

## Sparkle Model for the Calculation of Lanthanide Complexes: AM1 Parameters for Eu(III), Gd(III), and Tb(III)<sup>†</sup>

Ricardo O. Freire, Gerd B. Rocha, and Alfredo M. Simas\*

Departamento de Química Fundamental, Universidade Federal de Pernambuco, 50.740-540 Recife, PE, Brazil

Received October 19, 2004

Our previously defined Sparkle model (*Inorg. Chem.* **2004**, *43*, 2346) has been reparameterized for Eu(III) as well as newly parameterized for Gd(III) and Tb(III). The parameterizations have been carried out in a much more extensive manner, aimed at producing a new, more accurate model called Sparkle/AM1, mainly for the vast majority of all Eu(III), Gd(III), and Tb(III) complexes, which possess oxygen or nitrogen as coordinating atoms. All such complexes, which comprise 80% of all geometries present in the Cambridge Structural Database for each of the three ions, were classified into seven groups. These were regarded as a “basis” of chemical ambience around a lanthanide, which could span the various types of ligand environments the lanthanide ion could be subjected to in any arbitrary complex where the lanthanide ion is coordinated to nitrogen or oxygen atoms. From these seven groups, 15 complexes were selected, which were defined as the parameterization set and then were used with a numerical multidimensional nonlinear optimization to find the best parameter set for reproducing chemical properties. The new parameterizations yielded an unsigned mean error for all interatomic distances between the Eu(III) ion and the ligand atoms of the first sphere of coordination (for the 96 complexes considered in the present paper) of 0.09 Å, an improvement over the value of 0.28 Å for the previous model and the value of 0.68 Å for the first model (*Chem. Phys. Lett.* **1994**, *227*, 349). Similar accuracies have been achieved for Gd(III) (0.07 Å, 70 complexes) and Tb(III) (0.07 Å, 42 complexes). Qualitative improvements have been obtained as well; nitrates now coordinate correctly as bidentate ligands. The results, therefore, indicate that Eu(III), Gd(III), and Tb(III) Sparkle/AM1 calculations possess geometry prediction accuracies for lanthanide complexes with oxygen or nitrogen atoms in the coordination polyhedron that are competitive with present day *ab initio*/effective core potential calculations, while being hundreds of times faster.

### Introduction

Lanthanide chemistry has been experiencing an upsurge in research activities since the late 1980s, gaining further momentum from Lehn's proposal in 1990 that lanthanide complexes could be regarded as light conversion molecular devices,<sup>1</sup> thus leading to the discovery of numerous new compounds, with many applications. Indeed, luminescent and electroluminescent materials,<sup>2</sup> luminescent sensors for chemical species,<sup>3</sup> UV dosimeters,<sup>4</sup> antireflection coatings for solar

cells,<sup>5</sup> fluorescent lighting,<sup>6</sup> complexes as diagnostics tools,<sup>7</sup> molecular and supramolecular polymetallic functional assemblies,<sup>8</sup> liquid crystals, and surfactants<sup>9</sup> are all being developed with lanthanides, at an increasing pace. The ability to efficiently and accurately model all of these molecular systems and interactions is, therefore, an open area of research.

More specifically, modeling the influence of the chemical ambience on the 4f<sup>n</sup> configuration is of significance in the investigation of magnetic and spectroscopic properties of

\* Author to whom correspondence should be addressed. E-mail: simas@ufpe.br.

<sup>†</sup> Part of the material in this paper has been presented by Ricardo O. Freire in his M.Sc. thesis, Universidade Federal de Pernambuco, Brazil, 2004.

(1) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 304.

(2) Kido, J.; Okamoto, Y. *Chem. Rev.* **2002**, *102*, 2357.

(3) Parker, D.; Kanthi-Senanayake, P.; Williams, J. A. G. *J. Chem. Soc., Perkin Trans.* **1998**, *2*, 2129.

(4) Gameiro, C. G.; da Silva, E. F., Jr.; Alves, S., Jr.; de Sá, G. F.; Santa-Cruz, P. A. *Mater. Sci. Forum* **1999**, *315*, 249.

(5) de Sá, G. F.; Alves, S., Jr.; da Silva, B. J. P.; da Silva, E. F., Jr. *Opt. Mater.* **1998**, *11*, 23.

(6) Ronda, C. R. *J. Alloys Compd.* **1995**, *225*, 534.

(7) Thunus, L.; Lejeune, R. *Coord. Chem. Rev.* **1999**, *184*, 125.

(8) Bünzli, J.-C. G.; Pighet, C. *Chem. Rev.* **2002**, *102*, 1897.

(9) Binnemans, K.; Görller-Walrand, C. *Chem. Rev.* **2002**, *102*, 2303.

lanthanide compounds. Semiempirical lanthanide intermediate neglect of differential overlap parameterizations have been achieved by Zerner's group,<sup>10,11</sup> which proved useful, for example, in the investigation of electronic and photoelectron spectra of lanthanide complexes.<sup>11,12</sup> Similarly, the description of ligand field effects is central in the design of new ligands capable of forming stable and highly luminescent complexes,<sup>13,14</sup> where the aim is to achieve strong ligand-to-metal energy transfer rates and intense metal-centered emission.

The characterization of the interaction between the ligands and the central ion can be done through the ligand field parameters,  $B_q^k$ , which can be calculated provided that the coordination geometry is known. Within the simple overlap model,<sup>15,16</sup> the values of  $B_q^k$  depend mainly on the interatomic distances between the ligand atoms and the central lanthanide ion. This dependence goes with the third, the fifth, and even the seventh power of the ligand–lanthanide interatomic distances, thus amplifying any inaccuracies. Such interatomic distances are the most sensitive geometric variables impacting upon the description of the effect of the surrounding chemical scenery on the lanthanide ion  $4f^n$  configuration, so creation of a method to accurately predict the geometries of lanthanide complexes from theoretical calculations would be of great advantage, more especially so in light of the fact that obtaining single crystals of lanthanide complexes of the appropriate size and optical quality for crystallographic structure determinations is difficult.<sup>17,18</sup>

In 1994, the Sparkle model for the Eu(III) ion was conceived and a first version<sup>19</sup> was presented within AM1.<sup>20</sup> In it, the 4f orbitals are contracted toward the nucleus and shielded from fields outside the ion by the outermost 5s and 5p closed shells, and the lanthanide ion was represented by a sparkle, that is, by a Coulombic charge of +3  $e$  superimposed to a repulsive exponential potential of the form  $\exp(-\alpha r)$ , which accounts for the size of the ion. At the same time, the then two variables of the model were parameterized using only one complex, tris(acetylacetonate)-(1,10-phenanthroline) of europium (III). That original Sparkle model was a new concept in lanthanide complex modeling, although its accuracy could only be regarded as qualitative;

the unsigned mean error for all interatomic distances between the europium (III) ion and the ligand atoms of the first sphere of coordination for the complexes considered was on the order of 0.68 Å.

Recently<sup>21</sup>, the Sparkle model was extended by including the lanthanide atomic mass, by introducing into it Gaussian functions in the core–core repulsion energy term to make it compatible with AM1, and by reparameterizing it within AM1, this time, using all distances of the coordination polyhedron of tris(acetylacetonate)(1,10-phenanthroline) of europium(III). The model<sup>21</sup> was improved to the point that it could be regarded as semiquantitative. Indeed, the unsigned mean error for all interatomic distances between the europium(III) ion and the ligand atoms of the first sphere of coordination for the complexes considered was lowered from 0.68 Å to 0.28 Å. In addition, it improved the  $\beta$ -diketones, complexes of the type used in the parameterization.

On the other hand, ab initio calculations of lanthanide complexes have been appearing in the literature using various types of effective core potentials (ECPs).<sup>22–28</sup> ECPs replace the chemically inert core electrons of the lanthanide with a potential acting on the valence electrons, which can also be derived to take into account relativistic effects. However, such ECP calculations still demand a large amount of CPU time, rendering high-quality calculations on systems of real chemical interest impractical. Reported ECP ab initio calculations have been performed on small systems, such as lanthanide trihalides,<sup>29–32</sup> the  $\text{Pr}^{3+}-\text{Cl}^-$  system,<sup>33</sup>  $\text{GdO}$ ,<sup>34</sup>  $\text{YbO}$ ,<sup>35</sup>  $\text{Ln}_2\text{O}_3$ ,<sup>36</sup> and  $\text{Ce}^{3+}$  oxide and fluoride compounds,<sup>37</sup> as well as on larger systems such as lanthanocenes.<sup>38–41</sup> Density functional theory (DFT) calculations have also been reported on larger complexes such as the triamido complexes  $\text{Ln}[\text{N}(\text{SiMe}_3)_2]_3$ ,<sup>42</sup> lanthanide(III) texaphyrins,<sup>43</sup> and lantha-

