



**HAL**  
open science

**An improved description of the interactions between  
rare earth elements and humic acids by modeling:  
PHREEQC-Model VI coupling**

Remi Marsac, Mélanie Davranche, Gérard Gruau, Martine Bouhnik-Le Coz,  
Aline Dia

► **To cite this version:**

Remi Marsac, Mélanie Davranche, Gérard Gruau, Martine Bouhnik-Le Coz, Aline Dia. An improved description of the interactions between rare earth elements and humic acids by modeling: PHREEQC-Model VI coupling. *Geochimica et Cosmochimica Acta*, 2011, 75 (19), pp.5625 - 5637. 10.1016/j.gca.2011.07.009 . hal-01904313

**HAL Id: hal-01904313**

**<https://hal.science/hal-01904313>**

Submitted on 24 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **An improved description of the interactions between rare earth**  
2 **elements and humic acids by modeling:**  
3 **PHREEQC-Model VI coupling**  
4

5 Rémi Marsac\*, Mélanie Davranche, Gérard Gruau, Martine Bouhnik-Le Coz and Aline Dia

6

7 Geosciences Rennes, UMR CNRS – 6118, University of Rennes 1, Campus Beaulieu, 35042

8

Rennes Cedex, France

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23 \*Corresponding author: Tel +33 223 235 395; Fax: +33 223 236 090

24 E-mail address: remi.marsac@univ-rennes1.fr

25

26 **Abstract-** The Humic Ion Binding Model VI (Model VI) - previously used to model the  
27 equilibrium binding of rare earth elements (REE) by humic acid (HA) - was modified to  
28 account for differences in the REE constant patterns of the HA carboxylic and phenolic  
29 groups, and introduced into PHREEQC to calculate the REE speciation on the HA binding  
30 sites. The modifications were shown to greatly improve the modeling. They allow for the first  
31 time to both satisfactorily and simultaneously model a large set of multi-REE experimental  
32 data with the same set of equations and parameters. The use of PHREEQC shows that the  
33 light rare earth elements (LREE) and heavy rare earth elements (HREE) do not bind to HA by  
34 the same functional groups. The LREE are preferentially bound to carboxylic groups, whereas  
35 the HREE are preferentially bound to carboxy-phenolic and phenolic groups. This binding  
36 differentiation might lead to a fractionation of REE-HA patterns when competition between  
37 REE and other metals occur during complexation. A survey of the available data shows that  
38 competition with  $Al^{3+}$  could lead to the development of HREE-depleted HA patterns. This  
39 new model should improve the hydrochemical modeling of the REE since PHREEQC takes  
40 into account chemical reactions such as mineral dissolution/precipitation equilibrium and  
41 redox reactions, but also models kinetically controlled reactions and one-dimensional  
42 transport.

43

44 Keywords: rare earth elements, humic acid, complexation, modeling, PHREEQC, Model VI,  
45 site heterogeneity.

46

## 1. INTRODUCTION

47 Dissolved humic acids (HA) are complex macro-molecules that contain a large variety  
48 of functional groups giving rise to strong binding capacities with cations. Because of HA

49 ubiquity, a lot of effort has been made in the past 20 years to set up thermodynamic models  
50 that are suitable for predicting cation complexation by HA in natural waters (Marinsky and  
51 Ephraim, 1986; Tipping and Hurley, 1992; Koopal et al., 1994; Kim and Czerwinski, 1996;  
52 Tipping, 1998; Sasaki et al., 2008). The two major types of HA-cation binding surface groups  
53 are the carboxylic and phenolic groups. Each type covers a large range of cation stability  
54 constants ( $\log K$ ). The two most used cation-HA binding models are Model VI (Tipping,  
55 1998), in which both the carboxylic and phenolic groups are discrete entities displaying  
56 specific stability constants, and NICA-Donnan (Kinniburgh et al., 1996), in which continuous  
57 distributions of  $\log K$  are considered. In both models, the variety of the HA binding sites is  
58 taken into account, as the  $\log K$  distribution is calculated from the cation-carboxylic and  
59 cation-phenolic specific binding parameters. By fitting the datasets for each of these two  
60 models, Tipping (1998) and Milne et al. (2003) were able to determine the generic HA  
61 binding parameters for a wide variety of cations. They found that cation-phenolic model  
62 parameters can be mathematically correlated to cation-carboxylic parameters. These  
63 mathematical correlations are very useful to estimate phenolic binding parameters from  
64 carboxylic binding parameters, notably when experimental data are limited or non-existent.  
65 For Model VI, Tipping (1998) chose to impose a linear relationship between the carboxylic  
66 and phenolic cation binding parameters ( $\log K_{MA}$  and  $\log K_{MB}$ , respectively). Thus, only one  
67 parameter has to be adjusted in order to describe the interaction of cations with humic acids  
68 using Model VI, i.e.  $\log K_{MA}$ .

69 Rare earth elements (REE) are a highly coherent series of elements, whose chemical  
70 properties vary regularly along the series. Their relative affinity for ligands can be visualized  
71 in the plot of the intrinsic stability constants ( $\log K$ ) or the partition coefficient between two  
72 phases ( $\log K_d$ ) versus the REE atomic number. Different and specific shapes of  $\log K$

73 patterns are therefore displayed depending on the ligands (Byrne and Li, 1995). Moreover, the  
74 REE are particularly suitable to study cation-HA interactions. Experimental and field studies  
75 showed that when HAs are present in natural waters, most of the REE occur as REE-HA  
76 complexes (Bidoglio et al., 1991; Takahashi et al., 1997; Viers et al., 1997; Dia et al., 2000;  
77 Tang and Johannesson, 2003; Johannesson et al., 2004; Gruau et al., 2004; Davranche et al.,  
78 2005; Sonke and Salters 2006; Pourret et al., 2007a, b; Pédrot et al., 2008). Recent  
79 experimental studies on REE-HA binding have shown that the metal loading (REE/HA) can  
80 modify the REE partitioning between an aqueous solution and HA (Marsac et al., 2010;  
81 Yamamoto et al., 2010). At high loading, a middle-REE (MREE) downward concavity,  
82 typical of REE binding to low affinity carboxylic sites, was observed. However, under low  
83 REE/HA conditions, the log  $K_d$  pattern exhibited a regular increase from La to Lu, which is  
84 typical of REE binding to strong affinity multidentate sites (Marsac et al., 2010; Yamamoto et  
85 al., 2010). These two types of log  $K_d$  patterns correspond to the log  $K$  patterns obtained for  
86 the corresponding simple organic ligands (namely acetic acid for the carboxylic sites, and  
87 EDTA for the multidentate sites). Their successive occurrence is plausible since the site  
88 density of the HA high affinity sites is much lower than the density of the low affinity sites.  
89 Marsac et al. (2010) showed that in order to properly model this data with Model VI, it was  
90 necessary to let the parameter accounting for the high affinity sites,  $\Delta LK_2$ , vary among the  
91 REE series instead of keeping it constant, as commonly assumed. However, although there is  
92 a good agreement between the experimental and modeled data, the HA sites defined in Model  
93 VI are still not fully satisfactory for the REE series. The assumption of a similar log  $K$  pattern  
94 shape for the carboxylic and phenolic sites, imposed by the Log  $K_{MA}$ -Log  $K_{MB}$  linear  
95 relationship, is not consistent with the log  $K$  pattern observed for the REE-catechol complex -  
96 an analog of phenol - which exhibits a regular increase in heavy-REE (HREE) log  $K$  values  
- 4 -

97 (IUPAC, Stability Constants Database). Because phenolic sites have a lower acidic constant  
98 ( $K_a$ ) than carboxylic sites and occur in lower density (Ritchie and Perdue, 2003), their log K  
99 pattern could theoretically be directly determined by performing REE-HA binding  
100 experiments at  $\text{pH} > 7$  and low REE/HA ratios. However, these kinds of experiments are  
101 technically hard to perform with regard to the low aqueous REE concentration imposed by the  
102 pH values. Marsac et al. (2010) and Yamamoto et al. (2010) therefore carried out their  
103 experimental studies at a low pH (3 and 4.7, respectively) in order to be able to measure the  
104 REE aqueous concentration at low REE/HA ratios. Pourret et al. (2007b) explored higher pH  
105 values (up to 10.5), but the high imposed REE/HA ratio prevented the study of the phenolic  
106 log K patterns. This problem might be circumvented by studying REE-HA binding through  
107 competition ligand experiments with EDTA, as previously done by Sonke and Salters (2006).  
108 However, as stated by Yamamoto et al. (2010), the increasing log K(REE-EDTA) values from  
109 La to Lu implies that each REE binding is not studied at the same metal loading, thereby  
110 preventing the comparison of the log K patterns with the patterns for the conditional stability  
111 constants.

112 In this study, a modeling approach was used to evaluate whether the HA phenolic sites  
113 differ from the HA carboxylic sites in their ability to complex REE. The stability constants of  
114 the REE-HA phenolic sites were thus independently estimated. For this purpose, Model VI  
115 equations were introduced into PHREEQC (Parkhurst and Appelo, 1999), following Liu et al.  
116 (2008). The PHREEQC/Model VI model was then used to fit the HA-REE binding datasets  
117 published by Sonke and Salters (2006), Pourret et al. (2007b) and Marsac et al. (2010).

118

## 2. MATERIALS AND METHODS

### 119 2.1. Humic Ion binding Model VI

120 The humic-ion binding Model VI was developed by Tipping (1998) and introduced  
121 into the WHAM 6 program. A thorough description of Model VI can be found in Tipping  
122 (1998). The model is a discrete binding site model which takes electrostatic interactions into  
123 account. Eight sites are considered and divided into an equal number of type A sites (the weak  
124 acidic sites, commonly associated with carboxylic functional groups) and type B sites (the  
125 strong acidic sites, commonly associated with phenolic functional groups). There are  $n_A$  (mol  
126  $g^{-1}$ ) type A sites and  $n_A/2$  type B sites. Proton binding is described by the two median intrinsic  
127 constants ( $pK_A$  or  $pK_B$ ) and two parameters defining the spread of the equilibrium constants  
128 around the median ( $\Delta pK_A$  or  $\Delta pK_B$ ). The intrinsic equilibrium constants for cation binding are  
129 defined by two median constants ( $\log K_{MA}$  and  $\log K_{MB}$ ), together with parameters ( $\Delta LK_{1A}$   
130 and  $\Delta LK_{1B}$ ) that define the spread of the values around the medians. By considering the  
131 results from many datasets, a universal average value of  $\Delta LK_1$  was obtained for all cations  
132 (2.8) and a linear relationship was established between the  $\log K_{MB}$  and  $\log K_{MA}$  parameters  
133 (Tipping, 1998):

$$134 \quad \log K_{MB} = 3.39 \times \log K_{MA} - 1.15 \text{ (Eq. 1)}$$

135 Therefore, only one single adjustable parameter ( $\log K_{MA}$ ) is necessary to describe metal  
136 complexation by HA in Model VI. In Model VI, cations first hydrolysis products (e.g.  
137  $LaOH^{2+}$ ) are also considered to bind to HA with the same parameters than the corresponding  
138 aquo ions.

