

# RM1 Model for the Prediction of Geometries of Complexes of the Trications of Eu, Gd, and Tb

Manoel A. M. Filho,<sup>†</sup> José Diogo L. Dutra,<sup>†</sup> Higo L. B. Cavalcanti,<sup>‡</sup> Gerd B. Rocha,<sup>‡</sup> Alfredo M. Simas,<sup>§</sup> and Ricardo O. Freire<sup>\*,†</sup>

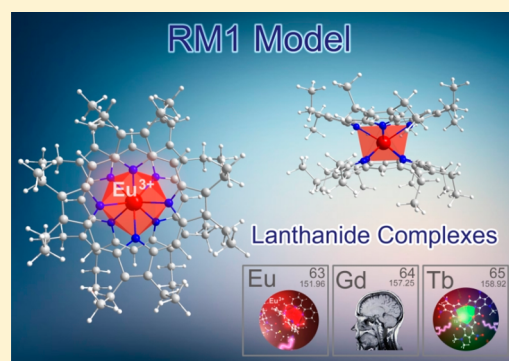
<sup>†</sup>Pople Computational Chemistry Laboratory, Departamento de Química, Universidade Federal de Sergipe, 49.100-000 São Cristóvão, SE, Brazil

<sup>‡</sup>Departamento de Química, CCEN, Universidade Federal da Paraíba, 58.059-970 João Pessoa, PB, Brazil

<sup>§</sup>Departamento de Química Fundamental, Universidade Federal de Pernambuco, 50.740-540 Recife, PE, Brazil

## S Supporting Information

**ABSTRACT:** All versions of our previous Sparkle Model were very accurate in predicting lanthanide–lanthanide distances in complexes where the two lanthanide ions directly face each other, and mainly lanthanide–oxygen, and lanthanide–nitrogen distances, which are by far the most common ones in lanthanide complexes. In this article, we are advancing for the first time the RM1 model for lanthanides. Designed to be a much more general NDDO model, the RM1 model for lanthanides is capable of predicting geometries of lanthanide complexes for the cases when the central lanthanide trication is directly coordinated to any other atoms, not only oxygen or nitrogen. The RM1 model for lanthanides is defined by three important attributes: (a) the orbitals, the lanthanide ion has now three electrons and a NDDO basis set made of 5d, 6s, and 6p functions; (b) the parametrization, via cluster analysis and an adequate sampling; and (c), the statistical validation of the parameters to make sure the errors behave as random around a mean. All three aspects are described in detail in the article. Results indicate that the RM1 model does extend the accuracy of the previous Sparkle Models to types of coordinating bonds other than Ln–O and Ln–N; the most common ones for Eu, Gd, and Tb, being Ln–C, Ln–S, Ln–Cl, and Ln–Br. Overall, these other coordinating bonds are now predicted within 0.06 Å of their correct values. Therefore, the RM1 model here presented is capable of predicting geometries of lanthanide complexes, materials, metal–organic frameworks, etc., with useful accuracy.



## INTRODUCTION

Lanthanide chemistry is an active area of research with several technological applications due to the enormous diversity of properties of the various complexes of lanthanide ions: some absorb in the UV and emit visible light, others absorb visible light and emit in the near-infrared, still others find many applications in contrast agents for MRI; others must be present in zeolites for these minerals to be stable at high temperatures and act as catalysts in the petroleum industry, etc.<sup>1–4</sup> In all these applications, the geometries of the complexes, of the metal–organic frameworks, and of the nanostructures, are one of the most important properties. For example, lanthanide luminescence depends on an effective absorption of energy by the ligands, and to a transfer of this energy to an excited state of the metal ion, which will then emit its typical radiation. The transfer of energy from the excited state of the ligands to the excited state of the metal depends on the Judd–Ofelt intensity parameters  $\Omega_{\lambda}^{\text{ed}}$  (forced electric dipole contribution only). These parameters are dependent on the odd-rank ligand field parameters,  $\gamma_p^t$ , and these, in turn, are calculated from some integrals that depend on the second, fourth, sixth, and even

eighth inverse power of the distances between the lanthanide ion and its coordinating atoms.<sup>5</sup> Consequently, any inaccuracies in the geometries will be amplified. That is why lanthanide complexes geometry prediction is essential for the design of highly luminescent complexes. A similar situation occurs with the design of gadolinium complexes for magnetic resonance imaging. In order to enhance or differentiate the MRI signal, one must add a gadolinium complex with at least one water molecule coordinated to it, accessible to the water solvent, which then works as a contrast agent by catalytically relaxing the solvent water hydrogen atoms to which they come close, generating image contrast. The magnitude by which a contrast agent affects the relaxation rate of the hydrogen atoms of the water molecules of the tissues is called relaxivity. The Gd-induced longitudinal relaxation is a dipolar process that depends on the inverse sixth power of the distance between the hydrogen atom of the coordinated water and the Gd(III) ion. So, once again, geometric factors are important in complex

Received: October 18, 2013

Published: June 12, 2014

design,<sup>6</sup> including the design of the so-called “smart” MRI contrast agents, whose relaxivity is magnified by their recognition of specific biomolecules.<sup>7</sup> Besides, lately, ill side effects related to administration of Gd-based MRI contrast agents have been discovered such as nephrogenic systemic fibrosis.<sup>8</sup> Therefore, a quality molecular modeling approach must be available for the design of very effective, and at the same time nontoxic, Gd MRI agents. Other applications of lanthanides are also heavily dependent on good geometries for the correct understanding of the phenomena involved, such as inclusion processes, supramolecular arrangements in metal–organic frameworks,<sup>9</sup> etc.