- (10) Culbertson, J. C.; Knappe, P.; Rosch, N.; Zerner, M. C. *Theo. Chim. Acta* **1987**, *71*, 21.
- (11) Cory, M. G.; Kostlmeier, S.; Kotzian, M.; Rosch, N.; Zerner, M. C. *J. Chem. Phys.* **1994**, *100*, 1353.
- (12) O'Brien, T. A.; Zerner, M. C.; Rack, P. D.; Holloway, P. H. *J. Electrochem. Soc.* **2000**, *147*, 792.
- (13) 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.
- (14) Faustino, W. M.; Rocha, G. B.; Silva, F. R. G.; Malta, O. L.; de Sá, G. F.; Simas, A. M. *J. Mol. Struct. (THEOCHEM)* **2000**, *527*, 245.
- (15) Malta, O. L. *Chem. Phys. Lett.* **1982**, *87*, 27.
- (16) Malta, O. L. *Chem. Phys. Lett.* **1982**, *88*, 353.
- (17) Thompson, L. C. In *Handbook on the Physics and Chemistry of Rare-Earths*; Gschneider, K. A., Eyring, L., Ed.; North-Holland: Amsterdam, The Netherlands, 1979.
- (18) Mondry, A.; Bukietyńska, K. *J. Alloys Compd.* **2004**, *374*, 27.
- (19) de Andrade, A. V. M.; da Costa, N. B., Jr.; Simas, A. M.; de Sá, G. F. *Chem. Phys. Lett.* **1994**, *227*, 349.
- (20) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

- (21) Rocha, G. B.; Freire, R. O.; da Costa, N. B., Jr.; de Sá, G. F.; Simas, A. M. *Inorg. Chem.* **2004**, *43*, 2346.
- (22) Dolg, M.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1989**, *90*, 1730.
- (23) Dolg, M.; Stoll, H.; Savin, A.; Preuss, H. *Theor. Chim. Acta*, **1989**, *75*, 173.
- (24) Cundari, T. R.; Stevens, W. *J. Chem. Phys.* **1993**, *98*, 5555.
- (25) Ross, R. B. *J. Chem. Phys.* **1994**, *100*, 8145.
- (26) Dolg, M. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; NIC series 1; John von Neumann Institute for Computing: Jülich, Germany, 2000; p 479.
- (27) Tsuchiya, T.; Nakajima, T.; Hirao, K.; Seijo, L. *Chem. Phys. Lett.* **2002**, *361*, 334.
- (28) Cao, X.; Dolg, M. *J. Mol. Struct. (THEOCHEM)* **2002**, *581*, 139.
- (29) Cundari, T. R.; Sommerer, S. O.; Strohecker, L. A.; Tipett, L. *J. Chem. Phys.* **1995**, *103*, 7058.
- (30) Lanza, G.; Fragalà, I. L. *Chem. Phys. Lett.* **1996**, *255*, 341.
- (31) Adamo, C.; Maldivi, P. *Chem. Phys. Lett.* **1997**, *268*, 61.
- (32) Tsuchiya, T.; Taketsugu, T.; Nakano, H.; Hirao, K. *J. Mol. Struct. (THEOCHEM)* **1999**, *461–462*, 203.
- (33) Reid, M. F.; Ng, B.; Newman, D. J. *J. Less-Common Met.* **1989**, *148*, 219.
- (34) Dolg, M.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1990**, *174*, 208.
- (35) Dolg, M.; Stoll, H.; Flad, H. J.; Preuss, H. *J. Chem. Phys.* **1992**, *97*, 1162.
- (36) Hirotsaki, N.; Ogata, S.; Kocer, C. *J. Alloys Compd.* **2003**, *351*, 31.
- (37) Dorenbos, P.; Andriessen, J.; van Eijik, C. W. E. *J. Solid State Chem.* **2003**, *171*, 133.
- (38) Kaupp, M.; Schleyer, P. V.; Dolg, M.; Stoll, H. *J. Am. Chem. Soc.* **1992**, *114*, 8202.
- (39) Dolg, M.; Fulde, P.; Stoll, H.; Preuss, H.; Chang, A.; Pitzer, R. M. *Chem. Phys.* **1995**, *195*, 71.
- (40) Liu, W. J.; Dolg, M.; Fulde, P. *J. Chem. Phys.* **1997**, *107*, 3584.
- (41) Liu, W. J.; Dolg, M.; Fulde, P. *Inorg. Chem.* **1998**, *37*, 1067.
- (42) Eisenstein, O.; Maron, L. *J. Organomet. Chem.* **2002**, *647*, 190.

nocenes.<sup>44</sup> More recently, Perrin et al. reported full DFT and QM/MM calculations on lanthanide complexes with cyclopentadienyl ligands in order to study their chemical reactivities.<sup>45</sup> Cosentino's group have also reported full geometry optimization calculations on larger complexes such as [Gd-(H<sub>2</sub>O)<sub>9</sub>],<sup>3+</sup> gadolinium(III) complexes with poly-amino carboxylate ligands,<sup>47</sup> gadolinium(III) with 1,4,7,10-tetraazacyclodecane-N,N',N'',N'''-tetraacetic acid (DOTA) and diethylenetriamine pentaacetic acid ligands,<sup>48</sup> four [Ln(DOTA)(H<sub>2</sub>O)]<sup>-</sup> systems,<sup>49</sup> and four complexes of europium(III) with bis(R-amide) diethylenetriamine pentacetate ligands.<sup>50</sup> Because of their large size, the geometry optimization calculations on lanthanide complexes carried out by Cosentino's group have been primarily restricted to restricted Hartree-Fock (RHF) calculations using the small basis set 3-21G, which seems to provide reliable geometries because their results indicate that the use of better ligand basis sets are not counterbalanced by a significant improvement in the calculated geometries.

In the present paper, 21 ab initio ECP full geometry optimizations are presented on larger lanthanide complex systems, the largest complex so far to have its geometry fully RHF/6-31G\*/ECP optimized being a dieuropium complex with 134 atoms.

Ab initio methods, even at the RHF/3-21G/ECP level, are already time-consuming and expensive for research in complex design, where many calculations must be carried out on several different complexes simultaneously by varying the ligands and studying the impact of these structural changes in their luminescence and in other properties. Therefore, reliable, accurate, and fast quantum chemical models for predicting geometries are urgently needed.

The computing times for Hartree-Fock ab initio techniques normally scale as  $N^4$ , where  $N$  is the number of basis functions used. As such, by doubling the size of the system, the calculation is expected to take 16 times longer; this severely limits the size of the systems that can be studied. Calculations taking into account electron correlation, such as configuration interaction, Möller-Plesset, coupled cluster, and others, impose an even more severe penalty as the system size increases. Conventional neglect of diatomic differential overlap semiempirical methods scale as  $N^3$ , which is better but still leads to performance problems, particularly when applied to large biomolecules. However, Stewart<sup>51</sup> has developed the MOZYME algorithm, which has permitted calculations on systems as large as 20 000 atoms. Recently, Anikin et al.<sup>52</sup> developed a truly linear scaling technique for

semiempirical methods, called LocalSCF, which now allows AM1 calculations on systems as large as 120 000 atoms on a personal computer. As a result of these developments in linear scaling, combined semiempirical/DFT calculations are starting to appear<sup>53</sup> in lieu of the more widely used QM/MM methods,<sup>54</sup> so that the application of semiempirical techniques to giant biomolecular systems is beginning to be implemented, with important consequences such as locating enzyme active sites,<sup>55</sup> estimating protein  $pK_a$  values,<sup>56,57</sup> and conducting electronic property studies of mechanisms of DNA binding sites in transcription.<sup>58</sup> This demonstrates that there is value in extending the parameterization of AM1 to lanthanide ions because they also play an important role in molecular recognition and chirality sensing of biological substrates.<sup>59</sup>

In the absolute vast majority of all of their complexes, europium, gadolinium, and terbium ions are directly coordinated to nitrogen or oxygen atoms of the ligands. Indeed, 80% of all Eu(III), Gd(III), and Tb(III) complexes whose structure have been deposited in the Cambridge Structural Database 2003 (CSD)<sup>60-62</sup> possess only nitrogen or oxygen atoms in their first sphere of coordination. Other types of ligand coordinating species, which are less common, are carbon atoms and halide ions, each present in 9% of the lanthanide complexes deposited in the CSD<sup>60-62</sup>. Complexes with other ligand coordinating atoms, such as S, P, Se, Te, Si, and so forth, are extremely rare.

To address this, a much more sophisticated parameterization of our Sparkle model is reported here, with the same number of parameters as earlier versions<sup>21</sup> but being carried out in a much more extensive way, aiming at producing a method that can be regarded as quantitative for lanthanide complexes with oxygen or nitrogen atoms in the coordination polyhedron.

New parameterizations are presented not only for Eu(III) but also for Gd(III) and Tb(III). Europium(III) and terbium(III) were chosen because their complexes display bright red and green emissions,<sup>2</sup> respectively. Europium(III) and terbium(III) complexes, indeed, function as emitting materials with many uses, such as being probes in biophysical applications.<sup>63</sup> Gadolinium(III), a highly paramagnetic ion,

(43) Cao, X.; Dolg, M. *Mol. Phys.* **2003**, *101*, 2427.

(44) Perrin, L.; Maron, L.; Eisenstein, O.; Schwartz, D. J.; Burns, C. J.; Andersen, R. A. *Organometallics* **2003**, *22*, 5447.

(45) Perrin, L.; Maron, L.; Eisenstein, O. *New J. Chem.* **2004**, *10*, 1255.

(46) Cosentino, U.; Moro, G.; Pitea, D.; Calabi, L.; Maiocchi, A. *J. Mol. Struct. (THEOCHEM)* **1997**, *392*, 75.

(47) Cosentino, U.; Moro, G.; Pitea, D.; Villa, A. *J. Phys. Chem. A* **1998**, *102*, 4606.

(48) Villa, A.; Cosentino, U.; Pitea, D. *J. Phys. Chem. A* **2000**, *104*, 3421.

(49) Cosentino, U.; Villa, A.; Pitea, D.; Moro, G.; Barone, V.; Maiocchi, A. *J. Am. Chem. Soc.* **2002**, *124*, 4901.