139 Some of the type A and type B monodentate sites can form bi- and tridentate sites. The  
140 stability constants of these bi- and tridentate sites are defined by the sum of the  $\log K$  values

141 of the monodentate sites of which they are constituted. Cation-HA binding can occur through  
142 carboxylic groups (CG), carboxy-phenolic groups (CPG) or phenolic groups (PG). A small  
143 part of the stability constants of multidentate groups are increased by the  $\Delta LK_2$  parameter, the  
144 so-called "strong binding site term" (Tipping, 1998).  $\Delta LK_2$  is attributed to both favorable  
145 spatial arrangements of the ligands, which stabilizes the complex, and to the chelate effect or  
146 participation of ligands, such as the nitrogen-containing groups.  $\Delta LK_2$  is the parameter of the  
147 HA site heterogeneity. An elevation of  $\Delta LK_2$  increases the range of the stability constants  
148 covered by the HA sites, thus increasing the heterogeneity of the HA sites. The total  
149 heterogeneity of the HA surface sites in Model VI is defined by 80 different sites.

150 An electrical double layer, where only counter-ions can accumulate, is defined. The  
151 double layer thickness is set by the Debye-Hückel parameter  $\kappa$ . The distribution of ions  
152 between the diffuse layer and the bulk solution is calculated by a simple Donnan model. The  
153 Donnan volume ( $V_D$ ) is the volume adjacent to the surface. The electrostatic correction is  
154 represented by an empirical equation which mimics the Boltzmann factor:

$$155 \quad \text{Exp}(-zF\psi/RT) = \exp(-2PzZ \log_{10}I) \text{ (Eq. 2)}$$

156 where P is an adjustable parameter, z is the ion charge, Z is the net humic acid charge  
157 (eq/g), I is the ionic strength ( $\text{mol L}^{-1}$ ),  $\psi$  is the surface potential (V), T is the temperature (K),  
158 F is the Faraday constant and R is the gas constant. The Model VI parameters used in the  
159 present study are summarized in Table 1.

## 160 **2.2. PHREEQC/Model VI coupling**

161 PHREEQC is based on an ion-association aqueous model and was designed to perform  
162 speciation and saturation-index calculations (Parkhurst and Appelo 1999). The cation-HA  
163 binding reaction with the 80 types of HA sites defined in Model VI was written and  
- 7 -

164 introduced into the PHREEQC “minteq.v4” database. Before removing the constraint  
165 imposed by the linear relationship between  $\log K_{MA}$  and  $\log K_{MB}$ , a first version of  
166 PHREEQC/Model VI using the equations defined in WHAM 6/Model VI was tested for its  
167 ability to give identical results to WHAM 6/Model VI. Humic acid was defined as the  
168 SOLUTION\_MASTER\_SPECIES, SOLUTION\_SPECIES and PHASES in the “minteq.v4”  
169 database. The 80 types of sites considered by Model VI were defined as the  
170 SURFACE\_MASTER\_SPECIES, and their respective stability constants as the  
171 SURFACE\_SPECIES in the “minteq.v4” database. This procedure is similar to that used  
172 previously by Liu et al. (2008) to model the binding of Eu by HA in clay pore water using the  
173 same PHREEQC-Model VI combination, except that in this study, we used  $\Delta LK_2$  values  $\neq 0$ .  
174 Therefore, the 80 binding sites defined by Tipping (1998) are taken into account versus only  
175 32 sites in Liu et al. (2008). In order to model similar REE binding by each specific HA sites  
176 with WHAM 6/Model VI and PHREEQC/Model VI, REE-HA binding simulations were  
177 compared with the corresponding REE-simple ligands complexes in solution. The results  
178 obtained differed between both models. The  $\Delta LK_1$  parameter (=2.8) should actually be  
179 defined as  $\Delta pK_A - \Delta LK_{1A}$ , for carboxylic groups, and as  $\Delta pK_B - \Delta LK_{1B}$ , for phenolic groups  
180 (i.e.  $\Delta LK_{1A} = -0.7$  and  $\Delta LK_{1B} = 0.8$ ). In order to model the non-specific binding in a similar  
181 fashion (i.e. ions accumulated around HA molecules by electrostatic attraction) as well as the  
182 electrostatic correction in WHAM 6/Model VI and PHREEQC/Model VI, the Gouy-Chapman  
183 type double layer defined in PHREEQC was coupled with a Donnan phase, excluding the  
184 presence of co-ions. The Debye-Hückel parameter  $\kappa$  was calculated from the relationship:  $\kappa =$   
185  $(3.29 \times 10^9 \times I^{1/2})^{-1}$  (Appelo and Postma, 2005). In PHREEQC, the surface area (SA), upon  
186 which the Boltzmann factor depends, is an adjustable parameter. Marsac et al. (2010)

187 determined Model VI REE-HA specific binding parameter with  $P = -330$ . To satisfy equation  
188 2, the PHREEQC SA was simply determined by fitting the HA surface charge ( $Q_{WHAM}$  in eq  
189  $g^{-1}$ ) generated by WHAM 6/Model VI as a function of pH, at varied NaCl concentrations and  
190 10 mg of HA. The best fit was obtained for a SA equal to 40000, 19000 and 13000  $m^2 g^{-1}$  at  
191 an ionic strength 0.001, 0.01 and 0.1, respectively. By this method we found  $Q_{PHREEQC} = 0.92$   
192  $\times Q_{WHAM} + 3 \times 10^{-4}$  ( $R^2 = 1$ ). Since SA decreases with P, our results are consistent with those of  
193 Liu et al. (2008) who determined  $SA = 30000 m^2 g^{-1}$  at an ionic strength of  $0.015 mol L^{-1}$  and  
194  $P = -120$ .

### 195 **2.3. Experimental data**

196 Three datasets were used to calibrate the PHREEQC/Model VI parameters for REE  
197 binding to HA (Sonke and Salters, 2006; Pourret et al., 2007b; Marsac et al., 2010). These  
198 datasets were preferentially chosen because they provide data for the entire REE group and  
199 they covered a large range of pHs, metal loading and organic ligand competition. The  
200 experimental conditions of the three datasets used are given below.

#### 201 2.3.1. Pourret et al.'s. dataset (2007b)

202 Pourret et al. (2007b) performed REE-HA binding experiments for the 14 REE  
203 simultaneously over pHs ranging from 2 to 10.5 at an ionic strength (IS) of  $10^{-3} M$ . The HA  
204 used was purified Aldrich Humic Acid. The experiments were performed with 50 ppb of each  
205 REE at three HA concentrations: 5, 10 and 20  $mg L^{-1}$  ( $5 \times 10^{-3} < REE/HA < 2 \times 10^{-2} mol REE/mol$   
206 C). The separation of HA from the inorganic complexes was performed by ultrafiltration. The  
207 REE concentrations were measured by Inductively Coupled Plasma Mass Spectrometry (ICP-  
208 MS). The results showed an increase in the REE-HA binding with increasing pH and

209 increasing HA concentrations. The log K REE-HA pattern is marked by a downward, MREE  
210 concavity. These data therefore provide a dataset for REE-HA binding at high metal loading  
211 and over a large pH range where the REE are bound to HA by carboxylic or phenolic sites,  
212 depending on the pH. Note however, that the data at  $\text{pH} > 7$  were not considered in the model  
213 calibration, since 100% of the REE were bound to HA in Pourret et al. (2007b), making the  
214 data useless for the calibration of the model.

### 215 2.3.2. Marsac et al.'s dataset (2010)

216 Marsac et al. (2010) used Pourret et al.'s method (2007b) to study the metal loading  
217 effect on REE binding to HA ( $4 \cdot 10^{-4} < \text{REE}/\text{HA} < 2.7 \cdot 10^{-2}$  mol REE/mol C). In order to obtain  
218 measurable REE concentrations in the solution, Marsac et al. (2010) performed all their  
219 experiments under acidic conditions ( $\text{pH} = 3$ ). The ionic strength was fixed at  $10^{-2}$  M. Marsac  
220 et al. (2010) observed an increase in the REE-HA binding and a change in the log K REE-HA  
221 pattern with decreasing metal loading. At high metal loading, the log K REE-HA pattern  
222 exhibited a MREE-downward concavity and at low metal loading, a regular increase from La  
223 to Lu was observed. These data therefore provide an REE-HA complexation dataset at  
224 variable metal loading under acidic pH conditions. At pH 3 and low metal loading, the REE  
225 are bound to HA by low density strong multidentate sites, whereas at high loading, the REE  
226 are primarily bound to high density weak sites.

### 227 2.3.3. Sonke and Salters' dataset (2006)

228 Sonke and Salters (2006) performed REE-HA binding experiments for each REE  
229 separately. The HA used was purified Leonardite HA. These authors used EDTA as a  
230 competitive ligand of HA for the REE. The experiments were carried out at low metal loading

231 ( $2 \cdot 10^{-4}$  mol REE/mol C), slightly acidic to basic pH (6 to 9) and an IS = 0.1 M. The separation  
232 and measurement of the EDTA-REE and HA-REE complexes were done by capillary  
233 electrophoresis coupled with ICP-MS. This study provides single REE-HA binding data for  
234 high pH and low metal loading, two conditions that cannot be obtained by the methods  
235 developed by Pourret et al. (2007b) and Marsac et al. (2010). To simulate Sonke and Salters'  
236 experiments (2006), EDTA was introduced into the PHREEQC “minteq.v4” database as the  
237 SOLUTION\_MASTER\_SPECIES and SOLUTION\_SPECIES. The four EDTA protonation  
238 constants and the fourteen REE-EDTA stability constants for IS = 0 given in Sonke and  
239 Salters (2006) were also introduced as the SOLUTION\_SPECIES.

#### 240 **2.4. Modeling strategy**

241 Linear free energy relationships (LFER) are commonly used to determine the cation-  
242 HA binding parameters (Tipping, 1998). This method was used by several authors to estimate  
243 the log  $K_{MA}$  parameter in Model VI (or the previous version, Model V) in order to describe  
244 the REE-HA binding (Tang and Johannesson, 2003; Sonke, 2006; Pourret et al., 2007b). In  
245 particular, Tang and Johannesson (2003) and Pourret et al. (2007b) demonstrated that REE-  
246 HA binding could be reasonably well modeled by using the existing LFER between log  $K_{MA}$   
247 ( $pK_{MA}$  in Model V) and log K for metal complexation by acetic acid. However, Marsac et al.  
248 (2010) demonstrated that log  $K_{MA}$  for the HREE are lower than those estimated by LFER  
249 when log  $K_{MA}$  and  $\Delta LK_2$  were optimized simultaneously. The decrease in log  $K_{MA}$   
250 compensates for the high  $\Delta LK_2$  values obtained for the HREE. Here, the new REE log  $K_{MA}$   
251 were optimized, thus keeping a linear relationship with log  $K_{MA}$  reported in Marsac et al.  
252 (2010). This method allows preserving the same pattern. Log  $K_{MA}$ , the REE-HA carboxylic  
253 binding parameter, mainly influences the calculation of REE-HA binding at acidic to neutral

254 pHs since the carboxylic groups have high proton dissociation constants. Therefore, the log  
255  $K_{MA}$  values were first optimized by fitting the "acidic" experimental data from Pourret et al.  
256 (2007b) and Marsac et al. (2010).

