

Sparkle/PM7 Lanthanide Parameters for the Modeling of Complexes and Materials

José Diogo L. Dutra,[†] Manoel A. M. Filho,[†] Gerd B. Rocha,[‡] Ricardo O. Freire,[†] Alfredo M. Simas,^{*,§} and James J. P. Stewart^{||}

[†]Departamento de Química, Universidade Federal de Sergipe, 49.100-000 – São Cristóvão, SE, Brazil

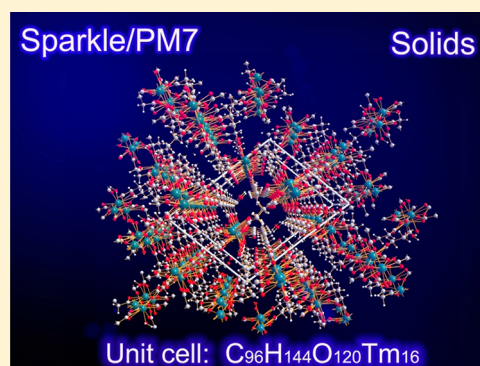
[‡]Departamento de Química, CCEN, Universidade Federal da Paraíba, 58.059-970 – João Pessoa, PB, Brazil

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

^{||}Stewart Computational Chemistry, 15210 Paddington Circle, Colorado Springs, Colorado 80921, United States

S Supporting Information

ABSTRACT: The recently published Parametric Method number 7, PM7, is the first semiempirical method to be successfully tested by modeling crystal structures and heats of formation of solids. PM7 is thus also capable of producing results of useful accuracy for materials science and constitutes a great improvement over its predecessor, PM6. In this article, we present Sparkle model parameters to be used with PM7 that allow the prediction of geometries of metal complexes and materials which contain lanthanide trications. Accordingly, we considered the geometries of 224 high-quality crystallographic structures of complexes for the parametrization set and 395 more for the validation of the parametrization for the whole lanthanide series, from La(III) to Lu(III). The average unsigned error for Sparkle/PM7 for the distances between the metal ion and its coordinating atoms is 0.063 Å for all lanthanides, ranging from a minimum of 0.052 Å for Tb(III) to 0.088 Å for Ce(III), comparable to the equivalent errors in the distances predicted by PM7 for other metals. These distance deviations follow a gamma distribution within a 95% level of confidence, signifying that they appear to be random around a mean, confirming that Sparkle/PM7 is a well-tempered method. We conclude by carrying out a Sparkle/PM7 full geometry optimization of two spatial groups of the same thulium-containing metal organic framework, with unit cells accommodating 376 atoms, of which 16 are Tm(III) cations; the optimized geometries were in good agreement with the crystallographic ones. These results emphasize the capability of the use of the Sparkle model for the prediction of geometries of compounds containing lanthanide trications within the PM7 semiempirical model, as well as the usefulness of such semiempirical calculations for materials modeling. Sparkle/PM7 is available in the software package MOPAC2012, at no cost for academics and can be obtained from <http://openmopac.net>.



Lanthanide ions find several applications, mainly due to their luminescent and magnetic properties. Their luminescence leads, for example, to their use in imaging with chemical sensors,^{1,2} time-resolved luminescence immunoassays,³ biomedical analysis and imaging,⁴ organic–inorganic hybrid phosphors,⁵ functional materials for the biosciences,⁶ photoemitting materials,⁷ upconverting fluorescent nanoparticles⁸ used in bioconjugation and bioimaging,⁹ as well as other uses. In addition, their near-infrared luminescence has found uses in telecommunications, bioanalyses, and solar energy conversion.¹⁰ More recently, their magnetic properties are gaining renewed interest due to the discovery that slow relaxation of the magnetization at liquid nitrogen temperatures can occur in mononuclear complexes of these ions;^{11,12} previously, in single molecule magnets, the highest blocking temperatures lay in the liquid helium region.¹³ Also, in recent years, lanthanide ions have found many uses in metal–organic frameworks: as magnetic resonance imaging probes,¹⁴ as contrast agents,¹⁵ in biomedical imaging and drug delivery,¹⁶ as luminescent multifunctional lanthanides-based metal–organic

frameworks,^{17,18} etc. Consequently, the design of ligands for functional metal–organic frameworks is an active area of research.¹⁹

All these applications require accurate quantum chemical methods that are both capable of modeling molecular structures containing hundreds of atoms and are general enough as to be able to handle, besides the lanthanides, many other elements of the periodic table.

Semiempirical methods have proven their usefulness for several applications^{20,21} and have evolved in a continuous basis over the last 30 years. Indeed, within this period, the most widely used methods have been, successively, MNDO,^{22,23} AM1,²⁴ AM1*,²⁵ PM3,^{26–28} RM1,²⁹ and PM6.³⁰

The sparkle model for lanthanides, originally introduced in 1994,^{31,32} represented an innovative manner of predicting the geometries of europium complexes via AM1²⁴ semiempirical calculations.

Received: November 17, 2012

Published: June 19, 2013

Table 1. Parameters for the Sparkle/PM7 Model for All Lanthanide Trications, from La(III) to Lu(III)^a

Sparkle/PM7					
	La ³⁺	Ce ³⁺	Pr ³⁺	Nd ³⁺	Pm ³⁺
α	2.72878366	3.67207094	4.56658977	4.70304554	2.62831297
GSS	51.00521418	56.61996758	56.12411820	51.86812242	56.60105705
a_1	0.27941646	0.23721649	0.27951368	0.37179588	0.25254367
b_1	13.44993801	9.05598877	14.17545394	11.01756863	11.32762969
c_1	3.03636844	3.05700455	2.96980969	2.94211575	2.91051149
a_2	0.22388397	2.16960030	2.01816742	2.01934079	0.42096080
b_2	13.33085266	8.57474567	5.58007643	11.90273308	9.59944282
c_2	3.44240129	2.97137088	3.13582779	3.31388588	3.51758897
	Sm ³⁺	Eu ³⁺	Gd ³⁺	Tb ³⁺	Dy ³⁺
α	2.48035044	2.64501092	4.11940733	2.71687807	2.82756925
GSS	56.91979708	57.16819322	55.80448082	57.37215107	55.61069514
a_1	0.27745438	0.25290496	0.33731573	0.30634054	0.35812416
b_1	11.39510076	12.42815090	9.55199105	10.88125130	8.98941131
c_1	2.89944391	2.88466231	2.90273188	2.85828879	2.86815735
a_2	0.38744285	0.56909102	2.93935956	0.66155895	0.40611392
b_2	9.47981871	9.75538134	8.93461822	9.13341532	8.95555466
c_2	3.73339418	3.00378806	3.32168472	3.41583594	3.69733260
	Ho ³⁺	Er ³⁺	Tm ³⁺	Yb ³⁺	Lu ³⁺
α	2.44068845	3.79457816	1.95626166	3.82324875	5.33103337
GSS	56.68096021	58.52071878	55.97220279	56.8376004	55.14010416
a_1	0.27966404	0.35009817	0.36171208	0.40711494	0.34482438
b_1	11.42737486	9.49506162	10.49976955	9.29829027	10.07236769
c_1	2.83804623	2.81837640	2.81614888	2.82195825	2.80387860
a_2	2.52779706	0.78014002	0.63407320	0.52341070	1.30538617
b_2	9.75311832	8.47946317	8.17764346	10.33065656	7.21222726
c_2	3.41599480	3.59076847	3.52086186	3.73059950	2.96294393

