Metal-Humate Interaction: 2.: Application and Comparison of Models

ARTICLE *in* ENVIRONMENTAL SCIENCE AND TECHNOLOGY · JULY 1986

Impact Factor: 5.33 · DOI: 10.1021/es00149a005 · Source: PubMed

3 AUTHORS, INCLUDING:

William Fish
Portland State University
29 PUBLICATIONS **892** CITATIONS
SEE PROFILE

David A. Dzombak
Carnegie Mellon University
200 PUBLICATIONS **5,297** CITATIONS
SEE PROFILE

Environ. Sci. Technol. 1986, 20, 676-683

Metal–Humate Interactions. 2. Application and Comparison of Models

William Fish,† David A. Dzombak, and François M. M. Morel*

Ralph M. Parsons Laboratory, 48-425, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

■ Discrete and continuous multiligand models for metal–humate interactions are compared and analyzed by using both synthetic and experimental data. Discrete ligands (typically two or three) are shown to be a simple and accurate means of predicting metal–humate binding within the range of calibrating titrations. Selection of discrete ligand parameters is best achieved via nonlinear regression. The continuous affinity spectrum model is highly sensitive to experimental error, and thus its usefulness as an aid for selection of discrete ligands is limited. As anticipated, only the weakest, most abundant ligands in a ligand mixture can be identified with the continuous stability function model. The continuous normal ligand distribution model is capable of fitting metal–humate binding data with only three parameters, but only the stronger ligands in the assumed distribution are important for fitting observed data. The discrete ligand approach is probably the most useful way to model metal–humate binding because of the ease with which discrete ligands can be incorporated into chemical equilibrium computer programs.

Introduction

Many models for metal–humate interactions have been proposed in the literature in recent years. However, there has been little or no comparison of the various models to determine their relative utility or the circumstances for which one model is better suited than others. In part 1 (*1*), we critically examined the theoretical basis of the discrete ligand and the continuous affinity spectrum, stability function, and normal distribution models. Here, we assess the behavior of these models when applied to a consistent group of data sets.

To examine the capabilities and limitations of the models in detail, we devised three sets of data to which the models were applied. A calculated, or "synthetic", data set was prepared which duplicated experimental titration data but which was actually based on a defined set of hypothetical, discrete ligands. This perfectly defined data set was used especially for evaluation of the continuous distribution models. To study the effect of a known source of systematic error, a second data set was prepared by subjecting the synthetic data to the arithmetic manipulation and roundoff typical of experimental data. Finally, each of the models was applied to a set of experimental Cu(II)–humate titration data.

Methods

Data Set 1: Synthetic Data. A set of synthetic metal–humate binding data was generated by calculating the equilibrium speciation of a hypothetical six-ligand system (Table I) at various total metal concentrations. The stability constants and concentrations of the ligands were selected so that the computed titration curve was similar to an experimental Cu(II) titration of fulvic acid. Calculations were performed with the chemical equilibrium program MINEQL (*2*). For this synthetic data set the chemical interactions responsible for the "observed" metal binding are completely understood. For real metal–humate titration data, nothing is known of the underlying chemical speciation. In addition, the synthetic data can be specified to any desired degree of precision and accuracy.

Although the number of ligands in the hypothetical system was arbitrary, we selected the concentrations and stability constants listed in Table I for specific reasons. Ligands 2–4 in Table I define metal binding behavior typical of fulvic acid over the range of metal concentrations commonly used in titrations. A strong ligand at low concentration (ligand 1) and two weak but relatively abundant ligands (ligands 5 and 6) were added to test the sensitivity of the models to very strong and very weak sites suspected to exist in humic material but not experimentally verified. The simulated titration of the six-ligand mixture is presented in Figure 1.

The formation function ($\bar{\nu}$ = [ML]/L_T vs. pM = −log [M]) for the synthetic data is plotted in Figure 2. Computed bound-metal concentrations were normalized with respect to the *known* L_T for the hypothetical six-ligand system (22.1 μM). Note that at pM = 4, L_T is only slightly over half-saturated for the synthetic data. This is frequently the highest cupric ion activity attainable in titrations for which true solutions of humate ions are maintained (*3*).

Data Set 2: Truncated Synthetic Data. To evaluate the sensitivity of the models to measurement error, we imposed limits on the degree of accuracy of the synthetic data by extracting only the total metal concentrations to three significant digits and the logarithms of the free metal activities to two significant digits in the mantissa. The resulting data possess the accuracy obtainable from a metal–humate titration, as monitored by an ion-selective electrode. In all subsequent calculations these data are treated as if they were experimentally obtained. Bound-metal concentrations were computed by subtracting the free metal activities from the corresponding total metal concentrations. Bound metal was then plotted (Figure 3) as a function of the total metal added. The total ligand concentration, L_T, was estimated to be 15 μM, corresponding to the maximum concentration of bound metal observed in Figure 3. While this commonly used procedure probably underestimates L_T, it has the advantage that uniform L_T estimates for a given data set can be obtained by different investigators. Estimating L_T from an inflection in a plot of bound metal vs. pM is particularly imprecise because of the weakly inflected character of the curves.

The computed bound-metal concentrations were normalized to the empirically determined L_T, and the formation function was plotted (Figure 2). Note that the scatter in the truncated synthetic data evident in Figures 2 and 3 is due *only* to roundoff error in the free metal activities and the resulting propagation of error in the bound-metal concentrations; experimental error is not included. The difference in the position of the truncated synthetic data formation curve compared to that of the synthetic data (Figure 2) is the result of normalizing to different L_T values. For the ideal, synthetic data, L_T was known exactly; for the truncated synthetic data, which were intended to reproduce realistic analytical limitations, L_T was approximated as it would be for an actual titration.

†Present address: Oregon Graduate Center, Beaverton, OR 97006.

