Microscopic Protonation Equilibria of Oxidized Glutathione

Béla Noszál*

Department of Pharmaceutical Chemistry, Semmelweis University, H-1092 Budapest, Högyes E. u. 9, Hungary

Zoltán Szakács

Department of Inorganic and Analytical Chemistry, Loránd Eötvös University, H-1117 Budapest, Pázmány Péter sétány 1/a, Hungary

Received: January 13, 2003

The first complete microequilibrium analysis of a tetrabasic acid is presented. A total of 18 protonation microconstants of glutathione disulfide (GSSG) were determined, 16 of them quantitate the overlapping equilibria in acidic medium, where 4 carboxylate sites protonate. The related, pH-dependent concentrations of 16 coexisting microspecies were also obtained, including the 10 different ones and their 6 identical twins of symmetry origin. The GSSG acid−base chemistry in basic medium is also characterized. It is shown for tetra- and further multibasic acids that the number of traditional microconstants mandatorily exceeds that of the pieces of independent information available by any state-of-the-art methodology, making the complete microequilibrium resolution not only highly complex but theoretically impossible on the basis of hitherto reported principles. The GSSG microspeciation was made feasible here by introducing cumulative microconstants, a new equilibrium parameter, and the invariance of the interactivity coefficient, a plausible simplifying principle, on the experimental basis of ^{1}H NMR-pH titrations in H_2O/D_2O 9/1 media. The methods and results were verified by corresponding data of GSSG-(glycyl)-diethyl-ester, a model compound of reduced microequilibrium complexity. The obtained log $k^U = 2.44$ and log $k^Y = 3.59$ "core" microconstants of the respective glutamyl and glycyl carboxylates of GSSG are close to those of the sulfhydryl (GSH) form of glutathione. The GSSG intermoiety interactions were found to be weak, apparently of Coulombic nature. The pH-dependent distribution of the 18 microspecies is depicted.

1. Introduction

The tripeptide glutathione (γ-L-lutamyl-L-cysteinyl-glycine, GSH) plays a key role in a number of biochemical processes in the human body.[1] The GSH/GSSG system is the most important intracellular oxidation−reduction buffer, providing protection against carcinogenic, radical agents of oxidative stress and lipid peroxidation.[2] Glutathione detoxifies a variety of compounds by forming glutathione conjugates.[3−5] It regulates the SH status of proteins and other biological thiols by homo- and heterodisulfide formation equilibria.[6−10] GSH possesses several metal-binding sites[11,12] and participates in transport and elimination of metal ions,[13,14] such as Hg^{2+} or $(CH_3)Hg^+$. Glutathione has been shown to react with its ionized thiolate group in all of the mentioned reactions. The pH dependence of the observed kinetic constants and equilibrium constants could be quantitated and interpreted using the microscopic (site-specific) protonation constants of GSH.[6−10,14,16,17]

The acid−base chemistry of multidentate (bio)ligands are usually characterized in terms of macroscopic protonation (or dissociation) constants, log *K* (or p*K*), which are composites of the microscopic constants (log *k* or p*k*) for the individual groups.[18−21] The protonation microconstants are, by definition, specific basicity parameters of a particular molecular subunit, in a defined protonation state for all other moieties in the molecule. Due to the instantaneous character of the protonation reactions, microspecies are in incessant interconversion. The coexistence and short individual lifetime preclude them from direct analytical determination by any of the existing, fast experimental techniques. For the same reason, the microspecies-specific intensive properties (microscopic molar absorptivity, ellipticity, NMR chemical shifts, etc.) cannot be determined directly either. Thus, the physicochemical parameters and the analytical concentrations can only be obtained by microspeciation, a composite, indirect methodology that uses at least two experimental techniques, and/or auxiliary compounds, followed by an appropriate evaluation method.

The most widely used combination of experimental techniques is pH-metry and solution spectroscopy. The former yields the number of protons bound to the molecule, whereas the latter (e.g., UV or NMR) monitors the site-specific proton binding as a function of pH. The PC-controlled hyphenation of the above techniques and the resulting automatic acquisition of NMR-pH and UV-pH titration curves have recently been reported.[22,23] Other approaches utilize a priori simplifying assumptions, such as (i) the invariance of interactivity between corresponding sites in analogous moieties of related compounds,[24] (ii) the use of the group constant concept,[20] and (iii) the introduction of basicities from derivative compounds with reduced number of functional groups into the related subunit of the parent molecule.[19,21,25,26]

The applications of microspeciation lie primarily in biology and molecular medicine, since the binding of bio- and drug molecules to receptors, enzymes, and plasma proteins are well-

* Corresponding author. Fax: +361-2170891. E-mail: NOSBEL@hogyes.sote.hu.

10.1021/jp0340820 CCC: $25.00
Published on Web 05/02/2003

known to be specific processes, which take place via complementary microforms of the interacting molecules.[21] Due to its biological significance, GSH itself was the very first tetradentate molecule to be completely characterized in terms of protonation microequilibria.[11] The inherent basicity difference between the amino/thiolate pair and the carboxylate groups allowed a reduction of the tetraprotic microspeciation problem into two biprotic ones. Moreover, metal complex[11,12] and heterodisulfide[6−10] formation equilibria of GSH have also been thoroughly characterized at the microscopic level.

Nevertheless, literature data on the microforms of the glutathione thiol−disulfide system are highly one-sided. Though GSSG has been investigated in thousands of biochemical and physicochemical studies, no report appeared on its site-specific carboxylate protonation, thus hampering the description of reactions of GSSG at the microscopic level. This is obviously due to the complexity of the tetraprotic microspeciation problem.

The vast majority of microspeciation studies deals with compounds of two proton-binding sites.[11,21−23,27,28] Attempts to characterize the microequilibria of ligands with three or more donor groups are still sporadic.[24,27,29,30] There are at least three reasons for this scarcity. (1) In a molecule of n basic sites, the protonation of at least $n - 1$ sites needs to be selectively monitored, which is often a demanding experimental task, even in the NMR age. Nevertheless, meeting the above criterion is only a necessary, but not sufficient precondition in several cases of microspeciation, as it will be discussed in the Results and Discussion section. (2) Inherently, not many molecules contain three trilaterally interacting sites. (3) The major reason is certainly the mounting complexity of the microequilibrium equations, as the number of the binding sites grows. In fact, a linear increase in the number of sites generates an exponential increase in the number of microspecies, and the number of microconstants is even more prolific.[20]

A fourth principal difficulty has only most recently been revealed.[31] It was shown that in molecules of four or more sites and low level of symmetry, the minimum amount of independent information for exact microspeciation mandatorily exceeds the amount of the experimentally available data. Symmetry elements and reduced strictness in the treatment may allow, however, the elucidation of microconstants for such molecules as well. Relationships between the number of sites and symmetry types with respect to feasibility and theoretical details have been established.[31] To date, however, no theoretically flawless, experiment-based microspeciation study appeared on any tetradentate ligand.

