Semiempirical Quantum Chemistry Model for the Lanthanides: RM1 (Recife Model 1) Parameters for Dysprosium, Holmium and Erbium

Manoel A. M. Filho¹, José Diogo L. Dutra¹, Gerd B. Rocha², Alfredo M. Simas³, Ricardo O. Freire¹*

¹ Pople Computational Chemistry Laboratory, Departamento de Química, Universidade Federal de Sergipe, São Cristóvão, SE, Brazil; ² Departamento de Química, CCEN, Universidade Federal da Paraíba, João Pessoa, PB, Brazil; ³ Departamento de Química Fundamental, Universidade Federal de Pernambuco, Recife, PE, Brazil

Abstract

Complexes of dysprosium, holmium, and erbium find many applications as single-molecule magnets, as contrast agents for magnetic resonance imaging, as anti-cancer agents, in optical telecommunications, etc. Therefore, the development of tools that can be proven helpful to complex design is presently an active area of research. In this article, we advance a major improvement to the semiempirical description of lanthanide complexes: the Recife Model 1, RM1, model for the lanthanides, parameterized for the trications of Dy, Ho, and Er. By representing such lanthanide in the RM1 calculation as a three-electron atom with a set of 5 d, 6 s, and 6 p semiempirical orbitals, the accuracy of the previous sparkle models, mainly concentrated on lanthanide-carbon and lanthanide-nitrogen distances, is extended to other types of bonds in the trications' coordination polyhedra, such as lanthanide-carbon, lanthanide-chlorine, etc. This is even more important as, for example, lanthanide-carbon atom distances in the coordination polyhedra of the complexes comprise about 30% of all distances for all complexes of Dy, Ho, and Er considered. Our results indicate that the average unsigned mean error for the lanthanide-carbon distances dropped from an average of 0.30 Å, for the sparkle models, to 0.04 Å for the RM1 model for the lanthanides; for a total of 509 such distances for the set of all Dy, Ho, and Er complexes considered. A similar behavior took place for the other distances as well, such as lanthanide-chlorine, lanthanide-bromine, lanthanide, phosphorus and lanthanide-sulfur. Thus, the RM1 model for the lanthanides, being advanced in this article, broadens the range of application of semiempirical models to lanthanide complexes by including comprehensively many other types of bonds not adequately described by the previous models.

Introduction

Lanthanide complexes, as is well known, have a wide range of high technology applications. Of particular importance is the discovery that, due to their slow magnetization relaxation, lanthanide mononuclear complexes may function as single-molecule magnets [1,2], the ultimate size limit for spin-based devices. Dysprosium complexes, in particular, will be very important in the development of magnetic materials because of recent results leading to the highest relaxation energy barriers for multinuclear clusters [3,4], the highest temperature at which hysteresis has been observed for any single complex [5], and a record magnetic blocking temperature of 8.3 K at a sweep rate of 0.08 Ts-1 [6]. Future research, for example, might be directed towards the design of dysprosium complexes that may operate as single-molecule magnets capable of preserving their magnetization at higher and more practical temperatures [6]. Dysprosium complexes are therefore promising for optical storage and memory.

Not only that, both dysprosium and holmium complexes can also effectively function in magnetic resonance imaging, MRI, as negative contrast agents at high magnetic fields, producing darker images, and as agents for susceptibility-induced enhancement at low magnetic fields [7]. Indeed, they are complementary to gadolinium complexes, which act as positive contrast agents, which brighten the image. Indeed, through the simultaneous applications of gadolinium and dysprosium based contrast agents to the MRI diagnosis of conditions such as ischemic heart disease, unprecedented details can now be revealed [8,9]. Future efforts will likely be intensified towards the design of such MRI contrast agents for the imaging of cellular molecular events involved in normal and pathological processes, including site specific macromolecular and particulate delivery systems [7].

Holmium is also employed in cancer therapeutics due to the characteristics of its ¹⁶⁶Ho isotope and of its complexes, like ¹⁶⁶Ho-DOTMP which has been used in combination with chemotherapy in the treatment of myeloma because it concentrates in metastases of the skeleton and irradiates bone marrow [10].

Erbium (III) luminesces at 1.55 µm, essentially at the center of the third telecommunication window at 1.540 nm. Hence, erbium has been used in long-distance optical transmissions, power amplifiers, repeaters, etc. However, inorganic materials doped
with erbium, display a very narrow full width at half maximum [11]. In order to increase the band width, erbium complexes have been used in order to both protect the erbium ion from vibrational
coupling, at the same time enhancing the absorption of light
through the so-called antenna effect. Indeed, erbium complexes
have been prepared that exhibit a much broader full width at half
maximum of 68 nm [12], a significant broadening when
compared to erbium implanted silica which has a typical value
of 11 nm for its most intense peak.

Thus, the design of lanthanide complexes towards enhancement
of the property of interest, while seeking to avoid eventual side
effects to the health of the subject (where applicable) is an active
area of research, which may largely benefit from quantum
chemical tools that attempt to predict several of the physical,
chemical and even pharmacological [13] properties of the
conjectural new structures being considered; structures which
might be sketched by assembling around the lanthanide ion,
ligands, selected from a library of ligands in a combinatorial
manner. And the most important information, from which

essentially all quantum chemical property predictions derive, is
an accurate geometry of the molecular structure of the complex.