^a α , is in Å⁻¹. GSS, is in eV. a_i is dimensionless. b_i is in Å⁻². c_i is in Å.

In it, the metal ion was replaced by the Coulomb potential caused by a +3 point charge, superimposed on a repulsive radial potential of the form $\exp(-\alpha r)$, where α effectively attributes a size to the combined spherical potential, preventing the ligand atoms from collapsing toward the lanthanide ion. Simultaneously, the three electrons of the lanthanide are given to the lowest unoccupied orbitals of the ligands, which are then usually somewhat polarized toward the metal ion. We also added, as the heat of formation of the gaseous lanthanide trication, the enthalpy of sublimation of the lanthanide atoms plus the sum of its first three ionization potentials. The success of the Sparkle model can be attributed to the fact that the chemical bonds between the lanthanide ion and the ligands are essentially electrostatic. In 1997, we introduced a technique to obtain UV-vis spectra of the complexes due to the ligands by replacing the lanthanide ion by a point of charge +3, followed by a ZINDO³³ calculation.³⁴

Later, in 2004,³⁵ we improved the formalism of the Sparkle model by adding Gaussian functions to the expression which computes the core-core energy, in order to make the model compatible with AM1, and also added the mass of europium, allowing the calculation of infrared spectra. However, the full capability of the Sparkle model was only realized in 2005 with an improved parametrization for Eu(III), Gd(III), and Tb(III) within AM1,³⁶ which has proven popular [for one of its first applications, see Faustino et al.³⁷]. The sparkle concept was then applied to alkaline and alkaline earth complexes as well.^{38,39} Subsequently, the model was extended to all other lanthanide ions within AM1^{40–48} and was later parametrized for PM3^{49–55} and for PM6.⁵⁶ In 2006, we published an article in

which we compared Sparkle calculations with ab initio effective core potential calculations.⁵⁷ In that study, we showed that the most accurate chemistry model for predicting the geometry of the coordination polyhedron of lanthanide complexes was RHF/STO-3G using the ECP of Dolg et al.⁵⁸ These ab initio ECP calculations turned out to be atypical: both augmenting the basis set and/or including electron correlation generally tended to enlarge the deviations and worsen the quality of the predicted coordination polyhedron geometry, when compared to the corresponding X-ray data. Sparkle model calculations had the added advantage of being much faster than any ab initio ECP calculation, as well as describing not only the coordination polyhedron but also the ligands as well.⁵⁷ A recent thorough independent assessment of the Sparkle model accuracy has been recently published by Seitz and Alzakhem,⁵⁹ in which they calculated the geometries of over 650 lanthanide complexes with Ln–OH₂ motifs and found them, especially Sparkle/AM1, to be “surprisingly accurate for the estimation of the average bond lengths between the lanthanide ion and the water oxygen for the technologically important central lanthanides Eu(III), Gd(III), and Tb(III)”. Indeed, since its conception, the Sparkle model has been successfully applied to several research problems, such as the study of energy transfer in Ir(III)/Eu(III) dyads,⁶⁰ investigation of a Sm membrane sensor and membrane interfaces,^{61–63} modeling metal organic frameworks,⁶³ chiral complexes,⁶⁵ Tb(III) light conversion molecular devices,⁶⁶ and a tetramer complex of Eu(III) containing 181 atoms⁶⁷ and other problems.

The recently published Parametric Method Number 7, PM7,⁶⁸ is a new semiempirical method with a better representation of

long-range interactions, resulting in a much improved prediction of crystal structures and heats of formation of solids. Indeed, PM7 contains improvements in the description of dispersion and hydrogen bonds which were incorporated into the method before the parametrization was carried out, to better describe noncovalent interactions. PM7 also allows for unrestricted Hartree–Fock partial open shells and possesses a specific parametrization to reproduce barrier heights. Finally, PM7 also corrected a fault in earlier NDDO methods that resulted in infinite errors when solids were modeled⁶⁸ but which was otherwise negligible. Thus, PM7 represents a major improvement over PM6 and is, in principle, capable of producing results for materials science research of useful accuracy. Indeed, for organic solids, the PM7 average unsigned errors (AUE) are 33% smaller for the geometries and 60% smaller for the enthalpies of formation, when compared to the same quantities for PM6.

In this article, we extend the Sparkle model to PM7 and report parameters for all 15 lanthanide trications, from La(III) to Lu(III). Finally, we carry out a sample solid state calculation on a unit cell of a metal organic framework containing 376 atoms, of which 16 are Tm(III), and compare the result with crystallographic data for the same compound.

RESULTS AND DISCUSSION

The Sparkle model in PM7 retains the same mathematical structure used in AM1, PM3, and PM6. That is, it includes in the core–core interaction the same sum of two spherical Gaussian functions, each with three adjustable coefficients.³⁵ Therefore, as before, Sparkle PM7 possesses eight parameters for each lanthanide trication: α , the core–core repulsion term, in \AA^{-1} ; GSS, the s – s atomic orbital one-center two-electron repulsion integral, in eV; and the six parameters for the two Gaussian functions: height, a_p , dimensionless; inverse broadness, b_p , in \AA^{-2} ; and displacement, c_p , in \AA , as in $G(R) = \sum_{i=1}^2 a_i e^{[b_i(R - c_i)^2]}$, where R , in \AA , is the interatomic distance between the lanthanide and the other atom.