(50) Cosentino, U.; Pitea, D.; Moro, G.; Barone, V.; Villa, A.; Muller, R. N.; Botteman, F. *Theor. Chem. Acc.* **2004**, *111*, 204.

(51) Stewart, J. J. P. *Int. J. Quantum Chem.* **1996**, *58*, 133.

(52) Anikin, N. A.; Anisimov, V. M.; Bugaenko, V. L.; Bobrikov, V. V.; Adreyev, A. M. *J. Chem. Phys.* **2004**, *121*, 1266.

(53) Titmuss, S. J.; Cummins, P. L.; Bliznyuk, A. A.; Rendell, A. P.; Gready, J. E. *Chem. Phys. Lett.* **2000**, *320*, 169.

(54) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1996; p 119.

(55) Ohno, K.; Kamiya, N.; Asakawa, N.; Inoue, Y.; Sakurai, M. *J. Am. Chem. Soc.* **2001**, *123*, 8161.

(56) Ohno, K.; Kamiya, N.; Asakawa, N.; Inoue, Y.; Sakurai, M. *Chem. Phys. Lett.* **2001**, *341*, 387.

(57) Nakajima, S.; Ohno, K.; Inoue, Y.; Sakurai, M. *J. Phys. Chem. B* **2003**, *107*, 2867.

(58) Kurita, N.; Sengoku, Y.; Sekino, H. *Chem. Phys. Lett.* **2003**, *372*, 583.

(59) Tsukube, H.; Shinoda, S. *Chem. Rev.* **2002**, *102*, 2389.

(60) Allen, F. H. *Acta Crystallogr., Sect. B* **2002**, *58*, 380-388.

(61) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr., Sect. B* **2002**, *58*, 389-397.

(62) Allen, F. H.; Motherwell, W. D. S. *Acta Crystallogr., Sect. B* **2002**, *58*, 407-422.

(63) Selvin, P. R. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 275.

was chosen because of important applications of its complexes when used as contrast agents in magnetic resonance imaging.<sup>64–67</sup>

### Implementation of the Sparkle Model in the Mopac93r2 Package

To implement the Sparkle model in the software package Mopac93r2, a previously described procedure<sup>21</sup> was followed, where only one lanthanide ion is implemented at a time via modifications in subroutines block.f, calpar.f, and rotate.f. Because position 103 of Mopac's various arrays is used to define the original Mopac +2e Sparkle, it was used to define the Sparkle representing the lanthanide ions, either Eu(III), Gd(III), or Tb(III) ions. As such, the Sparkle core charge, CORE(103), was changed from +2e to +3e for all three cases.

The heats of formation of the lanthanide ions were obtained by adding their first three ionization potentials to their respective atomic heats of atomization,<sup>68</sup> yielding 1006.6 kcal mol<sup>-1</sup> for europium, 991.4 kcal mol<sup>-1</sup> for gadolinium, and 999.0 kcal mol<sup>-1</sup> for terbium, which were assigned to EHEAT(103).

The Sparkle parameters that were optimized are all related to the core–core repulsion energy between atoms A and B,  $E_N(A,B)$ , which, in the semiempirical method AM1, is<sup>69</sup>

$$E_N(A,B) = Z_A Z_B \langle S_A S_A | S_B S_B \rangle (1 + e^{-\alpha_A R_{AB}} + e^{-\alpha_B R_{AB}}) + \left( \frac{Z_A Z_B}{R_{AB}} \right) \left( \sum_{k=1}^{n_{GA}} a_{k_A} e^{-b_{k_A} (R_{AB} - c_{k_A})^2} + \sum_{k=1}^{n_{GB}} a_{k_B} e^{-b_{k_B} (R_{AB} - c_{k_B})^2} \right) \quad (1)$$

where  $\langle S_A S_A | S_B S_B \rangle$  is parameter GSS;  $\alpha_A$  and  $\alpha_B$  are ALP, a parameter representing the hardness of the lanthanide ion core;  $Z_A$  and  $Z_B$  are CORE, the core charges associated with nuclei A and B, respectively;  $R_{AB}$  is the interatomic distance between atoms or sparkles A and B;  $a_k$ ,  $b_k$ , and  $c_k$  are, respectively, the intensity, width, and position of the  $k^{\text{th}}$  Gaussian function; and, finally,  $n_{GA}$  and  $n_{GB}$  are the number of Gaussian functions centered in atoms A and B, respectively.

As in the previous version of the Sparkle model,<sup>21</sup> two spherical Gaussian functions were added to the core–core repulsion energy, each one containing three adjustable coefficients,  $a_k$ ,  $b_k$ , and  $c_k$ .

Therefore, the eight AM1 parameters to be optimized for each lanthanide ion sparkle are GSS, ALP,  $a_1$ ,  $b_1$ ,  $c_1$ ,  $a_2$ ,  $b_2$ , and  $c_2$ .

### General Parameterization Procedure

The parameterization procedure is a nonlinear minimization of an eight-dimension response function, using a combination of Simplex and Newton–Raphson methods, aimed at finding one of its local minima, which, ideally, should be the global minimum and should make chemical sense.

The experimental crystallographic structures used were all taken from the CSD.<sup>60–62</sup> The traditional figure of merit for crystal structures is the crystallographic R factor, which provides a measure of how well the refined structure agrees with the experimental model. In the present study, only structures of high quality were considered, that is, structures with R factors of less than 5%.

(64) Aime, S.; Botta, M.; Fasano, M.; Terreno, E. *Chem. Soc. Rev.* **1998**, 27, 19.

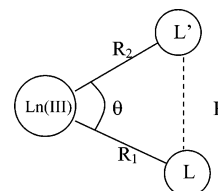
(65) Caravan, P.; Ellison, J. J.; McMurry, T. J.; Lauffer, R. B. *Chem. Rev.* **1999**, 99, 2293.

(66) Adding, L. C.; Bannenberg, G. L.; Gustafsson, L. E. *Cardiovasc. Drug Rev.* **2001**, 19, 41.

(67) Atar, E. *Isr. Med. Assoc. J.* **2004**, 6, 412.

(68) *Handbook of Chemistry and Physics* [CD-ROM]; Lide, D. R., Ed.; CRC Press: New York, 2002.

(69) Equation 1 applies to all atom pairs in AM1 except N–H and O–H.



**Figure 1.** Drawing representing the main variables used in the response function, Ln(III)–L interatomic distances and L–Ln(III)–L' bond angles. Ln(III) stands for the central lanthanide ion, and L and L' are ligand atoms of the coordination polyhedron.

For the current work, 15 different structures of complexes for each lanthanide ion were also considered for parameterization. The response function,  $F_{\text{resp}}$ , was thus defined as

$$F_{\text{resp}} = \sum_{i=1} \left\{ \sum_{j=1} [100(R_{i,j}^{\text{CSD}} - R_{i,j}^{\text{calc}})]^2 + \sum_{k=1} \left[ \frac{2}{3} (\theta_{i,k}^{\text{CSD}} - \theta_{i,k}^{\text{calc}}) \right]^2 \right\} \quad (2)$$

where index  $i$  runs over all of the different complexes; 100 and  $2/3$  are coefficients taken from the response function originally used to parameterize MNDO;<sup>70</sup> index  $j$  runs over all of the distances,  $R$ , of the lanthanide ion to each of the directly coordinated atoms from the ligands; superscripts CSD and calc refer to experimental and calculated quantities, respectively; and index  $k$  runs over all of the  $\theta$  angles formed by all combinations of two of the directly coordinated atoms from the ligands with the europium(III) ion in its vertex, as in Figure 1. By adjusting  $R_1$ ,  $R_2$ , and the  $\theta$  angle, in Figure 1, the L–L' interatomic distance, which belongs to the coordination polyhedron, was indirectly adjusted.

The next step was to define the set of complex structures to be used in the parameterization procedure; this set, specific for each lanthanide ion, was called the parameterization set. It is not simple to classify dozens of structures into smaller representative groups, from which one or two structures can be picked to include in the parameterization set, so that a cluster analysis of all complexes available in the CSD for each lanthanide with oxygen or nitrogen as coordinating atoms could be done. The cluster analysis was run with Statistica 6.0 software, using the Euclidean distances with complete linkage to cluster the complexes. As variables, the number of atoms directly coordinated to the lanthanide ion for each of the following types of ligands were used:  $\beta$ -diketone, nitrate, monodentate, bidentate, tridentate, and polydentate; the dilanthanides were considered as a separate group. Moreover, the average unsigned mean error for each complex  $i$ ,  $\text{UME}_i$ , was also included and is defined as

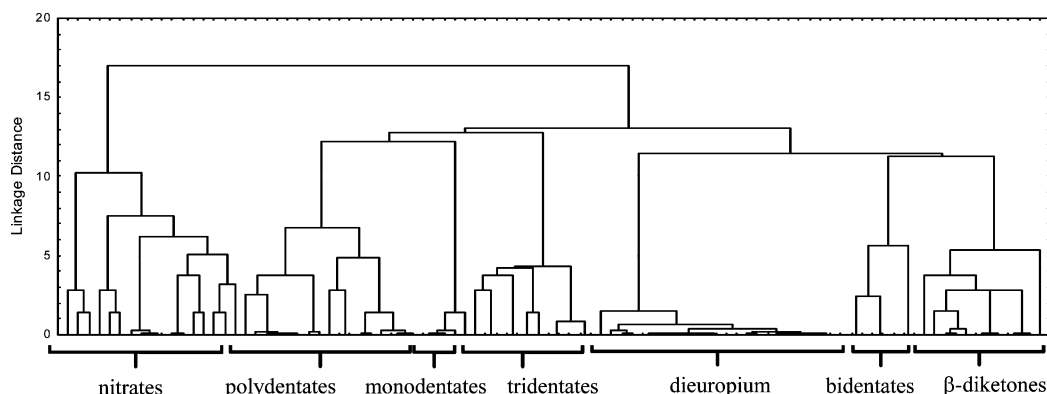
$$\text{UME}_i = \frac{1}{n} \sum_{j=1}^n |R_{i,j}^{\text{CSD}} - R_{i,j}^{\text{calc}}| \quad (3)$$

where  $n$  is the number of ligand atoms directly coordinating the lanthanide ion. The  $\text{UME}_i$  values were obtained from calculations using an initial guess for all eight Sparkle parameters, as will be described in the specific sections below. As a result, all available complexes in CSD for each lanthanide ion were classified accordingly.