0013-936X/86/0920-0676$01.50/0

Table I. Stability Constants and Ligand Concentrations for Discrete Ligand Data Sets

		1	2	3	4	5	6	L_T (μM)
three-ligand set								
Scatchard fit of exptl data	log K_i		10.83	8.80	6.48			
	log L_i		−6.95	−6.17	−5.17			7.55
FITEQL fit of exptl data	log K_i		10.72	8.29	5.94			
	log L_i		−6.84	−5.87	−5.04			10.6
six-ligand set								
synthetic data	log K_i	15.0	10.8	8.8	6.5	4.0	3.0	
	log L_i	−10.0	−7.0	−6.2	−5.2	−5.0	−5.3	22.1

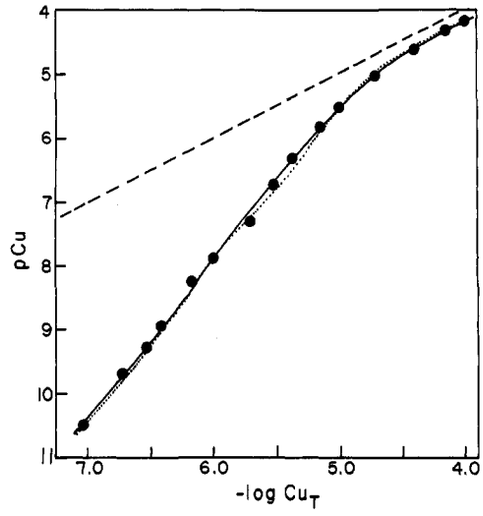

Figure 1. Discrete ligand fitting of ion-selective electrode titration results; pCu vs. total Cu(II) added. Grassy Pond fulvic acid, 5.0 mg C/L at pH 6.0 (●); three discrete ligand model (—); synthetic data (---).

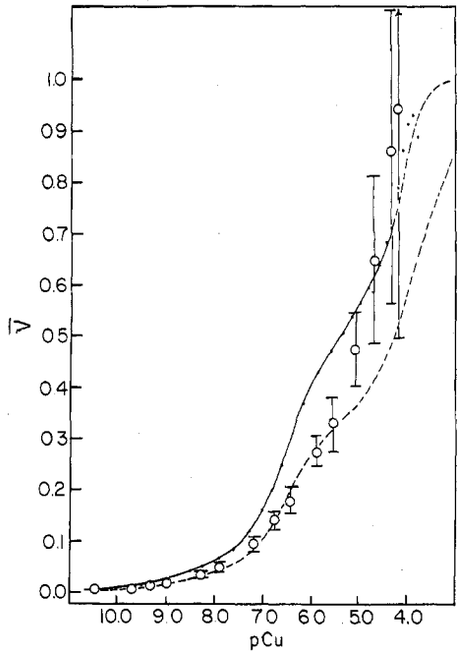

Figure 2. Formation functions ($\bar{\nu}$ vs. −log [Cu^{2+}]) for synthetic data (---), truncated synthetic data (—), and experimental data (O). Error bars indicate ± 1 standard deviation, calculated by the methods of Fish and Morel (*3*).

Data Set 3: Experimental Data. Experimental metal–humate binding data were obtained from a Cu(II) titration of a 5.0 mg C/L solution of fulvic acid extracted according to the methods of Thurman and Malcolm (*4*) from Grassy Pond in Bedford, MA. Grassy Pond fulvic acid has been used in several other studies of trace metal binding (*3*, *5*) and iron photochemistry (*6*). We used a low

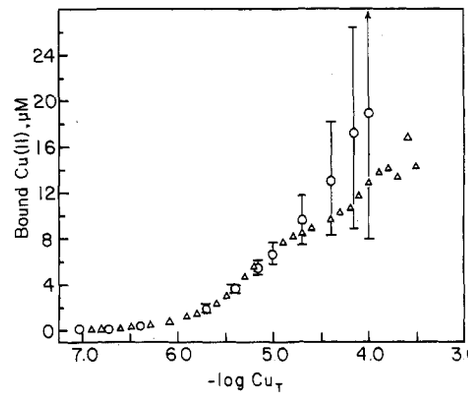

Figure 3. Concentrations of bound metal vs. −log (total metal) for experimental data (O) and the truncated synthetic data (△). Error bars indicate ± 1 standard deviation.

humate concentration to mimic natural waters and to avoid flocculation problems. The solution was maintained at pH 6.0, and the cupric ion activity was measured with a cupric ion selective electrode. Details of the methodology are discussed elsewhere (*3*, *5*); results of the titration are presented in Figure 1.

In Figure 3 the experimental data are replotted as bound Cu(II) vs. −log Cu_T for which bound Cu(II) was determined by subtracting the cupric ion activity and calculated hydroxo- and carbonato–Cu(II) complexes from the total Cu(II) concentration. The total ligand concentration, L_T, for the experimental data was estimated to be 20 μM (Figure 2) corresponding to the maximum observed concentration of bound Cu(II). This estimate for L_T is very imprecise because of the large error associated with bound-metal concentrations determined by the ion-selective electrode at high M_T (*3*). The experimental data were also transformed into a formation function by normalizing the bound-metal data to L_T (Figure 2).

Results and Discussion

Discrete Ligand Model. As noted in part 1 (*1*), there are several methods available for selecting discrete ligands to reproduce titration data. The most commonly applied graphical method is a linearization of titration data introduced by Scatchard (*7*) for use in studies of ion binding by proteins. The Scatchard method is based on the formula

$$\frac{[ML]}{[M]} = \frac{\sum[ML_i]}{[M]} = \sum K_i[L_i^f] = \sum(K_iL_i) - \sum(K_i[ML_i]) \quad (1)$$

where the symbols are the same as defined in part 1 (*1*). For a single ligand, a plot of [ML]/[M] vs. [ML] has slope K and horizontal intercept L_T. When more than one ligand is present, this plot—known as a Scatchard plot—takes on a hyperbolic shape with two limiting linear regions (see Figure 4). The slopes and intercepts for the two linear regions correspond to the parameter combinations indi-

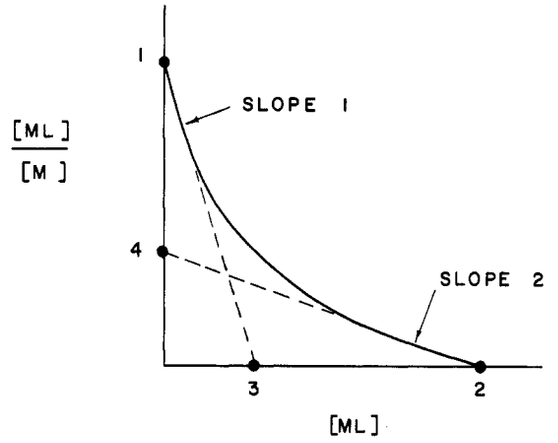

Figure 4. Schematic plot of [ML]/[M] vs. [ML] for metal binding to a multitude of independent humate site types (after Klotz and Hunston (*8*)).