For the parametrization training set, we used the same set of 224 complexes of high quality (R-factor < 0.5%) structures taken from the Cambridge Structural Database^{69,70} (CSD) that

Table 2. Means, Standard Deviations, Two- σ Intervals, and p Values for the γ Distribution Fits for the UME(Ln–L)’s Computed for the N Complexes, for Each Lanthanide Trication

Ln ³⁺	N	UME(Ln–L)				p value
		mean (\AA)	σ (\AA)	mean + 2σ (\AA)	mean – 2σ (\AA)	
La ³⁺	72	0.0626	0.0306	0.1237	0.0015	0.2557
Ce ³⁺	37	0.0828	0.0499	0.1826	0.0000	0.3822
Pr ³⁺	44	0.0781	0.0497	0.1775	0.0000	0.0773
Nd ³⁺	52	0.0675	0.0370	0.1414	0.0000	0.1121
Pm ³⁺	15	0.0512	0.0266	0.1043	0.0000	0.0605
Sm ³⁺	44	0.0597	0.0402	0.1400	0.0000	0.1455
Eu ³⁺	91	0.0537	0.0375	0.1288	0.0000	0.3141
Gd ³⁺	72	0.0580	0.0361	0.1303	0.0000	0.4679
Tb ³⁺	43	0.0483	0.0267	0.1017	0.0000	0.9725
Dy ³⁺	28	0.0533	0.0300	0.1133	0.0000	0.1670
Ho ³⁺	27	0.0542	0.0358	0.1258	0.0000	0.6119
Er ³⁺	35	0.0529	0.0540	0.1608	0.0000	0.7236
Tm ³⁺	14	0.0665	0.0288	0.1241	0.0088	0.8634
Yb ³⁺	45	0.0591	0.0351	0.1292	0.0000	0.9418
Lu ³⁺	30	0.0522	0.0326	0.1173	0.0000	0.2557

were used in previous parametrizations, such as Sparkle/PM6,⁵⁶ 15 representative complexes for each of the 15 lanthanide trications, from La(III) to Lu(III), and carried out the procedure as described before.³⁶ The Sparkle/PM7 parameters are presented in Table 1.

For the validation set, we added 425 more complexes for a total of 679 coordination compounds involving all 15 lanthanide ions.

As before, in order to assess the accuracy of the method, we considered the unsigned mean errors, UME, defined for all complexes as

$$\text{UME} = \frac{1}{n} \sum_{j=1}^n |R_j^{\text{CSD}} - R_j^{\text{calc}}|$$

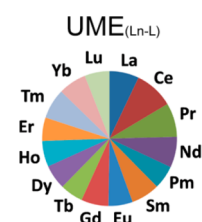


Figure 1. Pie plot of the various unsigned mean errors UME(Ln–L) for sets of complexes for all lanthanides, in order to allow a comparison among them, showing that Sparkle/PM7 is a balanced model, across the series, in terms of its accuracy of lanthanide-coordinating atom distances.

Table 3. Means, Standard Deviations, Two- σ Intervals, and p Values for the γ Distribution Fits for the UMEs Computed for the N Complexes for Each Lanthanide Trication

Ln ³⁺	N	UMEs				p value
		mean (\AA)	σ (\AA)	mean + 2σ (\AA)	mean – 2σ (\AA)	
La ³⁺	72	0.1946	0.0946	0.3837	0.0055	0.6017
Ce ³⁺	37	0.2106	0.1015	0.4136	0.0076	0.6700
Pr ³⁺	44	0.2209	0.1085	0.4380	0.0038	0.2586
Nd ³⁺	52	0.2135	0.1204	0.4543	0.0000	0.9561
Pm ³⁺	15	0.1485	0.0806	0.3098	0.0000	0.3845
Sm ³⁺	44	0.1798	0.0916	0.3631	0.0000	0.7558
Eu ³⁺	91	0.1787	0.0846	0.3479	0.0095	0.7389
Gd ³⁺	72	0.1710	0.0853	0.3416	0.0003	0.9780
Tb ³⁺	43	0.1674	0.0851	0.3377	0.0000	0.8046
Dy ³⁺	28	0.1934	0.0993	0.3919	0.0000	0.8413
Ho ³⁺	27	0.1735	0.0843	0.3421	0.0049	0.6900
Er ³⁺	35	0.1759	0.1028	0.3815	0.0000	0.6105
Tm ³⁺	14	0.1783	0.0940	0.3663	0.0000	0.8138
Yb ³⁺	45	0.1533	0.0694	0.2920	0.0146	0.6255
Lu ³⁺	30	0.1624	0.0923	0.3470	0.0000	0.6017

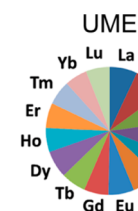


Figure 2. Pie plot of the various unsigned mean errors UME for sets of complexes for all lanthanides, in order to allow a comparison among them, showing that Sparkle/PM7 is a balanced model, across the series, in terms of its accuracy of coordination polyhedra geometries.

Table 4. Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, and Sparkle/PM7 Unsigned Mean Errors for All Distances Involving the Central Lanthanide Ion, Ln, and the Ligand Atoms of the Coordination Polyhedron, L, L', and the Specific Cases When L Is Either Oxygen, or Nitrogen, for All Lanthanide Complexes Considered