### Sparkle/AM1 Parameters for Eu(III)

The initial estimate for the eight Eu(III) Sparkle parameters was taken from the earlier parameterization,<sup>21</sup> from which all 96 CSD-available Eu(III) complex structures with R factors of less than 5% were calculated. Figure 2 shows the

(70) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, 99, 4899.



**Figure 2.** Cluster analysis of 96 Eu(III) complexes, in terms of both the UMEs and the number of atoms directly coordinated to the lanthanide ion, for each of the various types of ligands. The UMEs are calculated as the sum of all of the absolute values of the differences between the experimental and calculated interatomic distances between all of the atoms of the coordination polyhedron and between all of these and the central Eu(III) ion.

**Table 1.** Number of Structures of Lanthanide Complexes Classified into Each Ligand Group by Cluster Analysis, for Each of the Three Ions Considered in This Paper

ligand group		number of structures		
number	name	Eu(III)	Gd(III)	Tb(III)
1	$\beta$ -diketone	13	5	4
2	nitrate	17	8	4
3	monodentate	5	10	12
4	bidentate	7	3	2
5	tridentate	13	2	3
6	polydentate	17	30	8
7	dilanthanide	24	12	9

**Table 2.** Parameters for the Sparkle/AM1 Model for the Eu(III), Gd(III), and Tb(III) Ions

	Sparkle/AM1		
	Eu(III)	Gd(III)	Tb(III)
GSS	55.605 912 203 3	55.708 324 761 8	55.724 595 690 4
ALP	2.124 718 861 3	3.652 548 457 6	2.341 888 909 5
$a_1$	0.569 512 247 5	0.701 351 205 9	0.773 445 798 6
$b_1$	7.468 020 764 2	7.545 448 254 4	7.651 052 576 8
$c_1$	1.731 972 985 5	1.776 195 267 4	1.703 346 395 5
$a_2$	0.328 661 904 6	0.129 309 357 7	0.393 623 343 0
$b_2$	7.800 977 959 9	8.343 799 146 5	7.926 145 660 2
$c_2$	2.964 128 549 0	3.011 031 971 5	3.013 295 134 5

dendrogram obtained from the subsequent cluster analysis that was performed, from which the number of complexes in each of the seven groups can be detected, and these complexes are presented in Table 1.

On the basis of the dendrogram, 15 structures were chosen to constitute the Eu(III) parameterization set by picking three from the  $\beta$ -diketone group, two from the nitrate group, four from the tridentate group, three from the polydentate group, and three from the dieuropium group. These were selected because eight of them also contained monodentate ligands, and similarly, two of them also contained bidentate ligands.

The optimized Sparkle parameters obtained from the nonlinear minimization of the response function are presented in Table 2. To assess the predictive power of the Sparkle model, Figure 3 plots the UMEs for each of the 96 Eu(III) complexes, grouped according to the cluster analysis and the group numbers classified in Table 1.

Parts a and b of Figure 3 consider UMEs for all interatomic distances of the coordinated polyhedron, including distances from the center of Eu(III) to all of its vertexes. From Figure

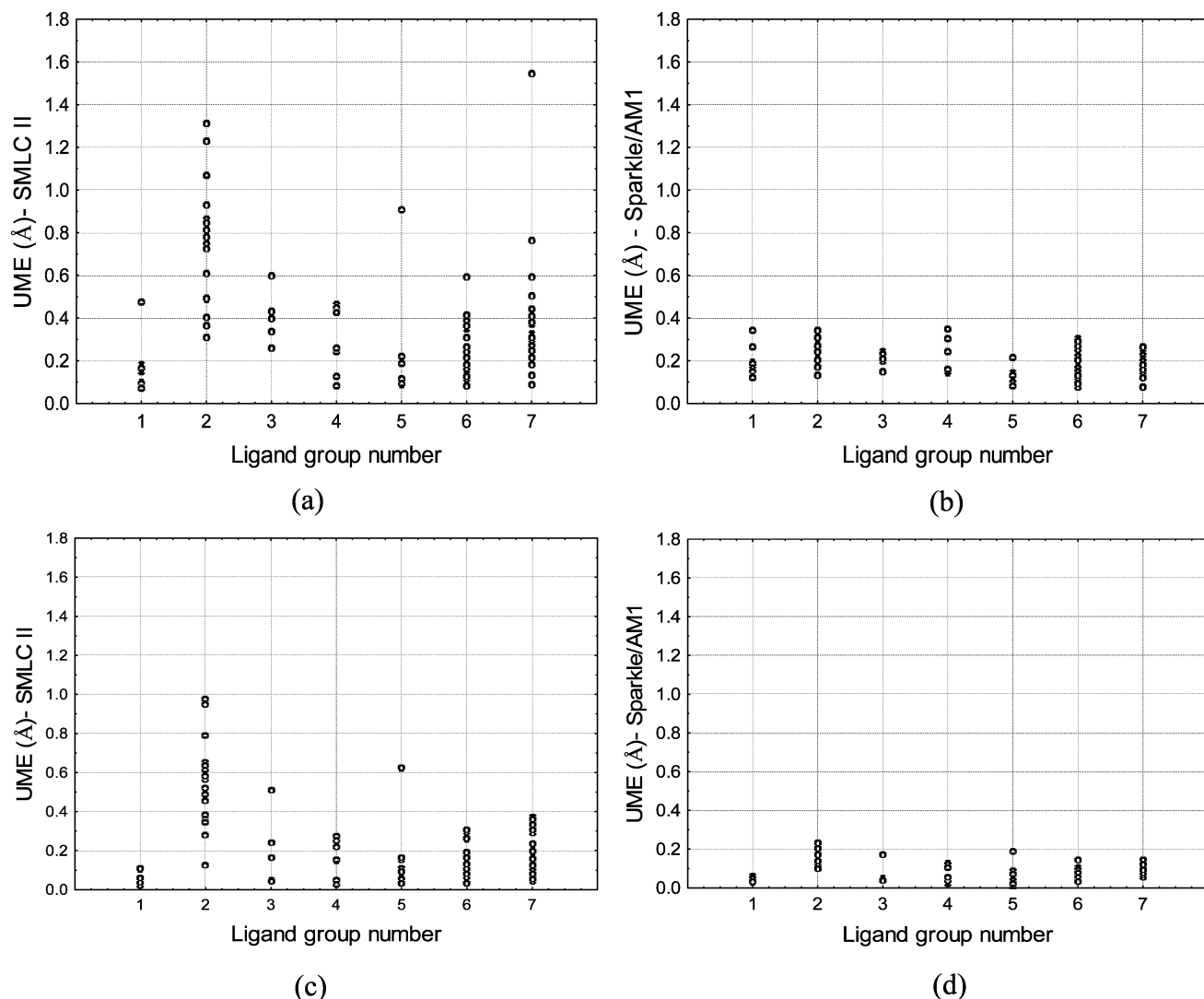
3, it can be seen that the present parameterization is able to predict the geometry of the ground state of the various types of Eu(III) complexes with an accuracy similar to the one displayed by the previous SMLC II method for  $\beta$ -diketones (Group 1). As such, upon moving from SMLC II to the present parameterization, we can obtain a generalization of the Sparkle model. Moreover, some problems encountered in version II have also been solved in the present version, such as, for example, the case of the structures with nitrate ligands (Group 2 in Figure 3). In SMLC II, nitrates were predicted to coordinate to the europium(III) ion in a monodentate manner, causing a large error in the predicted geometry for this class of ligands, Figure 3a. In the present version, nitrates coordinate in the correct manner, that is, as bidentates, thus reducing the associated UMEs, Figure 3b.

Parts c and d of Figure 3 consider UMEs only for the distances from the Eu(III) ion to all vertexes of the coordinated polyhedron. Again, an improvement can be demonstrated by moving from SMLC II to the present parameterization, suggesting that, in the present study, a better generalization could be obtained for the Sparkle model, in terms of ligand types.

Another innovative aspect of the present version is the parameterization of the distance between two Eu(III) ions when facing each other in a complex. Figure 4 shows a plot of the UMEs for each of the 24 calculated dieuropium structures. The UMEs for the dieuropium complexes, being below 0.3 Å, are compatible with the UMEs obtained for all of the other ligand types, as shown in Figure 3b.

To assess the usefulness of the present Sparkle method, it is important to compare its computing time and accuracy with those from ab initio methods of the type normally used nowadays by researchers in the field of lanthanide chemistry. Accordingly, in lieu of the lanthanide atoms, almost all ab initio calculations use the effective core potential developed by Dolg et al.<sup>23</sup> available in Gaussian 98.<sup>71</sup> This ECP includes 46 + 4f<sup>n</sup> electrons of the lanthanide in the core, leaving the outermost electrons to be treated explicitly by a [5s4p3d]-GTO valence basis set.