Predictions of geometries of lanthanide complexes from ab initio
calculations are not so easy due not only to significant relativistic
effects, a consequence of their high atomic numbers, but also to
the complex manifold of microstates due, not only to a partially
filled f-shell, but also from possibly partially filled 5 d 6 s and 6 p
shells [14]. Therefore, full geometry optimizations from such first
principles calculations are essentially unfeasible for the technologically
useful complexes, which usually exhibit sizes of the order of
100 atoms or more. As a consequence, effective core potentials
arise as a practical and very efficient manner of circumventing the
complexity, while retaining important characteristics of ab initio
calculations. Of these, the most widely used are the relativistic
pseudopotentials of Dolg [15,16] which represent an excellent
compromise between accuracy and usage of computational
resources, mainly computing time. So far, the most thorough
study of the geometry prediction accuracy of these relativistic
potentials has been carried out by our research group in 2006
when full geometry optimizations were carried out on 52 different
lanthanide complexes, including complexes of dysprosium (III),
holmium (III) and erbium (III). [17] The counterintuitive results
obtained indicated that the best combination of method with basis
set when using the MVB pseudopotential was RHF/STO-3G
when the intent of the calculation was to predict the geometry of
the coordination polyhedron – very important for any subsequent
ligand field application. Moreover, either increasing the basis set,
or adding the quality of the obtained coordination polyhedron via RHF/
STO-3G was very good, that could not be said of the geometry of the
attached organic ligands.

In 1994, we introduced the Sparkle Model for the calculation of
lanthanide complexes [18,19], a semiempirical model within the
framework of the AM1 semiempirical model [20], which replaced
the lanthanide ion by a 4+e charge, with the corresponding
Coulomb field superimposed to a repulsive potential of the form
exp(−r), with being a parameter designed to somewhat delineate
the size of the lanthanide ion, preventing the implosion of the
ligands towards it. A very useful method of obtaining absorption
spectra of lanthanide complexes was subsequently published [21].
Later [22], Gaussians were added to the core-core repulsion of
the sparkle-ligand atom to make the Sparkle Model more consistent
with the AM1. In 2005, based on a parameterization scheme
employed for europium, gadolinium and terbium [23], the first
useful and accurate semiempirical model for dysprosium was
defined [24], followed by holmium [25]; and in 2006 for erbium
[26]. These models were defined for AM1, and became later
available in MOPAC2007 [27], the overall model being called
Sparkle/AM1. So far, most applications of the Sparkle Model are
related to luminescence research [28–30]. But since different
semiempirical models possess different accuracies and eventually
develop particular niches of applications, it soon became a
necessity to extend the Sparkle Model to others, giving rise to
Sparkle/PM3 [31,32], Sparkle/PM6 [33], Sparkle/PM7 [34],
targeted to solids, and Sparkle/RM1 [35].

However, none of the above mentioned Sparkle Models
attaches semiempirical atomic orbitals to the lanthanide ion.
Nevertheless, these models are all very accurate to describe
lanthanide-ligand atom distances when the coordinating atom of
the ligands is another lanthanide, oxygen or nitrogen. By moving
towards other types of lanthanide-ligand atom bonds, however, the
accuracy of the Sparkle Models starts to wane. All that points out
to the fact that there is some degree of overlap between the orbitals

Table 1. Parameters* for the RM1 model for the trications of Dy, Ho and Er.

| Parameter | Unit | Dy3⁺ | Ho3⁺ | Er3⁺ |
|-----------|------|-------|-------|-------|
| U₀       | eV   | −20.92623973 | −22.05745867 | −21.97839904 |
| U₂p     | eV   | −7.66730575 | −7.59563761 | −7.60784986 |
| U₂d     | eV   | −17.94081525 | −18.00040589 | −17.97684107 |
| ξ₁,b     | bohr | 1.29527540 | 1.33055043 | 1.34775672 |
| ξ₁,p     | bohr | 1.91210659 | 1.77955393 | 1.80648084 |
| ξ₂,b     | bohr | 1.41339670 | 1.53652417 | 1.46618905 |
| ξ₂,p     | bohr | 1.13524616 | 1.05957341 | 1.08839064 |
| β₁       | eV   | −7.80670536 | −5.84522644 | −5.83471034 |
| β₂       | eV   | 1.96173362 | 0.00653676 | −0.01897203 |
| β₃       | eV   | −4.36852734 | −4.31289917 | −4.25067889 |
| β₄        | eV   | 8.30543139 | 8.24056943 | 8.25732681 |
| G²w       | eV   | 1.31036509 | 1.24543189 | 1.28474510 |
| r₀        | bohr | 1.62505051 | 1.71955962 | 2.71713627 |
| r²        | Å    | 1.34825876 | 1.33007543 | 1.32010273 |
| ξ₁,b     | bohr | 1.37236617 | 1.49038444 | 1.44675714 |
| ξ₁,p     | bohr | 1.07407253 | 1.96749739 | 1.97388315 |
| ξ₂,b     | bohr | 0.81914360 | 0.66302146 | 0.65046083 |
| σ₁      | none | 1.13071544 | 1.09070576 | 1.17417665 |
| σ₁      | Å ²  | 7.71195583 | 7.57151625 | 7.58325164 |
| σ₁      | Å ²  | 1.53668519 | 1.49095411 | 1.50354881 |
| δ₂      | none | 0.06845575 | 0.00164914 | 0.00864571 |
| δ₂      | Å ²  | 7.50653990 | 7.79969366 | 7.81378785 |
| c₂      | Å ²  | 3.23417102 | 3.25425084 | 3.23359665 |

All these parameters are as defined in the formalisms and equations of the RM1 model.

