ELSEVIER

Biophysical Chemistry 88 (2000) 35–45

Biophysical Chemistry

www.elsevier.nl/locate/bpc

A simple chemical example of hierarchical thermodynamic interactions: the protonation equilibria of inorganic polyprotic acids

William G. Gutheil*

Division of Pharmaceutical Sciences, School of Pharmacy, University of Missouri-Kansas City, 5005 Rockhill Road, Kansas City, MO 64110, USA

Received 23 May 2000; received in revised form 25 July 2000; accepted 25 July 2000

Abstract

A general method for formulating complex thermodynamic systems in terms of hierarchical interactions has been developed, and has been applied in a previous analysis to hemoglobin oxygen binding data. Polyprotic acids can be considered a simple chemical model of thermodynamic interaction between ligand binding events. To further illustrate the hierarchical interaction approach it is applied to the analysis of the thermodynamic interactions between proton binding events in inorganic polyprotic acids. pK values for arsenate, carbonate, chromate, phosphate, phosphite, selenite, sulfide and sulfite were recast into hierarchical interaction terms. The intrinsic $K_{\mathrm{d},h}$ for protonation ranged from 8.8×10^{-13} (M) for phosphate to 1.3×10^{-6} (M) for chromate. Pairwise interactions ($K_{\mathrm{d},hh}$) between protonation events ranged from 1.3×10^{4} for phosphite to 9.4×10^{5} for carbonate. Third order interactions ($K_{\mathrm{d},hhh}$) were 0.91 and 0.51 for arsenate and phosphate, respectively, values relatively close to the no interaction value of 1. A principle feature of systems described by hierarchical interactions is that higher order interactions, representing more complex interactions, are less likely to be significant than lower order interactions, and this is further illustrated by these observations from polyprotic acids. The set of significant hierarchical interaction values can be used to predict values for as yet unobserved events, and projected pK values are made for all the polyprotic acids included in this study. Finally, application of this method to the protonation equilibria of water demonstrates a profound pairwise interaction between protonation events ($K_{\mathrm{d},hh} = 1.3 \times 10^{17}$), which is attributed to oxygen's small size and lack of polarizability. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Thermodynamics; Hierarchical interactions; Acids; Models/chemical; Allosterism

---

* Tel.: +1-816-235-2424; fax: +1-816-235-5190.
E-mail address: gutheilw@umkc.edu (W.G. Gutheil).

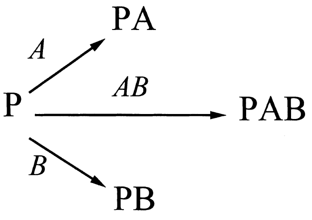
PII: S0301-4622(00)00198-8

1. Introduction

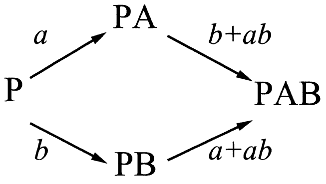
The inherent complexity of thermodynamic systems containing multiple interacting events has posed an impediment to the complete analysis of such systems since the advent of modern thermodynamic theory. An approach to such systems has been developed which explicitly treats the full *potential* complexity of a complex system as a hierarchy of interactions. The concept of hierarchical interactions is an extension of the well known concept of pairwise interactions [1–3] as illustrated in Scheme 1. This concept introduces a term, *ab*, describing the pairwise interaction between two events. The fact that the ΔG° values describing the assembly of the different complexes is a function of state (path independent) requires that the term *ab* occur on both pathways leading to the PAB complex in Scheme 1B. The concept of pairwise interaction cannot be used to treat the full potential complexity of systems with more than two interacting events. In order to extend this concept to the general case of N interacting ligands it is necessary to introduce the concept of higher than second order interactions [4,5]. The hallmark of this method is to take a system described by N ΔG° values of assembly and to recast them into N ΔG° values describing the hierarchy of potential interactions in the system.

PAB System

(a) Formulation in terms of free energies of assembly from free P, A, and B.