	unsigned mean errors for specific types of distances (Å)				UME(Ln–L) (Å)	UME (Å)
	Ln–Ln	Ln–O	Ln–N	L–L'	Ln–L and Ln–Ln	Ln–L, Ln–Ln, and L–L'
lanthanum						
Sparkle/AM1	0.2118	0.0871	0.0479	0.2098	0.0781	0.1847
Sparkle/PM3	0.1016	0.0628	0.0831	0.1856	0.0690	0.1634
Sparkle/PM6	0.2428	0.0798	0.0621	0.2385	0.0774	0.2079
Sparkle/PM7	0.1275	0.0686	0.0582	0.2276	0.0666	0.1970
cerium						
Sparkle/AM1	0.2161	0.0859	0.0749	0.2217	0.0848	0.1946
Sparkle/PM3	0.2180	0.0810	0.0692	0.1823	0.0798	0.1619
Sparkle/PM6	0.2102	0.0866	0.0703	0.2379	0.0840	0.2074
Sparkle/PM7	0.1743	0.0943	0.0666	0.2582	0.0881	0.2245
praseodymium						
Sparkle/AM1	0.2417	0.0862	0.0700	0.2461	0.0853	0.2146
Sparkle/PM3	0.1587	0.0874	0.0784	0.2050	0.0867	0.1818
Sparkle/PM6	0.1480	0.0749	0.0735	0.2475	0.0758	0.2139
Sparkle/PM7	0.1427	0.0801	0.0593	0.2580	0.0767	0.2225
neodymium						
Sparkle/AM1	0.1076	0.0804	0.0532	0.2263	0.0742	0.1960
Sparkle/PM3	0.0953	0.0771	0.0497	0.1787	0.0708	0.1572
Sparkle/PM6	0.1717	0.0754	0.0746	0.2337	0.0758	0.2022
Sparkle/PM7	0.1795	0.0692	0.0669	0.2628	0.0693	0.2242
promethium						
Sparkle/AM1	0.3377	0.0559	0.0589	0.2094	0.0587	0.1773
Sparkle/PM3	0.4059	0.0473	0.0745	0.1643	0.0564	0.1413
Sparkle/PM6	0.3282	0.0535	0.0794	0.2141	0.0618	0.1817
Sparkle/PM7	0.0426	0.0542	0.0500	0.2030	0.0531	0.1711
samarium						
Sparkle/AM1	0.1379	0.0640	0.0933	0.2069	0.0738	0.1779
Sparkle/PM3	0.1843	0.0728	0.0812	0.1703	0.0773	0.1503
Sparkle/PM6	0.2334	0.0894	0.0577	0.2430	0.0825	0.2085
Sparkle/PM7	0.2857	0.0616	0.0616	0.2272	0.0643	0.1922
europium						
Sparkle/AM1	0.1714	0.0836	0.0703	0.2131	0.0827	0.1859
Sparkle/PM3	0.1727	0.0835	0.0498	0.1805	0.0782	0.1592
Sparkle/PM6	0.1802	0.0715	0.0888	0.2324	0.0777	0.2001
Sparkle/PM7	0.1766	0.0533	0.0490	0.2160	0.0551	0.1825
gadolinium						
Sparkle/AM1	0.1484	0.0620	0.0553	0.1882	0.0613	0.1614
Sparkle/PM3	0.1845	0.0586	0.0843	0.1559	0.0678	0.1373
Sparkle/PM6	0.3002	0.0517	0.0996	0.2119	0.0691	0.1818
Sparkle/PM7	0.1859	0.0585	0.0622	0.2039	0.0614	0.1739
terbium						
Sparkle/AM1	0.1867	0.0732	0.0460	0.2213	0.0700	0.1887
Sparkle/PM3	0.1816	0.0695	0.0508	0.1760	0.0679	0.1527
Sparkle/PM6	0.2542	0.0768	0.0540	0.2621	0.0753	0.2218
Sparkle/PM7	0.1978	0.0472	0.0548	0.2112	0.0520	0.1768
dysprosium						
Sparkle/AM1	0.2200	0.0680	0.0363	0.2177	0.0652	0.1860
Sparkle/PM3	0.1493	0.0645	0.0979	0.1683	0.0702	0.1479
Sparkle/PM6	0.1845	0.0739	0.0781	0.2585	0.0757	0.2205
Sparkle/PM7	0.1799	0.0561	0.0433	0.2292	0.0556	0.1931
holmium						
Sparkle/AM1	0.2385	0.0538	0.0436	0.2313	0.0550	0.1937
Sparkle/PM3	0.2153	0.0731	0.0595	0.1978	0.0728	0.1711
Sparkle/PM6	0.2035	0.0758	0.0779	0.2477	0.0785	0.2115
Sparkle/PM7	0.3208	0.0490	0.0700	0.202	0.0583	0.1714
erbium						
Sparkle/AM1	0.1373	0.0737	0.0365	0.2164	0.0690	0.1847
Sparkle/PM3	0.1797	0.0725	0.0551	0.1618	0.0715	0.1424

Table 4. continued

	unsigned mean errors for specific types of distances (Å)				UME(Ln–L) (Å)	UME (Å)
	Ln–Ln	Ln–O	Ln–N	L–L'	Ln–L and Ln–Ln	Ln–L, Ln–Ln, and L–L'
erbium						
Sparkle/PM6	0.1688	0.0595	0.0884	0.2307	0.0655	0.1952
Sparkle/PM7	0.0960	0.0534	0.0448	0.2128	0.0527	0.1783
thulium						
Sparkle/AM1	0.2110	0.0623	0.0591	0.1780	0.0646	0.1537
Sparkle/PM3	0.1814	0.0681	0.0808	0.1599	0.0738	0.1415
Sparkle/PM6	0.2347	0.0708	0.0770	0.2157	0.0758	0.1858
Sparkle/PM7	0.1314	0.0618	0.0799	0.2027	0.0680	0.1739
ytterbium						
Sparkle/AM1	0.1958	0.0764	0.0564	0.1889	0.0722	0.1626
Sparkle/PM3	0.1129	0.0754	0.0717	0.1563	0.0748	0.1379
Sparkle/PM6	0.2446	0.0619	0.0947	0.2109	0.0737	0.1800
Sparkle/PM7	0.2222	0.0625	0.0579	0.1890	0.0633	0.1607
lutetium						
Sparkle/AM1	0.2133	0.0845	0.0465	0.1742	0.0750	0.1529
Sparkle/PM3	0.1843	0.0827	0.0536	0.1485	0.0755	0.1328
Sparkle/PM6	0.1951	0.0876	0.0480	0.2138	0.0773	0.1845
Sparkle/PM7	0.1674	0.0535	0.0565	0.2001	0.0564	0.1692

where R_j^{CSD} and R_j^{calc} are, respectively, the j th distance from the CSD and from the semiempirical calculation. Again, we used two types of unsigned mean errors: UME(Ln–L) and UME. The first, UME(Ln–L), refers to a sum of all distances, for all complexes, from the central lanthanide ion to all of the atoms of its coordination polyhedron, including its distance to another lanthanide ion, when they face each other, and UME is the sum of all of these distances plus all lengths of the edges of the coordination polyhedra, a much more stringent criterion of accuracy.