In 1994, we advanced a new model, based on semiempirical quantum chemistry, to predict the geometries of europium complexes, we called the Sparkle Model.<sup>10,11</sup> The original Sparkle Model used AM1, replaced the europium ion by a point of charge +3e and superimposed to its Coulomb potential a spherically symmetric repulsive potential of the form  $\exp(-\alpha r)$ , where the inverse of  $\alpha$  can be interpreted as a measure of the size of the europium ion. Indeed, this potential must be present in order to prevent the ligands from getting suctioned by the point charge. This initial Sparkle Model was very successful, but its error in terms of distances between the europium ion and the coordinating atoms was, for a set of 96 complexes, 0.68 Å,<sup>12</sup> very high for today's standards. Nevertheless, a strategy was subsequently developed to allow a prediction of UV–vis absorption spectra from the ligands of the complexes, by simply replacing the europium ion by a point of charge +3e and by computing the spectra via ZINDO.<sup>13</sup> In 2004, the Sparkle Model was improved by taking into consideration the mass of the europium in the calculation and by adding to the semiempirical core–core term Gaussian functions that greatly reduced the errors to 0.28 Å.<sup>14</sup> Clearly the error was still at a very high level for the model to be considered useful. However, this version of the Sparkle Model was useful to study luminescent properties of europium systems,<sup>15</sup> for the development of models applied to elucidate the amount of water coordination molecules,<sup>16</sup> and for the design of highly luminescent europium complexes.<sup>17–20</sup> We then realized that the Sparkle Model defined in that manner was already adequate from the point of view of its formalism—what needed to be perfected was its parametrization. We then obtained the Cambridge Structural Database<sup>21,22</sup> and created a large enough collection of high crystallographic quality (R-factor <5%) structures of complexes of Eu(III), Gd(III), and Tb(III). We applied a cluster analysis to the set, from which we created a much smaller sets of complexes, one for each of the lanthanide trications, which we used as parametrization sets. We then parametrized the model within AM1, and the resulting unsigned average error for the distances between the lanthanide ions and the coordinating atoms then dropped to 0.09 Å, 0.07 Å, and to 0.07 Å, for Eu(III), Gd(III), and Tb(III) complexes, respectively.<sup>12</sup> At that point, the Sparkle Model was now accurate enough to be usefully applied to the design of new complexes and was very competitive with *ab initio* methods.<sup>23</sup> We then extended the model to all other lanthanides for AM1,<sup>12,24–32</sup> then for PM3,<sup>33–39</sup> to PM6,<sup>40</sup> to RM1,<sup>41</sup> and, more recently, to PM7,<sup>42</sup> with an emphasis on solid state calculations.

However, Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, Sparkle/RM1, and Sparkle/PM7 are very good for predicting the geometries of lanthanide complexes when they are coordinated to either oxygen or nitrogen, only. That is due to the reason

that in the vast majority of lanthanide complexes the ligands coordinate to the central lanthanide ion by means of oxygen or nitrogen atoms only.

Indeed, in the 2005 article, we mentioned that of all europium, gadolinium, and terbium complexes present in the CSD, 80% of all structures had only oxygen or nitrogen as directly coordinating atoms.<sup>12</sup> The second reason is that we discovered that it was not really possible to create a Sparkle Model with an arbitrary number of Gaussian functions in the core–core term that would be able to predict the distances for all of the different and possible coordinating atoms, O, N, C, Cl, Br, and S, simultaneously. Indeed, while the coordination bond between either O or N and the lanthanide cation seemed essentially electrostatic, with time it became increasingly apparent that the direct coordination by the other atoms involved a higher degree of covalency. Therefore, all Sparkle Models were developed for ligands with only oxygen or nitrogen as coordinating atoms only. The strategy was successful because, as expected, it satisfied most uses of the method.<sup>43–50</sup> However, now that the Sparkle Model has been correctly parametrized for all major semiempirical methods, is currently implemented in MOPAC,<sup>51</sup> and is being used in a worldwide manner, the moment has arrived in which a new model must be developed—a model for NDDO methods that will be able to calculate and predict to useful accuracy the geometry of any lanthanide complex, regardless of the nature of its directly coordinating ligand atoms. In this article, we advance precisely such an RM1 model for lanthanides. Unlike the previous Sparkle/RM1,<sup>41</sup> which is a Sparkle Model parametrized for RM1, the RM1 model for lanthanides is now a true RM1 parametrization for the lanthanide trications.