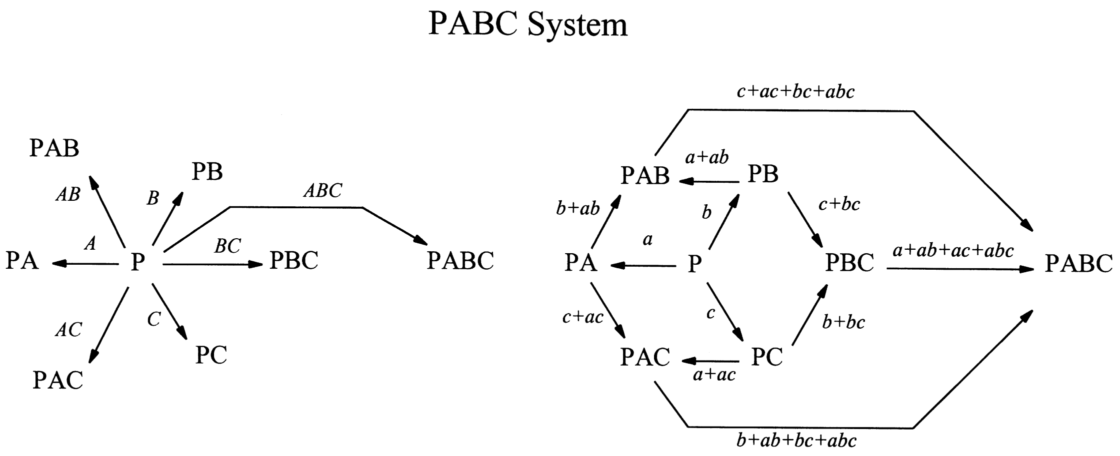
(b) Formulation using concept of interaction free energy.



Scheme 1. Alternative ways for formulating the PAB system [4]. (a) Uppercase letters are used to denote ΔG° values of assembly. (b) Lowercase italicized letters are used to denote hierarchical interaction ΔG° values. The free ligands A and B are not explicitly indicated in this scheme.

An example of a three ligand binding protein formulated using this method is shown in Scheme 2 [4]. The term *abc* is a third order interaction and represents the mutual interaction between the three ligands A, B and C. The term *abc* can be viewed equivalently as the effect of A binding on how B and C interact, the effect of B binding on how A and C interact, and the effect of C binding on how A and B interact (Scheme 3). The *abc* term is hierarchically analogous to the *ab* term used in Scheme 1, since *ab* represents equivalently the effect of A binding on B binding, and vice versa. Also note that the system shown in Scheme 2 has seven complexes, and, therefore, requires seven ΔG° values to describe its full potential complexity, a requirement which is satisfied using both ΔG° values of assembly and hierarchical interaction ΔG° values. There is a one-to-one transformation between ΔG° values of assembly and hierarchical interaction ΔG° values as illustrated in the bottom of Scheme 2 for the PABC system.

Casting a complex state system in terms of hierarchical interactions has significant conceptual, theoretical, and statistical advantages over alternative methods for formulating such systems. These advantages arise from the fact that higher order terms reflect increasingly complex interactions between successive events, and may or may not be physically significant. For example, the *abc* term (Scheme 2) represents the mutual interaction between the binding of A, B, and C. For the *abc* term to exist physically requires: (1) that the three second order interactions which it modulates (*ab*, *ac*, and *bc*) to exist physically, and (2) for the actual *abc* interaction to exist physically. If, for example, no interaction can occur between A and B, then both *ab* (the interaction between A and B) and *abc* (the effect of C on the interaction between A and B) will be equal to zero.

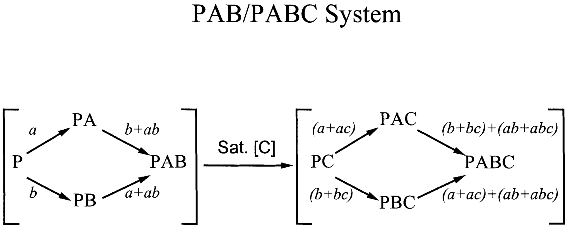


$A = a$
$B = b$
$C = c$
$AB = a+b+ab$
$AC = a+c+ac$
$BC = b+c+bc$
$ABC = a+b+c+ab+ac+bc+abc$

$a = A$
$b = B$
$c = C$
$ab = AB\text{-}A\text{-}B$
$ac = AC\text{-}A\text{-}C$
$bc = BC\text{-}B\text{-}C$
$abc = ABC\text{-}AB\text{-}AC\text{-}BC+A+B+C$

Scheme 2. Formulation of PABC system (seven complexes) in terms of seven ΔG^{0} values of assembly (top left) and in terms of seven hierarchical interaction formulated ΔG^{0} values (top right), and the algebraic definitions for the relationships between the parameters defining the two formulations (below) illustrating the one-to-one nature of this reformulation. Described in detail in Gutheil and McKenna [4].