During the statistical tests for the validation of Sparkle/PM7, we found that the presence of phosphorus atoms in the complex structure caused large errors in the Sparkle/PM7 geometry of the complexes and consistently gave very bad results. For example, for those cases where a phosphorus atom is bonded to a coordinating oxygen atom, the oxygen–lanthanide ion bond length was always predicted to be too large. However, phosphorus-containing complexes are rare. Only 30 were found in over 600 structures, and in none of these cases was a phosphorus atom directly coordinated to the lanthanide ion. We were unable to correct this fault, so we conclude that it is a systematic error of Sparkle/PM7. Thus, for those rare cases when phosphorus is present in the structure, users are advised to choose the Sparkle/PM6 model. For this work, we removed all 30 phosphorus-containing complexes from both the Tm(III) parametrization set and the validation set and chose to proceed without them. Nonetheless, in Table S16 of the Supporting Information, we list UME(Ln–L)'s and UMEs for all 30 of these phosphorus containing compounds. Moreover, not all elements of the periodic table are represented in the universe of lanthanide complexes. Indeed, out of the 679 complexes of the validation set, only 81 have ligands with atoms of elements different than H, C, N, O, P, and S, and none of these different atoms, which are B, Si, F, Cl, and Br, are directly coordinated to the lanthanide ion. Accordingly, we also included in the Supporting Information, Table S17, UME(Ln–L)'s and UMEs for all 81 complexes.

A precondition for a reliable model is that both UMEs must be randomly distributed about a mean; if true, then we can have more confidence that the Sparkle model captured the essence of the chemical bond between the lanthanide and its ligands for

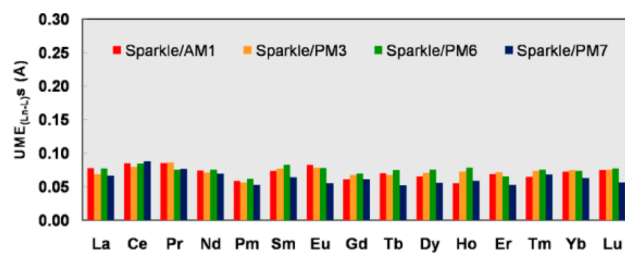


Figure 3. Unsigned mean errors UME(Ln–L), in Å, for all existing Sparkle models: Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, and Sparkle/PM7, for all sets of lanthanide complexes, one set for each lanthanide ion.

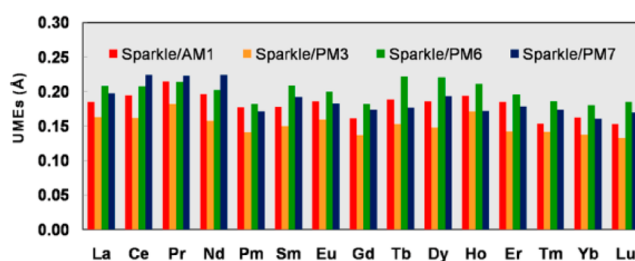


Figure 4. Unsigned mean errors UME, in Å, for all existing Sparkle models: Sparkle/AM1, Sparkle/PM3, Sparkle/PM6, and Sparkle/PM7, for all sets of lanthanide complexes, one set for each lanthanide ion.

the set of compounds studied. If the UMEs were randomly distributed about the mean, that would imply that they follow the probability density function of the γ distribution, since UMEs are positive and defined in the domain $(0, \infty)$. We then fitted the various UMEs for each lanthanide ion to a γ distribution, from which the mean and variance were obtained. Then, we tested the quality of the fit by means of the one-sample nonparametric Kolmogorov–Smirnov test.⁷¹ The p value of the test must be larger than 0.05 for the γ distribution fit to be justified within a 95% confidence interval—and for the mean and variance to be used as accuracy measures of the model.

Table 2 shows the means, variances, and p values for the $\text{UME}(\text{Ln}-\text{L})$'s for each lanthanide ion. We can see that these mean values are similar for all lanthanide ions, varying from 0.052 Å for Tm(III) to 0.088 Å for Ce(III), with an average of 0.063 Å.

Figure 1 shows a pie plot of these $\text{UME}(\text{Ln}-\text{L})$'s, which allows us to compare these values among themselves and verify that they are equivalent for all lanthanides, suggesting that Sparkle/PM7 is well balanced within the series.

Table 3 shows the same quantities for UMEs, which, as mentioned before, include not only the distances between the lanthanide cation and its coordinating atoms but also all lengths of the edges of the coordination polyhedron—a much more strict measure. Table S18 of the Supporting Information lists the complexes that presented the smallest as well as the largest $\text{UME}(\text{Ln}-\text{L})$ and UME measures, respectively, for each lanthanide trication.

Larger deviations are usually observed on the calculated values of the edges due to the fact that even small errors on the angle formed by the two coordinating atoms with the lanthanide ion in the center corner can cause relatively large deviations in the length of the edge, which is the line segment connecting the two coordinating atoms; that, of course, may occur, even when deviations in the distances between the lanthanide and the coordinating atoms are small. The UMEs range from 0.161 Å for Yb(III) to 0.224 Å for Ce(III), with an average of 0.187 Å for the whole series. Figure 2 shows a pie plot for the UMEs across the lanthanide series, again indicating that all parametrizations for the whole lanthanide series are essentially equivalent in accuracy.

What actually changes across the series are the types of ligands associated with each lanthanide. For each lanthanide, the ligands are optimized for certain properties, and these properties vary enormously throughout the series, so that a ligand that is useful for one lanthanide may not be advantageous to another. Indeed, whereas some complexes emit in the near-infrared, and need dyes as ligands, others emit in the visible, and so they need ultraviolet sensitizers as ligands. Still others are used in magnetic resonance imaging, so that at least one water molecule is needed in the coordination sphere and so on. Despite this, there are only minor variations in accuracy in going across the series.

Table 4 shows the most important unsigned mean errors for specific types of distances, as well as $\text{UME}(\text{Ln}-\text{L})$ and UME for Sparkle/PM7 and the other three already-published methods, Sparkle/AM1, Sparkle/PM3, and Sparkle/PM6, for the validation set of complexes, and for the 15 lanthanide ions.

From Table 4, it can be clearly noted that Sparkle/PM7 is equivalent in accuracy to all other methods, but it is still the best of all sparkle models for Ln–Ln bonds involving Ce(III), Pr(III), Pm(III), Er(III), Tm(III), or Lu(III). Of the Sparkle methods, Sparkle/PM7 is the most accurate for the prediction of Ln–O bonds in Nd(III), Sm(III), Eu(III), Tb(III), Dy(III), Ho(III), Er(III), Tm(III), and Lu(III); for Ln–L bonds in Ce(III), Pr(III), Pm(III), and Eu(III); and for the distance between coordinating atoms for Pm(III) and Ho(III). Sparkle/PM7 surpasses all other methods in the important $\text{UME}(\text{Ln}-\text{L})$ for Nd(III), Pm(III), Sm(III), Eu(III), Tb(III), Dy(III), Er(III), Yb(III), and Lu(III). In short, for complexes without phosphorus in their formula, Sparkle/PM7 is the most accurate of the Sparkle methods.