Here we report the complete resolution of the microscopic protonation equilibria of GSSG, which is the first example of a strict microspeciation analysis of a tetrabasic system.

2. Experimental Section

2.1. Materials. Glutathione disulfide (GSSG) and the (glycyl)-monoethylester of reduced glutathione (EtGSH), both of analytical grade, were purchased from Sigma and used without further purification. Other reagents were also of analytical grade. GSSG diethyl ester (EtGSSGEt) was synthesized by bubbling oxygen gas through a 0.005 mol dm^{-3} aqueous solution of EtGSH at a constant pH of 8.0. The progress of the reaction was followed by a capillary zone electrophoresis method, developed in our laboratory. Complete disulfide formation required 8-hr reaction time, and only minimal ester hydrolysis was observed.

2.2. Potentiometric Titrations. pH-Metric titrations were carried out at 25.0 ± 0.1 °C and 0.15 mol dm^{-3} ionic strength, adjusted with KCl. A Radiometer ABU 12 automatic buret and a pHm 64 digital research pH meter, attached to a Radiometer pHC2406 combined glass electrode, were used. The electrode chain was calibrated in terms of hydrogen ion concentration.[32] The peptide concentration varied between 0.002 and 0.010 mol dm^{-3}, the titrant was carbonate-free KOH. The concentration-based protonation macroconstants were evaluated by our nonlinear least-squares regression program PROTC.

2.3. NMR-pH Titrations. The pH-dependent series of ^{1}H NMR spectra was recorded on Bruker 500 MHz, 360 MHz, and Varian 200 MHz spectrometers at 22 °C. The solvent mixture was H_2O/D_2O, 9:1 by volume. The water resonance was diminished by a presaturation pulse before the observation pulse. The electrode chain and the composition of all solutions were identical with those used in potentiometry. The pH measurements were performed in 5 cm^3 vessels with proper stirring, before filling the peptide solutions into the NMR tube. This procedure resulted in much more accurate and reproducible pH data, in contrast to measuring the pH inside the NMR tube. Chemical shifts were referenced to sodium 3-(trimethylsilyl)-1-propanesulfonate (DSS). For control purposes, the samples contained also 0.001 mol dm^{-3} ethanol as a second internal reference. No change was observed in the ethanol chemical shifts (3.648 and 1.172 ppm) in the pH range 1−13.

2.4. Data Analysis. To fit our model equations to the experimental NMR-pH profiles by means of multiple nonlinear regression, C++ programs were developed on the basis of a standard Marquardt−Levenberg least squares parameter estimation algorithm.[33] Detailed listings of the results, demonstrating the statistical reliability of our procedure, are given as Supporting Information.

3. Results and Discussion

3.1. Potentiometric Titrations. The pH-metric titrations resulted in the following stepwise macroscopic protonation constants (uncertainties are estimates of standard deviation):

$$\log K_1 = 9.53 \pm 0.02 \quad \log K_2 = 8.83 \pm 0.01 \quad \log K_3 = 3.85 \pm 0.01$$

$$\log K_4 = 3.15 \pm 0.02 \quad \log K_5 = 2.32 \pm 0.03 \quad \log K_6 = 1.6 \pm 0.1$$

which are in good agreement with previously published data.[34,35] These macroconstants describe two well-separated protonation regions: the overlapping protonation of the two amino moieties in the range 8 < pH < 11 and that of the four carboxylates at pH < 5, respectively. However, the macroscopic constants quantitate the proton coordination of the ligand as a whole

$$H_{i-1}L^{(4-i+1)-} + H^+ \rightleftharpoons H_iL^{(4-i)-}$$

$$K_i = \frac{[H_iL^{(4-i)-}]}{[H_{i-1}L^{(4-i+1)-}][H^+]} \qquad i = 1, 2, ..., 6 \tag{1}$$

and, in principle, cannot be assigned to individual groups.[20,21]

3.2. ^{1}H NMR Spectroscopy. To characterize the basicities of the individual proton-binding sites, ^{1}H NMR-pH titrations have been carried out. Considering the 10 covalent bonds between the glutamyl and glycyl carboxylates, the protonation of these groups can well be assumed to be selectively reflected by the chemical shift changes of the adjacent carbon-bound

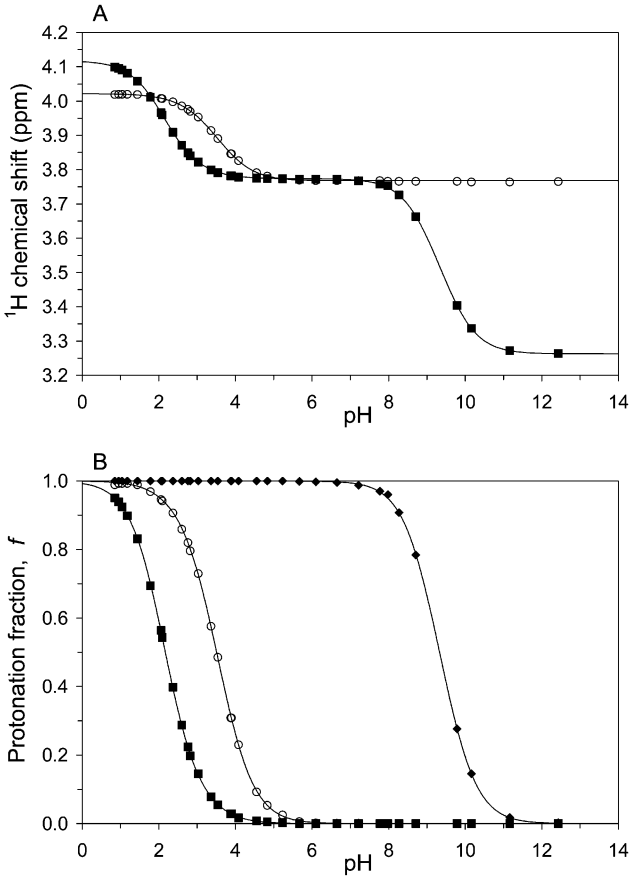


Figure 1. (A) ^{1}H NMR titration curves as a function of pH for the glutamyl methyne (■) and the glycyl methylene (○) protons. (B) Protonation fractions of individual groups in GSSG: glutamyl carboxylate (■), glycyl carboxylate (○), and amino group (♦). The computer fits are shown as solid lines.