Rules have been developed for reformulating complex systems in terms of hierarchical interactions [4,5]. The algebraically defined equilibrium constant equivalents to hierarchical interaction reformulated ΔG^{0} values can be used to algebraically analyze complex systems, which allow theoretical questions concerning the behavior of such systems to be appropriately framed and solved, as demonstrated by the theoretical analysis of cooperativity in a hypothetical dimeric protein [6], and the derivation of the mathematical relationships between monomer and multimer models of protein ligand binding isotherms [5]. Hierarchical interaction terms can also be used as stepwise regressors in the statistical analysis of data, as demonstrated by the analysis of hemoglobin oxygen binding data [7]. In this application higher order terms are incrementally included in the fitting of data, and the significance of the higher order terms determined by appropriate statistical tests. This analysis demonstrated that pairwise interactions were both necessary and sufficient to account for the oxygen binding



Scheme 3. The PABC system from Scheme 2 redrawn to emphasize effect of saturation with C on the observed interaction between A and B [4]. Note that the interaction between A and B with C bound on the right $(ab + abc)$ differs from the interaction between A and B on the left (ab) by the amount abc. Physically, abc represents the effect of A binding on the interaction between B and C, and the effect of B binding on the interaction between A and C, and the effect of C binding on the interaction between A and B (as illustrated here).

properties of hemoglobin. The conceptual approach of hierarchical reformulation can be applied to any state parameter, as demonstrated by the use of hierarchically reformulated extinction coefficients for the analysis of spectroscopically monitored hemoglobin oxygen binding data [7].

Interactions between proton binding events in organic polyprotic acids have been used to study electrostatic, dipolar and conformational effects between protons in such simple systems [8,9]. More recently these same systems have received attention as model systems for site specific thermodynamic interactions in more complex systems such as interactions between thermodynamic events in proteins, where site specific refers to knowledge of the behavior of a specific site when studied in the context of other sites [10]. These analyses have generally treated interactions as pairwise interactions, however, the possibility for more complex behavior, as revealed by third and higher order interactions, exists for polyprotic acids as it does for protein binding events. The application of the concept of hierarchical interactions to the analysis of the site specific thermodynamic properties of chemically simple polyprotic acids would further demonstrate this concept and its utility in analyzing interactions in 'complex' systems. This method is, therefore, used here to analyze the interactions between successive protonation events which occur in inorganic polyprotic acids.

2. Materials and methods

2.1. pK values

pK_d (dissociation) values for arsenate, carbonate, chromate, phosphate, phosphite, pyrophosphate, selenite, sulfide and sulfite were obtained from the *CRC Handbook of Chemistry and Physics* [11]. The apparent second pK_d for carbonate of 6.40 is composed of contributions from the $H_2CO_3 \rightleftarrows H_2O + CO_2$ equilibrium and the true pK_d for $H_2CO_3 \rightleftarrows H^+ + HCO^{3-}$ of 3.80 [12]. The true pK_d of 3.80 is used here.

2.2. Conversion to hierarchical interaction terms

Systematic rules for formulating complex systems in terms of hierarchical interactions have been developed [4,5], and will be summarized here for the present analysis. For multiple identical binding sites a statistical correction factor (SCF) is used to relate the macroscopically apparent behavior of the system to the microscopic properties of the individual interactions. For example, for a polyvalent anion such as phosphate there are four equivalent oxygens any one of which can associate with a proton to yield the singly protonated intermediate. If one oxygen has an intrinsic association equilibrium constant of K_i for binding a proton,

$$H^+ + \sim O^- \rightleftarrows \sim OH \quad K_i = \frac{[\sim OH]}{[H^+][\sim O^-]} \tag{1}$$

then for four such oxygens on PO_4^{3-} an apparent macroscopic association constant of $4K_i$ would be observed

$$H^+ + PO_4^{3-} \rightleftarrows HPO_4 \quad K_m = \frac{[HPO_4]}{[H^+][PO_4]} = 4K_i \tag{2}$$

This relationship is well known and easily demonstrated algebraically [13]. When working with ΔG^o values SCFs show up as $-RT\ln(\mathrm{SCF})$ terms, i.e.

$$\begin{aligned}\Delta G_m^o &= -RT\ln(K_m) = -RT\ln(4K_i)\\ &= -RT\ln(K_i) - RT\ln(4)\\ &= \Delta G_i^o - RT\ln(4)\end{aligned} \tag{3}$$