Figures 3 and 4 indicate in a pictorial manner the data in Table 4, where it can be readily seen that the values for

$\text{UME}(\text{Ln}-\text{L})$ are always much smaller than the values of UME for all lanthanides and that, in general, the errors in Sparkle/PM7 are similar to those in Sparkle/PM6.

Although PM7 has been parametrized for most elements, only two of the 15 lanthanides, La and Lu, were represented. This deficiency is corrected by the Sparkle/PM7 method, which provides parameters for all 15 lanthanides. For all geometry optimizations of lanthanide complexes calculated using MOPAC2012,⁷² the recommended method is Sparkle/PM7, activated by keyword SPARKLE. For complexes involving La or Lu, there is a choice of PM7 and Sparkle/PM7, with the latter method being recommended when geometries are of greatest importance.

SOLID STATE CALCULATION OF A MOF

In order to exemplify the power of the Sparkle/PM7 method, we chose to fully optimize the geometry of the unit cell of two new metal–organic frameworks, MOFs, of the same unit cell formula $\text{C}_{96}\text{H}_{144}\text{O}_{120}\text{Tm}_{16}$ that have been recently synthesized under different hydrothermal conditions and characterized by single-crystal X-ray diffraction.⁷³

The two polymorphs, one triclinic, the other monoclinic, form three-dimensional networks of Tm(III) ions and carboxylate dianions. Unit cell solid state unconstrained geometry optimization calculations were carried out for both compounds of 376 atoms, of which 16 are Tm(III). The experimental crystallographic structures and the optimized unit cell geometries for both compounds, monoclinic and triclinic, are shown in Figures 5 and 6, respectively.

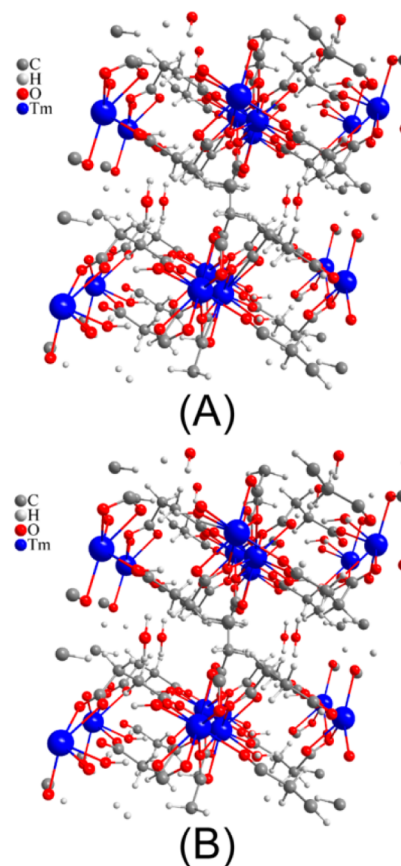


Figure 5. Unit cell for the compound of formula $\text{C}_{96}\text{H}_{144}\text{O}_{120}\text{Tm}_{16}$ crystallized in a monoclinic system. (a) Geometry from X-ray crystallography.⁷³ (b) Fully optimized Sparkle/PM7 solid state MOPAC⁷² calculation.

Table 5. Comparison between Crystallographic and Calculated Cell Parameters for the Monoclinic and Triclinic Systems of the Lanthanide-Organic Framework $C_{96}H_{144}O_{120}Tm_{16}$ ^a

structure	cell parameters		c	α	β	γ	volume
	a	b					
monoclinic system							
crystallographic	13.77	7.63	17.24	90.00	101.30	90.00	1776.62
Sparkle/PM7	13.69	7.87	17.27	89.64	103.62	89.95	1807.11
percent error	0.58%	3.15%	0.17%	0.39%	2.29%	0.06%	1.72%
triclinic system							
crystallographic	7.62	10.58	12.71	95.31	107.44	111.09	888.68
Sparkle/PM7	7.86	10.57	12.74	95.02	107.95	111.27	914.01
percent error	3.15%	0.09%	0.24%	0.30%	0.47%	0.16%	2.85%

^aParameters a , b , and c are in Å. α , β , and γ are in degrees. Volume is in Å³.

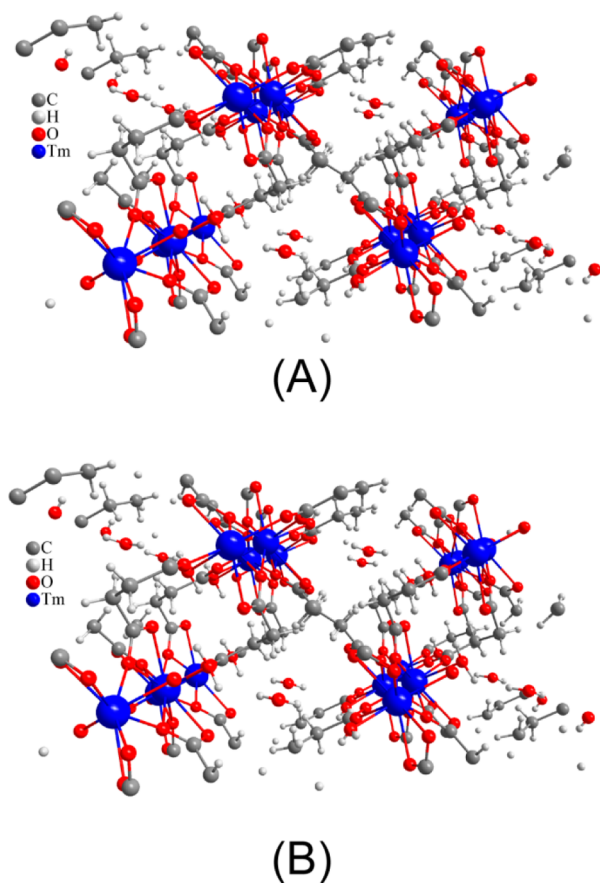


Figure 6. Unit cell for the compound of formula $C_{96}H_{144}O_{120}Tm_{16}$ crystallized in a triclinic system. (a) Geometry from X-ray crystallography.⁷³ (b) Fully optimized Sparkle/PM7 solid state MOPAC⁷² calculation.