Table II. Interpretation of Slopes and Intercepts in Figure 4 (after Klotz and Hunston (*8*))

intercept 1 = $\sum(K_iL_i)$
intercept 2 = $\sum L_i = L_T$
intercept 3 = $[\sum(K_iL_i)]^2/\sum(K_i^2L_i)$
intercept 4 = $[\sum L_i]^2/\sum(L_i/K_i) = L_T^2/\sum(L_i/K_i)$
slope 1 = $-\sum(K_i^2L_i)/\sum(K_iL_i)$
slope 2 = $-\sum L_i/\sum(L_i/K_i) = -L_T/\sum(L_i/K_i)$

cated in Table II (after Klotz and Hunston (*8*)). From the six equations given in Table II for the four intercepts and two slopes it is clear that the problem is determinate only for $i = 3$.

Linearized graphical methods such as the Scatchard plot are subject to two major sources of errors. The first is the error inherent in fitting a line to nonlinear data (see, for example, ref 9). The second and greater problem is that of equally weighing all data points, regardless of the relative error which varies throughout the titration (*3*, *10*). We demonstrated these problems by applying the Scatchard method to the experimental data set (data set 3).

Three orders of magnitude of bound metal concentrations were observed in the experimental titration data so we estimated that three ligands would best fit the data (*1*). The experimental data were plotted in the form of a Scatchard plot (Figure 5), and lines were fitted to the strong and weak binding data by linear regression. Values of K_i and L_i were then determined from the slopes and intercepts of these lines, yielding the set of constants in Table I. A calculated Cu(II) titration of the three ligands reproduces closely the experimental data in a plot of pCu vs. -log Cu_T (Figure 1). The deviation between model and experimental results at high Cu_T stems from the error associated with the data points used to evaluate the weakest ligand (Figure 3). Apparently small discrepancies in the pCu values at high Cu_T translate into large errors in the concentration of bound metal predicted in this range. The uncertainty in the last few bound-metal data points translates directly into uncertainty in the parameters fitted to the weakest ligand. Note in Table I that the Scatchard plot estimate of L_T (7.55 μM) greatly underestimates the L_T value estimated from the maximum observed binding (20 μM).

Other linear graphical techniques have been used to fit discrete ligands to ion binding data for macromolecules (*8*). The problems outlined above are applicable to all such procedures although the difficulties in defining L_T may not be as severe for well-defined biological macromolecules as they are for humic substances.

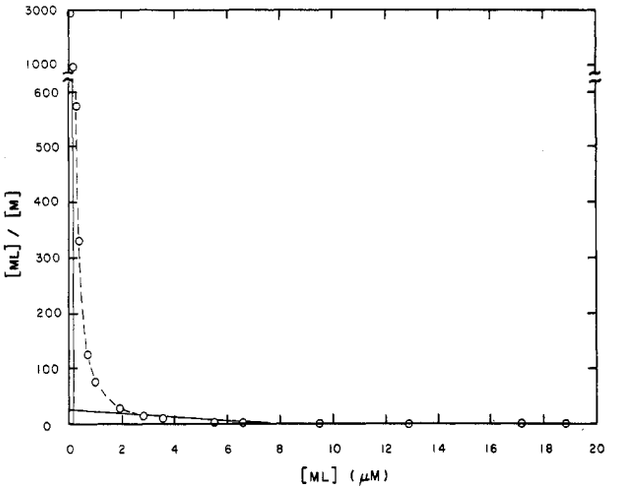

Figure 5. Scatchard plot for Cu titration of Grassy Pond fulvic acid.

Choosing discrete ligand parameters by nonlinear optimization offers several improvements over graphical methods. A computer program that simultaneously optimizes all parameters eliminates the arbitrariness of dividing a curve into quasi-linear subsections. Optimized parameters are more useful than Scatchard parameters for comparing the results of different titrations because they are reproducible; i.e., a set of data always yields the same parameter values as long as the same optimization algorithm is employed. Another advantage of computerized nonlinear optimization is the relative ease of statistical analysis. An optimization program can automatically calculate error propagation, apply weighting factors to the data, and calculate the standard deviation associated with each parameter. Accurate comparisons of titrations become possible because differences can be tested for statistical significance. The arbitrariness of graphical fitting precludes a precise quantitative evaluation of parameter error.

We applied the discrete ligand optimization program FITEQL (*11*) to the experimental data (data set 3) to obtain an optimal set of stability constants and concentrations for three ligands (Table I). The parameters for the first two ligands are very close to those obtained by the Scatchard method, a reflection of the precision of the data in the strong binding region. The weakest ligand is assigned a conditional stability constant by FITEQL that is about an order of magnitude smaller than the value obtained from the Scatchard plot. The concentration of weak ligand is somewhat higher when determined by FITEQL than by the Scatchard plot, but the value of L_T obtained (10.6 μM) is still only about half of the value estimated from the maximum observed bound-metal concentration.

Since ligands 2–4 of the hypothetical six-ligand mixture can reproduce the experimental Cu–fulvate titration data (see Table I and Figure 1), ligands 1, 5, and 6 do not significantly affect copper speciation in the range of experimentation. Thus, a Scatchard or FITEQL analysis of the synthetic data set over the pM range in Figure 1 would yield the same ligand parameters as obtained from the analyses of the experimental data; i.e., ligands 2–4 or something very close.