SCFs can be considered the entropic effect of multiple binding sites. When defining a system it is our convention to always use intrinsic constants and SCFs. The SCF for a step can easily be determined as the number of ways forward divided by the number of ways back. For example for phosphate with four oxygens to which a proton can bind we would have

$$\mathrm{PO_4^{3-}} + \mathrm{H^+} \overset{\mathrm{SCF}=4/1}{\rightleftarrows} \mathrm{HPO_4^{2-}} + \mathrm{H^+} \overset{\mathrm{SCF}=3/2}{\rightleftarrows} \mathrm{H_2PO_4^{1-}} + \mathrm{H^+} \overset{\mathrm{SCF}=2/3}{\rightleftarrows} \mathrm{H_3PO_4} \left(+ \mathrm{H^+} \overset{\mathrm{SCF}=1/4}{\rightleftarrows} \mathrm{H_4PO_4^{1+}} \right) \tag{4}$$

For the first step there are four ways forward and one way back. For the second step there are three ways forward and two ways back, etc. The final step is shown in parentheses since, to our knowledge, a fourth protonation for phosphate has not been observed but appears possible under sufficiently acidic conditions.

The second aspect of reformulating a complex system is the concept of hierarchical interactions [4]. The reformulation rule is that the association of the new ligand at a particular step is represented by a term equal to the primary (first order) term plus combinations of the new ligand with already bound ligands, first one at a time, then two at a time, etc., up to all the already bound ligands, and finally minus the statistical correction term $[-RT\ln(\mathrm{SCF})]$ for the step. To reformulate the thermodynamic model for example for phosphate (Table 1B), the first protonation has a macroscopic ΔG° (ΔG_1°) of h (the base term describing the intrinsic affinity of one of the oxygens for a proton) $-RT\ln 4$ (the SCF term for multiple oxygens). The second protonation step (ΔG_2°) is equal to the base term h, plus a term hh describing the possible pairwise (second order) interaction between the first and second proton binding event, minus the SCF term, $-RT\ln 3/2$. The third protonation step (ΔG_3°) is described by the base term h plus possible pairwise interactions with each of the already bound protons $2hh$, plus a third order interaction term of the incoming proton with the two already bound protons hhh, and finally minus the SCF term for the step, $-RT\ln 2/3$. A fourth protonation step (ΔG_4°) is possible on the remaining free oxygen (not previously observed to my knowledge) which is described by the base term h plus pairwise interactions with each of the already bound protons $3hh$, plus three third order interactions $3hhh$ (there are three ways of grouping the incoming proton with any two of the three already bound protons), plus a fourth order interaction term $hhhh$, minus the SCF term for the step, $-RT\ln 1/4$. This yields the model shown in Table 1B which is defined both in terms of hierarchical interaction free energies (ΔG° values) and corresponding hierarchical interaction K values.

2.3. Calculation of hierarchical interaction parameter values from pK_d values for phosphate

The hierarchically reformulated model for phosphate is shown in Table 1B, the unreformulated (stepwise) parameters describing the protonation of phosphate are summarized in Table 1C, and the calculations required to convert the unreformulated parameters into a corresponding set of hierarchically reformulated constants are summarized in Table 1D.

2.4. Calculation of the projected pK value for the fourth protonation of phosphate

Phosphate has four equivalent oxygen atoms, and it appears possible for a fourth protonation to occur to yield the cation $\mathrm{H_4PO_4^+}$ $[\mathrm{P(OH)_4^+}]$. As discussed above, an important feature of hierarchically reformulated parameters is that higher order terms are less likely to be physically and/or statistically significant than lower order terms. This is illustrated for phosphate where the third order dissociation interaction equilibrium constant has a value of $K_{\mathrm{d},hhh} = 0.51$, which is nearly equal to its null (no interaction) value of 1 (corre-

Table 1

Conversion of K_d values for phosphate protonations into hierarchical interactions[a]

(A) *Constants*

$T = 25°C$

$R = 8.314\ \mathrm{J/mol\ K}$

$RT = 2.479\ \mathrm{kJ/mol}$

(B) *Hierarchically reformulated model*

$$\mathrm{PO_4^{3-}} \quad \frac{h-RT\ln 4}{1} > \mathrm{HPO_4^{2-}} \quad \frac{h+hh-RT\ln 3/2}{2} > \mathrm{H_2PO_4^{-}} \quad \frac{h+2hh+hhh \atop -RT\ln 2/3}{3} > \mathrm{H_3PO_4} \quad \frac{h+3hh+3hhh \atop +hhhh-RT\ln 1/4}{4} > \mathrm{H_4PO_4^{+}}$$

(C) *Unreformulated parameters*

K_d (M)	2.20E − 13	6.23E − 08	7.52E − 03	?
K_a (M^{-1}) $= 1/K_d$	4.55E + 12	1.61E + 07	1.33E + 02	?
$pK_d = -\log(K_d)$	12.66	7.21	2.12	?
$\Delta G^o = -RT\ln(K_a)$	− 72.25	− 41.13	− 12.12	?