*Parameters are s, p, and d atomic orbital one-electron-one-center integrals Uᵢ, Uᵢp and Uᵢd the s, p and d Slater atomic exponents 〈ξ₁,s 〈ξ₁,p〉 and 〈ξ₂,s〉 and 〈ξ₂,p〉 the s, p and d atomic orbital one-electron-two-center resonance integral terms β₁, β₂, β₃ and β₄ the core-core repulsion term σ; the two-electron integrals Pᵢ,j,a,b the additive term r₀; the six parameters for the two Gaussian functions: height, a; inverse broadness, b; and displacement, c.

doi:10.1371/journal.pone.0006376.t001
of the lanthanide and those of the coordinating atoms – in short, there is a degree of covalence not taken into account by the Sparkle Model.

In this article, in order to considerably broaden the range of applications of semiempirical methods for lanthanide complexes, we introduce a new model with orbitals for the lanthanide trications of dysprosium, holmium, and erbium, within RM1 [36], which we call simply RM1 model for the lanthanides, a significantly more general model, not to be confused with Sparkle/RM1 [35] which does not have orbitals associated with the lanthanide ion.

| Complex* | RM1 | Complex* | RM1 |
|----------|-----|----------|-----|
|          | UME_{Dy-L} (Å) | UME (Å) | UME_{Dy-L} (Å) | UME (Å) |
| AMAQDY   | 0.0517 | 0.1503 | AKUKAT | 0.0216 | 0.1148 |
| BAFZUE   | 0.0470 | 0.1129 | BEXLIA | 0.0300 | 0.0972 |
| BIHLIN   | 0.0317 | 0.0979 | DANCEN | 0.0410 | 0.0953 |
| BUVXZO1  | 0.0658 | 0.0872 | DEKBEB | 0.0438 | 0.0832 |
| CECLIF   | 0.0522 | 0.0959 | DEKBEB01 | 0.0370 | 0.0843 |
| CECLIF10 | 0.0521 | 0.0960 | DEKCAY01 | 0.0632 | 0.1068 |
| DIBTIR   | 0.0313 | 0.0977 | KJFCOW10 | 0.0348 | 0.1042 |
| DIDBOH   | 0.0417 | 0.1567 | FIGXEZ | 0.0327 | 0.1403 |
| FOPNAZ   | 0.0138 | 0.1471 | FUXPAP01 | 0.1191 | 0.1026 |
| FUXARR   | 0.0401 | 0.1662 | HIWVEP | 0.0665 | 0.1222 |
| GAKYEW   | 0.0258 | 0.1273 | HOYCUU | 0.0942 | 0.0976 |
| GINPJO   | 0.0455 | 0.0813 | IMOXAJ | 0.0271 | 0.1036 |
| HANCEA   | 0.0553 | 0.1424 | KJZOUO2 | 0.0487 | 0.1307 |
| KITGEZ   | 0.0480 | 0.1391 | KUYBIP | 0.1338 | 0.1260 |
| LEYHUS   | 0.0296 | 0.2200 | LEZZOZ | 0.0343 | 0.0718 |
| MANHOY   | 0.0804 | 0.1208 | MECCUT | 0.0408 | 0.1836 |
| PALBIN   | 0.0205 | 0.4919 | NAMKDO | 0.0323 | 0.0713 |
| PQQDEM01 | 0.0328 | 0.0841 | NAPHAN | 0.0377 | 0.0883 |
| SETADY   | 0.0883 | 0.2516 | OHUYUM | 0.0286 | 0.0985 |
| TISQHI   | 0.0670 | 0.1240 | RABBBEX | 0.0825 | 0.2034 |
| TUQTOU   | 0.0601 | 0.0940 | ROCIN | 0.1118 | 0.2276 |
| TUQTUO10 | 0.0601 | 0.0943 | TESHEF | 0.0348 | 0.1448 |
| VOSBOU   | 0.0587 | 0.0879 | TESHOP | 0.0571 | 0.1010 |
| XAWVIA   | 0.0408 | 0.1087 | TESJHE | 0.0356 | 0.1150 |
| XEOQMAH  | 0.0967 | 0.1019 | TESJIL | 0.0564 | 0.1008 |
| XINFUD   | 0.1015 | 0.2004 | USEPEO | 0.0484 | 0.1090 |
| YAVSOD   | 0.0432 | 0.0726 | WAQZEU | 0.1499 | 0.1742 |
| ZAXSAS   | 0.0700 | 0.1357 | WAWOJTO1 | 0.1152 | 0.1538 |
| ZZZARG01 | 0.0602 | 0.0897 | WEDHUJ | 0.0282 | 0.1089 |
| AGUTOL   | 0.0719 | 0.1019 | XAYRIZ | 0.0416 | 0.0874 |
| AHANED   | 0.0817 | 0.1433 |          |         |         |

*The complexes are identified by their unique CSD codes [37–39].

doi:10.1371/journal.pone.0086376.t002

### Methods

The rationale of the RM1 model for the lanthanides starts with the following electron configuration for the lanthanide atoms: \( \{[Xe]\}^{n+}5d^16s^2 \) with \( n = 9, 10, 11 \) for Dy, Ho and Er, respectively. The semiempirical core of the atoms then becomes \( \{[Xe]\}^{n+} \). The semiempirical valence shells will now have three electrons and will be described by 5 d, 6 s and 6 p orbitals, for a total of 9 orbitals. Hence the model will work for trications only, because for dications there would be a need to parameterize another core of the form \( \{[Xe]\}^{n+} \) and assign two electrons to the valence shells, although they could still be described by another set of 5 d, 6 s, and 6 p orbitals. Since trications are by far the most common form of lanthanide ions, as before, we expect the present parameterization to be able to tackle essentially all cases relevant to technological applications.