The fits of data set 3 obtained with the FITEQL-derived ligands and the ligands derived with a Scatchard plot are very similar on a pCu vs. -log Cu_T plot but are clearly distinguished in the formation functions (Figure 6). The differences between the experimental, the FITEQL-based, and the Scatchard-based formation functions are *mostly*

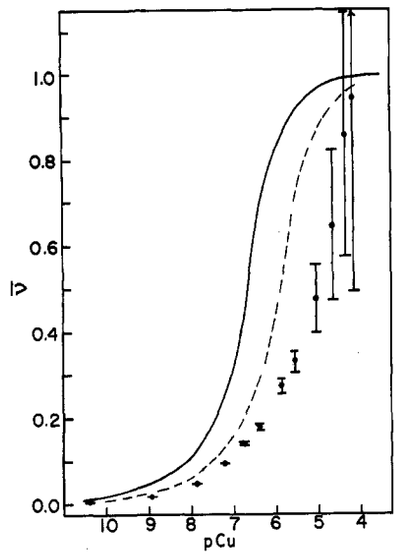

Figure 6. Formation functions of $\bar{\nu}$ vs. pCu. Grassy Pond fulvic acid, 5.0 mg C/L at pH 6.0 (●); error bars indicate ± 1 standard deviation; Scatchard, three-ligand fit of data (—); FITEQL, three-ligand fit (---).

attributable to differences in L_T and not to differences in the bound-metal concentrations. All of the formation functions are normalized to L_T, operationally defined either as the sum of ligand concentrations fitted to the data (for calculated data) or as the maximum amount of bound metal detected in the titration (for experimental data). Consequently, each formation function in Figure 6 is normalized to a different L_T value. This commonly used normalization which is intended to facilitate comparison among results actually makes comparisons difficult because L_T is so poorly defined. The situation is ameliorated by obtaining the most precise weak binding data possible so that L_T is determined with greater precision (*3*). Unfortunately the absence of a clear end point in most titrations of humic material precludes an *accurate* value of L_T. Therefore, L_T (or "binding capacity") is not an appropriate parameter for comparing different humates, and metal–humate binding data are best presented in plots of [ML] vs. log M_T (or log [M]).

Affinity Spectrum Model. The affinity spectrum for each of our three data sets was obtained by smoothing the formation function data (from the plots in Figure 2) with a cubic spline and using the smoothed data in a second-order solution of the fundamental integral equation (eq 10 in part 1 (*1*)) after the methods of Ninomiya and Ferry (*12*) as applied by Hunston (*13*), Thakur et al. (*14*), and Shuman et al. (*15*).

The affinity spectrum for the synthetic data set (data set 1) exhibits three smooth and clearly defined peaks at log K = 8.75, 6.50, and 3.95 (Figure 7A) corresponding to ligands 3, 4, and 5 (log K = 8.80, 6.48, and 4.00). Only three of the six ligands are identified because the affinity spectrum model indicates the importance of ligands with log K = pM and limits of pM 10.40–3.50 were imposed. It was thus not possible to detect the strongest ligands (log K_1 = 15.0, log K_2 = 10.83) or the weakest ligand (log K_6 = 3.00). Note that if the upper pM limit had been set at a more realistic value (e.g., pM 4), ligand 5 (log K = 4) would have been missed also.

The affinity spectrum for the truncated synthetic data (data set 2) shows a dramatic increase in the number of peaks compared to the six-ligand set on which it is based. At least 12 peaks or shoulders can be distinguished (Figure 7B). Recall that the truncated synthetic data differ from the synthetic data only in the inclusion of the roundoff

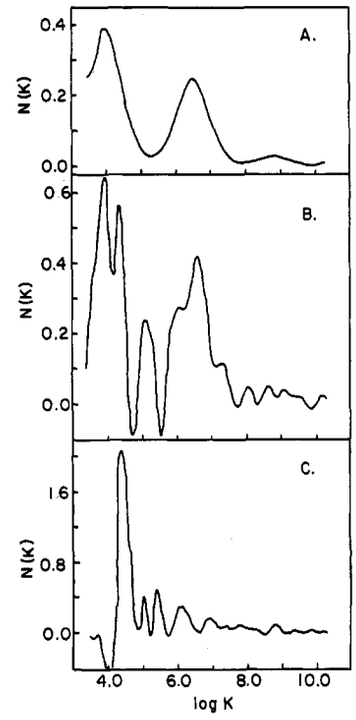

Figure 7. Affinity spectra for the synthetic data set (A), for the truncated synthetic data set (B), and for the experimental data (C).

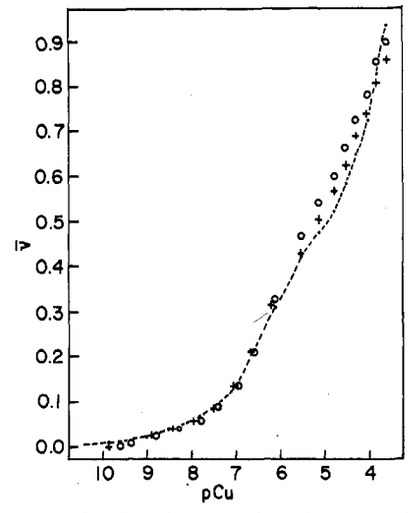

Figure 8. Formation functions for the truncated synthetic data set (---), the three-ligand, affinity spectrum fit (O), and the 17-ligand, affinity spectrum fit (+). Note, symbols represent continuous data; discrete points are used only for the purpose of clarity.

error typical of bound-metal data from ion-selective electrode titrations. This small, systematic error produces enough scatter to generate small inflections in the formation function and causes numerous spurious peaks in the affinity spectrum. Adjustments of the parameter that controls the sharpness of the peaks in the numerical integration method (*12*) had negligible effect on the noise in the spectrum, and adjusting the smoothing was of no value because we had no a priori knowledge of which inflections were artifacts and which were authentic. The choice of the normalization factor L_T dictates the total area under the curve, hence, the L_i values obtained, but has negligible effect on the shape of the spectrum.

Our knowledge of the components underlying the truncated synthetic data helped us sort the numerous peaks into three groups corresponding to the predominant ligands. The mean log K and L_i values for each group are in reasonable agreement with the actual component ligands

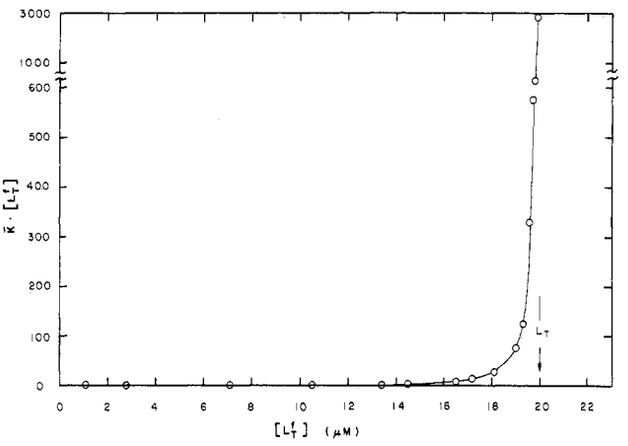

Figure 9. $\bar{K}\cdot[L_T^f]$ vs. $[L_T^f]$ (mirror image Scatchard plot) for the experimental data.

and yield a good fit of the formation function (Figure 8). An improved fit of the data may be expected if a large number of ligands are assigned from the spectrum, thereby approximating a continuous ligand distribution. Dividing the spectrum into 17 ligands greatly magnifies the computation necessary to calculate the speciation but improves the fit of the formation function only slightly (Figure 8).