(D) *Calculation of hierarchical parameter values*

ΔG^o_1	$= h - RT\ln 4$	$= -72.25$	
h[b]	$= \Delta G^o_1 + RT\ln 4$	$= -72.25 + 3.44$	$= -68.81$
$K_{a,h}$[c]	$= \exp(-h/RT)$	= 1.14E12 (M^{-1})	
$K_{d,h}$[d]	$= 1/K_{a,h}$	= 8.80E − 13 (M)	
$pK_{d,h}$	$= -\log(K_{d,h})$	= 12.06	
ΔG^o_2	$= h + hh - RT\ln 3/2$	$= -41.13$	
hh	$= \Delta G^o_2 - h + RT\ln 3/2$	$= -41.13 + 68.81 + 1.01$	$= 28.69$
$K_{a,hh}$	$= \exp(-hh/RT)$	= 9.42E − 6 (unitless)	
$K_{d,hh}$	$= 1/K_{a,hh}$	= 10600 (unitless)	
$pK_{d,hh}$	$= -\log(K_{d,hh})$	$= -5.03$	
ΔG^o_3	$= h + 2hh + hhh - RT\ln 2/3$	$= -12.12$	
hhh	$= \Delta G^o_3 - h - 2hh + RT\ln 2/3$	$= -12.12 + 68.81 - 2(28.69) - 1.01$	$= -1.69$
$K_{a,hhh}$	$= \exp(-hhh/\mathrm{RT})$	= 1.98 (unitless)	
$K_{d,hhh}$	$= 1/K_{a,hhh}$	= 0.51 (unitless)	
$pK_{d,hhh}$	$= -\log(K_{d,hhh})$	= 0.30	

E. *Calculation of projected value for pK_{d4}*

if $hhhh$	$= 0$ ($K_{hhhh} = 1$)		
ΔG^o_4	$= h + 3hh + 3hhh + hhhh - RT\ln 1/4$	$= -68.81 + 3(28.69) - 3(1.69) + 0 + 3.44$	$= 15.61$
K_{a4}	$= \exp(-\Delta G^o_4/RT)$	= 1.84E − 3 (M^{-1})	
K_{d4}	$= 1/K_{a4}$	= 543 (M)	
pK_{d4}	$= -\log(K_{d4})$	$= -2.74$	

[a] All ΔG^o values (including hierarchical interaction ΔG^o values) given in units of kJ/mol. Thermodynamic equilibrium constants are formally unitless. However, practical units do exist for equilibrium constants expressions [15] and are given in this Table in parentheses.

[b] Reformulated ΔG^o values are denoted with lowercase italicized letters.

[c] Reformulated equilibrium association constant.

[d] Reformulated equilibrium dissociation constant.

sponding to a ΔG° of 0), especially when compared with the second order interaction term value of $K_{d,hh} = 1.06 \times 10^5$. It is expected that the fourth order interaction term ($hhhh$) will have a value even closer to its null value. Based on this premise it is possible to project a best estimate value for the pK_d of the fourth protonation event. This calculation is summarized in Table 1E.

2.5. Hierarchically reformulated parameter values for other polyprotic acids, and projected pK values

The values for the reformulated parameters for the series of polyprotic acids listed in Table 2 were calculated from known pK values analogously as shown in Table 1 for phosphate, and are summarized in Table 2. Projected pK values are also summarized in Table 2.

2.6. Hierarchically reformulated model for the protic equilibria of water

Two pK_d values are known for water, a pK_d of -1.74 for protonation of water and a pK_d of 15.74 for protonation of the hydroxide ion [14]. In contrast to the other polyprotic acids treated above, the value for the basic interaction for $O^{2-} + H^+$, defining the base term h for this system, is not known independently. The method of hierarchical reformulation can be applied to situations of this type and this analysis is summarized in Table 3. In this analysis collections of hierarchically reformulated parameters are simply treated in the same way as individual parameters (cf. Scheme 3), for example the term $(h + hh)$, a combination of a first and second order term, can be considered as the apparent first order term (h^{app}) for the observable $HO^- \rightleftarrows H_2O \rightleftarrows H_3O^+$ equilibria. In this case $(hh + hhh)$ appears as the apparent second order term (hh^{app}) in this system, and $(hhh + hhhh)$ appears as the apparent third order interaction term (hhh^{app}). For example the term for the fourth protonation step in Table 3 can be rewritten as $(h + hh) + 2(hh + hhh) + (hhh + hhhh) \; -RT\ln 4$ corresponding to $h^{\text{app}} + 2hh^{\text{app}} + hhh^{\text{app}} - RT\ln 4$. Values for the aggregate terms $(h + hh)$ and $(hh + hhh)$ can therefore be derived from the known second and third pK_d values for water as shown in Table 3D. Best estimates for values of individual constants can be obtained if the third order term hhh is set equal to zero. This allows the projected values for hh, and then for h, to be derived (Table 3E). Finally, if $hhh = 0$ then the projected value for h provides an estimate for pK_{d1} for water, and if $hhh = 0$ and $hhhh = 0$, then a value for pK_{d4} can also be estimated (Table 3F).