The next step is to define the universes of complexes, one universe for each of the lanthanide ions under consideration. Accordingly, we selected from the Cambridge Crystallographic Database [37–39] all available complexes of high crystallographic quality (R <0.05), for a total of 61 of Dy(III), 40 of Ho(III), and 50 of Er(III).

We then proceeded to select sub-sets of complexes, the parameterization sets, according to some metric capable of guaranteeing that these sub-sets are representative of the universe of complexes with respect to some accuracy measure. Assuming
that any difficulties Sparkle/AM1 might be having in describing the coordination polyhedron of the complexes is a reasonable first order approximation to the eventual overall difficulty which the present model will encounter, we defined the following $R_i$ metric for each one of the $i$ complexes of the universe:

$$R_i = \sum_j \sum_k \frac{1}{a_{j,k}} |d_{j,k}^{CSD} - d_{j,k}^{Calc}| + \sum_j \frac{1}{\sigma_{j,k}} |\theta_{j,k}^{CSD} - \theta_{j,k}^{Calc}| \quad (1)$$

where $j$ runs over all types of bonds, e.g. Ln-N, Ln-O, Ln-C, Ln-S, Ln-P, etc; $k$ runs over all bonds of type $j$; $a_{j,k}$ is the standard deviation of all crystallographic bond lengths of type $j$ for all complexes of the universe; $d_{j,k}^{CSD}$ is the crystallographic $k$th bond distance of type $j$ for complex $i$; $d_{j,k}^{Calc}$ is the calculated value of the same bond; $\sigma_{j,k}$ 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; $\theta_{j,k}^{CSD}$ is the crystallographic $k$th bond angle of complex $i$; and $\theta_{j,k}^{Calc}$ is its calculated counterpart. The standard deviations were calculated from the experimental data only. We also found out that there was no need to split the angles into types, as they all formed a homogenous set. The divisions of the errors by their corresponding standard deviations make sure that the summations in Eq. (1) add comparable terms. To the set of $R_i$ values, each one associated with a different complex, we employed a hierarchical clustering analysis DIANA [40]. DIANA starts out with one large cluster containing all complexes. In the subsequent steps, the complexes that are the most dissimilar are split off into smaller clusters – a procedure which continues until each complex forms a cluster of itself. From the resulting dendogram, we chose two sets of complexes as parameterization sets: a smaller and a larger one. For Dy(III) these sets contained 13 and 26 complexes, respectively. The corresponding numbers for Ho(III) were 12 and 20, and for Er(III) 16, and 39.

The parameterization was carried out to minimize the sum of $R_i$s for all complexes of parameterization set, with the difference that the calculated distances and angles in Eq. (1), are now the ones calculated by the model being parameterized. For the parameterization, we used a combination of Simplex and generalized simulated annealing [41] algorithms. We started with the smaller parameterization sets. Once these preliminary optimizations converged, we then expanded the parameterization set to the larger ones and repeated the process until termination. Table 1 presents the final optimized parameters.

### Results and Discussion

In order to evaluate the quality of the optimized parameters, we devised two measures [23,42]. Both are based on the following formula:

$$UME_i = \frac{1}{n} \sum_{j=1}^{n} |R_i^{CSD} - R_i^{RM1}| \quad (2)$$

where $UME_i$ stands for the unsigned mean error; $i$ refers to a given complex; $n$ is the number of distances taken into consideration in the given complex; the superscript CSD indicates that the distance $R$ is an experimental crystallographic distance taken from CSD, and the superscript RM1 means that the distance was calculated from the present model. In the first measure we consider only

### Table 4. Unsigned mean errors, $UME_{Er-L}$, and UMEs, for Method RM1, as compared to the respective experimental crystallographic values, obtained from the Cambridge Structural Database, [37–39] for each of the 59 erbium (III) complexes.