The calculated and experimental structures are indeed so similar that it is difficult to detect differences between them with the naked eye; therefore, in Table 5, data for the crystallographic and calculated cell parameters for both compounds are presented.

Clearly, both unit cells are very similar, with their parameters agreeing within 4% in all cases. This is therefore a good illustration of the type of accuracy that can be achieved using the Sparkle/PM7 method, when used in modeling solids.

CONCLUSIONS

PM7 is the first semiempirical molecular orbital method parametrized using an entirely new approach for choosing training set reference data, one that better defines the parameter space and

aims at better characterizing its structure in the region of the error function minimum. Moreover, PM7 was successfully tested by modeling crystal structures and predicting heats of formation of solids.⁶⁸ Thus, PM7 has the potential to become the semiempirical method of choice for materials science, besides its applications to biochemistry, organic chemistry, pharmaceutical chemistry, etc. The extension of the Sparkle model for PM7, presented in this article, allows the calculation of lanthanide trication complexes, crystals of complexes, multidimensional metal–organic frameworks, nanostructures, et cetera, with useful accuracy. We also presented a full solid state geometry optimization of two metal–organic framework compounds of empirical formula $[Tm_2(C_4H_4O_4)_3(H_2O)_2] \cdot H_2O$ and demonstrated that all results agree very well with the X-ray diffraction geometries. With the sole exception of complexes containing phosphorus atoms, Sparkle/PM7 has been validated and can be used, if and when needed, with the expectation of obtaining accurate geometries, which can serve as a starting point to higher level calculations—mainly for complexes with ligands containing nitrogen and/or oxygen as coordinating atoms. Sparkle/PM7 is available in the software package MOPAC2012,⁷² which is free for academics and can be obtained from <http://openmopac.net>.

ASSOCIATED CONTENT

Supporting Information

Instructions on how to run lanthanide complexes Sparkle/PM7 calculations (<http://www.sparkle.pro.br>) on MOPAC2012 (<http://openmopac.net>), additional tables containing UME-(Ln–L)s and UMEs for all 649 complexes of the validation set, additional histograms comparing previous Sparkle models with respect to various classes of coordinating atoms, as well as sample MOPAC2012 input (.mop) and output (.arc) files for one representative complex for each of the lanthanide ions, from La(III) to Lu(III). MOPAC2012 input (.mop) and output (.arc) files for both the monoclinic and triclinic systems of $C_{96}H_{144}O_{120}Tm_{16}$. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: simas@ufpe.br.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We appreciate the financial support from the Brazilian agencies, institutes, and networks: CNPq, FACEPE (Pronex), FAPITEC-SE,

INAMI, and RENAMI. One of us (J.J.P.S.) thanks the National Institute of General Medical Sciences (NIGMS) for funding support via Grant No. 2 R44 GM067327-02. Finally, we gratefully acknowledge the Cambridge Crystallographic Data Centre for the Cambridge Structural Database.