3. Results and discussion

The results of the application of the method of hierarchical reformulation to the analysis of interactions in inorganic polyprotic acid equilibria are summarized in Table 2. Of the two acids with three available pK_d values, arsenate and phosphate, both clearly demonstrate small third order interactions. This observation demonstrates that pairwise interactions between protons dominate the observed equilibria. That third order interactions (hhh) are relatively insignificant demonstrates that one proton does not greatly influence how any two other protons interact. This result is expected for a rigid molecular system dominated by electrostatic interactions between protons. The polyprotic acids listed in Table 2 demonstrate a range of pairwise (hh) interaction values between subsequent protonations, with values for p$K_{d,hh}$ ranging from -4 to -6. Most of the acids listed in Table 2 are oxy-acids, and among these there is a rough correlation between the size of the central atom and the magnitude of the interaction, with carbonate being the smallest and showing the highest interaction in the series. However, it appears other factors are at work than simply distance since sulfide, where the incoming protons are expected to be in relatively close proximity to one another, shows substantially less interaction than for sulfite, where the incoming protons are further apart due to the intervening oxygens. The nature (polarizability) of the ligating atom (i.e. oxygen or sulfur) therefore plays an important role in determining the magnitude of interaction. These effects are dramatically illustrated by the profound interaction between subsequent oxygen binding events observed for water

Table 2
Summary of results for polyprotic acids[a]

Arsenate (n = 4)			
$pK_{d1} = 11.40$	$h = -60.20$	$K_{d,h} = 1.58E-11$	$pK_{d,h} = 10.80$
$pK_{d2} = 6.77$	$hh = 23.45$	$K_{d,hh} = 1.61E4$	$pK_{d,hh} = -4.21$
$pK_{d3} = 2.25$	$hhh = -0.23$	$K_{d,hhh} = 0.91$	$pK_{d,hhh} = 0.04$
Projected: $pK_{d4} = -2.30$			
Carbonate ($n = 3$)			
$pK_{d1} = 10.25$	$h = -55.79$	$K_{d,h} = 1.68E-10$	$pK_{d,h} = 9.77$
$pK_{d2} = 3.80$	$hh = 34.10$	$K_{d,hh} = 9.42E5$	$pK_{d,hh} = -5.97$
Projected: $pK_{d3} = -2.65$			
Chromate ($n = 4$)			
$pK_{d1} = 6.49$	$h = -33.63$	$K_{d,h} = 1.28E-6$	$pK_{d,h} = 5.89$
$pK_{d2} = 0.74$	$hh = 30.39$	$K_{d,hh} = 2.11E5$	$pK_{d,hh} = -5.32$
Projected: $pK_{d3} = -4.93$, $pK_{d4} = -10.68$			
Phosphate ($n = 4$)			
$pK_{d1} = 12.66$	$h = -68.81$	$K_{d,h} = 8.80E-13$	$pK_{d,h} = 12.06$
$pK_{d2} = 7.21$	$hh = 28.69$	$K_{d,hh} = 1.06E5$	$pK_{d,hh} = -5.03$
$pK_{d3} = 2.12$	$hhh = -1.69$	$K_{d,hhh} = 0.51$	$pK_{d,hhh} = 0.30$
Projected: $pK_{d4} = -2.74$			
Phosphite ($n = 3$)			
$pK_{d1} = 6.58$	$h = -34.04$	$K_{d,h} = 7.70E-7$	$pK_{d,h} = 6.10$
$pK_{d2} = 2.00$	$hh = 22.90$	$K_{d,hh} = 1.28E4$	$pK_{d,hh} = -4.11$
Projected: $pK_{d3} = -2.59$			
Selenite ($n = 3$)			
$pK_{d1} = 7.30$	$h = -38.03$	$K_{d,h} = 1.50E-7$	$pK_{d,h} = 6.82$
$pK_{d2} = 1.46$	$hh = 29.92$	$K_{d,hh} = 2.33E5$	$pK_{d,hh} = -5.37$
Projected: $pK_{d3} = -4.39$			
Sulfide ($n = 4$)			
$pK_{d1} = 11.96$	$h = -63.30$	$K_{d,h} = 4.40E-12$	$pK_{d,h} = 11.36$
$pK_{d2} = 7.04$	$hh = 25.04$	$K_{d,hh} = 3.10E4$	$pK_{d,hh} = -4.49$
Projected: $pK_{d3} = 2.20$, $pK_{d4} = -2.72$			
Sulfite ($n = 3$)			
$pK_{d1} = 6.99$	$h = -36.31$	$K_{d,h} = 3.06E-7$	$pK_{d,h} = 6.51$
$pK_{d2} = 1.68$	$hh = 26.90$	$K_{d,hh} = 6.69E4$	$pK_{d,hh} = -4.83$
Projected: $pK_{d3} = -3.31$			