| Complex | Method RM1 | Complex | Method RM1 |
|---------|------------|---------|------------|
|         | UME_{Er-L} | UME     | UME_{Er-L} | UME         |
|         | ()         | ()      | ()         | ()          |
| ARET502 | 0.0563     | 0.0895  | RIKTEK     | 0.0522      | 0.2208      |
| AKIYEOY | 0.0440     | 0.2700  | RIKTEKO1   | 0.0488      | 0.2197      |
| BAGBER  | 0.0462     | 0.1076  | ROCOSOS    | 0.0782      | 0.2444      |
| BEXLEW  | 0.0842     | 0.1678  | ROCTOT     | 0.0595      | 0.2138      |
| BOBWAQ  | 0.0421     | 0.1184  | RUNQOG     | 0.0513      | 0.1022      |
| BOMJUD  | 0.0388     | 0.1382  | SEGABAB    | 0.0691      | 0.1681      |
| BOWXOA  | 0.0450     | 0.1848  | SOKBID     | 0.0280      | 0.0513      |
| DEKCEC  | 0.0371     | 0.0914  | TACERB01   | 0.0262      | 0.0958      |
| DITJAJ  | 0.0369     | 0.1778  | TEFEU      | 0.1355      | 0.2048      |
| DIDCAU  | 0.0399     | 0.2601  | TEPKOO     | 0.0297      | 0.0715      |
| DUQAOY  | 0.0391     | 0.2321  | TMHDERER   | 0.0858      | 0.3224      |
| DUQIOJ  | 0.0310     | 0.1024  | TUMJEOJ    | 0.0583      | 0.1227      |
| DIYNNI  | 0.0260     | 0.1234  | UFRICK     | 0.0622      | 0.0716      |
| DOGKEP  | 0.0323     | 0.1867  | VEOPOOM    | 0.0382      | 0.0617      |
| GAKYOG  | 0.0266     | 0.1057  | VOSNOG     | 0.0375      | 0.0630      |
| GINRHE  | 0.0646     | 0.0950  | VUGSUL     | 0.0917      | 0.0941      |
| HANCER  | 0.0165     | 0.0605  | VUSHEE     | 0.0593      | 0.1107      |
| HEDVW   | 0.1052     | 0.1989  | WEIVFIM    | 0.0466      | 0.0843      |
| HENAEB  | 0.0283     | 0.1545  | XAXYAX     | 0.0593      | 0.0854      |
| KITJUP  | 0.0149     | 0.0628  | XEWWOKH    | 0.0442      | 0.1163      |
| KOZBUW  | 0.0144     | 0.0561  | XEWWUR     | 0.0472      | 0.0938      |
| LEYIII  | 0.0294     | 0.2276  | XQVIVAS    | 0.0851      | 0.2268      |
| MAGDOS  | 0.0449     | 0.1319  | XQXYXIS    | 0.0285      | 0.1590      |
| MEDGEE  | 0.0281     | 0.0853  | YEGFEEV    | 0.0504      | 0.1687      |
| NIVQUE  | 0.0499     | 0.2188  | YEMSIT     | 0.0767      | 0.1338      |
| NUVNUR  | 0.0838     | 0.2163  | YIKCIIW    | 0.0344      | 0.1527      |
| OHUZER  | 0.0269     | 0.0890  | YUYFWG     | 0.0196      | 0.0988      |
| OMATUS  | 0.0742     | 0.1597  | ZADWHV     | 0.1281      | 0.2265      |
| QIXKID  | 0.0742     | 0.2170  | ZUFSAU     | 0.0480      | 0.1304      |
| RELNIG  | 0.0530     | 0.1011  |             |             |

*The complexes are identified by their unique CSD codes [37–39].

doi:10.1371/journal.pone.0086376.t004

### Table 5. Means and variances for the $\gamma$ distribution fits of the $UME_{Lan-Ln}$ computed for the $N$ complexes for each lanthanide trication.

| Lanthanide | $N$ | mean (Å) | variance (Å$^2$) | $p$-value |
|------------|-----|-----------|-----------------|-----------|
| Dy$^{3+}$  | 61  | 0.0539    | 0.0032          | 0.7986    |
| Ho$^{3+}$  | 40  | 0.0602    | 0.0069          | 0.1292    |
| Er$^{3+}$  | 58  | 0.0506    | 0.0025          | 0.9082    |

doi:10.1371/journal.pone.0086376.t005
distances between the central lanthanide ion and its directly
coordinating atom distances, which we call UME (Ln-L). In the
second measure, which we call simply UME, we consider not only
the lanthanide ion-directly coordinating atoms as before, but also
all distances between all atoms of the coordinating polyhedron,
thus indirectly taking into account the angles within the
coordination polyhedron.

Tables 2–4 present UME(Ln-L) and UMEs for the universe set
of complexes for each of the lanthanide trications: Dy(III), Ho(III),
and Er(III), identified by their respective CSD codes.

We now proceed to the statistical validation of the model [43]. If
the parameterizations captured the essence of the coordinating
bonds, then the histograms of both UME(Ln-L) and UME must
follow gamma distribution functions since, by definition, the
UMEs can only have positive values. The gamma distributions are
then adjusted to reproduce the mean and variance of the UME(Ln-
L), for each of the parameterized trications. Finally, the qualities
of the gamma distribution fits of the data were then assessed via the
one-sample nonparametric Kolmogorov-Smirnov test [44]. If the
p-value of the test is larger than 0.05, then the gamma distribution
fit is justified within a 95% interval and the use of the mean and
variance of the data, as accuracy measures, is also statistically
justified. Accordingly, Tables 5 and 6 display the mean, variance,
and p-value of the test for each of the lanthanide ions, for both the
UME(Ln-L) and UMEs. All p-values are substantially larger than
0.05 and, therefore, the means and variances in Tables 5 and 6
can justifiably be taken as accuracy measures of the models for
Dy(III), Ho(III), and Er(III).