hydrogens.[11,36] The glutamyl α proton shows a triplet throughout the pH scale. The glycyl methylene protons exhibit a doublet at pH < 4.8, due to their coupling with the vicinal amide hydrogen, which exchanges with the H_2O protons slowly on the NMR time scale.[37] In the 4.8 < pH < 7.5 range, the methylene protons become nonequivalent (ABX pattern) and their signals collapse at pH > 7.5 into a doublet of doublets. The chemical shifts of these "indicator" α protons are depicted as a function of pH in Figure 1A. The chemical shift of the glycyl hydrogens does not undergo any tangible changes in the basic pH region, verifying that backbonewise the proton-binding sites at the glycyl and glutamyl residues are completely isolated by the 10 covalent bonds. The glycyl methylene chemical shift can thus be assumed to selectively reflect the protonation state of the glycyl carboxylate below pH 5. The glutamyl α protons monitor accordingly the protonation of the equivalent U and U′ sites.

3.3. Protonation of the Amino Groups. Due to symmetry, basicity of the equivalent amino moieties (denoted by N and N′ in Figure 2) can be quantitated by means of two identical pairs of microconstants, which can be readily calculated from the pH-potentiometric macroconstants:[18]

$$k^{N} = k^{N'} = K_1/2$$

and

$$k^{N'}_{N} = k^{N}_{N'} = 2K_2 \qquad (2)$$

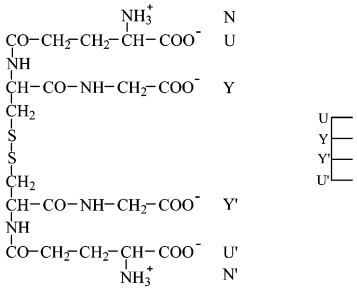


Figure 2. The diprotonated form H_2L^{2-} of oxidized glutathione (GSSG). The four-armed symbol represents the tetrabasic ligand in the carboxylate protonation region. In the bis(ethyl ester)-derivative (EtGSSGEt), the glycyl carboxylate groups (Y and Y′ $-COO^-$) are replaced by $-COOC_2H_5$.

TABLE 1: Microscopic Protonation Constants of Oxidized Glutathione (GSSG)[a]

microconstant	value ± std.dev.	microconstant	value ± std.dev.
$\log k^{U} = \log k^{U'}$	2.42 ± 0.02	$\log k^{Y} = \log k^{Y'}$	$3.59_3 \pm 0.006$
$\log k^{U}_{Y} = \log k^{U'}_{Y'}$	2.31 ± 0.02	$\log k^{Y}_{U} = \log k^{Y'}_{U'}$	3.48 ± 0.01
$\log k^{U'}_{Y} = \log k^{U}_{Y'}$	2.31 ± 0.02	$\log k^{Y'}_{U} = \log k^{Y}_{U'}$	3.48 ± 0.01
$\log k^{U'}_{U} = \log k^{U}_{U'}$	2.34 ± 0.02	$\log k^{Y'}_{Y} = \log k^{Y}_{Y'}$	3.45 ± 0.01
$\log k^{U'}_{UY} = \log k^{U}_{U'Y'}$	2.23 ± 0.02	$\log k^{Y'}_{UY} = \log k^{Y}_{U'Y'}$	3.34 ± 0.02
$\log k^{U'}_{UY'} = \log k^{U}_{U'Y}$	2.23 ± 0.02	$\log k^{Y'}_{U'Y} = \log k^{Y}_{UY'}$	3.34 ± 0.02
$\log k^{U}_{YY'} = \log k^{U'}_{YY'}$	2.20 ± 0.02	$\log k^{Y}_{UU'} = \log k^{Y'}_{UU'}$	3.37 ± 0.02
$\log k^{U'}_{YUY'} = \log k^{U}_{Y'U'Y}$	2.12 ± 0.03	$\log k^{Y'}_{UYU'} = \log k^{Y}_{UY'U'}$	3.23 ± 0.02
$\log k^{N} = \log k^{N'}$	9.41 ± 0.01	$\log k^{N'}_{N} = \log k^{N}_{N'}$	9.25 ± 0.02

[a] At 295 K, 0.15 mol dm^{-3} ionic strength, in H_2O/D_2O, 9:1 medium.

The ^{1}H NMR-pH titrations provide a control of the above data. Thus, the amino microconstants have been obtained by nonlinear regression to the alkaline inflection of the NMR-pH titration curve as well (see Table 1). The large number of intervening atoms prevents the inductive interaction between the two amino sites. The difference in macroconstants, $\Delta\log K = 0.7$, however, slightly exceeds the $\log 4 = 0.602$ "statistical value", expected on the basis of noninteracting moieties.[18,20] A plausible interpretation is the high flexibility of the GSSG molecule and the concomitant, slight Coulombic interaction of the $-NH_3^+$ groups.