[a] In the first column are pK values from the *CRC Handbook* [11], and subsequent columns are reformulated values calculated as in Table 1. The value of n for each acid denotes the number of protonation sites and is used in calculating SCFs. Projected values given in the last line for each acid were calculated from the reformulated parameters using the same procedure as outlined in Table 1 for phosphate. All the $K_{d,h}$ values in this Table have practical units of M.

($pK_{d,hh} = -17$, Table 3), and can be rationalized as a consequence both of oxygens small size and lack of polarizability.

Using the pairwise interaction values obtained with a given polyprotic acid and the premise that third and higher order interactions are relatively insignificant, as demonstrated by arsenate and phosphate (Table 2), allows pK_d values to be

Table 3
Analysis of the protonation equilibria of water in terms of hierarchical interactions[a]

A. *Constants*
$T = 25°\text{C}$
$R = 8.314\ \text{J/mol K}$
$RT = 2.479\ \text{kJ/mol}$

B. *Hierarchically reformulated model*

$$O^{2-} \quad \frac{h - RT\ln 4}{1} > HO^{-} \quad \frac{h + hh - RT\ln 3/2}{2} > H_2O \quad \frac{h + 2hh + hhh - RT\ln 2/3}{3} > H_3O^{-} \quad \frac{h + 3hh + 3hhh + hhhh - RT\ln 1/4}{4} > H_4O^{2+}$$

C. *Unreformulated parameters*

K_d (M)	?	1.82E − 16	55.0	?
K_a (M^{-1}) $= 1/K_d$	?	5.50E15	1.82E − 2	?
$pK_d = -\log(K_d)$	?	15.74	− 1.74	?
$\Delta G^o = -RT\ln(K_a)$	?	− 89.84	9.93	?

D. *Calculation of aggregate hierarchical parameter values*

$\Delta G^o{}_2$	$= h + hh - RT\ln 3/2$	$= -89.84$	
$(h + hh)$	$= \Delta G^o{}_2 + RT\ln 3/2$	$= -89.84 + 1.01$	$= -88.83$
$\Delta G^o{}_3$	$= h + 2hh + hhh - RT\ln 2/3$	$= (h + hh) + (hh + hhh) - RT\ln 2/3$	$= 9.93$
$(hh + hhh)$	$= \Delta G^o{}_3 - (h + hh) + RT\ln 2/3$	$= 9.93 + 88.83 - 1.01$	$= 97.76$

E. *Calculation of projected values for hh and h*
if $hhh = 0$

hh	$= (hh + hhh) - hhh$	$= 97.76 - 0$	$= 97.76$
$K_{d,hh}$	$= 1.32$ E17		
$pK_{d,hh}$	$=$	-17.12	
h	$= (h + hh) - hh$	$= -88.83 - 97.76$	$= -186.59$
$K_{d,h}$	$= 2.04$ E33		
$pK_{d,h}$	$= 32.69$		

F. *Calculation of projected values for pK_{d1} and pK_{d4}*
if $hhh = 0$

$\Delta G^o{}_1$	$= h - RT\ln 4$	$= -186.59 - 3.44$	$= -190.03$
K_{d1}	$= 5.13$ E34		
pK_{d1}	$= 33.29$		

if $hhh = 0$ and if $hhhh = 0$

$\Delta G^o{}_4$	$= h + 3hh + 3hhh + hhhh - RT\ln 1/4$	$= -186.59 + 3(97.76) - 3(0) + 0 + 3.44$	$= 110.12$
pK_{d4}	$= -19.29$		