The impact of random error is especially problematic when analyzing experimental data, as seen in the affinity spectrum for data set 3 (Figure 7C). Numerous small peaks appear, but we cannot distinguish between authentic binding sites and the sort of artifacts observed in the truncated synthetic data spectrum. Again we are faced with a choice between a few ligands representing arbitrary groups of peaks and a large number of discrete ligands that are computationally unwieldly and probably bear little relation to ligands actually present. Under ideal conditions, an affinity spectrum is a useful numerical method for selecting discrete ligands to fit titration data. Under experimental conditions, however, small, random errors cause fragmentation of the spectral peaks, and the choice of discrete ligands becomes arbitrary.

Continuous Stability Function Model. For each of the three data sets we determined the stability function $K([L_T^f])$ by differentiating numerically plots of $\bar{K}\cdot[L_T^f]$ vs. $[L_T^f]$ (e.g., Figure 9 for data set 3). As detailed in part 1 (*1*), the function $K([L_T^f])$ resulting from this differentiation (see Figure 10) represents the dominant stability constant at each titration point, and the corresponding ligand distribution is given by

$$L(K([L_T^f])) = \frac{1 + K([L_T^f])\,[M]([L_T^f])}{dK([L_T^f])/d[L_T^f]} \qquad (2)$$

The ligand distribution for the synthetic data set, shown in Figure 11, was determined by taking derivatives along Figure 10A and using eq 2. From Figure 11 it is evident that the continuous stability function model (*16*, *17*) is capable of identifying only the weakest, most abundant ligand in the system.

Normal Distribution Model. We applied the single mode, normal ligand distribution model to the three data sets. To determine the optimal normal distribution (in log K) for each data set, we employed the curve-matching technique of Karush and Sonenberg (*18*, *19*) and the nonlinear optimization program developed by Perdue and Lytle (*20*, *21*). For the curve-matching technique, the mean log K values were obtained directly from the formation functions in which μ (=log K_0) is given by the value of pM when $\bar{\nu} = 0.5$, as explained in part 1 (*1*). These μ values were substituted in eq 12 of part 1, and families of

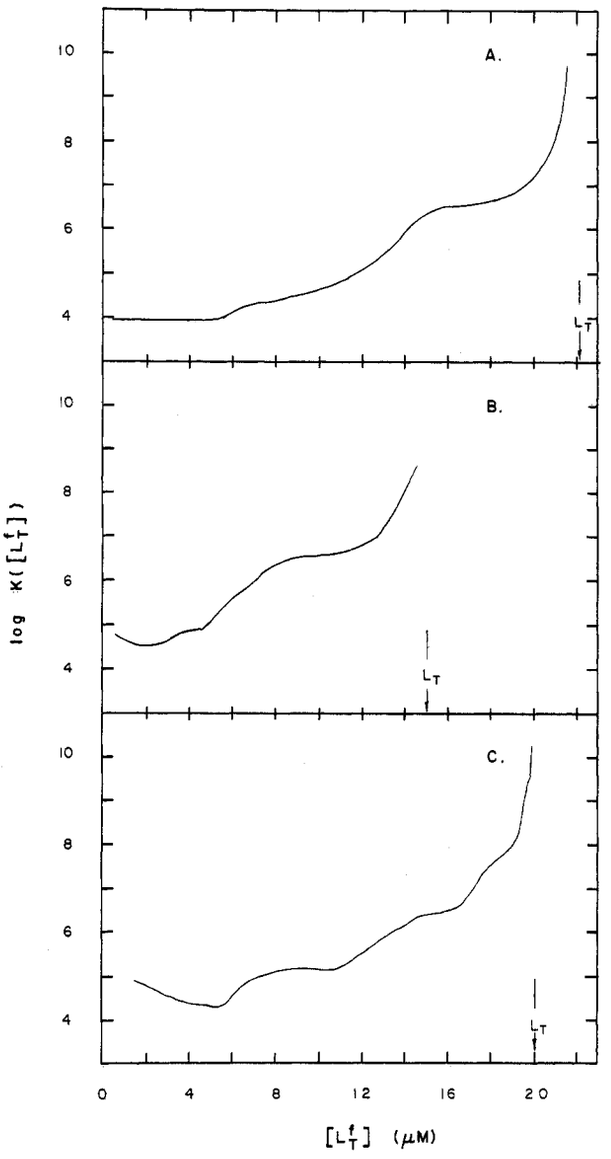

Figure 10. Continuous stability functions for the synthetic data set (A), for the truncated synthetic data set (B), and for the experimental data (C).

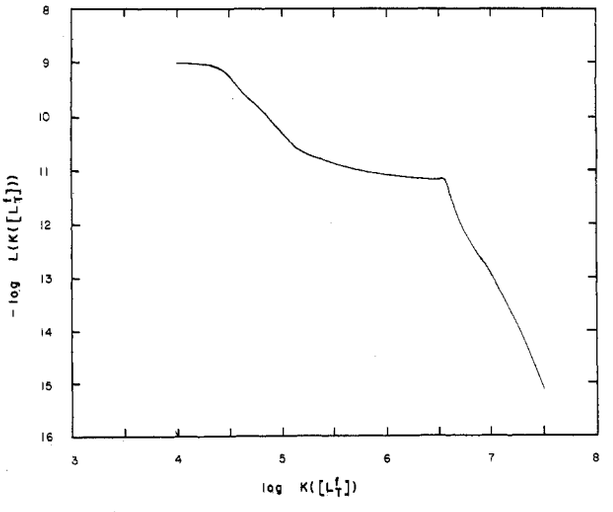

Figure 11. Ligand frequency distribution derived with the continuous stability function model for the synthetic data set.

theoretical binding curves were computed for various values of the standard deviation σ. The optimum value of σ was then determined by comparing the theoretical

Table III. Optimal Parameters for the Single-Mode Normal Distribution Fitted to the Three Data Sets

	curve matching			nonlinear optimization		
data set	μ	σ	RSS	μ	σ	RSS
1 (syn)	4.25	2.2	2.518E–6	4.45	1.96	1.254E–6
2 (t.syn)	5.35	1.5	2.797E–6	5.42	1.49	2.653E–6
2 (t.syn)[a]				4.41	1.99	1.618E–6
3 (expt)	5.00	1.6	3.341E–6	5.13	1.29	2.726E–6

[a] With L_T = 22.1 μM.