3.4. Protonation Microequilibria of the Carboxylates. By pH 7, protonation of the amino sites is practically complete. Hence, at pH < 7 GSSG can be treated as a symmetric, tetradentate ligand, characterized by the macroscopic protonation constants K_3, K_4, K_5, and K_6. The GSSG species with ionized glutamyl (U, U′) and glycyl (Y, Y′) carboxylate sites is represented by a four-armed symbol in Figure 2. A similar, separate treatment for the basic and acidic pH ranges of GSH has also been done.[11]

The microscopic protonation scheme of the GSSG carboxylates is depicted in Scheme 1. The 16 microspecies are denoted by *a*, *b*, ... *p* (ionic charges are omitted), while the symbols k^j_i stand for the microconstants. The superscripts of microconstants refer to the protonating functional group, while the subscript (if any) indicates the protonated site(s). For instance, microconstant $k^{Y'}_{UY}$ describes the process when microspecies *f*, holding protons at sites U and Y, binds a further hydrogen ion at its Y′ glycyl site:

$$k^{Y'}_{UY} = \frac{[l]}{[f][H^+]} \qquad (3)$$

In the general case, an aqueous solution of a tetradentate molecule contains $2^4 = 16$ microspecies and their equilibria can be expressed in terms of $4 \cdot 2^3 = 32$ microconstants.[20] In the case of the symmetrical GSSG molecule, some protonation isomers are indiscernible: $b \equiv e$, $c \equiv d$, $f \equiv k$, $g \equiv j$, $l \equiv o$, and $n \equiv m$. Thus, the number of chemically distinct microspecies is ten.[31] The complete microequilibrium characterization of this system needs nine pieces of independent information, which can be various combinations of macro- and microconstants.[31] For simplicity in the treatment and formalism, we have introduced the Hessian cumulative microconstant,[31] a novel type of microscopic parameter, denoted by κ, being partly analogous with the cumulative macroscopic constant. For example, for microspecies *f*, it reads:

$$\kappa_f = \frac{[f]}{[a][\mathrm{H}^+]^2} = k^{\mathrm{Y}}k^{\mathrm{U}}_{\mathrm{Y}} = k^{\mathrm{U}}k^{\mathrm{Y}}_{\mathrm{U}} \tag{4}$$

The term "Hessian" was chosen to express the thermodynamic nature of κ, because of its independence of the intermediate stages and ways of the microspecies formation. A possible set of cumulative microconstants for the formation of GSSG microspecies reads:

$$\kappa_a = 1 \tag{5}$$

$$\kappa_b = \kappa_e = k^{\mathrm{U}} \tag{6}$$

$$\kappa_c = \kappa_d = k^{\mathrm{Y}} \tag{7}$$

$$\kappa_f = \kappa_k = k^{\mathrm{U}}k^{\mathrm{Y}}_{\mathrm{U}} \tag{8}$$

$$\kappa_g = \kappa_j = k^{\mathrm{U}}k^{\mathrm{Y}'}_{\mathrm{U}} \tag{9}$$

$$\kappa_h = k^{\mathrm{U}}k^{\mathrm{U}'}_{\mathrm{U}} \tag{10}$$

$$\kappa_i = k^{\mathrm{Y}}k_{\mathrm{Y}}{}^{\mathrm{Y}'} \tag{11}$$

$$\kappa_l = \kappa_o = k^{\mathrm{U}}k^{\mathrm{Y}}_{\mathrm{U}}k^{\mathrm{Y}'}_{\mathrm{UY}} \tag{12}$$

$$\kappa_m = \kappa_n = k^{\mathrm{U}}k^{\mathrm{Y}}_{\mathrm{U}}k^{\mathrm{U}'}_{\mathrm{UY}} \tag{13}$$

$$\kappa_p = k^{\mathrm{U}}k^{\mathrm{Y}}_{\mathrm{U}}k^{\mathrm{U}'}_{\mathrm{UY}}k^{\mathrm{Y}'}_{\mathrm{UYU}'} \tag{14}$$

Using nine cumulative microconstants, instead of the nine stepwise ones, offers advantages both in the formalism and computation. The microspecies concentrations can thus be expressed in a compact way, as exemplified in the following equations for microspecies *a* and *f*

$$[a] = \alpha_a C_{\mathrm{L}} = \kappa_a C_{\mathrm{L}}/D \tag{15}$$

$$[f] = \alpha_f C_{\mathrm{L}} = \kappa_f [\mathrm{H}^+]^2 C_{\mathrm{L}}/D \tag{16}$$

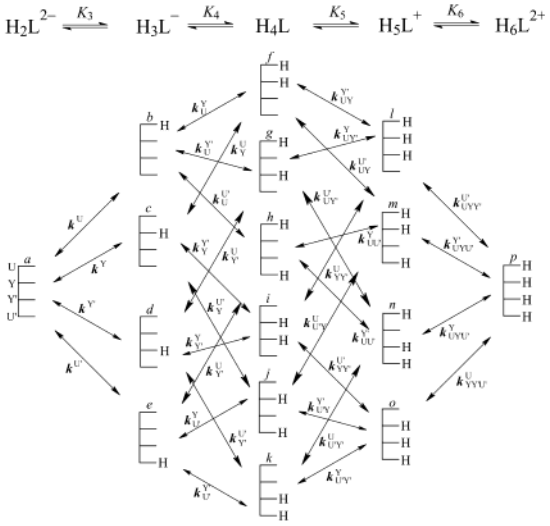
where α represents mole fraction, C_{L} is the total ligand (GSSG) concentration

$$C_{\mathrm{L}} = [a] + [b] + [c] + [d] + ... + [p] \tag{17}$$

and

$$D = 1 + K_3[\mathrm{H}^+] + K_3K_4[\mathrm{H}^+]^2 + K_3K_4K_5[\mathrm{H}^+]^3 + K_3K_4K_5K_6[\mathrm{H}^+]^4 \tag{18}$$

SCHEME 1: Macroscopic and Microscopic Protonation Equilibria of GSSG in the Carboxylate Protonation Region



The macroconstants should not be considered as further independent parameters during the calculations, as they can also be expressed in terms of the κ constants:

$$K_3 = 2(\kappa_b + \kappa_c) \tag{19}$$

$$K_3K_4 = 2(\kappa_f + \kappa_g) + \kappa_h + \kappa_i \tag{20}$$

$$K_3K_4K_5 = 2(\kappa_l + \kappa_m) \tag{21}$$

$$K_3K_4K_5K_6 = \kappa_p \tag{22}$$

Since each κ constant corresponds to the formation of a particular microspecies, the number of κ constants and that of the microspecies are identical. This is a great advantage over the traditional formalism, where the number of *k* microconstants becomes redundant in molecules of several protonation sites.

In the following paragraphs, extraction of the nine κ constants from the experimental NMR titration curves will be shown first. Decomposition of the κ constants into *k* microconstants will be discussed subsequently.