257 Log  $K_{MB}$  represents the binding parameter for the HA phenolic groups. However, no  
258 stability constants are currently available for REE complexation by phenol. Therefore,  
259 catechol was used as an analog. A linear relationship was applied between the log K values  
260 for REE-catechol (IUPAC stability constant database) and log  $K_{MB}$  in order to impose a  
261 phenolic-type pattern onto log  $K_{MB}$ . Because phenolic groups have low acidic constants ( $K_a$ ),  
262 the influence of log  $K_{MB}$  on REE-HA binding is mainly observed at neutral to basic pHs.  
263 Therefore, log  $K_{MB}$  was optimized by fitting the data of Sonke and Salters (2006).

264 As already stated,  $\Delta LK_2$  represents the HA heterogeneity parameter or the strong  
265 binding site term of Model VI.  $\Delta LK_2$  indicates the presence of multidentate sites, whose  
266 affinity for REE increases from La to Lu (Byrne and Li, 1995; Sonke and Salters, 2006;  
267 Pourret et al., 2007b; Marsac et al., 2010). Tipping (1998) suggests that the  $\Delta LK_2$  parameter is  
268 similar for the carboxylic and phenolic multidentate sites. However, preliminary experimental  
269 data fitting, carried out during the course of this study, showed that the three datasets used to  
270 optimize the HA binding parameters could not be satisfactorily modeled, except when using  
271 different  $\Delta LK_2$  values for the carboxylic and phenolic sites. In the same way that  
272 PHREEQC/Model VI allows independent log  $K_{MA}$  and log  $K_{MB}$  optimization, it also allows  
273 different  $\Delta LK_2$  values to be ascribed to the multidentate sites depending on their composition  
274 in carboxylic ( $\Delta LK_{2C}$ ) or phenolic ( $\Delta LK_{2P}$ ) sites. Note that the strong binding site terms of the  
275 carboxylic-phenolic multidentate sites are calculated as the  $\Delta LK_{2C}$  and  $\Delta LK_{2P}$  weighted  
276 average, depending on the chemical composition of the sites. Because the density of the

277 multidentate sites is low at the surface of HA, their control on REE binding occurs mainly  
278 under low metal loading conditions, and under low and high pH conditions for the carboxylic  
279 and phenolic sites, respectively. Consequently,  $\Delta\text{LK}_{2\text{C}}$  optimization was first performed with  
280 Marsac et al.'s data (2010) and  $\Delta\text{LK}_{2\text{P}}$  optimization with Sonke and Salters' data (2006). Both  
281  $\Delta\text{LK}_{2\text{C}}$  and  $\Delta\text{LK}_{2\text{P}}$  were constrained to increase progressively from La to Lu due to a linear  
282 correlation between the new  $\Delta\text{LK}_{2\text{C}}$  and  $\Delta\text{LK}_{2\text{P}}$  and the  $\Delta\text{LK}_2$  previously reported by Marsac  
283 et al. (2010).

284 Errors between the experimental data and the calculations were quantified by the root  
285 mean square error of the regression (rmse), i.e. the sum of the squares of the differences  
286 between the observed and calculated  $\log v$ , where  $v$  is the amount of REE bound to HA per  
287 gram of dissolved organic carbon (DOC). The fit was considered as being good when the total  
288 rmse calculated from the whole REE and experimental conditions was low and when the rmse  
289 values for each REE were close to each other. The modeling strategy here above described  
290 shows that some of WHAM 6/Model VI constrains on parameters were suppressed only to  
291 introduce other comparable constrains. Therefore, WHAM 6/Model VI and  
292 PHREEQC/Model VI fit goodness can be compared directly with rmse since the degree of  
293 freedom does not significantly varies from a model to another.

## 294 **3. RESULTS**

### 295 **3.1. Modeling parameters**

296 The new REE-HA binding parameters were obtained from the combination of  
297 PHREEQC/Model VI and the assumptions and strategy described above. They were adjusted  
298 until a reasonable fit - quantified by the rmse - was reached. The results are presented in Table

299 2. In Figure 1, they are compared with the REE-HA binding parameters obtained by Marsac  
300 et al. (2010) with WHAM 6/Model VI and the corresponding model ligands (i.e. acetic acid  
301 for  $\log K_{MA}$  and catechol for  $\log K_{MB}$ ). The following equations for the linear relationships  
302 were used to determine the REE-HA specific binding parameters with PHREEQC/Model VI:

$$303 \quad \log K_{MA} = 2.12 \times \log K_{MA}(1) - 2.44 \quad (R^2 = 0.98) \quad (3)$$

$$304 \quad \log K_{MB} = 0.37 \times \log K(\text{REE-catechol}) + 0.86 \quad (R^2 = 0.95) \quad (4)$$

$$305 \quad \Delta LK_{2C} = 0.26 \times \Delta LK_2(1) + 1.74 \quad (R^2 = 0.97) \quad (5)$$

$$306 \quad \Delta LK_{2P} = 0.72 \times \Delta LK_2(1) + 3.60 \quad (R^2 = 1) \quad (6)$$

307 where  $\log K_{MA}(1)$  and  $\Delta LK_2(1)$  refer to the  $\log K_{MA}$  and  $\Delta LK_2$  values determined by  
308 Marsac et al. (2010).

309 The removal of the constraint imposed by the linear relationship between  $\log K_{MA}$  and  
310  $\log K_{MB}$  in WHAM6/Model VI and the use of REE-catechol binding parameters results in a  
311 significant increase of the  $\log K_{MA}$  values (3.29 versus 2.7, on average, in Marsac et al. 2010).  
312 By contrast, the  $\log K_{MB}$  values are greatly reduced: from 8.02 to 4.93, on average, in Marsac  
313 et al. (2010) and as reported in this study, respectively. The  $\log K_{MB}$  values produced  
314 independently with PHREEQC/Model VI are very different from the  $\log K_{MB}$  values  
315 calculated with the  $\log K_{MB}$ - $\log K_{MA}$  linear correlation established by Tipping (1998) in  
316 WHAM6/Model VI (Eq. 1). The present  $\log K_{MA}$  and  $\log K_{MB}$  are not related to each other by  
317 a linear correlation, but their relative order of magnitude order is preserved,  $\log K_{MB} > \log$   
318  $K_{MA}$ . Tipping (1998) obtained values of  $\log K_{MA}(\text{Eu}) = 3.09$  and  $\log K_{MB}(\text{Eu}) = 5.33$ , when  
319 they independently optimized the  $\log K_{MA}$  and  $\log K_{MB}$  values for Eu during the construction  
320 of Model VI. These values are close to the values obtained here in this study:  $\log K_{MA}(\text{Eu}) =$   
321  $3.36$  and  $\log K_{MB}(\text{Eu}) = 4.98$ ; (Table 2). The  $\Delta LK_{2C}$  and  $\Delta LK_{2P}$  values obtained here are

322 higher than the  $\Delta\text{LK}_2$  values reported in Marsac et al. (2010) (denoted as  $\Delta\text{LK}_2(1)$  in Table 2)  
323 for all REE. Hence, HA appears more heterogeneous with respect to REE binding when the  
324 constraint of a single  $\Delta\text{LK}_2$  value is deleted. The  $\Delta\text{LK}_{2C}$  and  $\Delta\text{LK}_{2P}$  patterns also differ from  
325 the  $\Delta\text{LK}_2$  patterns given in Marsac et al. (2010). The major change concerns the  $\Delta\text{LK}_{2C}$   
326 pattern which presents a lower increase for HREE than Marsac et al. (2010)  $\Delta\text{LK}_2$  (Fig. 1c).  
327 This lower increase for HREE is the consequence of both the introduction of REE-phenolic  
328 site binding constants that increase from La to Lu, and the propagation of this increase to the  
329 phenolic containing-multidentate site binding constants. As a result, the strong multi-phenolic  
330 sites, which depend on  $\Delta\text{LK}_{2P}$ , are no longer constrained to accommodate so much HREE,  
331 which explains why the increase in  $\Delta\text{LK}_{2P}$  for HREE is less pronounced as compared to  $\Delta\text{LK}_2$   
332 presented in Marsac et al. (2010) (Fig. 2c).

333 In Model VI, the  $\Delta\text{LK}_2$  are thought to depend on either the geometry of the carboxylic  
334 and phenolic multidentate sites ( $\Delta\text{LK}_2$  accounts for favorable arrangement of the spatial  
335 ligands that stabilize REE complexes), or the involvement of additional groups, notably  
336 nitrogen-containing groups (Tipping, 1998). Tipping (1998) suggested that  $\Delta\text{LK}_2$  could be  
337 estimated from the metal stability constants with  $\text{NH}_3$ . Therefore,  $\Delta\text{LK}_2$  is linearly correlated  
338 with the stability constant of metal- $\text{NH}_3$  complexes in WHAM 6/Model VI. The  $\Delta\text{LK}_2$  values  
339 obtained for the REE using this linear relationship are very low (e.g. 0.29 for Eu), and much  
340 lower than the  $\Delta\text{LK}_2$  values calculated in Marsac et al. (2010), and the  $\Delta\text{LK}_{2C}$  and  $\Delta\text{LK}_{2P}$   
341 values obtained in the present study. This may suggest that the participation of N-containing  
342 groups only plays a small role in the strong, bi- and tridentate binding sites. The strength of  
343 these sites has to be mainly due to the development of favorable spatial arrangements of  
344 carboxylic and phenolic ligands as well as the effect of chelation. It is interesting to note that

345 the high  $\Delta\text{LK}_{2P}$  values stemming from the new simulations compensate for the decrease in the  
346  $\log K_{\text{MB}}$  values generated by this simulation. Because phenolic groups have high pKa (Ritchie  
347 and Perdue, 2003), only a small fraction of the REE is able to bind to monodentate phenolic  
348 ligands, as the majority remain bound to specific multidentate, phenol-containing groups.