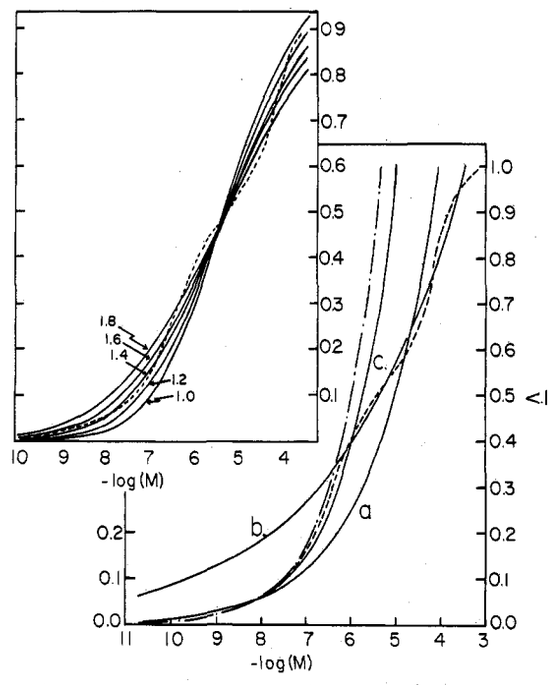

Figure 12. (Upper left panel) Family of curves (solid lines) used to find normal distribution parameters by matching with the formation function of the truncated synthetic data set (- - -). Value of μ is fixed as described in the text. Trial values of σ are 1.0, 1.2, 1.4, 1.6, and 1.8. (Lower right panel) Comparison of formation functions obtained with the Sips model. Truncated synthetic data (- - -); optimal fit from a family of curves (– • –). Fits obtained from the linearization method: regression of all data (a); regression of weak binding data (b); regression of strong binding data (c).

binding curves with the actual binding curve. The theoretical curve providing the closest fit to the most precise binding data ($\bar{\nu} < 0.3$, approximately) was selected as optimal. Acceptable fits were obtained for all data sets, and the optimal values of μ and σ are listed in Table III. The fit of the truncated synthetic data (data set 2) is given in Figure 12 as an example.

The fit obtained for data set 2 is similar in quality to those obtained for the other data sets by using the graphical parameter extraction method. In accord with the original fitting objective, the more precise low $\bar{\nu}$ data were fitted more closely than the less precise high $\bar{\nu}$ data. Note that in fixing $\mu = \log K_0$ the pivot point of the family of theoretical curves is also fixed. The location of this pivot point strongly influences the quality of fit achievable. Given the size of the error bars associated with the experimental binding data at $\bar{\nu} \cong 0.5$ (e.g., see Figure 2), and hence the range of possible values of μ, it seems probable that improved fits will result if both μ and σ are treated as adjustable parameters. Using the program developed by Perdue and Lytle (*20*, *21*), we indeed found this to be the case (see Table III).

With a nonlinear optimization computer program it is possible to adjust L_T, μ, and σ simultaneously. The results

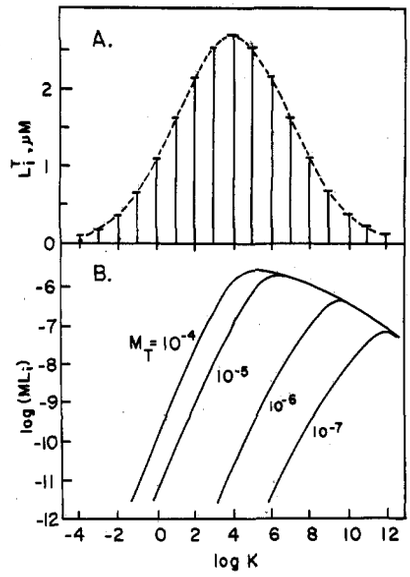

Figure 13. (A) Seventeen discrete ligands normally distributed with respect to log K. (B) log concentration of metal bound to each of the 17 normally distributed ligands. Bound-metal distributions are shown for $M_T = 10^{-7}$–10^{-4} M.

presented in Table III, however, were determined with L_T fixed according to our original estimates for each data set. We did search for better fits via optimization on all three parameters and found that though lower residuals could be achieved, the converged solution depended strongly on the initial estimates for the adjustable parameters. In fact, we found that we could obtain almost any result desired by changing the initial estimates. This occurred because, without a constraint on L_T, local minima abound.

The optimum normal distribution of log K values for data sets 1 and 2 (Table III) reflects the large concentrations of weak ligands in the hypothetical six-ligand system near log K = 4–5 and the successively smaller concentrations of strong ligands. The differences between the parameter values for the two data sets are due primarily to the different L_T values. If the correct value for L_T (22.1 μM) is used in the optimization for the truncated synthetic data (data set 2), the resulting model parameter values are nearly identical with those obtained for the synthetic data set (data set 1). The increased accuracy of data set 1 relative to data set 2 is reflected in the lower residual sum of squares for data set 1 even when the correct L_T value is used for data set 2.