3.5. Analysis of ^{1}H NMR-pH Titration Curves. Chemical shifts of carbon-bound protons undergo changes upon protonation of nearby basic sites. For example, the observed chemical shift of the glutamyl methyne proton can be expressed by the following equation:[11,27,36]

$$\delta_{\mathrm{U}} = \delta_{\mathrm{U,p}} f_{\mathrm{U}} + \delta_{\mathrm{U,d}}(1 - f_{\mathrm{U}}) \tag{23}$$

where $\delta_{\mathrm{U,p}} = 4.116$ ppm and $\delta_{\mathrm{U,d}} = 3.772$ ppm are the methyne chemical shifts in the protonated and deprotonated states of the adjacent glutamyl carboxylate. These values are readily available from NMR spectra at pH = 0 and pH = 7.0, respectively. f_{U} is the protonation fraction of the glutamyl carboxylate (U), as defined in eq 24:

$$f_{\mathrm{U}} = ([b] + [f] + [g] + [h] + [l] + [m] + [n] + [p])/C_{\mathrm{L}} \tag{24}$$

The protonation fraction of the glycyl carboxylate

$$f_{\mathrm{Y}} = ([c] + [f] + [i] + [j] + [l] + [m] + [o] + [p])/C_{\mathrm{L}} \tag{25}$$

is derived similarly from the experimental data,

$$\delta_Y = \delta_{Y,p} f_Y + \delta_{Y,d}(1 - f_Y) \quad (26)$$

with $\delta_{Y,p} = 4.021$ ppm and $\delta_{Y,d} = 3.767$ ppm. The experimental protonation fraction curves are depicted in Figure 1B. Due to symmetry, protonation fractions of equivalent groups are identical[31,38]

$$f_{U'} = f_U \text{ and } f_{Y'} = f_Y \quad (27)$$

The [*b*], [*c*], ..., [*p*] microspecies concentrations can be expressed from eq 16-type relationships. Their introduction into eqs 24 and 25, followed by reduction and using eqs 19 through 22, establishes relationships between the experimental f_U, f_Y, and $[H^+]$ values, and the unknown cumulative microconstants:

$$f_U = \{\kappa_b[H^+] + (\kappa_f + \kappa_g + \kappa_h)[H^+]^2 + (\kappa_l + 2\kappa_m)[H^+]^3 + \kappa_p[H^+]^4\}/\{1 + 2(\kappa_b + \kappa_c)[H^+] + (2\kappa_f + 2\kappa_g + \kappa_h + \kappa_i)[H^+]^2 + 2(\kappa_l + \kappa_m)[H^+]^3 + \kappa_p[H^+]^4\} \quad (28)$$

$$f_Y = \{\kappa_c[H^+] + (\kappa_f + \kappa_g + \kappa_i)[H^+]^2 + (\kappa_m + 2\kappa_l)[H^+]^3 + \kappa_p[H^+]^4\}/\{1 + 2(\kappa_b + \kappa_c)[H^+] + (2\kappa_f + 2\kappa_g + \kappa_h + \kappa_i)[H^+]^2 + 2(\kappa_l + \kappa_m)[H^{+\,3} + \kappa_p[H^+]^4\} \quad (29)$$

For further simplicity, coefficients of the $[H^+]$ polynomials in eqs 28 and 29 can be unified in seven *Q* terms as follows:

$$f_U = \{Q_1[H^+] + Q_2[H^+]^2 + Q_3[H^+]^3 + Q_4[H^+]^4\}/\{1 + 2(Q_1 + Q_5)[H^+] + (Q_2 + Q_6)[H^+]^2 + 0.666(Q_3 + Q_7)[H^+]^3 + Q_4[H^+]^4\} \quad (30)$$

$$f_Y = \{Q_5[H^+] + Q_6[H^+]^2 + Q_7[H^+]^3 + Q_4[H^+]^4\}/\{1 + 2(Q_1 + Q_5)[H^+] + (Q_2 + Q_6)[H^+]^2 + 0.666(Q_3 + Q_7)[H^+]^3 + Q_4[H^+]^4\} \quad (31)$$

Thus, the f_U and f_Y variables are functions of the hydrogen ion concentration and the seven *Q* terms. In other words, the f_U and f_Y functions allow the elucidation of the *Q* parameters, but not necessarily their nine κ components, as we have recently pointed out on a theoretical basis as well.[31] Since the complete description requires nine parameters, it is clearly impossible to calculate the concentration of each GSSG microspecies from the f_U and f_Y protonation mole fractions without further information.

Equations 28 and 29 demonstrate also that only the sum ($\kappa_f + \kappa_g$) can be unambiguously calculated from the coefficients Q_2 and Q_6. The protonation isomers *f* and *g* have the same contribution to the chemical shifts throughout the pH scale and, consequently, their concentration cannot be determined from the NMR-pH titration alone. Chemically modified (e.g., asymmetrically esterified) derivatives of GSSG could be used to make a distinction between these two microspecies. Based on theory, it has been shown that similar problems arise in microequilibria of all ligands with more than three groups, unless they are totally symmetrical.[31]

For the possible utmost quantitation of the GSSG carboxylate basicities, we have set four models of calculation. In all models, eqs 23 and 26 were fitted to the experimental chemical shift vs pH data. In models A and B, as many microconstants are obtained as possible without any a priori assumption about basicities. These models yield some of the microconstants and microspecies concentrations only. In models C and D, the minimal a priori assumption is introduced to obtain the complete resolution of the microscopic protonation scheme.

3.6. Partial Resolution of Microequilibria with No Approximations (Models A and B). The only assumption used here is a generally accepted one, namely, that protonation of the glutamyl and glycyl carboxylates can selectively be monitored by the adjacent α-protons.[11,27,36] In model A, eqs 30 and 31 are used to calculate the protonation fractions f_U and f_Y. Seven log *Q* parameters of these functions were optimized by nonlinear regression (for results, see Tables 1S-3S in Supporting Information). From these log *Q* values, the cumulative microconstants and concentrations of microspecies *a*, *b*, *c*, *l*, *m*, and *p* could be unambiguously calculated (see Table 13S). For the remaining four microspecies (*f*, *g*, *h*, and *i*), however, the NMR-pH profiles fail to provide microconstants, since only two further parameters (Q_2 and Q_6) are available. Consequently, the mole fractions of macrospecies H_4L cannot be decomposed into those of the protonation isomers *f* through *k*. If this undefined fashion of the mathematical system is not recognized and calculation of κ_f, κ_g, κ_h, and κ_i is also attempted, correlated, linearly codependent parameter estimates are obtained, with large standard deviations (see Tables 4S and 5S). In such cases, principal component analysis (PCA) of the parameter correlation matrix helps identify the parameter combinations that cause excess degree(s) of freedom in the parametrization (Table 6S).[33,39] A similar analysis of the parameter interrelations, which is inevitable in microspeciation of large systems would not be feasible using the nonlinearly coupled *k*-form microconstants.