We now proceed to analyze the performance of the models with
respect to specific types of distances for Dy(III), Ho(III), and

Table 6. Means and variances for the \( \gamma \) distribution fits for the
UMEs computed for the \( N \) complexes for each lanthanide
trication.

| Lanthanide ion | N | mean (Å) | variance (Å²) | p-value |
|---------------|---|----------|---------------|---------|
| Dy\(^{3+}\)  | 61 | 0.1193   | 0.0169        | 0.1578  |
| Ho\(^{3+}\)  | 40 | 0.1225   | 0.0290        | 0.1463  |
| Er\(^{3+}\)  | 58 | 0.1348   | 0.0228        | 0.5425  |

doi:10.1371/journal.pone.0086376.t006

Figure 1. UME(Ln-L)\(^{\delta}\) obtained using the RM1 model for the 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 universe set for each of the lanthanide
trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated Ln-
Ln interatomic distances, summed up for all complexes, for each of the lanthanides.
doi:10.1371/journal.pone.0086376.g001

Figure 2. UME(Ln-O)\(^{\delta}\) obtained using the RM1 model for the 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 universe set for each of the lanthanide
trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated Ln-
O interatomic distances, summed up for all complexes, for each of the lanthanides.
doi:10.1371/journal.pone.0086376.g002
Figure 3. UME\textsubscript{(Ln-N)} obtained using the RM1 model for the 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 universe set for each of the lanthanide trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated Ln-N interatomic distances, summed up for all complexes, for each of the lanthanides. doi:10.1371/journal.pone.0086376.g003

Table 7. RM1, Sparkle/RM1, Sparkle/PM7, Sparkle/PM6, Sparkle/PM3, and Sparkle/AM1 unsigned mean errors for different types of distances of dysprosium(III) complexes.

| Type of distances | unsigned mean errors for specific types of distances (Å) |
|-------------------|----------------------------------------------------------|
|                   | N  | RM1   | Sparkle/RM1 | Sparkle/PM7 | Sparkle/PM6 | Sparkle/PM3 | Sparkle/AM1 |
| Dy - Dy           | 16 | 0.2118 | 0.2397      | 0.5950      | 0.5025      | 0.1784      | 0.2531      |
| Dy - O            | 283| 0.0637 | 0.0760      | 0.0692      | 0.1259      | 0.0685      | 0.0740      |
| Dy - N            | 105| 0.0594 | 0.0510      | 0.2084      | 0.0992      | 0.0741      | 0.0500      |
| Dy - C            | 315| 0.0341 | 0.2567      | 0.4616      | 0.2346      | 0.2161      |
| Dy - S            | 21 | 0.0824 | 0.5016      | 0.9833      | 0.5359      | 0.4996      | 0.5019      |
| Dy - P            | 3  | 0.0273 | 0.1591      | 0.2319      | 0.3934      | 0.4205      | 0.3918      |
| Dy - Cl           | 20 | 0.0525 | 0.2858      | 0.1757      | 0.2318      | 0.2475      | 0.2546      |
| Dy - Br           | 5  | 0.0311 | 0.4564      | 1.3418      | 0.3986      | 0.4209      | 0.4320      |
| Dy - L            | 768| 0.0539 | 0.1408      | 0.2209      | 0.2846      | 0.1598      | 0.1527      |
| L-L               | 3499| 0.1336 | 0.2370      | 0.2931      | 0.3449      | 0.2433      | 0.2621      |
| Dy-L, Dy-Dy and L-L | 4267| 0.1193 | 0.2197      | 0.2801      | 0.3341      | 0.2283      | 0.2424      |

doi:10.1371/journal.pone.0086376.t007

Table 8. RM1, Sparkle/RM1, Sparkle/PM7, Sparkle/PM6, Sparkle/PM3, and Sparkle/AM1 unsigned mean errors for specific types of distances for holmium(III) complexes.

| Type of distances | unsigned mean errors for specific types of distances (Å) |
|-------------------|----------------------------------------------------------|
|                   | N  | RM1   | Sparkle/RM1 | Sparkle/PM7 | Sparkle/PM6 | Sparkle/PM3 | Sparkle/AM1 |
| Ho - Ho           | 4  | 0.1301 | 0.2083      | 0.2144      | 0.2751      | 0.1958      | 0.2747      |
| Ho - O            | 219| 0.0475 | 0.0604      | 0.0554      | 0.1021      | 0.0639      | 0.0557      |
| Ho - N            | 58 | 0.0786 | 0.0696      | 0.1592      | 0.1732      | 0.0585      | 0.0469      |
| Ho - C            | 98 | 0.0752 | 0.2256      | 0.4544      | 0.5380      | 0.2537      | 0.2655      |
| Ho - Cl           | 28 | 0.0585 | 0.3055      | 0.1262      | 0.2777      | 0.2480      | 0.2679      |
| Ho - L            | 407| 0.0602 | 0.1198      | 0.1727      | 0.2310      | 0.1228      | 0.1217      |
| L - L'            | 1748| 0.1371 | 0.2326      | 0.3289      | 0.3462      | 0.2324      | 0.2585      |
| Ho-L, Ho-Ho and L-L | 2155| 0.1225 | 0.2113      | 0.2994      | 0.3245      | 0.2117      | 0.2327      |

doi:10.1371/journal.pone.0086376.t008
Table 9. RM1, Sparkle/RM1, Sparkle/PM7, Sparkle/PM6, Sparkle/PM3, and Sparkle/AM1 unsigned mean errors for specific types of distances for erbium complexes.