### 349 **3.2. Experimental data fit**

350 The three experimental datasets used to optimize the Model VI parameters were  
351 modeled separately. The results of the simulations with WHAM 6/Model VI and  
352 PHREEQC/Model VI are compared in Figure 2 for (i) Eu binding by HA at pH 2 to 7 and  
353 under high metal loading conditions (Pourret et al.'s experimental dataset (2007b); Fig. 2a),  
354 (ii) La and Lu binding by HA at pH = 3 under variable metal loading conditions (Marsac et  
355 al.'s dataset (2010); Fig. 2b), and (iii) Eu binding by HA at pH 6 to 9 with EDTA, and under  
356 low metal loading conditions (Sonke and Salters's dataset (2006); Fig. 2c). The quality of the  
357 fits can also be evaluated by comparing the measured and simulated REE distribution patterns  
358 ( $\log K_d$ ) for the data of Pourret et al. (2007b) and Marsac et al. (2010) (Figs. 2d and 2e,  
359 respectively), or the measured and simulated patterns of the conditional REE-HA stability  
360 constants ( $\log \beta_{\text{HS}}$ ) for Sonke and Salters' data (2006) (Fig. 2f). The  $\log \beta_{\text{HS}}$  values were  
361 corrected from the REE-EDTA interaction (Sonke and Salters 2006). The new hypotheses in  
362 PHREEQC/Model VI model significantly improve the quality of the fits as the total rmse (i.e.  
363 calculated for the whole REE and conditions studied) decreased from 0.10 to 0.07 (Table 2).  
364 There is a moderate improvement in the fit for Pourret et al.'s results (2007b) (from 0.10 to  
365 0.07), which is primarily seen in high pH conditions at low metal loading (Figs. 2a and 2d).  
366 By contrast, the improvement of the fit is very significant for Sonke and Salters' data (2006),  
367 as the rmse is divided by a factor of around 2: 0.10 in the new calculations versus 0.26 in

368 WHAM 6/Model VI. This improvement is particularly significant for the LREE at high pH  
369 (e.g. at pH 9 in Fig. 2f). The strong overestimation of the  $\log \beta_{HS}$  values obtained for the  
370 LREE under this pH condition with WHAM 6/Model VI is no longer apparent. At a high pH  
371 and low REE/HA ratio, the REE are strongly bound to the HA phenolic groups. The LREE  
372 overestimation of the latter with WHAM 6/Model VI is the consequence of the  $\log K_{MA}$ - $\log$   
373  $K_{MB}$  linear relationship, the HA phenolic sites present high stability constants with LREE.  
374 Rmse increases slightly for Marsac et al.'s results (2010), from 0.02 with WHAM 6/Model VI  
375 to 0.04 with the new parameters, but remains very low.

## 376 **4. DISCUSSION**

377 The independent optimization of the HA carboxylic and phenolic binding parameters  
378 (i.e. removal of the  $\log K_{MA}$ - $\log K_{MB}$  linear relationship and  $\Delta LK_2$  dissociation for the  
379 carboxylic and phenolic groups) allowed for the first time to satisfactorily and simultaneously  
380 model a large set of multi-REE experimental data with the same set of equations and  
381 parameters. The following discussion will show that these parameter modifications not only  
382 improve the fit of the data, but that the 80 HA sites present REE  $\log K$  patterns that are more  
383 consistent with the data available for simple organic ligands (Byrne and Li, 1995). We will  
384 also discuss the implications of these parameter modifications on REE speciation of the  
385 different HA sites. This was possible for the first time by the introduction of Model VI into  
386 PHREEQC.

### 387 **4.1. Assessing the consistency of the new modeling parameters**

388 Figure 3 presents a plot comparing the  $\log K_{Lu}/K_{La}$  ratio, which represent the  
389 fractionation of the REE  $\log K$  pattern with the average REE stability constants determined

390 for (i) 101 organic ligands studied independently (Byrne and Li, 1995) and (ii) the 80 HA  
391 sites defined in WHAM 6/Model VI and PHREEQC/Model VI. Since the REE-ligand log K  
392 values were available at IS = 0.1M, the REE-HA log K values were calculated at IS = 0.1M  
393 using the Davies equation. The log  $K_{Lu}/K_{La}$  ratio shows a positive correlation when reported  
394 against the average REE log K value ( $R^2=0.77$ ). This correlation indicates that the LREE and  
395 MREE have more affinity for weak carboxylic ligands, such as acetic acid, than the HREE  
396 whereas HREE have more affinity for strong multidentate ligands, such as EDTA, than LREE  
397 and MREE. This general increase in the log  $K_{Lu}/K_{La}$  ratio with increasing stability constants  
398 can be used to assess the consistency and accuracy of the new optimized parameters obtained  
399 for the 80 HA sites. It is remarkable that the REE-HA log K values defined by the new  
400 optimized parameters show a similar positive correlation ( $R^2=0.75$ ) as do the data for the 101  
401 ligands, whereas the parameters obtained in WHAM 6/Model VI lead to a dispersion of the  
402 log ( $K_{Lu}/K_{La}$ ) versus the average stability constants in Figure 3 (no observed correlation).  
403 More particularly, because of the log  $K_{MA}$ -Log  $K_{MB}$  linear relationship in WHAM 6/Model  
404 VI, strong HA sites (i.e. average REE log K > 15) present  $K_{La} > K_{Lu}$ , which is not observed  
405 among either the 101 organic ligands or the HA sites defined with the new PHREEQC/Model  
406 VI parameters. Consequently, PHREEQC/Model VI appears to be more realistic than WHAM  
407 6/Model VI.

#### 408 4.2. REE distribution on HA sites

409 This section compares the REE speciation on the HA surface obtained using the  
410 WHAM 6/Model VI hypotheses (i.e. using the log  $K_{MA}$ -log  $K_{MB}$  linear relationship and a  
411 single and constant  $\Delta LK_2$  value for all REE; Tipping, 1998) and the new hypotheses  
412 stemming from this study and involved in PHREEQC/Model VI.

413 In Model VI, 80 different sites are expected to participate in the cation-HA binding.  
414 To simplify the REE speciation on the HA surface, these 80 sites can be gathered in several  
415 groups based on two main characteristics: the chemical nature of the ligands (i.e. carboxylic  
416 groups: CG, phenolic groups: PG and carboxy-phenolic groups: CPG) and their denticity (i.e.  
417 mono-, bi- and tridentate sites).

418 Figure 4 displays the calculations performed at high metal loading using Pourret et  
419 al.'s experimental conditions (2007b), namely  $HA = 5\text{mg L}^{-1}$ , and over a wide pH range for  
420 La, Eu and Lu, respectively. At high metal loading, the few strong binding sites play a  
421 subordinate role in HA-REE complexation. The criterion that primarily differentiates the two  
422 simulations is the chemical nature of the binding sites (i.e. whether they are CG, CPG or PG  
423 sites). The WHAM 6/Model VI complexation hypotheses yield very similar speciation results  
424 for all of the REE. The complexation is dominated by CPG in each case ( $La_{CPG} = 50\text{-}80\%$ ;  
425  $Lu_{CPG} = 50\text{-}60\%$ ). The situation is clearly different if the new hypotheses and new modeling  
426 parameters are used. The speciation of La and Lu is no longer similar. The dominant REE-HA  
427 species for La is CG (80 to 100% vs. pH) instead of CPG when using the equations and  
428 parameters of WHAM 6/Model VI. The LREE speciation generated by the new equations and  
429 parameters (dominant role of the carboxylic groups) is consistent with the C 1s-NEXAFS  
430 analysis of Eu-HA complexes at high metal loading and low pH (pH = 5.1), which showed  
431 that Eu was bound to HA carboxylic sites (Plaschke et al., 2004; Naber et al., 2006). Note that  
432 the new equations and parameters do not induce significant changes in Lu-HA speciation.

433 With regards to the effects of site denticity under high metal loading and varying pH  
434 conditions (data not shown), the main consequence of the new equations and parameters is to  
435 reduce the relative amount of the REE bound to monodentate sites to almost 0%, whereas it  
436 could reach 15% at pH = 7 with the WHAM 6/Model VI parameters. REE-HA binding occurs  
- 19 -

437 mainly through bidentate sites (80-90%) under high metal loading conditions, regardless of  
438 the REE and pH. Only a small proportion (10 to 20%, depending on the REE) is bound to the  
439 tridentate sites. Another consequence of the new equations and parameters is a drastic  
440 decrease of the REE bound to tridentate sites at low metal loading (Marsac et al. 2010;  
441 experimental conditions). The WHAM 6/Model VI equations and parameters led to a (i) large  
442 increase of the REE bound to HA tridentate sites with decreasing metal loading at pH = 3  
443 (20% at  $2 \cdot 10^{-2}$  metal loading to 60% at  $4 \cdot 10^{-4}$  metal loading for La) (Fig. 5) and (ii)  
444 fractionation between La and Lu: the relative amount of La bound to tridentate sites increased  
445 much more strongly than that of Lu with decreasing metal loading. The new equations and  
446 parameters clearly modify the REE speciation on HA sites. Firstly, the increase of the REE  
447 bound to tridentate sites with decreasing metal loading is less pronounced. Secondly, the  
448 amount of La bound to tridentate sites always remains lower than that of Lu even under very  
449 low metal loading. This effect is clearly a consequence of the disappearance of multidentate  
450 sites combining REE  $\log K > 15$  and  $\log K_{La} > \log K_{Lu}$  in the new modeling strategy, a  
451 combination that is totally unrealistic based on Figure 3.

#### 452 **4.2. Consequences for cation competition**

453 In natural waters, the REE compete with many other dissolved cations during HA  
454 complexation. Among the cations present in natural waters, several cations, such as  $Fe^{3+}$  and  
455  $Al^{3+}$ , are known to complex strongly with HA (Tanizaki et al., 1992; Takahashi et al., 1997;  
456 Dupré et al., 1999; Olivié-Lauquet et al., 1999; Allard et al., 2004; Tipping et al., 2002;  
457 Tipping, 2005; Lofts et al., 2008), and are therefore a strong potential competitor of REE for  
458 HA binding. Experimental studies of a single REE (Lippold et al., 2005; 2007) and a  
459 modeling study of the whole REE series (Tang and Johannesson, 2003) showed that Al and

460 Fe could significantly reduce the amount of the REE bound to HA. Therefore, the question to  
461 be asked is if metal competition could modify the REE-HA pattern. This cation-REE  
462 competition can be accessed through WHAM 6/Model VI simulations. Although the  $\log K_{MA}$ -  
463  $\log K_{MB}$  linear relationship used in WHAM 6/Model VI implies that all cations will exhibit  
464 the same carboxylic to phenolic HA site affinity, the competitive metal tendency to form  
465 strong multidentate complexes (simulated by the  $\Delta LK_2$  parameter) can modify the REE-HA  
466 pattern. For instance, the  $\Delta LK_2$  value is equal to 0.46 for  $Al^{3+}$  and 2.2 for  $Fe^{3+}$  (Tipping,  
467 1998). Therefore, according to WHAM 6/Model VI's definition of HA,  $Al^{3+}$  and  $Fe^{3+}$  will  
468 compete more efficiently for the LREE and HREE, respectively, since the  $\Delta LK_2$  parameter  
469 increases from La to Lu. However, the results obtained here demonstrate that the LREE and  
470 HREE are not complexed to the same sites. The LREE are preferentially complexed to  
471 carboxylic sites, whereas the HREE are preferably bound to carboxy-phenolic sites (Fig. 4).  
472 Therefore, a metal will compete with the REE in different ways depending on whether it  
473 complexes itself preferentially with the carboxylic or phenolic sites. This unequal competition  
474 will produce REE-HA fractionation patterns that are different from the WHAM 6/Model VI  
475 simulations. The occurrence of competing metals with strong affinities for phenolic sites will  
476 lead to a release of part of the HREE bound to HA. As part of the HA phenolic groups are  
477 unavailable for REE binding, the resulting REE-HA pattern will present the specific feature of  
478 carboxylic ligands such as acetic acid (Fig. 1a). Given that competing metals have strong  
479 affinities for carboxylic sites, this will lead to a relative decrease in the LREE bound to HA.  
480 The resulting REE-HA pattern will present the specific feature of phenolic ligands such as  
481 catechol (Fig. 1b).