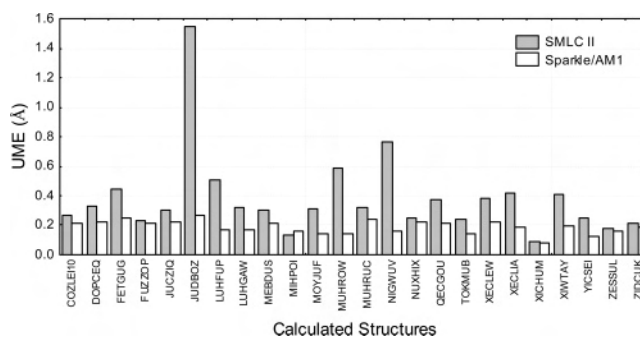
For that purpose, seven complexes were selected, one from each of the groups defined in Table 1 (Figure 5).



**Figure 3.** UMEs for each of the 96 Eu(III) complexes. The UMEs are grouped according to the cluster analysis shown in Figure 2, and the group numbers are defined in Table 1. Parts a and b present these UMEs using SMLC II<sup>21</sup> and Sparkle/AM1, respectively, calculated by the sum of all absolute values of the differences between the experimental and calculated interatomic distances, between the atoms of the coordination polyhedron, and between all of these and the central Eu(III) ion. Parts c and d also present these UMEs using SMLC II<sup>21</sup> and Sparkle/AM1, respectively, but, this time, considering only the interatomic distances involving the Eu(III) ion.

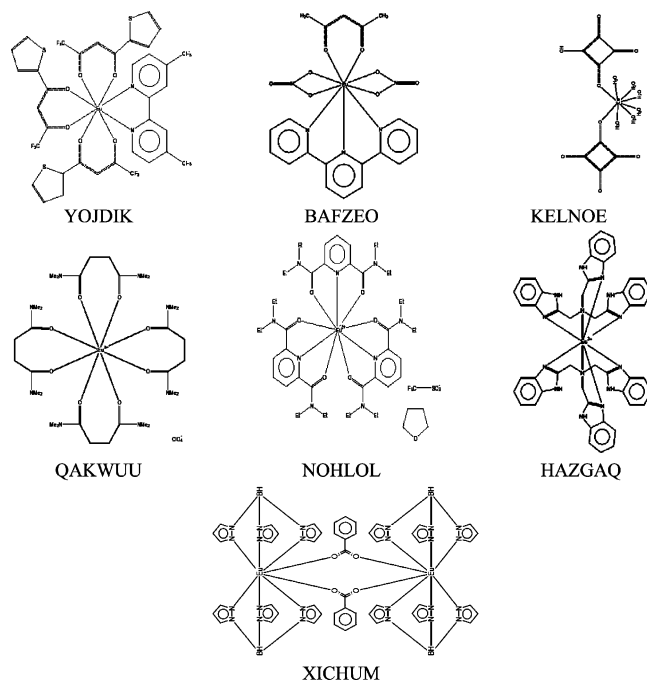
Let us consider, first, only the CPU time used in the single-point SCF calculation using either the present Sparkle/AM1 model or four ab initio calculations, all using the ECP, with the following bases: STO-3G, 3-21G, and 6-31G\*. Because of the length of the calculations, only two complexes, identified by their CSD codes of BAFZEO and XICHUM, were initially calculated. BAFZEO belongs to the nitrate group, although it also possesses a  $\beta$ -diketone ligand, as well

(71) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

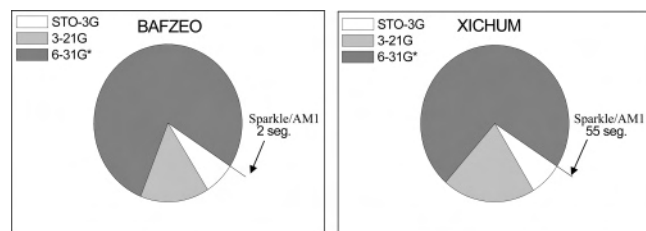


**Figure 4.** UMEs obtained using SMLC II<sup>21</sup> (previous version) and Sparkle/AM1 for all dieuropium (III) complexes that composed the test set that were obtained from the CSD.<sup>60–62</sup> The UMEs are calculated as the sum of all of the absolute values of the differences between the experimental and calculated interatomic distances between all of the atoms of the coordination polyhedron and between all of these and the central Eu(III) ion.

as a tridentate one. XICHUM is a dieuropium complex with two bidentate ligands bridging both europium(III) ions, with each ion also having two terminal tridentate ligands.



**Figure 5.** Schematic two-dimensional representation of the Eu(III) structures of complexes used for the comparison between Sparkle/AM1 and some ab initio models. The ab initio full geometry optimizations have been performed using the Hartree–Fock method with STO-3G, 3-21G, and 6-31G\* basis sets for all atoms, except for the Eu(III) ion, in which case we used the quasi-relativistic ECP of Dolg et al.<sup>23</sup>



**Figure 6.** Relative computational time spent in the single-point calculation of the complexes BAFZEO and XICHUM for each quantum chemical model considered. The structures used in this analysis were identified by their codes in the CSD.<sup>60–62</sup> All calculations have been carried out on an Athlon 1.53 GHz PC with 1 GB of RAM memory (DDR).

Figure 6 shows the relative SCF single-point CPU times for these complexes. Clearly, Sparkle/AM1 is a few hundred times faster than STO-3G, one thousand times faster than 3-21G, and several thousand times faster than 6-31G\*. For the larger XICHUM complex, the single-point Sparkle/AM1 took about 1 minute, whereas the single-point 6-31G\* took more than 2 days, on an Athlon 1.53 GHz PC with 1 GB of RAM memory (DDR).

The accuracy of the predicted geometrical variables for the complexes was then compared, between the present Sparkle/AM1 calculations and the ab initio RHF/STO-3G/ECP, RHF/3-21G/ECP (a commonly used standard of calculation for lanthanide complexes<sup>46–49</sup>), and RHF/6-31G\*/ECP calculations only. Figure 7 shows the UMEs for these calculations. Figure 7a considers the UMEs for all interatomic distances of the coordinated polyhedron, including the distances from Eu(III), at its center, to all of its vertexes, whereas Figure 7b considers the UMEs for all distances of the Eu(III) ion to all vertexes of the coordinated polyhedron.

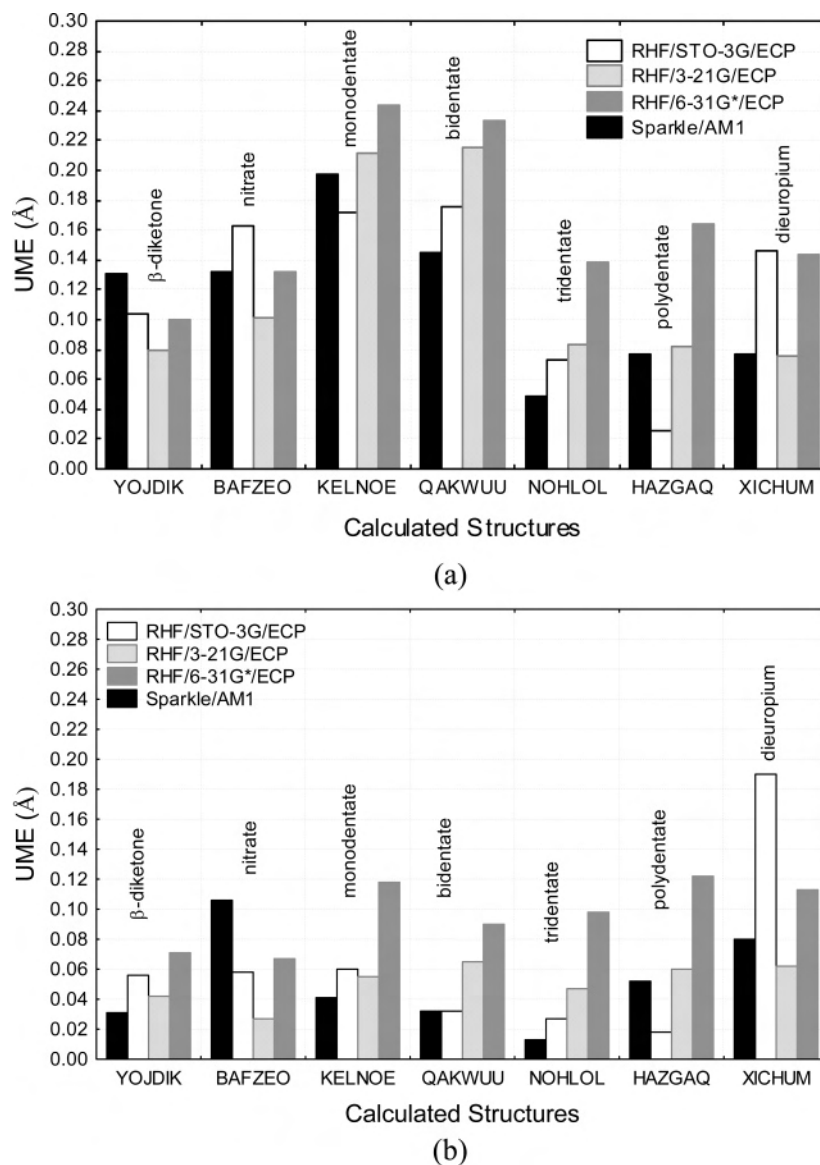
The complexes are ordered in Figure 7 according to the ligand group they belong to, as shown in Table 1. In both parts a and b of Figure 7, the Sparkle/AM1 calculations seem to be more accurate than the ab initio ones. Furthermore, moving from RHF/STO-3G/ECP to RHF/3-21G/ECP and up to RHF/6-31G\*/ECP does not seem to increase the accuracies of the predicted geometries in comparison to the crystallographic ones, as would normally be expected. In reality, for the complexes studied, improving the quality of the basis set may even worsen the quality of the predicted geometries, a fact that warrants further investigation. Indeed, both parts a and b of Figure 7 show that 6-31G\* UMEs are larger than 3-21G UMEs for all seven of the complexes considered. Moreover, no substantial differences are apparent between the RHF/STO-3G/ECP and RHF/3-21G/ECP results; STO-3G UMEs are smaller than 3-21G UMEs for exactly four complexes in Figure 7a, whereas the reverse is true in Figure 7b. Finally, Parts a and b of Figure 7 suggest that the Sparkle/AM1 model results follow a pattern of description of the chemical ambience around the lanthanide ion similar to the ab initio ones, with the same level of accuracy.