| Type of distances | unsigned mean errors for specific types of distances (Å) |
|-------------------|---------------------------------------------------------|
|                   | N     | RM1   | Sparkle/RM1 | Sparkle/PM7 | Sparkle/PM6 | Sparkle/PM3 | Sparkle/AM1 |
| Er - Er           | 6     | 0.1200 | 0.2639    | 0.5458    | 0.2124    | 0.2439    | 0.2626    |
| Er - O            | 336   | 0.0509 | 0.0730    | 0.0874    | 0.1285    | 0.0657    | 0.0689    |
| Er - N            | 77    | 0.0594 | 0.0484    | 0.0529    | 0.0846    | 0.0551    | 0.0434    |
| Er - C            | 96    | 0.0318 | 0.2004    | 0.4280    | 0.5474    | 0.2277    | 0.2177    |
| Er - S            | 12    | 0.1088 | 0.4802    | 1.2501    | 0.5234    | 0.5211    | 0.5212    |
| Er - Cl           | 33    | 0.0574 | 0.3243    | 0.3406    | 0.2753    | 0.2975    | 0.2928    |
| Er - Br           | 3     | 0.0463 | 0.4526    | 1.5964    | 0.4475    | 0.4146    | 0.4291    |
| Er - L            | 563   | 0.0511 | 0.1199    | 0.1967    | 0.2136    | 0.1205    | 0.1189    |
| L – L’            | 2259  | 0.1575 | 0.2197    | 0.2938    | 0.3519    | 0.2076    | 0.2371    |
| Er-L, Er-Er and L-L’ | 2822  | 0.1363 | 0.2000    | 0.2746    | 0.3243    | 0.1904    | 0.2137    |

Figure 4. UME\(_{\text{Ln-C}}\) obtained using the RM1 model for the 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 universe set for each of the lanthanide trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated Ln-C interatomic distances, summed up for all complexes, for each of the lanthanides.
doi:10.1371/journal.pone.0086376.g004

Figure 5. UME\(_{\text{Ln-S}}\) obtained using the RM1 model for the 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 universe set for each of the lanthanide trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated Ln-S interatomic distances, summed up for all complexes, for each of the lanthanides. There are no Ho-S distances in the universe of Ho(III) complexes considered.
doi:10.1371/journal.pone.0086376.g005
Er(III), Tables 7–9 and Figures 1–3 show UMEs for all types of Ln-L distances present in the universe of Ln(III) complexes, together with the corresponding values from the previous sparkle models for comparison. It must be noted, though, that in the original sparkle model articles, we only included complexes with exclusively Ln-O and Ln-N bonds. But, in the present article, we are considering a much larger set of complexes with other types of bonds such as Ln-C, Ln-Cl, etc. Indeed, here we now may have complexes which have not only Ln-O and/or Ln-N bonds, but also other types of bonds, such as Ln-C bonds, all in the same compound. Of course, these were not included as test cases for the previous Sparkle models, but are here taken in full consideration.

Figure 6. UME (Ln-Cl)s obtained using the RM1 model for the 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 universe set for each of the lanthanide trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated Ln-Cl interatomic distances, summed up for all complexes, for each of the lanthanides.

doi:10.1371/journal.pone.0086376.g006

Figure 7. UME (Ln-Br)s obtained using the RM1 model for the 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 universe set for each of the lanthanide trications: Dy(III), Ho(III) and Er(III). The RM1 model for lanthanides UME(Ln-Br) bars are very small. Besides, there are no Ho-Br distances in the universe of Ho(III) complexes considered.

doi:10.1371/journal.pone.0086376.g007

Figure 8. UME (L-L)9s obtained using the RM1 model for the 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 universe set for each of the lanthanide trications: Dy(III), Ho(III) and Er(III). The UMEs are calculated as the absolute value of the difference between the experimental and calculated interatomic distances between the coordinated atoms, L-L9, summed up for all complexes, for each of the lanthanides.

doi:10.1371/journal.pone.0086376.g008
And that is the reason why UMEs for the Ln-O and Ln-N types of bonds in the present article tend to be different, slightly larger, when set side by side to similar Ln-O and Ln-N UMEs of the original sparkle model articles. However, not to unnecessarily crowd the present article, in the tables, we only show numbers computed using the old models, but for the new test set.

Dy-Dy distances in dilanthanide complexes of Dy(III), Ho(III), and Er(III) lie in the range from 3.6 Å to 6.6 Å, whereas lanthanide–other ligand atom distances lie on average around 2.5 Å. That is why Ln-Ln UMEs are larger than other Ln-L UMEs. The previous Sparkle Models focused on these Ln-Ln, a also on Ln-O, and Ln-N distances only. Indeed, considering only Dy complexes (Table 7), there are 404 distances of these types, which represent 53% of all distances involving Dy(III) in its universe of complexes. The next most important types are Dy(III)-C distances, for which there are 315 of them making up 41% of the total.

By examining Table 7 and Figure 4, we can see a significant improvement in these next types of distances, with UME (Dy-C) for the RM1 model for the lanthanides being 0.03 Å, whereas the corresponding average value of the sparkle models is 0.27 Å, a value 9 times larger. That alone would justify the introduction of the RM1 model for the lanthanides because, in the case of dysprosium, almost half the extant Ln-L distances are significantly more accurately described by RM1.

The situation is less dramatic but still significant for the other trications being parameterized, when Ln-C distances represent 24% of the total for Ho(III), and 17% for Er(III). The RM1 model for the lanthanides is even further justified when we compare other types of less common distances like Ln-S, Ln-Cl, and Ln-Br, because it outperforms all previous sparkle models as shown in Tables 7–9 and Figures 5–7. In all these instances, the RM1 UMEs tend to be almost ten times smaller than the corresponding errors of all previous sparkle models.

Finally, we can have an idea of the accuracy of the angles by examining the distances between any two atoms of the coordination polyhedron, the L-L distances. For all three lanthanide trications, there was a significant reduction of these UMEs by a factor of two when compared to the corresponding UMEs for the previous sparkle models: from 0.26 Å to 0.13 Å, as can be inferred from Figure 8. This is indirect evidence that the angles are much better predicted in the RM1 model for the lanthanides.