[a] See footnote to Table 1.

projected for previously unobserved or unobservable pK_d values. These are summarized in Table 2 for polyprotic acids and in Table 3 for water. Some of the pK_d values projected in Table 2 may be within the realm of measurement. Determination of these projected values would either provide additional confirmation for the premise that third and higher order interactions are insignificant for the polyprotic acids, or demonstrate third and higher order interactions are significant and provide some basis for rationalizing these higher order interactions. (In some cases, e.g. carbonate, an additional protonation may result in a change in coordination, and the projected value should not in such cases be considered a reasonable expectation for the pK_d.) Sulfide is of particular interest since its projected third protonation pK_d of 2.2 is within the realm of measurement in aqueous solution.

The results presented here also provide a further illustration of a fundamental feature of hierarchically reformulated models — that higher order interactions are less likely to be physically significant than lower order interactions. As discussed in the introduction, for a higher order interaction to exist requires *all* lower order interactions which it modulates to exist. The method of hierarchical reformulation represents a one-to-one transformation from the N ΔG^o of assembly required to define N complexes relative to the free components in the system, and no information is gained or lost in performing this transformation. This method appears to be the only rational way in which such a one-to-one transformation can be made without losing or gaining information (i.e. N parameters to N parameters). Given the utility of this approach demonstrated here and in a number of previous applications [5–7], this method represents the best available method for examining interactions between events in complex state systems.

Site specific behavior is a term used to refer to the thermodynamic properties of a given functional group or site when contained within a network of interacting or potentially interacting sites [10]. The method of hierarchical reformulation allows the site specific properties of a system to either be determined directly from available data, as demonstrated for the pK_a values for the polyprotic acids listed in Table 2, or for these properties to be extrapolated from available data, as demonstrated by the projected pK_a values listed in Tables 2 and 3. The use of inorganic polyprotic acids provides a simple demonstration system for this approach. It is expected that this approach will also yield valuable insights when applied to the more complex organic polyprotic acids which have been the traditional subject of site specific behavior studies [10].

References

[1] G. Weber, Ligand binding and internal equilibria in proteins, Biochemistry 11 (1972) 864–878.

[2] G. Weber, Energetics of ligand binding to proteins, Adv. Prot. Chem. 29 (1975) 1–83.

[3] L. Pauling, The oxygen equilibrium of hemoglobin and its structural interpretation, Proc. Natl. Acad. Sci. USA 21 (1935) 186–191.

[4] W.G. Gutheil, C.E. McKenna, Unique and independent parameters (UIP) formulation for thermodynamic models of complex protein ligand systems, Biophys. Chem. 45 (1992) 171–179.

[5] W.G. Gutheil, The reformulation of thermodynamic systems with aggregation, and theoretical methods for the analysis of ligand binding in proteins with monomer-multimer equilibria, Biophys. Chem. 52 (1994) 83–95.

[6] W.G. Gutheil, Thermodynamic model of cooperativity in a dimeric protein: unique and independent parameters formulation, Biophys. Chem. 45 (1992) 181–191.

[7] W.G. Gutheil, Statistical analysis of data pertaining to complex state systems by stepwise regression with reformulated parameters. Application to spectroscopically monitored hemoglobin oxygen binding data, Biophys. Chem. 70 (1998) 185–201 (Erratum appears in Biophys. Chem. 74 (1998) 163.).

[8] F.H. Westheimer, J.G. Kirkwood, The electrostatic influence of substituents on the dissociation constants of organic acids. II, J. Chem. Phys. 6 (1938) 513–517.

[9] T.L. Hill, The electrostatic contribution to hindered rotation in certain ions and dipolar ions in solution. III, J. Chem. Phys. 12 (1944) 56–61.

[10] E. Di Cera, Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules, Cambridge University Press, Cambridge, 1995.

[11] CRC Handbook of Chemistry and Physics, 62nd ed, CRC Press, Boca Raton, FL, 1981, p. D144.

[12] D.D. Perin, Ionization Constants of Inorganic Acids and Basis in Aqueous Solution, Pergamon Press, New York, 1982, pp. ix–23.

[13] C. Tanford, Physical Chemistry of Macromolecules, Wiley, New York, 1961, p. 532.

[14] H. Rilbe, pH and Buffer Theory — a New Approach, Wiley, New York, 1996, pp. 139–143.

[15] K.J. Laidler, Physical Chemistry with Biological Applications, Benjamin/Cummings, San Francisco, 1978, pp. 223–225.