Table 3 shows the average distance UMEs for the seven europium(III) complexes presented in Figure 5, for all of the interatomic distances involving the Eu(III) ion, and for the atoms in the coordination polyhedra of the complexes. A close examination of Table 3 suggests that Sparkle/AM1 leads to results that are quite competitive with ab initio/ECP calculations. Indeed, although the UMEs for all of the distances from the RHF/3-21G/ECP calculations are only 4% higher than the corresponding UMEs for those of Sparkle/AM1, the same UMEs from the RHF/STO-3G/ECP calculations are 14% higher than the corresponding UMEs for the Sparkle/AM1 calculations. The larger error in the europium(III)–europium(III) distance for the RHF/STO-3G/ECP calculations is a result of its inability to describe such a distance in the XICHUM complex, as seen in Figure 7a. Consistent with the results present in Figure 7, RHF/6-31G\*/ECP UMEs are, on average, 48% larger than RHF/3-21G/ECP UMEs for all types of bonds. Hence, for the other ions presented in this paper, Gd(III) and Tb(III), we will present Sparkle/AM1 results only, because the heavy computational costs involved in the ab initio calculations do not seem to translate into increased geometry accuracy, at least not for comparison with the crystallographic data for these complexes.

Table 3 further shows the bond and dihedral angle UMEs for the same set of seven europium(III) complexes presented in Figure 5. The accuracy of the Sparkle/AM1 bond and dihedral angle results seem to be of similar quality to those of the RHF/3-21G/ECP and RHF/6-31G\*/ECP ab initio calculations. Sparkle/AM1 performs significantly better for the angles than RHF/STO-3G/ECP calculations.

### Sparkle/AM1 Parameters for Gd(III)

Because europium and gadolinium are neighbors in the periodic table, we used the Eu(III) parameters, as shown in Table 2, as the initial guess for the eight Gd(III) Sparkle parameters. Likewise, the same procedure as that for Eu(III)



**Figure 7.** UMEs (in Å) obtained from the Sparkle/AM1 and ab initio (RHF/STO-3G/ECP, RHF/3-21G/ECP, and RHF/6-31G\*/ECP) calculations of the ground state geometries for each of the seven Eu(III) complexes (a) considering all interatomic distances of the coordination polyhedron as well as all interatomic distances between all of these and the central Eu(III) ion and (b) considering only the distances involving the Eu(III) ion and the atoms of the coordination polyhedron.

was used to compute all 70 CSD-available Gd(III) structures with oxygen or nitrogen atoms in the coordination polyhedron. Figure 8 shows the dendrogram obtained from the subsequent cluster analysis that was performed, from which the number of complexes in each of the seven groups can be detected, and these complexes are presented in Table 1.

The dendrogram was very similar to the one obtained for Eu(III). Ligands of the  $\beta$ -diketone, tridentate, and digadolinium groups could be very well-represented by only two structures, but the nitrate group, which presented problems in their description, together with the polydentates, because of their large diversity, required more than one representative structure to properly describe the assortment of ligands. Accordingly, 15 structures were chosen to constitute the Gd(III) parameterization set by picking two from the  $\beta$ -diketone group, three from the nitrate group, one from the monodentate group, one from the bidentate group, two from the tridentate group, four from the polydentate group, and

two from the digadolinium group. The optimized Gd(III) Sparkle parameters, obtained from the nonlinear minimization of the response function, are presented in Table 2.

Figure 9 plots the UMEs for each of the 70 complexes, grouped according to the cluster analysis and the group numbers classified in Table 1. We kept the same vertical UME scale as that in Figure 3 to facilitate comparison. Again, the UME is shown to be generally lower than 0.3 Å for most of the complexes.

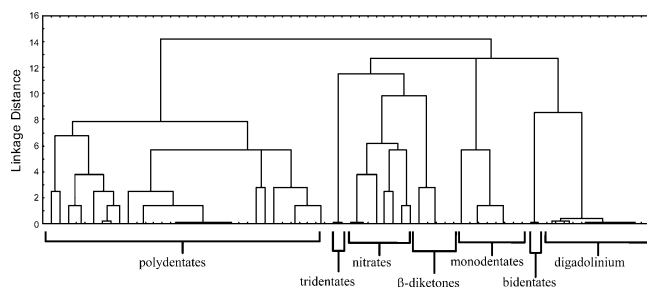
When Figure 9 is looked at more closely, it is noticeable that group 7, where the digadolinium structures are assembled, possesses two structures with UMEs above 0.4 Å, coded COSTAF and LASZIO by the CSD. In COSTAF, this increase is mainly due to difficulties in the description of the distance between two of the gadolinium atoms, yielding an error of 0.64 Å. In the case of LASZIO, the large error is caused by the presence of an atypical N–Gd(III) coordinating distance of 2.916 Å. Because this distance is outside



**Table 3.** UMEs for All Distances Involving the Central Eu(III) Ion and the Ligand Atoms of the Coordination Polyhedron and for All Bond and Dihedral Angles Involving Only the Ligand Atoms of the First Coordination Polyhedron and the Central Eu(III) Ion, for the Seven Complexes of Figure 5, for Some Quantum Chemical Models<sup>a</sup>

	Sparkle/ AM1	RHF/STO-3G/ ECP <sup>b</sup>	RHF/3-21G/ ECP <sup>b</sup>	RHF/6-31G*/ ECP <sup>b</sup>
Distances				
Eu–Eu	0.2662	1.2086	0.1500	0.1611
Eu–O	0.0480	0.0547	0.0378	0.0750
Eu–N	0.0499	0.0354	0.0642	0.1260
Eu–L	0.0520	0.0639	0.0505	0.0976
L–L'	0.1252	0.1425	0.1315	0.1785
Eu–L and L–L'	0.1092	0.1254	0.1138	0.1609
Angles				
O–Eu–O	6.94	7.35	7.23	6.93
N–Eu–O	2.66	3.27	2.35	3.14
N–Eu–N	2.23	2.16	1.56	2.21
all bond angles	3.15	3.60	2.96	2.99
dihedral angles <sup>c</sup>	8.73	11.62	8.42	8.71

<sup>a</sup> L and L' stand for either O or N as ligand atoms. <sup>b</sup> ECP refers to the quasi-relativistic effective core potential of Dolg et al.<sup>23</sup> <sup>c</sup> The Eu(III) ion is always the first atom for all dihedral angles considered.

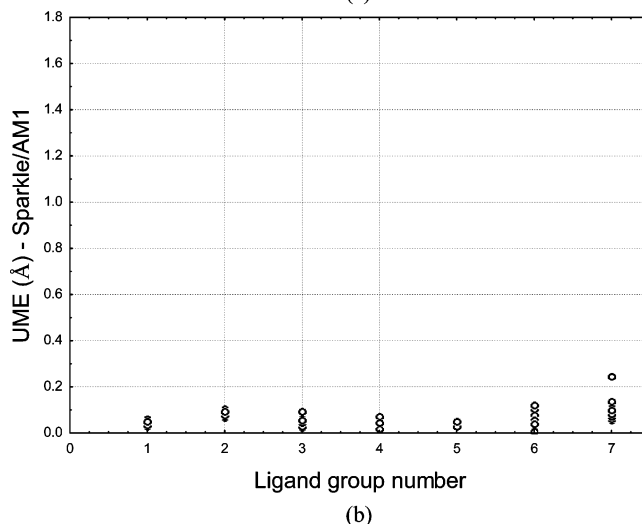
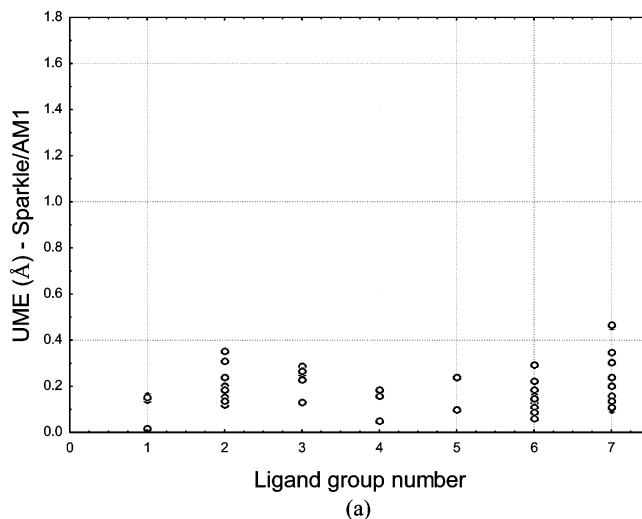
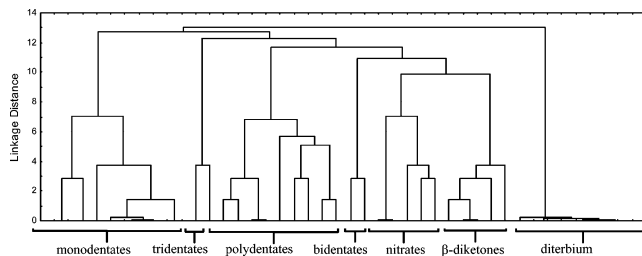
**Figure 8.** Cluster analysis of 70 Gd(III) complexes, in terms of both the UMEs and the number of atoms directly coordinated to the lanthanide ion, for each of the various types of ligands. The UMEs are calculated as the sum of all of the absolute values of the differences between the experimental and calculated interatomic distances between all of the atoms of the coordination polyhedron and between all of these and the central Gd(III) ion.

the range 2.2–2.7 Å, which Gd(III) Sparkle/AM1 reproduces well, the nitrogen is expelled, leading to an absolute error of 1.879 Å for each of the two Gd(III)–N distances in LASZIO.