482 To evaluate the expected effects of metal competition on the REE-HA binding  
483 patterns, the stability constants of the potential competing metals with acetic acid and catechol  
484 were compiled and compared both between them and with the stability constants of the REE  
485 (IUPAC, Stability Constants Database). Metals such as  $\text{Al}^{3+}$  and  $\text{Cu}^{2+}$ , with regards to their  
486 high affinity for phenol-bearing sites, should preferentially compete with the HREE (Fig. 6).  
487 The resulting REE pattern should exhibit a lower increase for HREE than that expected  
488 without any competition. Cadmium competition should result in a relative LREE depletion in  
489 the REE-HA pattern, as  $\text{Cd}^{2+}$  has more affinity for carboxylic groups (assumed to be  
490 analogous with acetic acid) than for phenolic groups (assumed to be analogous with catechol)  
491 when compared with the average REE. By contrast,  $\text{Fe}^{3+}$  competition should not result in a  
492 fractionation of the REE-HA pattern since  $\text{Fe}^{3+}$  presents a relative affinity for carboxylic and  
493 phenolic groups, equal to that of the average REE. Note that the effect of  $\text{Al}^{3+}$  and  $\text{Fe}^{3+}$  on the  
494 REE-HA pattern with regards to their affinity for carboxylic and phenolic ligands is opposed  
495 to what is expected with WHAM 6/Model VI. These presumed impacts, though plausible,  
496 must nevertheless be viewed with caution: only the affinity for carboxylic and phenolic  
497 ligands is discussed here. The metal affinity for multidentate ligands is also a binding  
498 mechanism thereby affecting the REE-HA pattern. As compared to the light rare earth  
499 elements, heavy rare earth elements are indeed more bound to the HA multidentate sites.  
500 Experimental studies should be conducted to confirm if this statement is both true and  
501 accurate. The present improvements in the modeling of REE-HA binding will therefore  
502 provide a valuable and powerful tool for the interpretation of future experimental results.

503 These expected effects of metal competition on the REE-HA pattern have two  
504 important implications. (i) The change in the REE pattern could provide a way to highlight  
505 the occurrence of competing processes. It might even help to clarify the affinity of a given

506 metal for the different HA sites. Several studies have shown the difficulty of calibrating  
507 cation-HA binding parameters, especially in cation competition experiments. An example of  
508 this is Al with the NICA-Donnan model (Kinniburgh et al., 1996). Because Al is a possible  
509 competitor to Pb, Kinniburgh et al. (1999) considered that Al was only bound to phenolic  
510 groups. However, since Al is able to compete with Cd and Pb, Pinheiro et al. (2000)  
511 considered that Al and Pb were only assigned to be bound with carboxylic groups. Our results  
512 suggest that the REE-HA pattern could be used as a probe of the metal affinity for HA  
513 binding sites in further REE-metal competition experiments, depending on whether the metal  
514 competes with LREE (i.e. a higher affinity for HA carboxylic sites), HREE (i.e. a higher  
515 affinity for HA phenolic sites) or every REE. (ii) The REE pattern variability in natural  
516 organic-rich waters is generally attributed to the competition between dissolved organic  
517 matter (DOM) and carbonate (Elderfield et al., 1990; Tang and Johannesson, 2010). A  
518 MREE-enrich pattern is observed when REE binding by DOM dominates, whereas a HREE-  
519 enrich pattern is observed (a REE-carbonate log K increase from LREE to HREE; Luo and  
520 Byrne, 2004) when REE binding by carbonate dominates. However, the LREE and HREE  
521 affinity for HA can also vary with regards to the considered metal loading (cation/DOM  
522 concentration ratio). Therefore, the present study suggests that not only the metal loading  
523 could account for the diversity of the REE patterns in organic-rich waters (Marsac et al.,  
524 2010), but also the “phenolic” and “carboxylic” competing cations, especially  $\text{Al}^{3+}$  and  $\text{Fe}^{3+}$ .

## 525 **5. CONCLUSION**

526 The results show that removal of the  $\log K_{\text{MA}}\text{-}\log K_{\text{MB}}$  linear relationship existing in  
527 Model VI, coupled with the introduction of two independent  $\Delta\text{LK}_2$  sets to describe the REE  
528 equilibrium complexation with HA low density strong sites, resulted in a strong improvement

529 of REE-HA interaction modeling. This new modeling approach allows, for the first time, to  
530 simultaneously and satisfactorily model a large set of multi-REE binding to HA experimental  
531 data. The increase for HREE in the log  $K_{MB}$  pattern, obtained by removing the linear  
532 relationship, is consistent with the phenolic nature of type B sites and REE-catechol binding  
533 constants, as catechol is considered as analogous with phenolic ligands. REE log K with the  
534 80 HA sites defined in Model VI, calculated from the newly optimized REE-HA Model VI  
535 parameters, are consistent with the REE stability constants for the 101 simple organic ligands,  
536 in contrast with the previous data. The introduction of the new Model VI equations and  
537 parameters into PHREEQC allows for the first time to highlight the speciation of the REE at  
538 the HA surface. The modeling data show that the LREE and HREE are not bound to similar  
539 HA sites. The LREE are preferentially bound to carboxylic groups, whereas HREE are  
540 preferentially complexed to carboxy-phenolic and phenolic groups. As a consequence, a  
541 different REE-HA pattern could arise from competition between cations and REE for HA  
542 binding, depending on the cation affinity for these different groups. A survey of available data  
543 shows that competition with  $Al^{3+}$  should result in a release of part of the HREE bound to HA,  
544 whereas competition with  $Fe^{3+}$  should not have any effect on the REE-HA pattern. Competing  
545 cations for phenolic and carboxylic HA sites may therefore account for the diversity of the  
546 REE patterns observed in natural organic-rich waters.

547 Further experimental studies have to be conducted to confirm the accuracy of these  
548 predictions, which should largely benefit from the performed improvements in the modeling  
549 of REE-HA interactions. The PHREEQC/Model VI model set-up provides a valuable and  
550 powerful tool for the interpretation of complex REE-competitor metal-HA interactions. The  
551 PHREEQC/Model VI coupling should also improve the hydrochemical modeling of the REE  
552 since PHREEQC/Model VI accounts for chemical reactions such as mineral

553 dissolution/precipitation, redox reactions, and the modeling of kinetically-controlled reactions  
554 and one-dimensional transport. Finally, it should be highlighted that the new modeling results  
555 argue for the development of fine spectroscopic measurements to confirm the exact nature of  
556 the HA sites involved in the complexation of the REE by HA.

557

#### 558 SUPPORTING INFORMATION AVAILABLE

559 Appendices containing PHREEQC input files, the PHREEQC database used in this study and  
560 the detailed calculation of the REE-HA stability constants are available free of charge. In the  
561 latter document a method is proposed to introduce another cation in PHREEQC/Model VI.

562

#### 563 ACKNOWLEDGEMENTS

564 We thank Dr C. Bruggeman and E. Tipping for their help in the PHREEQC/Model VI  
565 coupling. This research was funded by the French ANR, through the "Programme Jeunes  
566 Chercheuses - Jeunes Chercheurs"/"SURFREE: Rare earth elements partitioning at solid-  
567 water interface: Impact on REE geochemical behaviour and tracing properties". Dr. Sarah  
568 Mullin is acknowledged for post-editing the English style. We thank the two anonymous  
569 reviewers and AE, K. H. Johannesson, for their helpful comments.

570

#### 571 REFERENCES

572 Allard T., Menguy N., J. Salomon J., Calligaro T., Weber T., Calas G., and Benedetti M. F.  
573 (2004) Revealing forms of iron in river-borne material from major tropical rivers of the  
574 Amazon Basin (Brazil). *Geochim. Cosmochim. Acta* **68**, 3079-3094.  
575 Appelo C. and Postma D. (2005) *Geochemistry, groundwater and pollution* (2<sup>nd</sup> edition).  
576 Taylor & Francis, p. 595.

577 Bidoglio G., Grenthe I., Qi P., Robouch P. and Omentto N. (1991) Complexation of Eu and  
578 Tb with fulvic acids as studied by time-resolved laser-induced fluorescence. *Talanta* **38**  
579 (9), 999-1008.

580 Byrne R. H. and Li B. (1995) Comparative complexation behaviour of the rare earth elements.  
581 *Geochim. Cosmochim. Acta* **59**, 4575-4589.

582 Davranche M., Pourret O., Gruau G., Dia A. and Le Coz-Bouhnik M. (2005) Adsorption of  
583 REE(III)-humate complexes onto MnO<sub>2</sub>: experimental evidence for cerium anomaly  
584 and lanthanide tetrad effect suppression. *Geochim. Cosmochim. Acta* **69**, 4825-4835.

585 Dia A., Gruau G., Olivie-Lauquet G., Riou C., Molénat J. and Curmi P. (2000) The  
586 distribution of rare-earths in groundwater: assessing the role of source-rock  
587 composition, redox changes and colloidal particles. *Geochim. Cosmochim. Acta* **64**,  
588 4131-4151.

589 Dupré L., Viers J., Dandurand J.-L., Polve M., Bénézech P., Vervier P., and Braun J.-J. (1999)  
590 Major and trace elements associated with colloids in organic-rich river waters:  
591 Ultrafiltration of natural and spiked solutions. *Chem. Geol.* **160**, 63-80.

592 Elderfield H., Upstill-Goddard R. and Sholkovitz E.R. (1990) The rare earth elements in  
593 rivers, estuaries, and coastal seas and their significance to the composition of ocean  
594 waters. *Geochim. Cosmochim. Acta* **54**, 971-991.

595 Gruau G., Dia A., Olivie-Lauquet G., Davranche M. and Pinay G. (2004) Controls on the  
596 distribution of rare earth elements in shallow groundwaters. *Wat. Res.* **38**, 3576-3586.

597 IUPAC (2001) *IUPAC Stability Constants Database*. Version 5.4, IUPAC and Academic  
598 Software (acadsoft@bcs.org.uk).