## ■ THE RM1 MODEL FOR LANTHANIDE TRICATIONS

The RM1 model consists of three parts: (a) the equations of the model, (b) the parametrization scheme, and (c) the statistical validation of the obtained parameters. Let us consider first the equations of the model. Due to what we perceive as a non-negligible degree of covalency in the coordinating bonds between a lanthanide trication and carbon, sulfur, chlorine, and bromine, the lanthanide ion in the RM1 Model must have orbitals. Since we will not consider f-electrons, for the purpose of parametrizing the trications, we will assume that all lanthanides have the following configuration  $\{[\text{Xe}] 4f^n\} 5d^1 6s^2$ , where  $n$  varies from 6, in Eu, to 8, in Tb. Then, we take  $\{[\text{Xe}] 4f^n\}$  to be described by the effective core potential implicit in the semiempirical model and set the RM1 core charge of the lanthanide as +3e. The semiempirical basis set used is composed of 5d, 6s, and 6p orbitals, for a total of 9 orbitals. Thus, each lanthanide in the present model contributes with three electrons to the valence shell.

Our semiempirical model for lanthanide trications is inspired by the ECP (Effective Core Potential) by Dolg and co-workers,<sup>52</sup> where the electronic configurations for all lanthanides, in its oxidation state +3, is fixed as  $46 + 4f^n$  electrons in the “large core”, leaving the outermost 11 electrons ( $5s^2 5p^6 6s^2 5d^1$ ) to be treated explicitly. Dolg's ECPs is consistent with the effective core  $\{[\text{Kr}] 4d^{10} 4f^6\}$  for  $\text{Ln}^{3+}$ . The effective core potential to be used in compounds with  $\text{Ln}^{3+}$  ions is one that has an effective core potential charge of +11e and, therefore, initially contributes 11 electrons to the calculation. We leave the f electrons in the core and then move the  $5s^2$  and  $5p^6$  into the core as well. Then, in our RM1 model, the lanthanide metal is treated as a doublet in a formal sense when

counting electrons in terms of a  $6s^2 5d^1$  initial valence configuration. Singlets will arise for a neutral complex if we assume a  $\text{Ln}^{3+}$  and a charge of  $-3e$  on the ligands.

Also, we quote from the Dolg work: “Although the 4f-orbitals form an open shell they seem to have a core-like character, and the 5d- and 6s-valence orbitals (and their corresponding  $6s^2$ ,  $5d^1 6s^2$  or  $5d^2 6s^2$  valence subconfigurations) are responsible for the observed chemical behavior of the lanthanides” in support of our choice of a 3 electron valence space.

In addition, we note that Maron and Eisenstein<sup>53</sup> found that “The large core RECPs for the lanthanides, leading to formal  $\text{Ln}^{3+}$  complexes, are extracted from the  $f^{n-1} d^1 s^2$  atomic configuration and thus properly describe the participation of the 5d orbitals in the bonding” and further noted that “the large core RECPs used in this work describe properly the bonding at lanthanide through, in particular, the participation of the 5d orbitals”.

Lastly, our model is very similar to the earlier semiempirical model proposed by McNamara and co-workers for lanthanide trications.<sup>54</sup> In both models, there is a core of charge  $+3e$  and, consequently, three electrons in the valence space with a semiempirical basis set at the lanthanide center made of 6s, 5d, and 6p orbitals.

The present RM1 model has been designed to work for coordination compounds where the oxidation number of lanthanide is  $+3$  and has thus been parametrized only for such complexes.

Regular RM1 elements are parametrized to reproduce not only the geometry but also other quantities, such as enthalpies of formation, ionization potentials, and dipole moments, all at the same time. Nevertheless, in the RM1 Model for lanthanide compounds, we concern ourselves exclusively with geometry prediction. Of course, users must consider, as a meaningful result of such a calculation, only the geometry, understanding that values for all other properties will likely not be reliable.

Accordingly, the RM1 Model uses the 5d, 6s, and 6p orbitals already programmed in MOPAC2012 and thus has the following 22 parameters for each lanthanide: the s, p, and d atomic orbital one-electron one-center integrals  $U_{ss}$ ,  $U_{pp}$ , and  $U_{dd}$  (in eV); the s, p, and d Slater atomic orbital exponents  $\xi_s$ ,  $\xi_p$ , and  $\xi_d$  (in bohr<sup>-1</sup>); the s, p, and d atomic orbital one-electron two-center resonance integral terms  $\beta_s$ ,  $\beta_p$ , and  $\beta_d$  (in eV); the core–core repulsion term  $\alpha$  (in Å<sup>-1</sup>); the two-electron integrals  $F_{SD}^0$ ,  $G_{SD}^2$  (in eV); and the additive term  $\rho_{\text{core}}$  (in bohr) needed to evaluate core–electron and core–core nuclear interactions; the second set of exponents to compute the one-center integrals  $\xi'_s$ ,  $\xi'_p$ , and  $\xi'_d$  (in bohr<sup>-1</sup>); and the six parameters for the two Gaussian functions: height,  $a_i$ , dimensionless; inverse broadness,  $b_i$ , in Å<sup>-2</sup>; and displacement,  $c_i$ , in Å; as in  $G(R) = \sum_{i=1}^2 a_i e^{[b_i(R-c_i)^2]}$ , where  $R$ , in Å, is the interatomic distance between the lanthanide and the other atom.