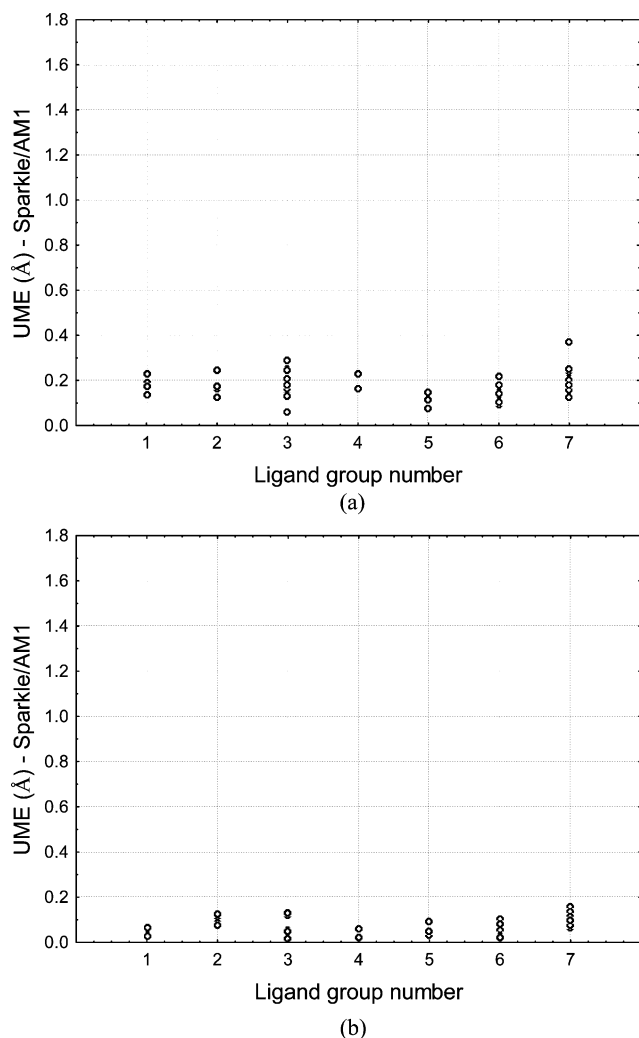
### Sparkle/AM1 Parameters for Tb(III)

Once more, as europium provided the initial guess for the Gd(III) parameters, the Gd(III) parameters presented in Table 2 were used as an initial guess for the eight Tb(III) Sparkle parameters. The same procedure was followed, computing all 42 CSD-available Tb(III) structures with oxygen or nitrogen atoms in the coordination polyhedron. Figure 10 shows the dendrogram obtained from the subsequent cluster analysis that was performed, from which the number of complexes in each of the seven groups can be detected, and these complexes are presented in Table 1. The dendrogram indicated seven different groups, as was the case for Eu(III) and Gd(III).

Fifteen structures were chosen to constitute the Tb(III) parameterization set by picking two from the  $\beta$ -diketone group, two from the nitrate group, one from the monodentate group, one from the bidentate group, three from the tridentate group, four from the polydentate group, and two from the diterbium group.

**Figure 9.** UMEs for each of the 70 Gd(III) complexes. The UMEs are grouped according to the cluster analysis shown in Figure 8, and the group numbers are defined in Table 1. Part a presents these UMEs from Sparkle/AM1, calculated as the sum of all of the absolute values of the differences between the experimental and calculated interatomic distances between all of the atoms of the coordination polyhedron and between all of these and the central Gd(III) ion. Part b also presents these UMEs using Sparkle/AM1 but, this time, considering only the interatomic distances involving the Gd(III) ion. The vertical scale is the same as that in Figure 3, to facilitate comparison.**Figure 10.** Cluster analysis of 42 Tb(III) complexes, in terms of both the UMEs and the number of atoms directly coordinated to the lanthanide ion, for each of the various types of ligands. The UMEs are calculated as the sum of all of the absolute values of the differences between the experimental and calculated interatomic distances between all of the atoms of the coordination polyhedron and between all of these and the central Tb(III) ion.

The optimized Tb(III) Sparkle/AM1 parameters obtained from the nonlinear minimization of the response function are presented in Table 2. Figure 11 plots the UMEs for each



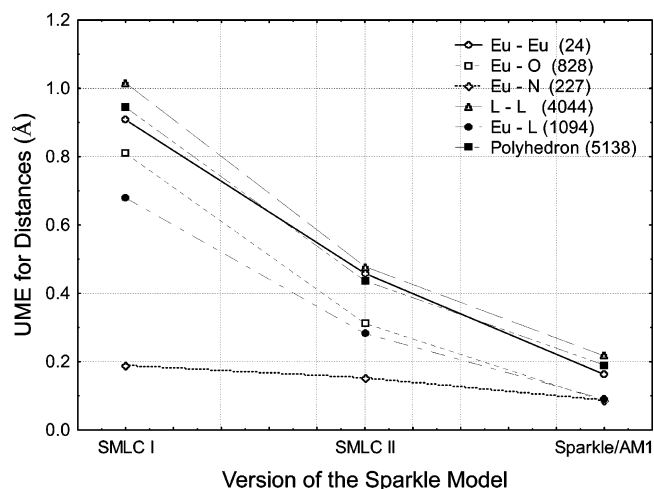
**Figure 11.** UMEs for each of the 42 Tb(III) complexes. The UMEs are grouped according to the cluster analysis shown in Figure 10, and the group numbers are defined in Table 1. Part a presents these UMEs from Sparkle/AM1, calculated as the sum of all of the absolute values of the differences between the experimental and calculated interatomic distances between all of the atoms of the coordination polyhedron and between all of these and the central Tb(III) ion. Part b also presents these UMEs using Sparkle/AM1 but, this time, considering only the interatomic distances involving the Tb(III) ion. The vertical scale is the same as that in Figure 3, to facilitate comparison.

of the 42 complexes, grouped according to the cluster analysis and the group numbers classified in Table 1. We kept the same vertical UME scale as that in Figures 3 and 9 to facilitate comparison. It is again observed that the UME is generally lower than 0.3 Å for most of the complexes.

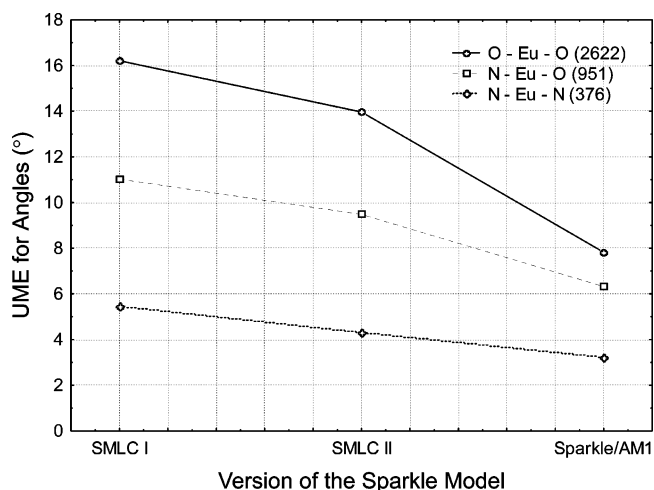
In Group 7, where the dimeric structures are collected, one structure with a UME slightly under 0.4 Å, coded COSVOV in the CSD, appears. However, no special features account for this behavior in COSVOV, other than the interatomic distance for Tb(III)–Tb(III), which is larger by 0.50 Å.