As shown in part 1 (*1*), only those ligands with stability constants greater than the mean and less than log $K_{crit} = \mu + 2.303\sigma^2$ influence metal speciation measurably. To demonstrate the dominance of the right side of a normal ligand distribution on the predicted metal speciation, we considered a hypothetical system of 17 ligands normally distributed in log K. The concentrations were selected assuming $\mu = 4$, $\sigma = 3$, and $L_T = 2 \times 10^{-5}$ M; the resulting normal distribution is depicted in Figure 13A. We calculated the metal speciation in this system for several concentrations of M_T, using MINEQL (*2*). When the concentration of metal bound to each ligand is plotted as a function of the ligand stability constants, we observe that most of the metal is bound only to the strong ligands, i.e., those with log $K > \mu$ (Figure 13B). Weaker ligands become increasingly important as M_T increases, but the left side of the distribution never affects metal speciation significantly. We confirmed the dominance of the right half of the normal distribution by reproducing *exactly* the results shown in Figure 13B using only the nine ligands to the

right of and including the mean.

These results can be understood by recalling that ligands are titrated in order of decreasing values of the product K_iL_i and increasing L_i (*1*). For the 17 ligand distribution depicted in Figure 13A, the order of the product K_iL_i is the same as the order of log K_i; hence, the strongest ligands bind most of the metal at all values of M_T. Below log K = 4 (and above log K_{crit} = 24.7), the ligand concentrations L_i decrease as K_iL_i decreases, and hence, these ligands can have no measurable effect on metal speciation.

Our analysis of the normal ligand distribution model shows that a single, relatively narrow mode centered at relatively low log K is adequate to describe metal–humate binding. If the strong binding (low M_T) portion of a titration curve can be fitted, a reasonable fit of the weak binding (high M_T) data will follow. The success of using a single mode means that the normal distribution model is an accurate three fitting parameter model for metal–humate binding. Since only the right side of a normal ligand distribution has any influence on metal ion activity, however, it is clear that a normal ligand distribution does not necessarily represent the actual distribution of ligands.

In part 1 (*1*) we noted that if the Sips approximation (*22–24*) of the normal distribution function is invoked, a binding expression can be derived analytically, namely

$$\bar{\nu} = \frac{K_s[M]^a}{1 + K_s[M]^a} \tag{3}$$

where $K_s = K_0{}^a$ and a is an index of heterogeneity ($0 \leq a \leq 1$). K_0 is an average binding constant analogous to the mean in the normal ligand distribution model. Parameter values for the Sips distribution model can be obtained by two graphical methods. The parameters K_s and a can be determined by changing eq 3 into logarithmic form and substituting for log K_s, which yields

$$\log \bar{\nu} = a(\log [M] + \log K_0) \tag{4}$$

If L_T is known or estimated, eq 4 contains only one fitting parameter (a) since log K_0 can be obtained directly from a formation function plot (as the value of pM for which $\bar{\nu}$ = 0.5). A family of theoretical binding curves can be hand calculated for various values of a and the optimum value ascertained by matching the curve with experimental results. The optimal fit obtained in this way for the truncated synthetic data is shown in Figure 12.

Alternatively, both K_s and a can be treated as adjustable parameters, and a graphical technique can be used to determine their values. For a linear plot of log $\bar{\nu}$ vs. log [M], the index a is given by the slope and K_s by the intercept. However, log $\bar{\nu}$ vs. log [M] plots for metal–humate binding typically comprise two linear regions, one in the high M_T (weak binding) region and one in the low M_T (strong binding) region. The regression for the low M_T data provides the best fit of the most precise portion of the formation function, i.e., the low $\bar{\nu}$ data. As shown in Figure 12c, this fit is slightly better than the fit obtained with the curve-matching technique. With either method, if the more precise strong binding data are fitted, the Sips model results in a single mode centered at a relatively low log K as is the case with the normal distribution model. The Sips model is convenient because it can be implemented without the aid of a computer.

Summary and Conclusions

The discrete ligand approach to modeling metal–humate interactions has been widely used because it is a simple and accurate means of reproducing titration data. Ligand parameters (K_i and L_i) can be fitted to the data by graphical techniques such as the Scatchard plot, but such linearizations can lead to substantial errors in the parameters fitted. Nonuniform data precision and error propagation are not easily accounted for in graphical techniques although these methods are adequate for quick estimates. Computer-based nonlinear optimization routines like FITEQL (*11*) minimize residuals, yielding the best fit possible with discrete ligands. The resulting parameters are reproducible if the same optimization algorithm is used. Error propagation calculations and weighting factors can also be incorporated in an optimization procedure. For most humic solutions two or three ligands will be required, resulting in four or six adjustable parameters. Titrations can be compared on the basis of these parameters if the fitting procedures are standardized.

It is important to note that discrete ligands fitted to titration curves should only be applied to systems in which the ratios of total metal to total fulvic acid are comparable to those for which the model was calibrated. The strongest ligand in the model is dictated by the lowest metal/ligand ratio in the calibrating titration, and extrapolations below this ratio may underestimate the strength of metal binding.

Because the discrete ligands used for modeling humates do not correspond to physically definable binding sites, they cannot accurately be treated as true chemical species. For example, the proton ionization constants obtained for a fulvic acid are difficult to assign to the ligands determined from a Cu(II) titration. Nevertheless, carefully defined discrete ligands can be useful for calculating the speciation of metals competing for humate binding sites (*5*). Modeling multimetal interactions with humic material by assigning stability constants for each metal to the same set of ligands yields a reasonable approximation of observed ion competition. However, the structure of computer programs such as MINEQL can easily lead the unwary modeler into incorrectly posing a complicated equilibrium problem by using a set of discrete ligands that may not account for competitive interactions among ions.

An affinity spectrum, while based on a postulated continuum of metal binding sites, does not represent an actual ligand frequency distribution but can be used to facilitate selection of discrete ligands. As we have shown, with ideally precise data the procedure is an improvement over the graphical Scatchard method for selecting binding parameters in the range log K < 8. For ligands stronger than this, the affinity spectrum is difficult to interpret, even when applied to ideally precise data. When indeterminant error is present in the titration data, the resulting small, random fluctuations in the slope of the formation function introduce sufficient noise into the affinity spectrum to make an unambiguous choice of fitting parameters difficult at best. In addition, the method requires a computer-based solution scheme even though the final selection of ligands involves a visual evaluation of the spectrum and trial and error adjustment of parameters.