The second part of the RM1 Lanthanide Model is its parametrization, a very critical one. For each lanthanide trication, we collected from the Cambridge Crystallographic Database all structures of high crystallographic quality (R factor  $< 0.5$ ). We then computed all of them with Sparkle/AM1 and attributed to each complex  $i$  a value  $R_i$  computed from the following formula:

$$R_i = \sum_j \sum_k \frac{1}{\sigma_j^{\text{dist}}} |d_{ij,k}^{\text{CSD}} - d_{ij,k}^{\text{calc}}| + \sum_l \frac{1}{\sigma_l^{\text{angle}}} |\theta_{il}^{\text{CSD}} - \theta_{il}^{\text{calc}}| \quad (1)$$

where  $j$  runs over all types of bonds (e.g., Ln–N, Ln–O, Ln–C, Ln–S, Ln–Br, and Ln–Cl);  $k$  runs over all bonds of type  $j$ ;  $\sigma_j^{\text{dist}}$  is the standard deviation of all crystallographic bond lengths of type  $j$  for all complexes considered;  $d_{ij,k}^{\text{CSD}}$  is the crystallographic  $k$ th bond distance of type  $j$  for complex  $i$ ;  $d_{ij,k}^{\text{calc}}$  is the calculated value of the same bond;  $\sigma_l^{\text{angle}}$  is the standard deviation of all crystallographic bond angles of the type A–Ln–B, with A/B = O, N, C, S, Cl, and Br;  $\theta_{il}^{\text{exp}}$  is the crystallographic  $l$ th bond angle of complex  $i$ ; and  $\theta_{il}^{\text{calc}}$  is its calculated counterpart. Note that we do not need to classify angles into types and that the standard deviations are computed from experimental values only, and their use as weighting factors is to guarantee that all terms of eq 1 can be indeed summed, as they then become comparable quantities. This way, all types of bond length errors become comparable, and likewise, they become comparable to the errors in the angles as well. In all three cases studied in this article, Eu(III), Gd(III), and Tb(III), we had 3.8 times more angles than distances, and we found out empirically that giving about four times more weight to the angles than to the distances led to a better method. Therefore, we decided that eq 1 was already naturally optimally balanced between distances and angles. After this process, each complex  $i$  will now be represented by a measure  $R_i$ . The set of  $R_i$  values,  $\{R_i\}$ , is now submitted to a hierarchical divisive cluster analysis DIANA<sup>55</sup> in order to cluster the complexes into homogeneous groups. We then sampled these groups to arrive to a small parametrization set, in average with about 11 complexes per lanthanide, chosen to be representative of the universe of complexes considered. When the parametrization converged for this small set, we then went back to the cluster analysis results and enlarged this parametrization set to a new set with an average of 30 complexes per lanthanide and refined the method by reoptimizing it. When the parametrization for this enlarged set converged, we ended the process and collected the resulting parameters.

The third and last part of the RM1 model is the statistical validation of the parameters. This is to guarantee that the essence of the chemical bond between the metal ion and the ligands has been captured by the model. For that, as before,<sup>12</sup> we consider the unsigned mean errors, UME, defined for the set of  $i$  complexes as

$$\text{UME} = \frac{1}{n_{\text{tot}}} \sum_i^{m_{\text{comp}}} \sum_{j=1}^{m_{i,\text{dist}}} |R_{ij}^{\text{CSD}} - R_{ij}^{\text{calc}}| \quad (2)$$

where  $R_{ij}^{\text{CSD}}$  and  $R_{ij}^{\text{calc}}$  are, respectively, the  $j$ th distance of complex  $i$ , from the Cambridge Crystallographic Database, CSD, and from the RM1 calculation, calc. The total number of distances from all complexes is  $n_{\text{tot}}$ ; the total number of complexes is  $m_{\text{comp}}$  and the number of distances in complex  $i$  is  $m_{i,\text{dist}}$ . As before,<sup>12</sup> we used two types of unsigned mean errors:  $\text{UME}_{(\text{Ln-L})}$  and  $\text{UME}$ . The first,  $\text{UME}_{(\text{Ln-L})}$ , refers to a sum of all distances from the central lanthanide ion to all of the atoms of its coordination polyhedron, including its distance to another lanthanide ion, when they face each other; and  $\text{UME}$  are the sum of all of these distances plus all lengths of the edges of the coordination polyhedron, which indirectly reflects the accuracies of the calculated angles. In short,  $\text{UME}_{(\text{Ln-L})}$  capture the accuracy of the predicted bond distances, whereas  $\text{UME}$



capture mostly the accuracy of the predicted angles. If the essence of the chemical bonds between the metal ions and the ligands has been captured by the model, then the UME must behave as random around a mean. Since, by definition, the UME are positive and belong to the domain  $[0, \infty)$ , their histogram should follow the probability density of the  $\gamma$  distribution,  $f(x; k, \theta)$

$$f(x; k, \theta) = x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)} \quad (3)$$

where  $x > 0$  stands for the UME,  $k > 0$  is the shape parameter,  $\theta > 0$  is the scale parameter of the  $\gamma$  distribution, and  $\Gamma(k)$  is the  $\gamma$  function of  $k$ . The expected value of the  $\gamma$  distribution is  $k\theta$  and its variance is  $k\theta^2$ . We then proceed with obtaining estimates of the shape and scale parameters by the method of maximum likelihood. Finally, the quality of the  $\gamma$  distribution fit is subject to the one-sample nonparametric Kolmogorov–Smirnov test<sup>56</sup> to verify statistically whether the distribution of the UME does indeed follow a gamma distribution indexed by the estimated  $k$  and  $\theta$  parameters. For the null hypothesis not to be rejected at the 5% level, the p-value of the test must be larger than 0.05. Moreover, the closer the p-value is to its maximum value of 1; the more the UME are random around the mean; the more certain we can be that the RM1 Model did capture the deterministic aspects of the chemical bond; and the more justifiable is the use of these statistical tools in the problem. Thus, if both UME<sub>(Ln–L)</sub> and UME pass this test, then we accept the model parameters found in part b, and use them to define the RM1 model<sup>57</sup> for the lanthanide being parametrized.