■ REFERENCES

- (1) Schaeferling, M. *Angew. Chem., Int. Ed.* **2012**, *51* (15), 3532–3554.
- (2) Cable, M. L.; Levine, D. J.; Kirby, J. P.; Gray, H. B.; Ponce, A. *Adv. Inorg. Chem.* **2011**, *63*, 1–45.
- (3) Hagan, A. K.; Zuchner, T. *Anal. Bioanal. Chem.* **2011**, *400*, 2847–2864.
- (4) Bünzli, J.-C. G. *Chem. Rev.* **2010**, *110*, 2729–2755.
- (5) Carlos, L. D.; Ferreira, R. A. S.; Bermudez, V. Z.; Julián-López, B.; Escibano, P. *Chem. Soc. Rev.* **2011**, *40*, 536–549.
- (6) Eliseeva, S. V.; Bünzli, J.-C. G. *Chem. Soc. Rev.* **2010**, *39*, 189–227.
- (7) Armelao, L.; Quici, S.; Barigelletti, F.; Accorsi, G.; Bottaro, G.; Cavazzini, M.; Tondello, E. *Coord. Chem. Rev.* **2010**, *254*, 487–505.
- (8) Chatterjee, D. K.; Zhang, Y. *Small* **2010**, *6* (24), 2781–2795.
- (9) Mader, H. S.; Kele, P.; Saleh, S. M.; Wolfbeis, O. S. *Curr. Opin. Chem. Biol.* **2010**, *14*, 582–596.
- (10) Bünzli, J.-C. G.; Eliseeva, S. V. *J. Rare Earths* **2010**, *28* (6), 824–842.
- (11) Giraud, R.; Wernsdorfer, W.; Tkachuk, A. M.; Barbara, D. M. B. *Phys. Rev. Lett.* **2001**, *87* (057203), 1–4.
- (12) Sorace, L.; Benelli, C.; Gatteschi, D. *Chem. Soc. Rev.* **2011**, *40*, 3092–3104.
- (13) Milios, C. J.; Vinslava, A.; Wernsdorfer, W.; Moggach, S.; Parsons, S.; Perlepes, S. P.; Christou, G.; Brechin, E. K. *J. Am. Chem. Soc.* **2007**, *129*, 2754–2755.
- (14) Peters, J. A.; Djanashvili, K. *Eur. J. Inorg. Chem.* **2012**, *12*, 1961–1974.
- (15) Rocca, J. D.; Lin, W. *Eur. J. Inorg. Chem.* **2010**, *24*, 3725–3734.
- (16) Rocca, J. D.; Liu, D.; Lin, W. *Acc. Chem. Res.* **2011**, *44* (10), 957–968.
- (17) Rocha, J.; Carlos, L. D.; Paz, F. A. A.; Ananias, D. *Chem. Soc. Rev.* **2011**, *40*, 926–940.
- (18) Cui, Y.; Yue, Y.; Qian, G.; Chen, B. *Chem. Rev.* **2012**, *112*, 1126–1162.
- (19) Paz, F. A. A.; Klinowski, J.; Vilela, S. M. F.; Tomé, J. P. C.; Cavaleiro, J. A. S.; Rocha, J. *Chem. Soc. Rev.* **2012**, *41*, 1088–1110.
- (20) Stewart, J. J. P. *J. Mol. Model.* **2009**, *15*, 765–805.
- (21) Stewart, J. J. P. *J. Mol. Model.* **2008**, *14*, 499–535.
- (22) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- (23) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907–4917.
- (24) Winget, P.; Horn, A. H. C.; Selçuki, C.; Martin, B.; Clark, T. J. *Mol. Model.* **2003**, *9*, 408–414.
- (25) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (26) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (27) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221–264.
- (28) Stewart, J. J. P. *J. Comput. Chem.* **1991**, *12*, 320–341.
- (29) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1101–1111.
- (30) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (31) de Andrade, A. V. M.; da Costa, N. B., Jr.; Simas, A. M.; de Sá, G. F. *Chem. Phys. Lett.* **1994**, *227*, 349–353.
- (32) de Andrade, A. V. M.; da Costa, N. B., Jr.; Simas, A. M.; de Sá, G. F. *J. Alloys Compd.* **1995**, *225*, 55–59.
- (33) Zerner, M. C. *ZINDO Manual*; QTP, University of Florida: Gainesville, FL, 1990.
- (34) Andrade, A. V. M.; da Costa, N. B., Jr.; Malta, O. L.; Longo, R. L.; Simas, A. M.; de Sá, G. F. *J. Alloys Compd.* **1997**, *250*, 412–416.
- (35) Rocha, G. B.; Freire, R. O.; Costa, N. B., Jr.; de Sá, G. F.; Simas, A. M. *Inorg. Chem.* **2004**, *43*, 2346–2354.
- (36) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Inorg. Chem.* **2005**, *44*, 3299–3310.
- (37) Faustino, W. M.; Malta, O. L.; Teotonio, E. E. S.; Brito, H. F.; Simas, A. M.; de Sá, G. F. *J. Phys. Chem. A* **2006**, *110*, 2510–2516.
- (38) Galle, M.; Puchta, R.; van Eikema Hommes, N. J. R.; van Eldik, R. Z. *Phys. Chem.* **2006**, *220*, 511–523.
- (39) Puchta, R.; van Eldik, R. *Eur. J. Inorg. Chem.* **2007**, 1120–1127.
- (40) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Organomet. Chem.* **2005**, *690*, 4099–4102.
- (41) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Chem. Phys. Lett.* **2005**, *411*, 61–65.
- (42) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Comput. Chem.* **2005**, *26*, 1524–1528.
- (43) Costa, N. B., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *Polyhedron* **2005**, *24*, 3046–3051.
- (44) Costa, N. B., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *Inorg. Chem. Commun.* **2005**, *8*, 831–835.
- (45) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Chem. Theory Comput.* **2006**, *2*, 64–74.
- (46) Bastos, C. C.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Photochem. Photobiol. A* **2006**, *117*, 225–237.
- (47) Freire, R. O.; Monte, E. V.; Rocha, G. B.; Simas, A. M. *J. Organomet. Chem.* **2006**, *691*, 2584–2588.
- (48) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Phys. Chem. A* **2006**, *110*, 5897–5900.
- (49) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Chem. Phys. Lett.* **2006**, *425*, 138–141.
- (50) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Chem. Phys. Lett.* **2007**, *441*, 354–357.
- (51) Costa, N. B., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Phys. Chem. A* **2007**, *111*, 5015–5018.
- (52) Simas, A. M.; Freire, R. O.; Rocha, G. B. *Lect. Notes Comput. Sci.* **2007**, *4488*, 312–318.
- (53) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Chem. Theory Comput.* **2007**, *3*, 1588–1596.
- (54) Simas, A. M.; Freire, R. O.; Rocha, G. B. *J. Organomet. Chem.* **2008**, *693*, 1952–1956.
- (55) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Braz. Chem. Soc.* **2009**, *20*, 1638–1645.
- (56) Freire, R. O.; Simas, A. M. *J. Chem. Theory Comput.* **2010**, *6*, 2019–2023.
- (57) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Mol. Model.* **2006**, *12*, 373–389.
- (58) Dolg, M.; Stoll, H.; Savin, A.; Preuss, H. *Theor. Chim. Acta* **1989**, *75*, 173–194.
- (59) Seitz, M.; Alzakhem, N. J. *Chem. Inf. Model.* **2010**, *50*, 217–220.
- (60) Sykes, D.; Tidmarsh, I. S.; Barbieri, A.; Sazanovich, I. V.; Weinstein, J. A.; Ward, M. D. *Inorg. Chem.* **2011**, *50*, 11323–11339.
- (61) Rezaei, B.; Askarpour, N.; Hadadzadeh, H. *IEEE Sens. J.* **2011**, *11* (9), 2077–2083.
- (62) Rezaei, B.; Hemati, A.; Hadadzadeh, H. *J. Colloid Interface Sci.* **2011**, *354*, 268–274.
- (63) Rezaei, B.; Askarpour, N.; Hadadzadeh, H. *Electroanalysis* **2011**, *23* (4), 1029–1037.
- (64) Rodrigues, M. O.; Paz, F. A.; Freire, R. O.; de Sá, G. F.; Galembeck, A.; Montenegro, M. C.; Araujo, A. N.; Alves, S. J. *Phys. Chem. B* **2009**, *113*, 12181–12188.
- (65) Li, D.-P.; Li, C.-H.; Wang, J.; Kang, L.-C.; Wu, T.; Li, Y.-Z.; You, X.-Z. *Eur. J. Inorg. Chem.* **2009**, 4844–4849.
- (66) Biju, S.; Reddy, M. L. P.; Freire, R. O. *Inorg. Chem. Commun.* **2007**, *10*, 393–396.
- (67) Souza, A. P.; Paz, F. A. A.; Freire, R. O.; Carlos, L. D.; Malta, O. L.; Alves, S., Jr.; de Sá, G. F. *J. Phys. Chem. B* **2007**, *111*, 9228–9238.
- (68) Stewart, J. J. P. *J. Mol. Model.* **2013**, *19*, 1–32.
- (69) Allen, F. H. *Acta Crystallogr., Sect. B* **2002**, *58*, 380–388.

- (70) 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.
- (71) William, J. C. *Practical Nonparametric Statistics*; John Wiley & Sons: New York, 1971; pp 440–484.
- (72) Stewart, J. J. P. *MOPAC2012; Stewart Computational Chemistry*; Colorado Springs, CO, 2012. [HTTP://OpenMOPAC.net](http://OpenMOPAC.net).
- (73) Oliveira, C. A. F.; Silva, F. F.; Malvestiti, I.; Malta, V.R. S.; Dutra, J. D. L.; da Costa, N. B., Jr.; Freire, R. O.; Alves, S., Jr. *J. Sol. State Chem.* **2013**, *197*, 7–13.