Conclusions

The RM1 model for the lanthanides represents a significant improvement in the theoretical semiempirical modeling of lanthanide complexes, which started with the sparkle model in 1994 [18,19] which and attained maturity with the introduction of Sparkle/AM1 in 2005 [23], and was extended to Sparkle/PM3 [31,32], Sparkle/PM6 [33], to Sparkle/PM7 [34] and up to Sparkle/RM1 [35], the last two in 2013.

There is, however, a cost associated with the improvement. The RM1 model for the lanthanides adds nine more orbitals per lanthanide to the calculation, whereas all sparkle models add none. Thus, for a single SCF calculation of a complex of about 60 atoms, for example, an RM1 model for the lanthanides calculation takes about twice the computing time of a Sparkle/RM1 calculation. Such cost can become even weightier if the molecular structure contains more than one lanthanide ion, as is usually the case of metal-organic frameworks. Since the performance of both Sparkle/RM1 and RM1 model for the lanthanides is essentially equivalent for Ln-Ln, Ln-O, and Ln-N, the user may still benefit from the speed of the sparkle models if the structure of interest contains only these types of bonds, as it takes place in the majority of cases.

In conclusion, via the introduction of a set of 5 d, 6 s, and 6 p semiempirical atomic orbitals, the RM1 model for the lanthanides thus extends the Sparkle Models’ capabilities of correctly describing Ln-Ln, Ln-O, and Ln-N distances, to other types of distances, such as Ln-C, Ln-S, Ln-P, Ln-Cl, and Ln-Br, while simultaneously improving the coordinating bond angles.

Supporting Information

File S1 Instructions on how to run the RM1 model for the lanthanides in MOPAC12 [45], and MOPAC sample input and output files for complexes of Dy(III), Ho(III), and Er(III).

Acknowledgments

We gratefully acknowledge the Cambridge Crystallographic Data Centre for the Cambridge Structural Database.

Author Contributions

Conceived and designed the experiments: ROF GBR AMS. Performed the experiments: MAMF JDLD ROF. Analyzed the data: MAMF JDLD ROF. Wrote the paper: AMS ROF. Coded the model into MOPAC: GBR. Conceived the parameterization techniques: AMS GBR ROF.

Conceived the sampling of the reference structures and the statistical validation of the model: AMS.

References

1. Ishikawa N, Sugita M, Ishikawa T, Koshihara S, Kaizu Y (2003) Lanthanide Double-Decker Complexes Functioning as Magnets at the Single-Molecular Level. Journal of the American Chemical Society 125: 8694–8695.
2. Zhang P, Guo Y-N, Tang J (2013) Recent advances in dysprosium-based single molecule magnets: Structural overview and synthetic strategies. Coordination Chemistry Reviews 257: 1728–1763.
3. Ishikawa N, Minoura Y, Takamatsu S, Ishikawa T, Koshihara S-Y (2008) Effects of chemically induced contraction of a coordination polyhedron on the dynamical magnetism big/pthalocyaninate/dysprosium, a single-4f-ionic single-molecule magnet with a Kramers ground state. Inorganic Chemistry 47: 10217–10219.
4. AllDamen MA, Cardona-Serra S, Clemente-Juan JM, Coronado E, Gaia-Artejo A, et al. (2009) Mononuclear lanthanide single molecule magnets based on the polyoxometalates [LnWO$_4$O$_{17}$$^{2-}$] and [Lu$_2$Sb$_2$O$_{13}$$^{8-}$]. [LaWO$_6$O$_{19}$$^{4-}$]–[La$_2$O$_8$W$_2$O$_{19}$$^{9-}$]–[La$_2$O$_6$W$_2$O$_{19}$$^{10-}$]–[LaW$_6$O$_{19}$$^{11-}$]–[TaW$_6$O$_{19}$$^{12-}$]–[AlO$_4$]–[Ta=W]-[Ta=W]-[Ta=W]-[Ta=W]. Angewandte Chemie International Edition 48: 9489–9492.
5. Rinehart JD, Fang M, Evans WJ, Long JR (2011) Strong exchange and magnetic blocking in N2 32-radical-bridged lanthanide complexes. Nature Chemistry 3: 336–342.
6. Wikstrom M (1992) MR imaging of preserved porcine myocardial infarction. An experimental study using Gd-DTPA and Dy-DTPA-BMA. Acta Radiologica 37: 27–35.
7. Wikstrom M (1992) MR imaging of experimental myocardial infarction. Acta radiologica. Supplementum 379: 1–30.
8. Bayouth JE, Macey DJ, Kasi LP, Garlich JR, McMillan K, et al. (1995) Pharmakoekinetik, dosimetry and toxicity of Holmium-166-DOTMP for bone marrow ablation in multiple myeloma. Journal of Nuclear Medicine 36: 730–737.
9. Sun L-N, Zhang H-J, Fu L-S, Liu F-Y, Meng Q-G, et al. (2005) A New Sol-Gel ateral Doped with an Erbium Complex and Its Potential optical-Amplification Application. Advanced Functional Materials 15: 1041–1046.
10. Park OH, Seo SY, Jung JH, Bae JY (2003) Photoilluminescence of mesoporous silica films impregnated with an erbium complex, Journal of Materials Research 18: 1039–1042.
11. Peter E (2003) Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and