## RESULTS AND DISCUSSION

The first parametrization set for each of the lanthanides contained, respectively, 10, 18, and 16 complexes. After the parameters converged, the sets were increased to, respectively, 24, 40, and 24 complexes, and the parameters were refined. In sequence, they were subjected to the statistical tests. This process was repeated a number of times until when, eventually, the parameters became acceptable. Table 1 presents the parameters found for the technologically important lanthanides Eu (III), Gd(III), and Tb(III).

Table 2 presents the means and variances for the  $\gamma$  distribution fits for the UME<sub>(Ln–L)</sub> for the three metal ions. It is noteworthy to observe that the average of the three mean values is 0.056 Å. If we compare this value with the average of the crystallographic bond lengths themselves for all three lanthanide ions, which is 2.52 Å, clearly, the error of the RM1 method is, in average, 2%.

Table 3 presents the means and variances for the  $\gamma$  distribution fits, this time for the UME for the three metal ions. Again, the parameters passed the statistical test, indirectly indicating that the errors in the angles are also random around a mean.

Table 4 now presents unsigned mean errors for all types of metal-coordinating atom distances existing in the complexes of Europium considered. The error in Eu–Eu distances are larger because Eu–Eu distances are larger, in average 4.06 Å. Clearly, the accuracy of RM1, with respect to Eu–Eu, Eu–O, and Eu–N, are similar to the accuracies of the previous Sparkle Models: Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, Sparkle/PM7, and Sparkle/RM1, because these were parametrized precisely for these distances. However, when we consider the Eu–C, Eu–S, Eu–Cl, and Eu–Br distances, the Sparkle Model errors are on

**Table 1. Parameters for the RM1 model for Europium, Gadolinium, and Terbium**

	RM1		
	Eu <sup>3+</sup>	Gd <sup>3+</sup>	Tb <sup>3+</sup>
$U_{ss}$	−18.54 409 746	−19.90 876 775	−20.92 699 065
$U_{pp}$	−6.93 694 146	−7.75 830 394	−7.75 242 523
$U_{dd}$	−19.87 088 096	−18.96 507 720	−19.97 111 674
$\zeta_s$	1.35 034 212	1.27 277 575	1.21 005 179
$\zeta_p$	1.73 371 439	1.90 812 202	1.92 151 399
$\zeta_d$	1.49 412 230	1.51 590 540	1.52 812 251
$\beta_s$	−8.16 807 552	−7.58 869 921	−7.57 509 749
$\beta_p$	0.88 070 943	−2.04 108 977	−2.05 642 477
$\beta_d$	−3.98 746 964	−4.28 126 742	−4.37 308 135
$F_{SD}^0$	8.01 723 371	8.20 476 687	8.21 270 284
$G_{SD}^2$	3.09 515 011	1.63 956 989	1.33 691 737
$\rho_{core}$	1.75 805 984	1.52 188 028	1.52 506 485
$\alpha$	1.33 437 094	1.29 675 431	1.29 831 556
$\zeta'_s$	0.67 987 995	0.98 649 202	1.20 626 018
$\zeta'_p$	1.57 857 694	2.04 302 591	2.06 882 428
$\zeta'_d$	0.79 856 058	0.97 618 566	0.85 874 958
$a_{11}$	0.23 330 379	1.03 024 882	1.03 703 172
$b_{21}$	7.75 970 789	7.37 597 186	7.59 912 344
$c_{31}$	1.74 703 473	1.54 947 307	1.57 019 577
$a_{12}$	0.01 949 759	0.03 590 437	0.04 649 592
$b_{22}$	9.97 138 936	7.59 682 640	7.61 829 651
$c_{32}$	3.22 931 221	3.03 620 292	3.12 817 730

**Table 2. Means and Variances of the  $\gamma$  Distribution Fits for the UME<sub>(Ln–L)</sub> Computed for the N Complexes for Each Lanthanide Trication**

lanthanide ion	N	UME <sub>(Ln–L)</sub>		
		mean (Å)	variance (Å <sup>2</sup> )	p-value
Eu <sup>3+</sup>	103	0.0597	0.0090	0.5291
Gd <sup>3+</sup>	71	0.0513	0.0018	0.4981
Tb <sup>3+</sup>	43	0.0568	0.0029	0.7864

**Table 3. Means and Variances of the  $\gamma$  Distribution Fits for the UME Computed for the N Complexes for Each Lanthanide Trication**

lanthanide ion	N	UME		
		mean (Å)	variance (Å <sup>2</sup> )	p-value
Eu <sup>3+</sup>	103	0.1516	0.041	0.3098
Gd <sup>3+</sup>	71	0.1135	0.017	0.0815
Tb <sup>3+</sup>	43	0.1217	0.017	0.8832

average 0.494 Å, about 7.5 times the average error of the RM1 model for these same distances. The effect of the presence of the new types of distances also affects the more general averages, L–L, especially Eu–L. Clearly, the RM1 model for lanthanides is an improvement with respect to the previous Sparkle Model because it can be applied to more general situations.

