) 387.
[18] M. Gilson and B. Honig, J. Comp. Chem. 9 (1988) 327.
[19] D. Bashford and M. Karplus, Biochemistry 29 (1990) 10219.
[20] R. Sampogna and B. Honig, Biophys. J. 66 (1994) 1341.
[21] E. Mehler and T. Solmajer, J. Chim. Phys. and Phys. Chim. Biol. 88 (1991) 2411.
[22] T. Simonson, G. Perahia, G. Bricogne and A. Brunger, J. Chim. Phys. and Biol. 88 (1991) 2701.
[23] V. Spassov, A. Karshikov and B. Atanasov, Biochim. Biophys. Acta 999 (1989) 1.
[24] C. Bjurulf, Eur. J. Biochem. 30 (1972) 33.
[25] W. Pfeil and P. Privalov, Biophys. Chem. 4 (1976) 23.
[26] A. Karshikov, R. Engh, W. Bode and B. Atanasov, Eur. Biophys. J. 17 (1989) 287.
[27] F. Bernstein, T. Koetzle, J. Williams, E. Meyer, M. Brice, J. Rogers, O. Kennard, T. Shimanouchi and M. Tasumi, J. Mol. Biol. 112 (1977) 535.
[28] B. Lee and F. Richards, J. Mol. Biol. 55 (1975) 379.
[29] R. Benoit, C. Louis and M. Frechette, Thermochim. Acta 171 (1990) 115.
[30] L. Chiche, L. Gregoret, F. Cohen and P. Kollman, Proc. Natl. Acad. Sci. (USA) 87 (1990) 3240.
[31] T. Takano, Refinement of Myoglobin and Cytochrome *c*. in S. Hall and T. Ashida (Eds.), Methods and Applications in Crystallographic Computing, Oxford University Press, Oxford, England (1984).
[32] H. Pelletier and J. Kraut, Science 258 (1992) 1748.
[33] G. Bushnell, G. Louie and G. Brayer, J. Mol. Biol. 214 (1990) 585.
[34] A. Wlodawer, L. Svensson, L. Sjolin and G. Gilliand, Biochemistry 27 (1988) 2705.
[35] O. Herzberg and J. Sussman, J. Appl. Crystallogr. 16 (1983) 144.
[36] J. Birktoft and D. Blow, J. Mol. Biol. 68 (1972) 187.
[37] R. Kuroki, K. Nitta and K. Yutani, J. Biol. Chem. 267 (1992) 24297.
[38] I. Tunon, E. Silla and J. Pascual-Ahuir, Protein Eng. 5 (1992) 715.
[39] V. Fedotov, Y. Feldman, A. Krushelnitsky and I. Ermolina, J. Mol. Struct. 219 (1990) 293.
[40] A. Berghuis, J. Guillemette, G. McLendon, F. Sherman, M. Smith and G. Brayer, J. Mol. Biol. 236 (1994) 786.
[41] D. Bhattacharyya, A. Basu and S. Aditya, J. Indian. Chem. Soc. 61 (1984) 956.
[42] Yu. Koragin, V. Shormanov and G. Krestov, Koord. Khim. (Russian) 11 (1985) 1046.
[43] R. Saxena and S. Bhatia, Indian J. Chem. 13 (1975) 837.
[44] C. Rochester and D. Wilson, J. Chem. Soc. Faraday Trans. 1 73 (1977) 569.
[45] G. Parsons and C. Rochester, J. Chem. Soc. Faraday Trans. 1 71 (1975) 1069.
[46] Z. Pawlak and A. Hampton, Thermochim. Acta 59 (1982) 313.
[47] Y. Akhadov, Dielectric properties of binary solutions (A Data Handbook), Pergamon Press (1981).
[48] F. Rodante and F. Fantauzzi, Thermochim. Acta 144 (1989) 275.
[49] V. Vassilev, L. Kochergina and V. Gavarin, Zh. Obsh. Khim. (Russian) 55 (1985) 2780.
[50] J. Steinhardt and J. Reynolds, Multiple equilibria in proteins, Academic Press (1969) p.178.
[51] J. Larson and L. Hepler, in J. Coetze and C. Ritchie (Eds.), Solute–solvents interactions, Marcal Dekker, New York, NY (1969) p.1.
[52] M. Marini and R. Berger, Anal. Biochem. 43 (1971) 188.
[53] R. Izatt and J. Christensen, in H. Sober (Ed.), Handbook of Biochemistry, J-49. Chemical Rubber Co., Cleveland (1968).