- 599 Johannesson K.H., Tang J., Daniels J.M., Bounds W.J. and Burdige D.J. (2004) Rare earth  
600 element concentrations and speciation in organic rich blackwaters of the Great Dismal  
601 Swamp, Virginia, USA. *Chem. Geol.* **209**, 271-294.
- 602 Kim J.I. and Czerwinski K.R. (1996) Complexation of metal ions with humic acid: metal ion  
603 charge neutralisation model. *Radiochim. Acta* **73**, 5-10.
- 604 Kinniburgh D. G., van Riemsdijk W. H., Koopal L. K., Borkovec M., Benedetti M. F., and  
605 Avena M. J. (1999) Ion binding to natural organic matter: competition, heterogeneity,  
606 stoichiometry and thermodynamic consistency. *Colloid Surf. A* **151**, 147-166.
- 607 Kinniburgh D.G., Milne C.J., Benedetti M.F., Pinheiro J.P., Filius J., Koopal L. and Van  
608 Riemsdijk W.H. (1996) Metal ion binding by humic acid: application of the NICA-  
609 Donnan model. *Environ. Sci. Technol.* **30**, 1687-1698.
- 610 Koopal L.K., van Riemsdijk W.H., de Wit J.C.M and Benedetti M.F. (1994) Analytical  
611 isotherm equations for multicomponent adsorption to heterogeneous surfaces. *J. Coll.*  
612 *Int. Sci.* **166**, 51-60.
- 613 Lippold H., Evans N.D.M., Warwick P. and Kupsch H. (2007) Competitive effect of iron(III)  
614 on metal complexation by humic substances: Characterisation of ageing processes.  
615 *Chemosphere* **67**, 1050-1056.
- 616 Lippold H., Mansel A. and Kupsch H. (2005) Influence of trivalent electrolytes on the humic  
617 colloid-borne transport of contaminant metals: competition and flocculation effects. *J.*  
618 *Contam. Hydro.* **76**, 337-352.
- 619 Liu D.J, Bruggeman C. and Maes N. (2008) The influence of natural organic matter on the  
620 speciation and solubility of Eu in Boom Clay porewater. *Radiochim. Acta* **96**, 711-720.
- 621 Lofts S., Tipping E. and Hamilton-Taylor J. (2008) The Chemical Speciation of Fe(III) in  
622 Freshwaters. *Aquat. Geochem.* **14**, 337-358.

623 Luo Y. R. and Byrne R. H. (2004) Carbonate complexation of yttrium and the rare earth  
624 elements in natural waters. *Geochim. Cosmochim. Acta* **68**, 691-699.

625 Marinsky J.A. and Ephraim J. (1986) A unified physicochemical description of the  
626 protonation and metal ion complexation equilibria of natural organic acids (humic and  
627 fulvic acids). 1. Analysis of the influence of polyelectrolyte properties on protonation  
628 equilibria in ionic media: fundamental concept. *Environ. Sci. Technol.* **20**, 349-354.

629 Marsac R., Davranche M., Gruau G. and Dia A. (2010) Metal loading effect on rare earth  
630 element binding to humic acid: Experimental and modelling evidence. *Geochim.*  
631 *Cosmochim. Acta* **74**, 1749-1761.

632 Milne C.J, Kinniburgh D.G., van Riemsdijk W.H. and Tipping E. (2003) Generic NICA-  
633 Donnan model parameters for metal-ion binding by humic substances. *Environ. Sci.*  
634 *Technol.* **37**, 958-971.

635 Naber A., Plaschke M., Rothe J., Hofmann H. and Fanghänel T. (2006) S Scanning  
636 transmission X-ray and laser scanning luminescence microscopy of the carboxyl group  
637 and Eu(III) distribution in humic acid aggregates. *J. Electron Spectrosc. Relat. Phenom.*  
638 **153**, 71-74.

639 Olivie-Lauquet G., Allard T., Benedetti M. and Muller J.-P. (1999) Chemical distribution of  
640 trivalent iron in riverine material from a tropical ecosystem: a quantitative EPR study.  
641 *Wat. Res.* **33**, 2726-2734.

642 Parkhurst D.L. and Appelo C.A.J. (1999) User's guide to PHREEQC (Version 2) - a computer  
643 program for speciation, batch reaction, one-dimensional transport and inverse  
644 geochemical calculation. Water-resources Investigation Report 99-4259, USGS, Denver,  
645 Colorado, p. 312.

- 646 Pédrot M., Dia A., Davranche M., Bouhnik-Le Coz M., Henin O. and Gruau G. (2008)  
647 Insights into colloid-mediated trace element release at soil/water interface. *J. Coll. Int.*  
648 *Sci.* **325**, 187-197.
- 649 Pinheiro J.P., Mota A.M. and Benedetti M.F. (2000) Effect of aluminum competition on lead  
650 and cadmium binding to humic acids at variable ionic strength. *Environ. Sci. Technol.*  
651 **34**, 5137-5143.
- 652 Plaschke M., Rothe J., Denecke M. A. and Fanghänel T. (2004) Soft X-ray spectromicroscopy  
653 of humic acid europium(III) complexation by comparison to model substances. *J.*  
654 *Electron Spectrosc. Relat. Phenom.* **135**, 53-65.
- 655 Pourret O., Davranche M., Gruau G. and Dia A. (2007a) Organic complexation of rare earth  
656 elements in natural waters: evaluating model calculations from ultrafiltration data.  
657 *Geochim. Cosmochim. Acta* **71**, 2718-2735.
- 658 Pourret O., Davranche M., Gruau G. and Dia A. (2007b) Rare earth complexation by humic  
659 acid. *Chem. Geol.* **243**, 128-141.
- 660 Ritchie J.D. and Perdue E.M. (2003) Proton-binding study of standard and reference fulvic  
661 acids, humic acids, and natural organic matter. *Geochim. Cosmochim. Acta* **67**, 85-96.
- 662 Sasaki T., Kobayashi T., Tagagi I. and Moriyama H. (2008) Discrete fragment model for  
663 complex formation of europium(III) with humic acid. *J. Nucl. Sci. Technol.* **45**(8), 718-  
664 724.
- 665 Sonke J.E. (2006) Lanthanide–humic substances complexation. II. Calibration of humic ion-  
666 binding model V. *Environ. Sci. Technol.* **40**, 7481-7487.
- 667 Sonke J.E. and Salters V.J.M. (2006) Lanthanide–humic substances complexation. I.  
668 Experimental evidence for a lanthanide contraction effect. *Geochim. Cosmochim. Acta*  
669 **70**, 1495–1506.

- 670 Takahashi Y., Minai Y., Ambe S., Makide Y., Ambe F. and Tominaga T. (1997)  
671 Simultaneous determination of stability constants of humate complexes with various  
672 metal ions using multitracer technique. *Sci. Tot. Environ.* **198**, 61-71.
- 673 Tang J. and Johannesson K.H. (2003) Speciation of rare earth elements in natural terrestrial  
674 waters: assessing the role of dissolved organic matter from the modeling approach.  
675 *Geochim. Cosmochim. Acta* **67**, 2321-2339.
- 676 Tang J. and Johannesson K.H. (2010) Ligand extraction of rare earth elements from aquifer  
677 sediments: Implications for rare earth element complexation with organic matter in  
678 natural waters. *Geochim. Cosmochim. Acta* **74**, 6690-6705.
- 679 Tanizaki Y., Shimokawa T., and Nakamura M. (1992) Physicochemical speciation of trace  
680 elements in river waters by size fractionation. *Environ. Sci. Technol.* **26**, 1433-1444.
- 681 Tipping E. (1998) Humic ion-binding model VI: an improved description of the interactions  
682 of protons and metal ions with humic substances. *Aquat. Geochem.* **4**, 3-48.
- 683 Tipping E. (2005) Modelling Al competition for heavy metal binding by dissolved organic  
684 matter in soil and surface waters of acid and neutral pH. *Geoderma* **127**, 293-304.
- 685 Tipping E. and Hurley M.A. (1992) A unifying model of cation binding by humic substances.  
686 *Geochim. Cosmochim. Acta* **56**, 3627-3641.
- 687 Tipping E., Rey-Castro C., Bryan S. E., and Hamilton-Taylor J. (2002) Al(III) and Fe(III)  
688 binding by humic substances in freshwaters and implications for trace metal speciation.  
689 *Geochim. Cosmochim. Acta* **66**, 3211-3224.
- 690 Viers J., Dupré B., Polvé M., Schott J., Dandurand J.-L. and Braun J.J. (1997) Chemical  
691 weathering in the drainage basin of a tropical watershed (Nsimi–Zoetele site,  
692 Cameroon): comparison between organic poor and organic-rich waters. *Chem. Geol.*  
693 **140**, 181-206.

694 Yamamoto Y., Takahashi Y. and Shimizu H. (2010) Systematic change in relative stabilities  
695 of REE-humic complexes at various metal loading levels. *Geochem. J.* **44**, 39-63.

696

697

## 698 **Table and Figure Captions**

699

700 **Table 1.** Parameters used in this study to model REE binding by humic acid.

701

702 **Table 2.** Log  $K_{MA}$ , log  $K_{MB}$ ,  $\Delta LK_{2C}$  and  $\Delta LK_{2P}$  values fitted from the experimental data using  
703 PHREEQC/Model VI. The quality of the fit is determined by the rmse, calculated as

704  $\sqrt{\text{mean}(\log v_{\text{exp}} - \log v_{\text{calc}})^2}$ , where  $v$  is the amount of REE bound to HA per gram of DOC

705 for the experimental and modeled data, respectively. The log  $K_{MA}$ , log  $K_{MB}$ ,  $\Delta LK_2$  (where  
706  $\Delta LK_2 = \Delta LK_{2C} = \Delta LK_{2P}$ ) data published by Marsac et al. (2010) using WHAM 6/Model VI  
707 are shown for comparison.

708

709 **Figure 1.** Comparison of the REE patterns for the HA binding parameters obtained by Marsac  
710 et al. (2010) with WHAM 6/Model VI and the patterns obtained with PHREEQC/Model VI:

711 (a) log  $K_{MA}$  compared with log  $K$  (REE-acetic acid), (b) log  $K_{MB}$  compared with log  $K$ (REE-  
712 catechol) and (c)  $\Delta LK_{2C}$  and  $\Delta LK_{2P}$  (this study) compared with  $\Delta LK_2$  (Marsac et al., 2010;

713  $\Delta LK_2 = \Delta LK_{2C} = \Delta LK_{2P}$ ).

714

715 **Figure 2.** Comparison between published experimental data for REE binding with HA and  
716 the results of simulations of the same data using WHAM 6/Model VI and PHREEQC/Model  
717 VI, respectively. The experimental data are from Pourret et al. (2007b) (a), Marsac et al.  
718 (2010) (b) and Sonke and Salters (2006) (c). The "patterns shown" correspond to the REE-HA  
719 patterns, presented for the experiments demarcated with an open symbol, in (d), (e) and (f).

720

721 **Figure 3.** Log ( $K_{Lu}/K_{La}$ ) versus the average REE Log  $K_{REE}$  values obtained for 101 organic  
722 ligands (Byrne and Li, 1995) and for the 80 ligands defined in WHAM 6/Model VI and  
723 PHREEQC/Model VI. Log ( $K_{Lu}/K_{La}$ ) is used as an indicator of the REE-ligand pattern. The  
724 linear regression for the 101 organic ligands (Byrne and Li, 1995) and for the 80 HA sites  
725 obtained here with PHREEQC/Model VI are also presented.

726

727 **Figure 4.** Speciation calculation results for La (a), Eu (b) and Lu (c) depending on the  
728 complexation hypotheses used. The simulated experimental data are those from Pourret et al.  
729 (2007b). CG: carboxylic groups; CPG: carboxy-phenolic groups; PG: phenolic groups.