Tables 5 and 6 show the same results for Gd and Tb. The pattern in the data is the same. That is, RM1 expands the accuracy of the Sparkle Model to other types of bonds. We could find neither Gd(III) complexes nor Tb(III) complexes of high crystallographic quality with sulfur directly coordinated to them. Likewise, we could not find terbium complexes directly coordinated to Br. That is why there are no averages for Gd–S bonds in Table 5 or Tb–S, and Tb–Br in Table 6.

**Table 4.** Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, Sparkle/PM7, Sparkle/RM1, and RM1 Unsigned Mean Errors of the Europium

types of distances	unsigned mean errors (Å)						
	N	RM1	Sparkle/RM1	Sparkle/PM7	Sparkle/PM6	Sparkle/PM3	Sparkle/AM1
Eu–Eu	20	0.1558	0.2387	0.2260	0.1462	0.1357	0.2014
Eu–O	783	0.0494	0.0746	0.0738	0.0839	0.0833	0.1091
Eu–N	203	0.0772	0.0658	0.0724	0.0731	0.0658	0.1169
Eu–C	76	0.0894	0.4144	0.6204	0.4677	0.4215	0.4818
Eu–S	12	0.0997	0.3657	1.5888	0.4504	0.3936	0.4683
Eu–Cl	12	0.0672	0.2700	0.1192	0.2604	0.2298	0.3015
Eu–Br	3	0.0058	0.4401	1.2720	0.4432	0.3918	0.4749
L–L	4353	0.1750	0.2397	0.2795	0.2170	0.2001	0.2751
Eu–L	1109	0.0597	0.1055	0.1339	0.1162	0.1102	0.1447
Eu–L, Eu–Eu, and L–L'	5462	0.1516	0.2124	0.2499	0.1965	0.1818	0.2486

**Table 5.** Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, Sparkle/PM7, Sparkle/RM1, and RM1 Unsigned Mean Errors of the Gadolinium

types of distances	unsigned mean errors (Å)						
	N	RM1	Sparkle/RM1	Sparkle/PM7	Sparkle/PM6	Sparkle/PM3	Sparkle/AM1
Gd–Gd	10	0.1566	0.1631	0.2947	0.1443	0.1443	0.1191
Gd–O	478	0.0441	0.0597	0.0564	0.0550	0.0553	0.0841
Gd–N	171	0.0560	0.0523	0.0938	0.0778	0.0768	0.0799
Gd–C	33	0.0859	0.2152	0.4002	0.2070	0.2075	0.2095
Gd–Cl	10	0.0925	0.3001	0.1422	0.2822	0.2458	0.2571
Gd–Br	3	0.0612	0.4819	1.5389	0.4891	0.4141	0.4449
L–L	2694	0.1298	0.1889	0.2294	0.1614	0.1607	0.2275
Gd–L	705	0.0513	0.0719	0.0925	0.0740	0.0732	0.0934
Gd–L, Gd–Gd, and L–L'	3399	0.1135	0.1646	0.201	0.1433	0.1425	0.1997

**Table 6.** Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, Sparkle/PM7, Sparkle/RM1, and RM1 Unsigned Mean Errors of the Terbium

types of distances	unsigned mean errors (Å)						
	N	RM1	Sparkle/RM1	Sparkle/PM7	Sparkle/PM6	Sparkle/PM3	Sparkle/AM1
Tb–Tb	10	0.1076	0.2332	0.3274	0.2826	0.2342	0.2426
Tb–O	344	0.0538	0.0707	0.0574	0.0793	0.0730	0.0783
Tb–N	62	0.0442	0.0495	0.0575	0.0594	0.0545	0.0422
Tb–C	28	0.0750	0.2175	0.3587	0.1943	0.2201	0.2388
Tb–Cl	13	0.0956	0.3334	0.3374	0.2802	0.2951	0.3360
L–L	1723	0.1390	0.2252	0.2328	0.2746	0.2027	0.2450
Tb–L	458	0.0568	0.0920	0.0942	0.0969	0.0929	0.0985
Tb–L, Tb–Tb, and L–L'	2181	0.1217	0.1971	0.2036	0.2371	0.1796	0.2141

Tables S4, S5, and S6, presented in the Supporting Information, contain unsigned mean errors for the angles  $UME_{(L'-Ln-L)}$ , where Ln stands for Eu, Gd, and Tb, obtained using the RM1 model for lanthanides and all five versions of the Sparkle Model: Sparkle/RM1, Sparkle/PM7, Sparkle/PM6, Sparkle/PM3, and Sparkle/AM1 for all complexes of the validation set, for each of the lanthanide trications.