If a normal frequency distribution of metal binding sites is assumed, only three fitting parameters are needed to model a wide range of metal–humate binding data. This concise description of a data set makes comparisons among results easier. The numerical integration required by the model necessitates a computer, although the Sips approximation of the normal distribution function enables an approximate analytical solution. While it is not known whether a Gaussian ligand distribution has any intrinsic physical meaning, previous analysis (*1*) and a simulated metal titration presented here demonstrate that only the right side of the distribution has any influence on metal ion activity.

In contrast to the discrete ligand model, the normal distribution model postulates the presence of ligands with stronger stability constants than are experimentally verifiable. In fact, the model includes small but finite concentrations of ligands with nearly infinite stability constants. Perdue and Lytle (*21*) showed that when the normal distribution model is extrapolated beyond its calibration range, it predicts stronger binding than does a set of discrete ligands, but this binding may not reflect the correct metal speciation. Thus, the normal distribution model should also be applied only within the range of calibrating titrations.

Researchers interested only in concisely describing the results of a single metal titration of a humic material should consider using the three-parameter, normal distribution model of metal binding. Inclusion of metal–humate interactions in a general equilibrium model of chemical speciation is most easily accomplished by discrete ligands, however. With either approach, care must be taken to ensure that these empirically based models are not used to calculate ion speciation in systems for which they have not been calibrated.

In applying any of the models examined here to metal–humate titration data, a fundamental problem is determining the total ligand concentration L_T. The concentration L_T, which cannot be measured unambiguously, is important in that values for all other model parameters depend on it. Several L_T estimation methods have been proposed; for the sake of reproducibility we recommend estimating L_T operationally as the maximum bound-metal concentration achievable in a metal–humate titration. Because L_T is poorly known, we strongly recommend against the normalization of binding data to L_T (i.e., the use of formation functions). The clearest way to present the results of a metal–humate titration is in a plot of bound metal ([ML]) vs. log M_T or log [M].

While no single model of metal–humate interactions has yet emerged that is clearly superior for all purposes, discrete ligands are probably the most useful. Continuous distribution models currently offer few advantages over discrete ligand representations. The fact that discrete ligands can be used to predict competitive interactions among metal ions (*5*) is a powerful argument for the use of discrete rather than continuous distributions of ligands in models of solute transport or biological phenomena. The choice of a metal binding model should be guided by a careful consideration of the uses for which the model is intended, the quality and quantity of titration data at hand, and the computational resources that are available.

Acknowledgments

We thank K. Farley, T. D. Waite, and R. Hudson for their helpful discussion of this work and E. M. Perdue for use of his computer program and for his extensive review of the manuscript.

Literature Cited

(1) Dzombak, D. A.; Fish, W.; Morel, F. M. M. *Environ. Sci. Technol.*, preceding paper in this issue.
(2) Westall, J. C.; Zachary, J. L.; Morel, F. M. M. "MINEQL: A Computer Program for the Calculation of Chemical Equilibrium Composition of Aqueous Systems"; R. M. Parsons Laboratory Technical Note 18; Massachusetts Institute of Technology: Cambridge, MA, 1976.
(3) Fish, W.; Morel, F. M. M. *Can. J. Chem.* **1985**, *63*, 1185–1193.
(4) Thurman, E. M.; Malcolm, R. L. *Environ. Sci. Technol.* **1981**, *15*, 463–466.
(5) Fish, W. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1984.
(6) Waite, T. D.; Morel, F. M. M. *Anal. Chim. Acta* **1984**, *162*, 263–274.
(7) Scatchard, G. *Ann. N.Y. Acad. Sci.* **1949**, *51*, 660–672.
(8) Klotz, I. M.; Hunston, D. L. *Biochemistry* **1971**, *10*, 3065–3069.
(9) Bevington, P. R. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill: New York, 1969.
(10) Munson, P. J.; Rodbard, D. *Science (Washington, D.C.)* **1983**, *220*, 979–981.
(11) Westall, J. C. "FITEQL: A Program for the Determination of Chemical Equilibrium Constants from Experimental Data"; Technical Report, Department of Chemistry; Oregon State University: Corvallis, OR, 1982.
(12) Ninomiya, K.; Ferry, J. D. *J. Colloid Sci.* **1959**, *14*, 36–48.
(13) Hunston, D. L. *Anal. Biochem.* **1975**, *63*, 99–109.
(14) Thakur, A. K.; Munson, P. J.; Hunston, D. L.; Rodbard, D. *Anal. Biochem.* **1980**, *103*, 240–254.
(15) Shuman, M. S.; Collins, B. J.; Fitzgerald, P. J.; Olsen, D. L. In *Aquatic and Terrestrial Humic Materials*; Christman, R. F.; Gjessing, E. T., Eds.; Ann Arbor Science: Ann Arbor, MI, 1983; pp 349–370.
(16) Gamble, D. S.; Underdown, A. W.; Langford, C. H. *Anal. Chem.* **1980**, *52*, 1901–1908.
(17) Gamble, D. S.; Schnitzer, M.; Kerndorff, H.; Langford, C. H. *Geochim. Cosmochim. Acta* **1983**, *47*, 1311–1323.
(18) Karush, F.; Sonenberg, M. *J. Am. Chem. Soc.* **1949**, *71*, 1369–1376.
(19) Karush, F. In *Advances in Immunology*; Taliaferro, W. H.; Humphrey, J. H., Eds.; Academic: New York, 1962; Vol. 2, pp 1–40.
(20) Perdue, E. M.; Lytle, C. R. *Environ. Sci. Technol.* **1983**, *17*, 654–660.
(21) Perdue, E. M.; Lytle, C. R. In *Aquatic and Terrestrial Humic Materials*; Christman, R. F.; Gjessing, E. T., Eds.; Ann Arbor Science: Ann Arbor, MI, 1983; pp 295–313.
(22) Nisonoff, A.; Pressman, D. *J. Immunol.* **1958**, *80*, 417–428.
(23) Posner, A. M. In *Transactions of the 8th International Congress of Soil Science*; Bucharest, Romania, 1964; Part II, pp 161–174.
(24) Sips, R. *J. Chem. Phys.* **1948**, *16*, 490–495.

Received for review July 9, 1984. Revised manuscript received July 26, 1985. Accepted February 17, 1986. This study was supported by NSF Grant 0CE81-18103, NOAA Grant NA 79AA-D-00077, U.S. EPA Grant CR-811181-01-01, and a grant from the International Copper Research Association (Project 346).