730

731 **Figure 5.** The proportion of La and Lu bound to tridentate sites versus the REE/HA ratios.  
732 The displayed evolutions depend on whether the new equations and parameters or those  
733 formerly installed in WHAM6/ModelVI are used. The simulated experimental data are from  
734 Marsac et al. (2010).

735

736 **Figure 6.** Literature compilation of metal-acetic acid or -catechol binding constants. The  
737 straight line passing through the mean REE value indicates metals that should not compete  
738 differently with the LREE and MREE. Conversely, metals below this line, such as  $Cd^{2+}$ ,

739 should compete primarily with the LREE, whereas those above this line, such as  $Al^{3+}$ , should  
740 be mainly in competition with the HREE. Data sources: IUPAC Stability constants database.  
741

Parameter	Description	Values
$n_A$	Amount of type A sites ( $\text{mol g}^{-1}$ )	$3.3 \cdot 10^{-3}$
$n_B$	Amount of type B sites ( $\text{mol g}^{-1}$ )	$0.5 \times n_A$
$\text{pK}_A$	Intrinsic proton dissociation constant for type A sites	4.1
$\text{pK}_B$	Intrinsic proton dissociation constant for type B sites	8.8
$\Delta\text{pK}_A$	Distribution term that modifies $\text{pK}_A$	2.1
$\Delta\text{pK}_B$	Distribution term that modifies $\text{pK}_B$	3.6
$\log K_{MA}$	Intrinsic equilibrium constant for metal binding at type A sites	Fitted from experimental data
$\log K_{MB}$	Intrinsic equilibrium constant for metal binding at type B sites	Fitted from experimental data
$\Delta\text{LK}_1$	Distribution term that modifies $\log K_{MA}$ and $\log K_{MB}$	2.8
$\Delta\text{LK}_2$	Distribution term that modifies the strengths of bidentate and tridentate sites	Fitted from experimental data
SA	HA surface area ( $\text{m}^2 \text{g}^{-1}$ )	Fitted from WHAM 6/Model VI simulations
$\kappa$	Debye-Hückel parameter (m)	$(3.29 \times 10^9 \times I^{1/2})^{-1}$

743

744

**Table 1**

745

746

747

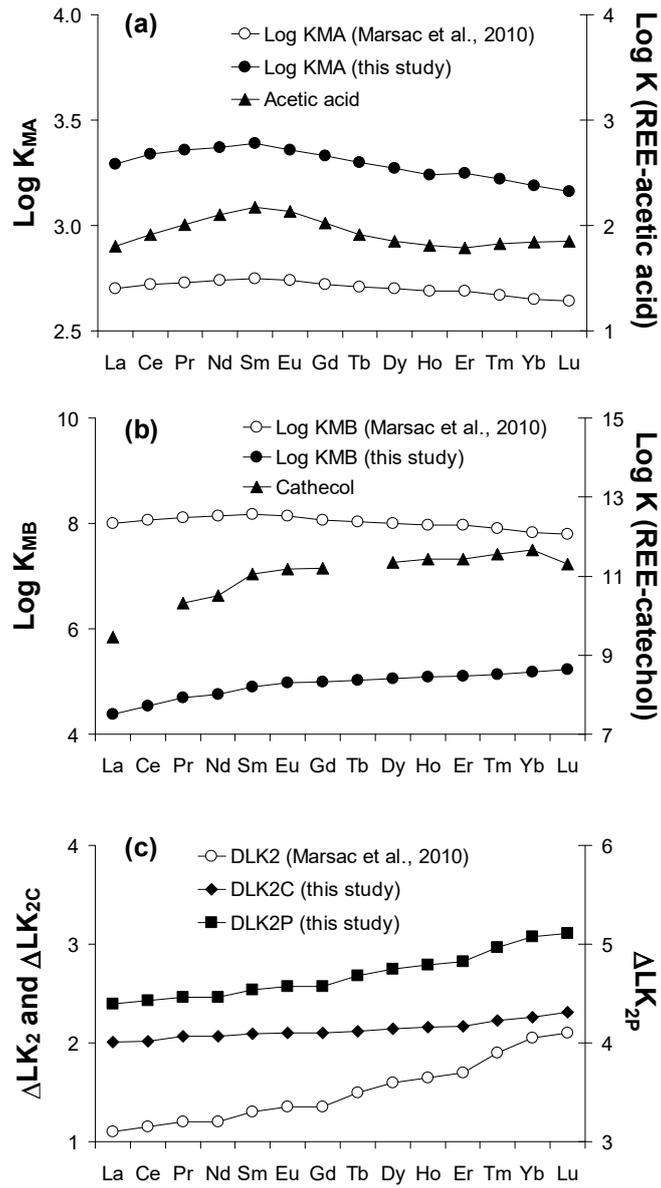
REE	$\log K_{MA}^{(1)}$	$\log K_{MA}^{(2)}$	$\log K_{MB}^{(1)}$	$\log K_{MB}^{(2)}$	$\Delta LK_2^{(1)}$	$\Delta LK_{2C}^{(2)}$	$\Delta LK_{2P}^{(2)}$	rmse <sup>(1)</sup>	rmse <sup>(2)</sup>
La	2.70	3.29	8.00	4.39	1.10	2.01	4.39	0.09	0.07
Ce	2.72	3.34	8.07	4.54	1.15	2.02	4.43	0.08	0.06
Pr	2.73	3.36	8.10	4.69	1.20	2.06	4.46	0.07	0.05
Nd	2.74	3.37	8.14	4.76	1.20	2.06	4.46	0.08	0.05
Sm	2.75	3.39	8.17	4.90	1.30	2.09	4.54	0.08	0.05
Eu	2.74	3.36	8.14	4.98	1.35	2.10	4.57	0.08	0.05
Gd	2.72	3.33	8.07	4.99	1.35	2.10	4.57	0.10	0.06
Tb	2.71	3.30	8.04	5.03	1.50	2.12	4.68	0.11	0.06
Dy	2.70	3.27	8.00	5.06	1.60	2.14	4.75	0.12	0.07
Ho	2.69	3.24	7.97	5.09	1.65	2.16	4.79	0.14	0.08
Er	2.69	3.25	7.97	5.09	1.70	2.17	4.82	0.15	0.09
Tm	2.67	3.22	7.90	5.14	1.90	2.22	4.97	0.15	0.09
Yb	2.65	3.19	7.83	5.18	2.05	2.26	5.08	0.15	0.10
Lu	2.64	3.16	7.80	5.23	2.10	2.31	5.11	0.16	0.10
Mean	2.70	3.29	8.02	4.93	1.51	2.13	4.69	0.11	0.07

748 *(1) Marsac et al. (2010)*749 *(2) This study*

750

751 **Table 2**

752



753

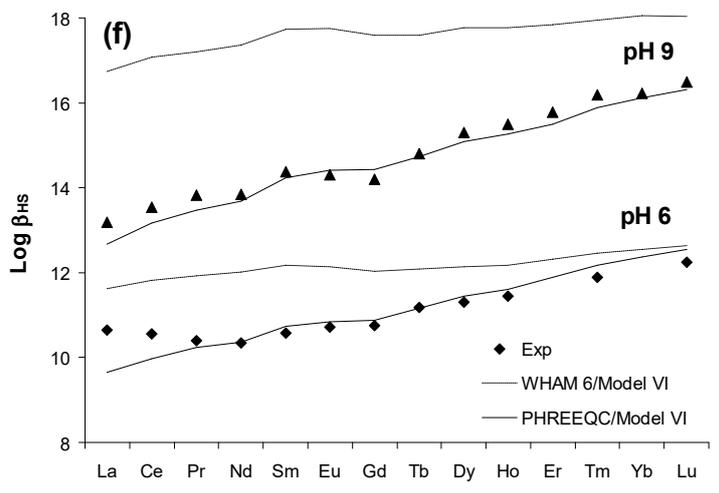
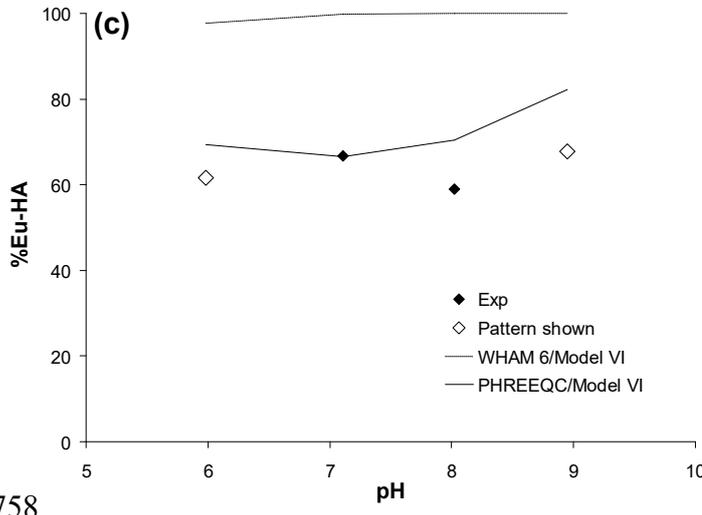
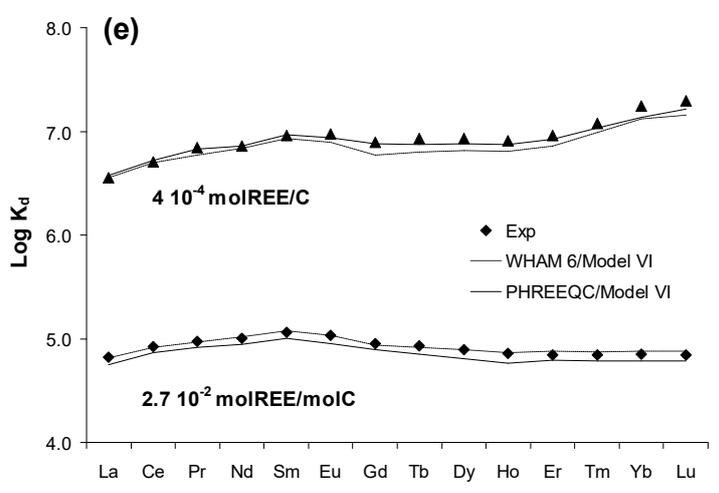
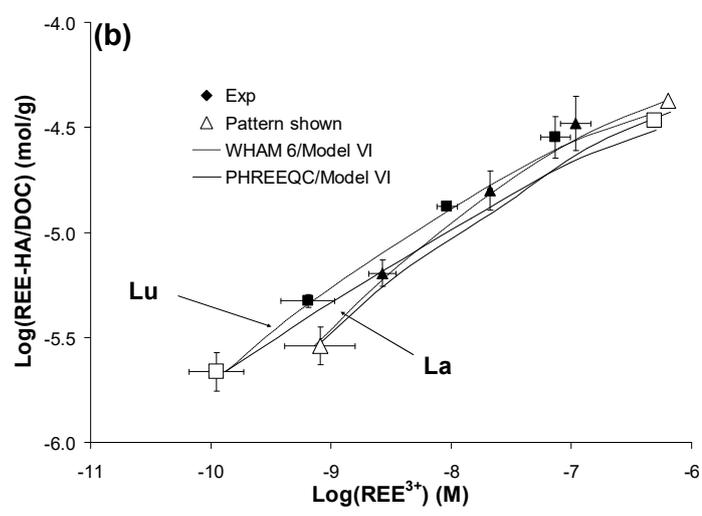
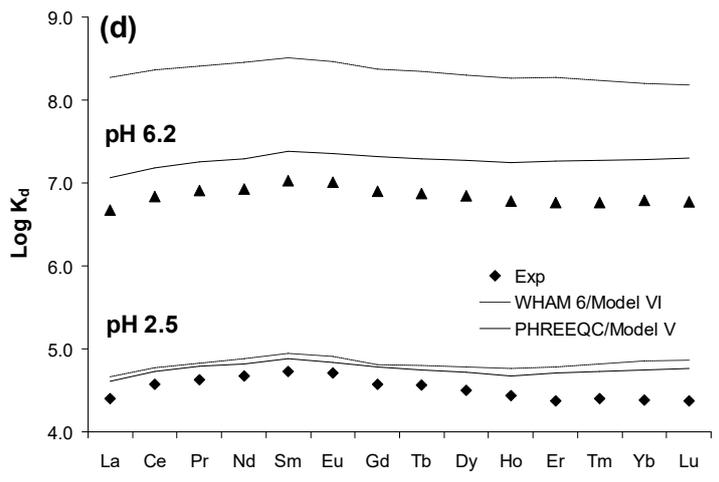
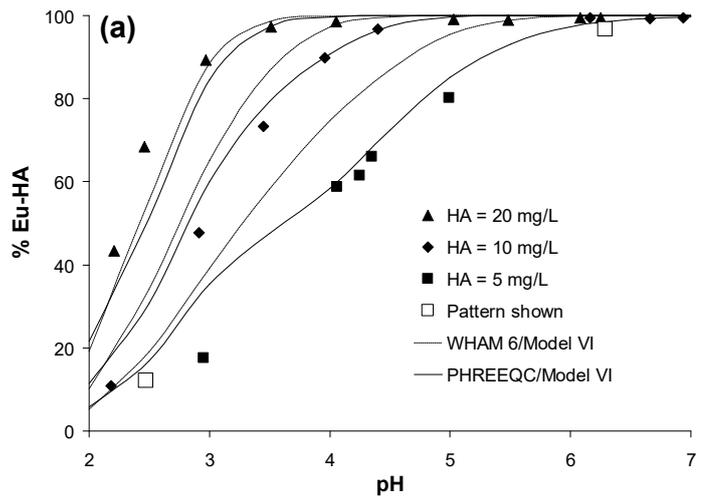
754