## CONCLUSION

The Sparkle Model started by replacing the lanthanide trication with the charge of the ion plus the repulsive potential with the  $\alpha$  parameter,<sup>10</sup> followed by the inclusion of the Gaussians and the mass of the lanthanide,<sup>14</sup> the better scheme of parametrization,<sup>12</sup> and statistical testing of the parameters.<sup>33</sup> The RM1 model now being defined adds orbitals to this picture, acquiring generality in terms of types of coordination bonds. Results indicate that the overall accuracy of the RM1 Model is about 2% for interatomic distances of coordination, a level of

accuracy that makes it useful for studies involving complexes, materials, and nanoparticles, as well as, for example, to the design of Eu(III) and Tb(III) luminescent complexes and Gd(III) contrast agents for magnetic resonance imaging, applications, which typically require accurate geometries.

## ASSOCIATED CONTENT

### Supporting Information

Instructions on how to run RM1 model for lanthanides calculations on MOPAC 2012 (<http://openmopac.net>). Additional tables containing  $UME_{(Ln-L)}$  and  $UME$  for all 217 complexes of the validation set. Additional histograms comparing the RM1 Model with all five previous Sparkle Models: Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, Sparkle/PM7, and Sparkle/RM1 with respect to various classes of coordinating atoms. Sample MOPAC 2012 input (.mop) and output (.arc) files for one representative complex for each of the three lanthanide ions: Eu(III), Gd(III), and Tb(III). This

material is available free of charge via the Internet at <http://pubs.acs.org>. Easy tutorials on how to use the Sparkle Models can be found at <http://www.sparkle.pro.br>.

## AUTHOR INFORMATION

### Corresponding Author

\*Email: [rfreire@ufs.br](mailto:rfreire@ufs.br).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We appreciate the financial support from the Brazilian agencies, institutes, and networks: FACEPE (Pronex), CNPq, CAPES, FAPITEC-SE, INAMI, and RENAMI. We gratefully acknowledge the Cambridge Crystallographic Data Centre for the Cambridge Structural Database.