3.7. Complete Resolution of Microequilibria Using Interactivity Parameters (Models C and D). It has long been recognized that protonation at one site decreases the basicity of other site(s) in polydentate molecules.[24,27] This effect can be quantified for each pair of groups in terms of interactivity parameters E_{ij}. For example, for carboxylates U and Y of the GSSG molecule this term is given as

$$\log E_{UY} = \log k_Y^U - \log k^U = \log k_U^Y - \log k^Y \quad (32)$$

The interactivity parameter is assumed to be largely invariant in analogous moieties of different compounds and also in various protonation states of the neighboring moiety in the same molecule.[24] For GSSG, the three further interactivity parameters, $E_{UY'}$, $E_{UU'}$, and $E_{YY'}$ can be defined analogously. Each microconstant can then be made up as product of k^U or k^Y, the "core" microconstants, multiplied appropriately by one or more of the E_{UY}, $E_{UY'}$, $E_{UU'}$, $E_{YY'}$ parameters. For instance, microconstant $k_{UY}^{Y'}$ is a product of k^Y and two interactivity parameters, as follows:

$$k_{UY}^{Y'} = k^Y E_{UY'} E_{YY'} \quad (33)$$

The use of interactivity parameters reduces the nine-parameter problem into a six-parameter one. The seven known *Q* values and the six unknown parameters make the mathematical system redundant. This is, however, a necessary, but not sufficient precondition of obtaining chemically sound microconstants. The extraction of a unique set of microconstants and interactivity parameters also depends on how the computation is conditioned, i.e., how the experimental data and the concomitant *Q* values are distorted. The question, whether this latter, practical criterion is also met, is answered by the correlation matrix and principal component analysis.

Applying the interactivity parameters, the cumulative microconstants can be reformulated:

$$\kappa_f = \kappa_k = k^{\mathrm{U}}k^{\mathrm{Y}}E_{\mathrm{UY}} \tag{34}$$

$$\kappa_g = \kappa_j = k^{\mathrm{U}}k^{\mathrm{Y}}E_{\mathrm{UY'}} \tag{35}$$

$$\kappa_h = (k^{\mathrm{U}})^2E_{\mathrm{UU'}} \tag{36}$$

$$\kappa_i = (k^{\mathrm{Y}})^2E_{\mathrm{YY'}} \tag{37}$$

$$\kappa_l = \kappa_o = (k^{\mathrm{Y}})^2k^{\mathrm{U}}E_{\mathrm{UY}}E_{\mathrm{UY'}}E_{\mathrm{YY'}} \tag{38}$$

$$\kappa_m = \kappa_n = (k^{\mathrm{U}})^2k^{\mathrm{Y}}E_{\mathrm{UY}}E_{\mathrm{UY'}}E_{\mathrm{UU'}} \tag{39}$$

$$\kappa_p = (k^{\mathrm{U}}k^{\mathrm{Y}}E_{\mathrm{UY}}E_{\mathrm{UY'}})^2E_{\mathrm{YY'}}E_{\mathrm{UU'}} \tag{40}$$

These new "building blocks" were introduced into eqs 28 and 29. The subsequent, six-parameter fitting (model C) resulted in the following values: $\log k^{\mathrm{U}} = 2.42$, $\log k^{\mathrm{Y}} = 3.59$, $\log E_{\mathrm{UY}} = -0.13$, $\log E_{\mathrm{UY'}} = -0.1$, $\log E_{\mathrm{UU'}} = -0.08$, and $\log E_{\mathrm{YY'}} = -0.14$ (see Table 7S for details). The correlation matrix and its principal component analysis showed still an excess degree of freedom in the parametrization, caused by the E_{UY} and $E_{\mathrm{UY'}}$ parameters (Tables 8S and 9S). This result corresponds to our above statement about the impossibility to assess [*f*] and [*g*] separately.

In our model D, the $E_{\mathrm{UY}} = E_{\mathrm{UY'}}$ equality constraint allowed a five-parameter fitting that resulted in virtually the same statistically and chemically sound values: $\log k^{\mathrm{U}} = 2.42 \pm 0.02$, $\log k^{\mathrm{Y}} = 3.593 \pm 0.006$, $\log E_{\mathrm{UY}} = \log E_{\mathrm{UY'}} = -0.11 \pm 0.01$, $\log E_{\mathrm{UU'}} = -0.08 \pm 0.01$, and $\log E_{\mathrm{YY'}} = -0.14 \pm 0.01$ (see also Tables 10S−12S). Using these results, all microconstants could be unambiguously calculated and they are summarized in Table 1.

These microconstants quantitate that the glycyl carboxylates are some fifteen times more basic than their glutamyl counterparts in every protonation state of the molecule. Highly similar basicity difference has been reported for the GSH carboxylates: $\log k^{\mathrm{U}} = 2.42$, $\log k^{\mathrm{Y}} = 3.45$, $k^{\mathrm{Y}}/k^{\mathrm{U}} = 10.7$.[11] Our GSSG interactivity parameters demonstrate that no specific interaction exists between the four carboxylate sites. Rather, the interactions are weak, presumably of through-space, electrostatic nature.

3.8. Comparison of Microspecies-Specific Basicities with Those of a Model Compound. It has long been shown[29] that methyl- and ethyl-carboxylates (−COOR) closely mimic the basicity-modifying effects of the carboxyl (−COOH) group, while they greatly differ from the ionic ($-COO^-$) carboxylates. For such comparison, EtGSSGEt, a model compound with two glutamyl carboxylate and two glycyl-ester sites, has been synthesized. In accordance with Eber's principle,[40] the k^{U} and $k_{\mathrm{U}}^{\mathrm{U'}}$ microconstants of EtGSSGEt can be hypothesized to be analogous with the $k_{\mathrm{YY'}}^{\mathrm{U}}$ and $k_{\mathrm{YUY'}}^{\mathrm{U'}}$ microconstants of GSSG. We carried out the ^{1}H NMR-pH titration of EtGSSGEt (Figure 3). The nonlinear curve-fitting to the NMR-pH profiles resulted in the $\log k^{\mathrm{U}} = 2.19 \pm 0.02$ and $\log k_{\mathrm{U}}^{\mathrm{U'}} = 2.10 \pm 0.02$ microconstants. These values are in excellent agreement with the corresponding microconstants of GSSG (2.18 ± 0.02 and 2.12 ± 0.02), providing an independent proof of our experimental and calculation procedure.