755

756

757

**Figure 1**



758

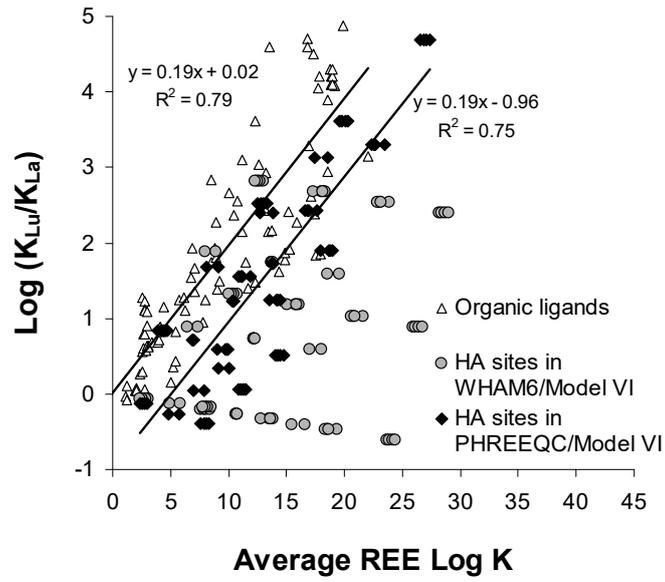
759

760

761

## Figure 2

762



763

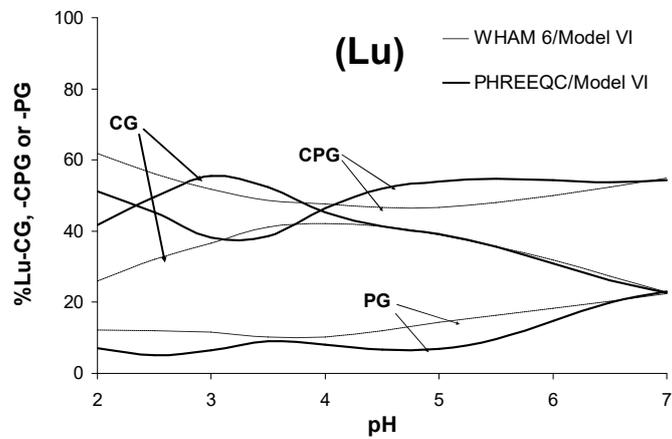
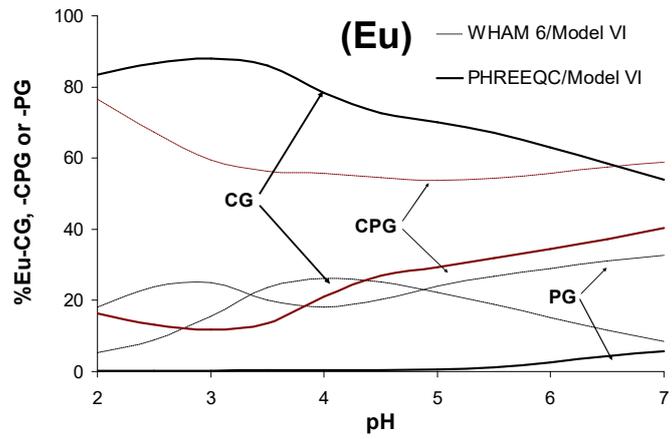
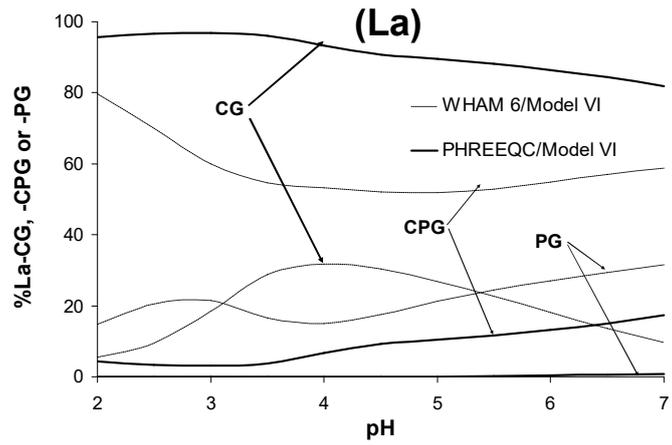
764

765

766

**Figure 3**

767



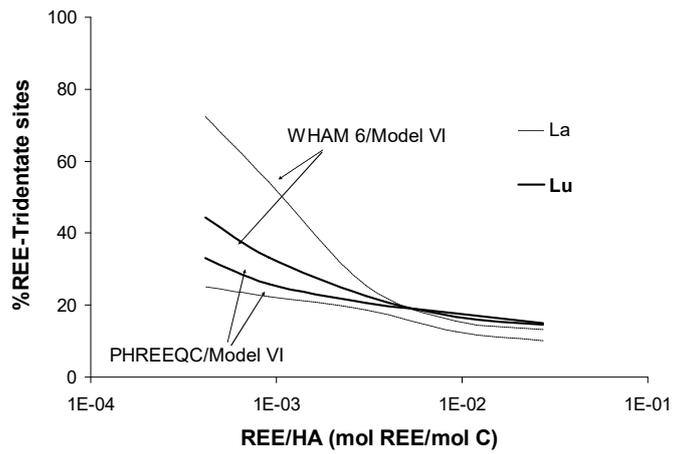
768

769

770

771

**Figure 4**



772

773

774

775

**Figure 5**

776

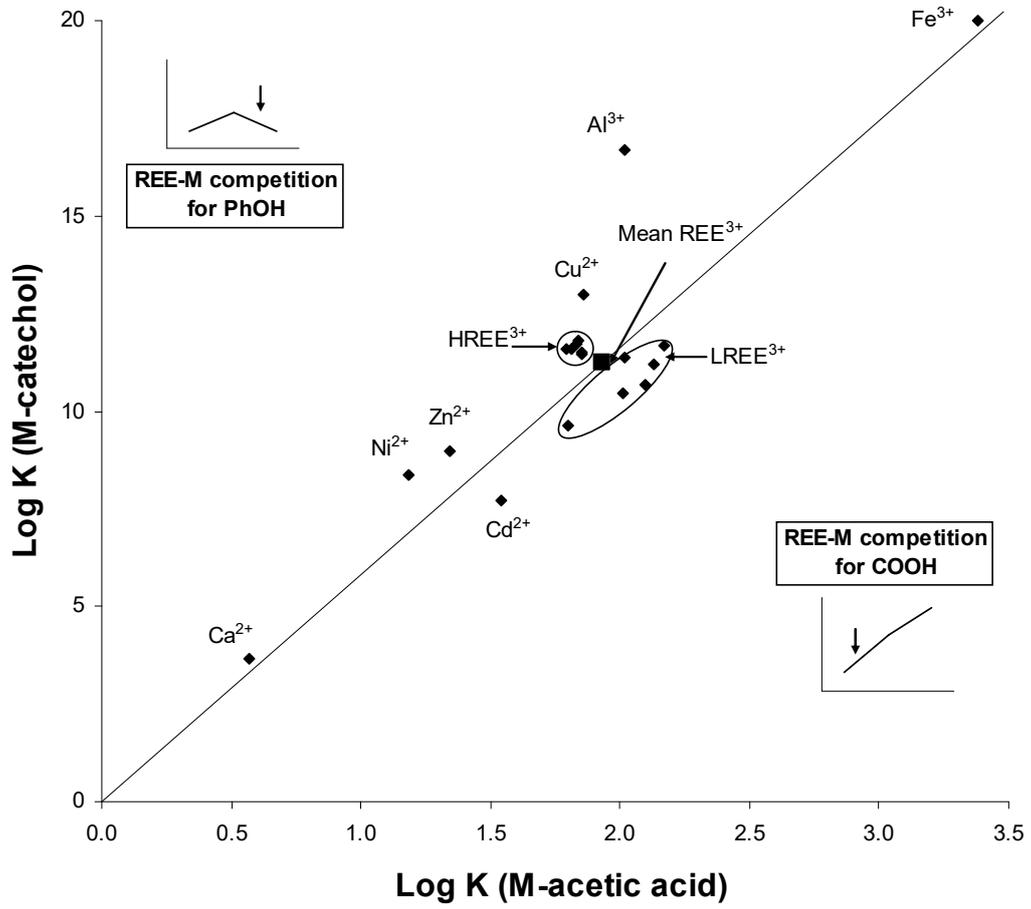
777

778

779

780

781



782

783

784

785

786

**Figure 6**