## REFERENCES

- (1) Ma, Y.; Wang, Y. *Coord. Chem. Rev.* **2010**, *254*, 972–990.
- (2) *Handbook on the Physics and Chemistry of Rare Earths*; Carnall, W. T.; Gschneidner, K. A., Jr.; Eyring, L., Eds.; North Holland Publ. Co.: Amsterdam, 1979; Vol. 3, p 172.
- (3) Bünzli, J. -C. G.; Comby, S.; Chauvin, A. S.; Vandevyver, C. D. B. *J. Rare Earths* **2007**, *25*, 257–274.
- (4) Munson, E. J.; Haw, J. F. *Anal. Chem.* **1990**, *62*, 2532–2536.
- (5) de Sá, G. F.; Malta, O. L.; Donega, C. M.; Simas, A. M.; Longo, R. L.; Santa-Cruz, P. A.; da Silva, E. F., Jr. *Coord. Chem. Rev.* **2000**, *196*, 165–195.
- (6) Dumas, S.; Jacques, V.; Sun, C.; Troughton, J. S.; Welch, J. T.; Chasse, J. M.; Schmitt-Willich, H.; Caravan, P. *Invest. Radiol.* **2010**, *45*, 600–612.
- (7) Hanaoka, K. *Chem. Pharm. Bull.* **2010**, *58*, 1283–1294.
- (8) Kitajima, K.; Maeda, T.; Watanabe, S.; Ueno, Y.; Sugimura, K. *Int. J. Urol.* **2012**, *19*, 806–811.
- (9) Eddaoudi, M.; Kim, J.; Vodak, D.; Sudik, A.; Wachter, J.; O’Keeffe, M.; Yaghi, O. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4900–4904.
- (10) de Andrade, A. V. M.; da Costa, N. B., Jr.; Simas, A. M.; de Sá, G. F. *Chem. Phys. Lett.* **1994**, No. 227, 349–353.
- (11) de Andrade, A. V. M.; da Costa, N. B., Jr.; Simas, A. M.; de Sá, G. F. *J. Alloys Compd.* **1995**, No. 225, 55–59.
- (12) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Inorg. Chem.* **2005**, *44*, 3299–3310.
- (13) de Andrade, A. V. M.; da Costa, N. B., Jr.; Malta, O. L.; Longo, R. L.; Simas, A. M.; de Sá, G. F. *J. Alloys Compd.* **1997**, No. 250, 412–416.
- (14) Rocha, G. B.; Freire, R. O.; da Costa, N. B., Jr.; de Sá, G. F.; Simas, A. M. *Inorg. Chem.* **2004**, *43*, 2346–2354.
- (15) Pavithran, R.; Reddy, M. L. P.; Alves, S., Jr.; Freire, R. O.; Rocha, G. B.; Lima, P. P. *Eur. J. Inorg. Chem.* **2005**, 4129–4137.
- (16) Albuquerque, R. Q.; Freire, R. O.; Malta, O. L. *J. Phys. Chem. A* **2005**, *109*, 4607–4610.
- (17) Freire, R. O.; Gonçalves e Silva, F. R.; Rodrigues, M. O.; de Mesquita, M. E.; Junior, N. B. C. *J. Mol. Model* **2005**, *12*, 16–23.
- (18) Freire, R. O.; Albuquerque, R. Q.; Alves, S., Jr.; Rocha, G. B.; de Mesquita, M. E. *Chem. Phys. Lett.* **2005**, *405*, 123–126.
- (19) Dutra, J. D. L.; Freire, R. O. *J. Photochem. Photobiol., A* **2013**, *256*, 29–35.
- (20) Dutra, J. D. L.; Bispo, T. D.; Freire, R. O. *J. Comput. Chem.* **2014**, *35*, 772–775.
- (21) Allen, F. H. *Acta Crystallogr. B* **2002**, *58*, 380–388.
- (22) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr. B* **2002**, *58*, 389–397.
- (23) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Mol. Model.* **2006**, *12*, 373–389.
- (24) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Organomet. Chem.* **2005**, *690*, 4099–4102.
- (25) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Chem. Phys. Lett.* **2005**, *411*, 61–65.
- (26) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Comput. Chem.* **2005**, *26*, 1524–1528.
- (27) Costa, N. B., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *Polyhedron* **2005**, *24*, 3046–3051.
- (28) Costa, N. B., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *Inorg. Chem. Commun.* **2005**, *8*, 831–835.
- (29) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Chem. Theory Comput.* **2006**, *2*, 64–74.
- (30) Bastos, C. C.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Photochem. Photobiol., A* **2006**, *117*, 225–237.
- (31) Freire, R. O.; Monte, E. V.; Rocha, G. B.; Simas, A. M. *J. Organomet. Chem.* **2006**, *691*, 2584–2588.
- (32) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Phys. Chem. A* **2006**, *110*, 5897–5900.
- (33) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Chem. Phys. Lett.* **2006**, *425*, 138–141.
- (34) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Chem. Phys. Lett.* **2007**, *441*, 354–357.
- (35) Costa, N. B., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Phys. Chem. A* **2007**, *111*, S015–S018.
- (36) Simas, A. M.; Freire, R. O.; Rocha, G. B. *Lect. Notes Computer Sci.* **2007**, *4488*, 312–318.
- (37) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Chem. Theory Comput.* **2007**, *3*, 1588–1596.
- (38) Simas, A. M.; Freire, R. O.; Rocha, G. B. *J. Organomet. Chem.* **2008**, *693*, 1952–1956.
- (39) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Braz. Chem. Soc.* **2009**, *20*, 1638–1645.
- (40) Freire, R. O.; Simas, A. M. *J. Chem. Theory Comput.* **2010**, *6*, 2019–2023.
- (41) Filho, M. A. M.; Dutra, J. D. L.; Rocha, G. B.; Freire, R. O.; Simas, A. M. *RSC Adv.* **2013**, *3*, 16747–16755.
- (42) Dutra, J. D. L.; Filho, M. A. M.; Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Chem. Theory Comput.* **2013**, *9*, 3333–3341.
- (43) Mesquita, M. E.; Junior, S. A.; Oliveira, F. C.; Freire, R. O.; Júnior, N. B. C.; de Sá, G. F. *Inorg. Chem. Commun.* **2002**, *5*, 292–295.
- (44) Faustino, W. M.; Malta, O. L.; Teotonio, E. E. S.; Brito, H. F.; Simas, A. M.; de Sá, G. F. *J. Phys. Chem. A* **2006**, *110*, 2510–2516.
- (45) Biju, S.; Reddy, M. L. P.; Freire, R. O. *Inorg. Chem. Commun.* **2007**, *10*, 393–396.
- (46) Souza, A. P.; Paz, F. A. A.; Freire, R. O.; Carlos, L. D.; Malta, O. L.; Alves, S., Jr.; de Sá, G. F. *J. Phys. Chem. B* **2007**, *111*, 9228–9238.
- (47) Allen, F. H. *Acta Crystallogr. B* **2002**, *58*, 380–388.
- (48) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr. B* **2002**, *58*, 389–397.
- (49) Allen, F. H.; Motherwell, W. D. S. *Acta Crystallogr. B* **2002**, *58*, 407–422.
- (50) Rodrigues, M. O.; Paz, F. A. A.; Freire, R. O.; de Sá, G. F.; Galembeck, A.; Montenegro, M. C.; Araujo, A. N.; Alves, S., Jr. *J. Phys. Chem. B* **2009**, *113*, 12181–12188.
- (51) Stewart, J. J. P. *MOPAC2012*; Stewart Computational Chemistry: Colorado Springs, CO, 2012; Available online: <http://openmopac.net>.
- (52) Dolg, M.; Stoll, H.; Savin, A.; Preuss, H. *Theor. Chim. Acta* **1989**, *75*, 173–194.
- (53) Maron, L.; Eisenstein, O. *J. Phys. Chem., A* **2000**, *104*, 7140–7143.
- (54) McNamara, J. P.; Berrigan, S. D.; Ian H. Hillier, I. H. *J. Chem. Theory Comput.* **2007**, *3* (3), 1014–1027.
- (55) *Finding Groups in Data: An Introduction to Cluster Analysis*, Kaufman, L., Rousseeuw, P. J., Eds.; John Wiley & Sons: Hoboken, NJ, 2009.

(56) William, J. C. *Practical Nonparametric Statistics*; John Wiley & Sons: New York, 1971.

(57) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1101–1111.