3.9. Microscopic Speciation Curves of GSSG. The microscopic protonation constants allow the plotting of all the microspecies distribution curves (Figure 4). Symbols *y* and *x* denote microspecies holding zero and one proton at the amino site(s), respectively. The overwhelmingly predominant microspecies under physiological conditions (pH 7.40) is *a*, with two ammonium and four carboxylate sites (Figure 4). The relative concentrations of those microspecies, which hold more proton(s) at the glutamyl than the glycyl site(s), are well below 2%. Models B and D, being different both in the number of parameters and functional forms, do not have to necessarily yield the same cumulative microconstants. The fact that the corresponding log κ values are essentially identical between models B and D (see Table 13S) provides

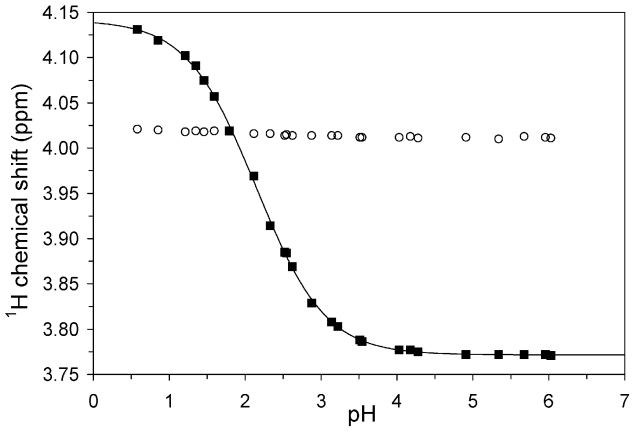


Figure 3. ^{1}H NMR titration curves of EtGSSGEt as a function of pH for the glutamyl methyne (■) and the glycyl methylene (○) protons.

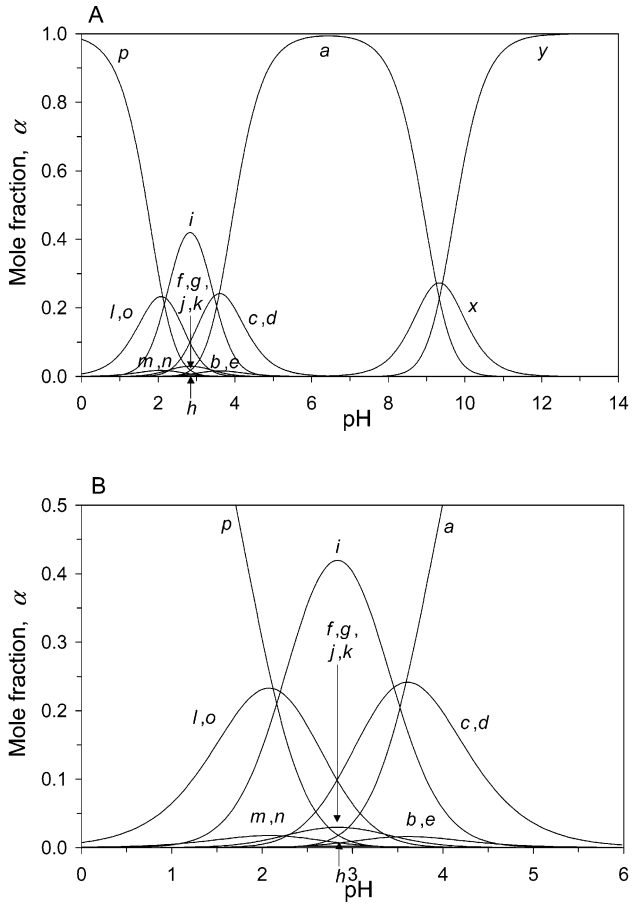


Figure 4. (A) Full microspeciation diagram of GSSG, based on the calculation by model B. For microspecies assignment see Scheme 1 and text. (B) The acidic pH range of the same diagram.

another justification of the use of the protonation stage invariant interactivity parameters, which was the key concept in models B and D.

3.10. Correlation of Protonation Constants by NMR and Potentiometry. Protonation macroconstants were also calculated from the NMR microconstants, using eqs 19−22. These values (K_{NMR}) are mixed constants, based on direct pH meter readings and involving hydrogen (deuterium) ion activity. Their conversion into concentration constants was not attempted. Nevertheless, the linearity with the pH-potentiometric K_{POT} macroconstants is excellent,

$$\log K_{NMR} = 0.11 + 1.00 \cdot \log K_{POT} \qquad (41)$$

with a correlation coefficient, $r = 0.9999$. This equation allows the conversion of the microconstants for H_2O/D_2O, 9:1 medium at 22 °C into those for 100% H_2O and 25 °C, and also a more direct comparison with results of studies in biological, analytical, and coordination chemistry.

4. Conclusions

The first complete microequilibrium analysis of a tetrabasic acid, exemplified by GSSG, the vitally important biomolecule of C_{2v} symmetry, resulted in 18 different microspecies concentrations and 18 different microscopic protonation constants. The two major conclusions drawn of the study are as follows. (1) A complete microspeciation for ligands of more than three binding sites and low symmetry is not only highly complex and technically difficult, but it is, in principle, impossible without a priori simplifying assumptions. Introducing, however, a minimal assumption on the carboxylate interactions, the microscopic protonation scheme could be completely resolved, without the use of any imported, auxiliary basicity data. (2) The 2 amino- and the 16 carboxylate microconstants indicate that the intrinsic basicity of these groups is virtually the same as in GSH, the reduced glutathione. This finding can readily be interpreted in terms of the large number of covalent bonds, isolating the proton-binding sites and the slight electrostatic interaction between them. These site-specific basicity data can be of predictive or interpretive value for studies on specific association reactions with other biomolecules.