## Discussion

Figure 12 shows the improvement in accuracy in UMEs from the first 1994 Sparkle model, SMLC I,<sup>19</sup> to SMLC II<sup>21</sup> and, finally, to the present Sparkle/AM1 parameterization for Eu(III) complexes, which is the only lanthanide ion



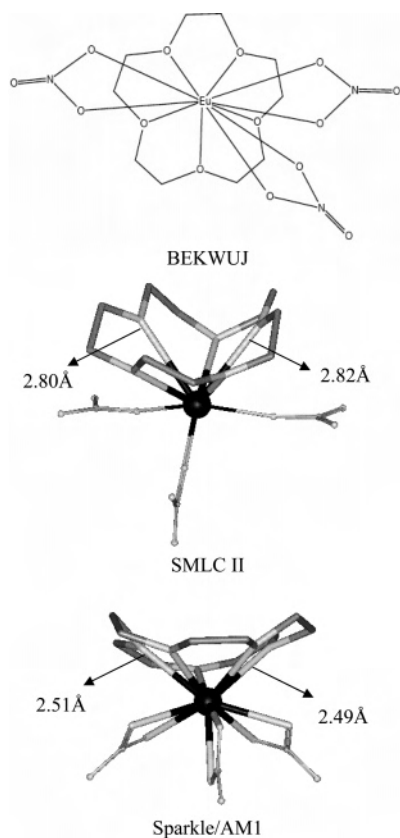
**Figure 12.** Evolution of the accuracy of the various versions of the Sparkle model, from SMLC I<sup>19</sup> to SMLC II<sup>21</sup> to the present Sparkle/AM1, in terms of various types of interatomic distance UMEs (as indicated in the labels) for all 96 complexes involving the Eu(III) ion. The UMEs are calculated, for each type of interatomic distance, as the sum of the absolute values of the differences between the experimental and calculated interatomic distances, added up for all complexes. L stands for a ligand atom coordinated to the Eu(III) ion, and polyhedron stands for all interatomic distances within the coordination polyhedron of the complex. The number of different interatomic distances from all 96 complexes considered, in each case, appears in parentheses.



**Figure 13.** Evolution of the accuracy of the various versions of the Sparkle model, from SMLC I<sup>19</sup> to SMLC II<sup>21</sup> to the present Sparkle/AM1, in terms of various types of bond angle UMEs (as indicated in the labels) for all 96 complexes involving the Eu(III) ion. The UMEs are calculated, for each type of bond angle, as the sum of the absolute values of the differences between the experimental and calculated bond angles, added up for all complexes. The number of different bond angles from all 96 complexes considered, in each case, appears in parentheses.

available in all three versions. Figure 12 shows that the improvement in the prediction of interatomic distances, involving the europium(III) ion, have dramatically improved, with the exception of the Eu(III)–N distance, which was already quite good in SMLC I. Figure 13 indicates that there was also an improvement in the prediction accuracy of the angles, with a marked improvement in the present Sparkle/AM1 version in the case of the O–Eu(III)–O angles.

Moreover, the improvement in accuracy also implies improvement in the qualitative description of bonds in the complexes. Take, for example, the case of the structures with



**Figure 14.** Example of qualitative improvement from SMLC II to the present Sparkle/AM1 model. In SMLC II,<sup>21</sup> middle figure, nitrates are coordinated to the Eu(III) ion through a single oxygen atom in a monodentate manner. In the present Sparkle/AM1, bottom figure, this situation is corrected and nitrates are coordinated in the correct bidentate manner. Bond lengths of the Eu(III) ion and oxygen atoms of the macrocyclic ligand are also indicated as another example of a quantitative improvement achieved going from SMLC II to Sparkle/AM1.

nitrate ligands totaling 29 complexes for all three lanthanide ions considered. In SMLC II, nitrates were predicted to coordinate to the Eu(III) ion in a monodentate manner, causing a large error in the predicted geometry for this class of ligands.

Observe Figure 14(top), where the SMLC II result for the BEKWUJ complex is shown. In the present version, nitrates coordinate in the correct manner, that is, as bidentates, reducing the associated UMEs, as shown in Figure 14 (bottom). In addition, BEKWUJ demonstrates the evolution of the Sparkle model since 1994. The length of the Eu(III)–O bond of the macrocyclic ligand was so overestimated in the first Sparkle model that the oxygen atoms could hardly be regarded as coordinated to the Eu(III) ion (distances above 3.0 Å). In SMLC II, the Eu(III)–O distances are still somewhat overestimated by 0.2 Å, being 2.8 Å instead of the correct experimental value of 2.6 Å. In the present Sparkle/AM1 version, the distances are slightly underestimated by 0.1 Å (Figure 14, bottom).

Table 4 summarizes the accuracies of the three parameterizations, displaying the UMEs for all distances involving the central lanthanide ion, Ln(III), and the ligand atoms of the coordination polyhedron, L, for all complexes considered in the present paper. The numbers indicate that the three models for Eu(III), Gd(III), and Tb(III) possess the same

**Table 4.** Sparkle/AM1 UMEs for All Distances Involving the Central Lanthanide Ion, Ln(III), and the Ligand Atoms of the Coordination Polyhedron, L, for All 96 Eu(III) Complexes, All 70 Gd(III) Complexes, and All 42 Tb(III) Complexes Considered

model	UME (Å)					
	Ln–Ln	Ln–O	Ln–N	L–L'	L–Ln	L–L'
Sparkle/AM1 Eu(III)	0.1624	0.0848	0.0880	0.2170	0.0900	0.1900
Sparkle/AM1 Gd(III)	0.1830	0.0600	0.0735	0.2082	0.0658	0.1781
Sparkle/AM1 Tb(III)	0.2251	0.0754	0.0440	0.2123	0.0746	0.1823

accuracies and that the parameterizations are equalized. The accuracies of the most important variables for the calculation of the ligand field parameters,  $B_q^k$ , are the interatomic distances between the ligand atoms and the central lanthanide ion. Indeed, as already mentioned above, within the simple overlap model,<sup>15,16</sup> the dependence of the  $B_q^k$  goes with the third, fifth, and seventh powers of these distances, thereby amplifying any inaccuracies present. The Sparkle/AM1 UMEs for these distances are the most accurate ones, being 0.09 Å for Eu(III) and 0.07 Å for both Gd(III) and Tb(III).

## Conclusion

An implication of the small errors in prediction of the environment of the lanthanides, for the absolute vast majority of Eu(III), Gd(III), and Tb(III) complexes where the directly coordinating atoms are either nitrogen or oxygen, is that the Sparkle model is valid. That is, the lanthanide trications behave like simple ions, without any angular steric properties. This behavior can be attributed to the fact that the outermost shells of the lanthanides are completely filled and, therefore, have spherical symmetry. The open shell, the 4f shell, would have angular asymmetry when the ion is involved in chemical bonding. Because this set of orbitals is shielded from the ligands by the closed shells, the conjecture has been made that the angular effects of the f orbitals are negligible. The good agreement between the predicted and observed environments of the ions is strong evidence for the validity of this idea.

The Sparkle model assumes complete ionization; that is, it assumes that the lanthanides exist as the simple, completely ionized 3+ ion. This assumption follows from the absence of a basis set on the metal. A consequence of this assumption is that covalent effects involving the ion are not modeled. Because the geometry of the ion is predicted with good accuracy, it is likely that covalent effects would be accounted for by adding to the model a simple basis set, possibly as simple as one “s” and three “p” orbitals.

Nothing has been said about the heats of formation because of the paucity of accurate thermochemical reference data on lanthanide complexes.

By moving from SMLC II<sup>21</sup> to the present parameterization, we obtained a generalization of the Sparkle model, both in terms of the different ligands that can now be treated reliably and in terms of lanthanide ions. By being able to establish a parameterization level of exactness for the UME for distances of 0.3 Å for the whole coordination polyhedron, including the central europium(III) ion, for at least 80% of

the complexes available in the CSD, we created a new paradigm named Sparkle/AM1 and used it to coin new parameters for Eu(III), Gd(III), and Tb(III).

Although Sparkle/AM1 was not parameterized for complexes with other coordinating atoms, such as carbon, sulfur, and so forth, calculations on such complexes are still possible, although larger deviations are expected. Hence, these should be regarded as either semiquantitative or constituting a first educated guess from which to start higher-level *ab initio* geometry optimizations.

The results suggest that Sparkle/AM1 seems to possess coordination polyhedron geometry prediction accuracies for lanthanide complexes with oxygen or nitrogen atoms in the coordination polyhedron that are competitive with present day *ab initio*/ECP calculations while being hundreds of times faster. An efficient method for larger complexes would be to optimize the geometry using the model described here, followed by a single-point high-level calculation to generate other properties.

Parameterizations for all other lanthanide ions using Sparkle/AM1 are presently being carried out.

**Acknowledgment.** We appreciate the financial support from CNPq and CAPES (Brazilian agencies) and also grants from the Instituto do Milênio de Materiais Complexos,

FACEPE (Programa Primeiros Projetos), and Construção do Conhecimento por Agrupamento de Dados (CoCADa). We also thank CENAPAD (Centro Nacional de Processamento de Alto Desempenho) at Campinas, Brazil, for having made available to us their computational facilities. Finally, we gratefully acknowledge the Cambridge Crystallographic Data Centre for the Cambridge Structural Database.

**Supporting Information Available:** Instructions and examples on how to implement the Sparkle/AM1 model in Mopac93r2. Parts of the codes of subroutines Block.f, Calpar.f, and Rotate.f that need to be changed, as well as their modified versions for Eu(III), Gd(III), and Tb(III). Examples of Mopac93r2 crystallographic geometry input (.dat) and optimized geometry summary output (.arc) files from Sparkle/AM1 calculations for (i) the Eu(III) complexes BAPXAR and DOPCEQ, (ii) the Gd(III) complexes PADEGA10 and CULNIG10, and (iii) the Tb(III) complexes CULSEH and ZUNCEQ. Additional figures showing schematic two-dimensional representations of the three sets of 15 complexes that constituted the parameterization training sets for Eu(III), Gd(III), and Tb(III). Additional tables with Sparkle/AM1 UMEs for all high-quality CSD<sup>60–62</sup> crystallographic structures considered: 96 with Eu(III), 70 with Gd(III), and 42 with Tb(III). This material is available free of charge via the Internet at <http://pubs.acs.org>.

IC048530+