Acknowledgment. The authors are grateful to Imre Tóth, Zoltán Berente (Kossuth University, Debrecen, Hungary), István Kövesdi (EGIS Pharmaceutical Works, Budapest), and Antal Csámpai (Eötvös University, Budapest) for their valuable contributions. Z. Sz. thanks Ernö Keszei for several inspiring discussions. This research was supported by grants OTKA T17570, MKM FKFP 1126/97, and NjM 447/96.

Supporting Information Available: Tables 1S−12S: detailed statistical information regarding the reliability of the estimated parameters (standard deviations, confidence bounds, parameter correlation matrix and its principal component analysis) in models A, B, C, and D, respectively; Table 13S: comparison of the cumulative microconstants (log κ), as calculated with model A. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Larsson, A.; Orrhenius, S.; Holmgrew, A.; Mannervik, B. *Functions of Glutathione*; Raven: New York, 1983.
(2) Proulx, M.; Souich, P. D. *J. Pharm. Pharmacol.* **1995**, *47*, 392−397.
(3) Radkowsky, A. E.; Kosower, E. M. *J. Am. Chem. Soc.* **1986**, *108*, 4527.
(4) Santala, T.; Fishbein, J. C. *J. Am. Chem. Soc.* **1992**, *114*, 8852.
(5) Novak, M.; Lin, J. *J. Am. Chem. Soc.* **1996**, *118*, 1302.
(6) Rabenstein, D. L.; Theriault, Y. *Can. J. Chem.* **1984**, *62*, 1672.
(7) Rabenstein, D. L.; Theriault, Y. *Can. J. Chem.* **1985**, *63*, 33.
(8) Keire, D. A.; Strauss, E.; Guo, W.; Noszál, B.; Rabenstein, D. L. *J. Org. Chem.* **1992**, *57*, 123.
(9) Rabenstein, D. L.; Yeo, P. L. *J. Org. Chem.* **1994**, *59*, 4223.
(10) Weaver, K. H.; Rabenstein, D. L. *J. Org. Chem.* **1995**, *60*, 1904.
(11) Rabenstein, D. L. *J. Am. Chem. Soc.* **1973**, *95*, 2797.
(12) Podányi, B.; Reid, R. S. *J. Am. Chem. Soc.* **1988**, *110*, 3805.
(13) Cheesman, B. V.; Arnold, A. P.; Rabenstein, D. L. *J. Am. Chem. Soc.* **1988**, *110*, 6359.
(14) Shoukry, M. M.; Cheesman, B. V.; Rabenstein, D. L. *Can. J. Chem.* **1988**, *66*, 3184.
(15) Neshvad, G.; Hoffman, M. Z. *J. Phys. Chem.* **1989**, *93*, 2445.
(16) Khan, N. M.; Malspeis, L. *J. Org. Chem.* **1982**, *47*, 2731.
(17) Patel, H. M. S.; Williams, D. L. H. *J. Chem. Soc., Perkin Trans. 2* **1990**, *1*, 37.
(18) Bjerrum, N. *Z. Phys. Chem., Stoechiom. Verwandtschaftsl.* **1923**, *106*, 219.
(19) Martin, R. B. *Introduction to Biophysical Chemistry*; McGraw-Hill: New York, 1964.
(20) Noszál, B. *J. Phys. Chem.* **1986**, *90*, 4104.
(21) Noszál, B. In *Biocoordination Chemistry*; Burger, K., Ed.; Ellis Horwood: Chichester, 1990, Chapter 11.
(22) Arendt, C.; Hägele, G. *Comput. Chem.* **1995**, *19*, 263.
(23) Hägele, G.; Holzgrabe, U. In *NMR Spectroscopy in Drug Development and Analysis*; Holzgrabe, U., Wavez, I., Dichl, B., Eds.; Wiley-VCH: Weinheim, 1999; pp. 61−76.
(24) Santos, M. A.; Esteves, M. A.; Vaz, M. C.; Fraústo da Silva, J. J. R.; Noszál, B.; Farkas, E. *J. Chem. Soc., Perkin Trans. 2* **1997**, 1977.
(25) Szilágyi, L.; Pusztahelyi, S. Z.; Jakab, S.; Kovács, I. *Carbohydr. Res.* **1993**, *247*, 99.
(26) Mazák, K.; Nemes, A.; Noszál, B. *Pharmacol. Res.* **1999**, *16*, 1757.
(27) Mernissi-Arifi, K.; Schmitt, L.; Schlewer, G.; Spiess, B. *Anal. Chem.* **1995**, *67*, 2567.
(28) D'Angelo, J. C.; Colette, T. W. *Anal. Chem.* **1997**, *69*, 1642.
(29) Martin, R. B. *J. Phys. Chem.* **1961**, *65*, 2053.
(30) Zekri, O. O.; Boudeville, P.; Genay, P.; Perly, B.; Braquet, P.; Jouenne, P.; Burgot, J. L. *Anal. Chem.* **1996**, *68*, 2598.
(31) Szakács, Z.; Noszál, B. *J. Math. Chem.* **1999**, *26*, 139.
(32) Irving, H. M.; Miles, M. G.; Pettit, L. D. *Anal. Chim. Acta* **1967**, *38*, 475.
(33) Valkó, P.; Vajda, S. *Advanced Scientific Computing in BASIC with applications in Chemistry, Biology and Pharmacology*; Elsevier: Amsterdam, 1989.
(34) *Critical Stability Constants*; Martell, A. E., Smith, R. M., Eds.; Plenum: New York, 1974; Vol. 1., p 404.
(35) Kozlowski, H.; Várnagy, K.; Sóvágó, I. *Polyhedron* **1990**, *9*, 831.
(36) Rabenstein, D. L.; Sayer, T. L. *Anal. Chem.* **1976**, *48*, 1141.
(37) Rabenstein, D. L.; Keire, D. A. In *Glutathione: Chemical, Biochemical and Medical Aspects*, Dolphin, D., Poulson R., Avramovic, O., Eds.; Wiley: New York, 1989; pp. 67−99.
(38) Sudmeier, J. L.; Reilly, C. N. *Anal. Chem.* **1964**, *36*, 1698.
(39) Chatfield, C.; Collins, A. J. *Introduction to Multivariate Analysis*; Chapman and Hall: New York, 1980; p. 61.
(40) Ebert, L. *Z. Phys. Chem.* **1926**, *121*, 385.