J. Phys. Chem. A **2007,** *111,* 5015−5018

Structure Modeling of Trivalent Lanthanum and Lutetium Complexes: Sparkle/PM3

Nivan B. da Costa, Jr.,[†] Ricardo O. Freire,[‡] Alfredo M. Simas,*[,‡] and Gerd B. Rocha[§]

Departamento de Química, CCET, UFS, 49100-000-Aracaju, SE, Brazil, Departamento de Química Fundamental, CCEN, UFPE, 50590-470 − Recife, PE, Brazil, and Departamento de Química, CCEN, UFPB, 58.059-970 − João Pessoa, PB, Brazil

Received: November 1, 2006; In Final Form: March 16, 2007

The recently defined Sparkle model for the quantum chemical prediction of geometries of lanthanum(III) and lutetium(III) complexes within AM1 (*J. Phys. Chem. A* **2006**, *110*, 5897) has been extended to PM3. As training sets, we used the same two groups, one for each lanthanide, of 15 high-crystallographic-quality (*R* factor < 0.05 Å) complexes as was previously chosen to parametrize Sparkle/AM1. Likewise, in the validation procedure, we used the same Sparkle/AM1 validation sets of 60 additional La(III) and 15 additional Lu(III) complexes. The Sparkle/PM3 unsigned mean errors for all interatomic distances between the metal ions and the ligand atoms of the first sphere of coordination proved to be random around the means of 0.068 Å for lanthanum(III) and 0.076 Å for lutetium(III), thus being comparable to the respective Sparkle/AM1 values of 0.078 and 0.075 Å. Furthermore, effective-core-potential ab initio calculations on smaller subsets of such complexes led to similar accuracies. Sparkle/PM3 and Sparkle/AM1 are therefore made available to the researcher who must decide which of the models to use based on considerations of the impact of either PM3 or AM1 on the description of the ligands and the consequence of such a choice on the properties of interest.

Introduction

Sparkle/AM1 was introduced in 2005[1] as a new paradigm for semiempirical molecular orbital calculations of lanthanide complexes with the specific aim of predicting the geometry of the lanthanide ion and its coordination polyhedron. Later, it was confirmed that the accuracy of Sparkle/AM1, for ligands with directly coordinating oxygen or nitrogen atoms, is indeed equivalent to the best present-day effective-core-potential ab initio full geometry optimization calculations, while being hundreds of times faster.[2] Subsequently, Sparkle/AM1 was parametrized for La(III) and Lu(III).[3]

PM3[4,5] was introduced by Stewart in 1989 as a more accurate semiempirical model that gives lower average errors than AM1,[6] mainly for the enthalpies of formation. Like AM1, which became a landmark quantum chemical semiempirical method, PM3 also became very popular.[7] PM3 is currently widely available in a variety of quantum chemical software programs, both commercial and noncommercial.[8−17] As an example, PM3 has been recently used for the calculation of five- and six-coordinate oxo iron(IV) porphyrin complexes.[18]

Lanthanum complexes display several interesting properties, such as cytotoxicity, and find a variety of applications, for example, as antitumor agents,[19] in ionic liquids research,[20] and in the spectrophotometric determination of anions,[21] among many others. Likewise, there are several applications involving lutetium complexes, such as in photodynamic therapy,[22,23] in a supporting role in the detection of solar neutrinos,[24] etc.

Therefore, as another option to research, especially in the modeling and design of lanthanide complexes, we advance in this article Sparkle/PM3 parameters for trivalent lanthanum and lutetium complexes.

Results and Discussions

The parametrization procedure used for La(III) and Lu(III) was essentially the same as that described in our previous work on Sparkle/AM1,[1] in which we used a parametrization set comprising 15 carefully chosen complexes, from those available in the Cambridge Structural Database (CSD),[25−27] and will not be repeated here. The Sparkle/PM3 parameters obtained for La(III) and Lu(III) are listed in Table 1. In the validation procedure, we used the same previously defined sets[3] of 75 and 30 coordination compounds carefully chosen from CSD for La(III) and Lu(III), respectively.

As a measure of accuracy, we used the average unsigned mean error for each complex i, UME$_i$, defined as

$$\mathrm{UME}_i = \frac{1}{n_i}\sum_{j=1}^{n}|R_{i,j}^{\mathrm{CSD}} - R_{i,j}^{\mathrm{calc}}| \qquad (1)$$

where n_i is the number of ligand atoms directly coordinating the lanthanide ion. The total UME is obtained by summing UME$_i$ over all 75 and 30 coordination compounds considered for La(III) and Lu(III), respectively.

Two cases were examined: (i) UME$_{(\mathrm{Ln-L})}$ values involving only the interatomic distances R_j between the lanthanum or lutetium central ion and the atoms of the coordination polyhedron and (ii) UME values involving the interatomic distances R_j between the lanthanum or lutetium central ion and the atoms of the coordination polyhedron, as well as the interatomic distances R_j between all atoms of the coordination polyhedron. Tables 1S and 2S of the Supporting Information present the UME$_{(\mathrm{Ln-L})}$ and UME values for both Sparkle/PM3 and Sparkle/AM1 for the 75 complexes of La(III) and 30 complexes of Lu(III).

Assuming that the sparkle model is a good representation of the lanthanide ion as it interacts with the ligands, the distribution

---

* Corresponding author. E-mail: simas@ufpe.br.
[†] CCET, UFS.
[‡] CCEN, UFPE.
[§] CCEN, UFPB.

10.1021/jp0672104 CCC: \$37.00

TABLE 1: Parameters for the Sparkle/PM3 Model for the La(III) and Lu(III) Ions

	Sparkle/PM3	
parameter	La(III)	Lu(III)
GSS	55.706 393 859 7	56.617 726 884 4
ALP	2.079 032 744 1	4.058 582 242 9
a_1	0.947 861 518 2	1.121 653 611 2
b_1	7.236 601 171 7	7.940 801 601 1
c_1	1.854 378 629 8	1.706 026 973 1
a_2	0.317 767 190 4	0.330 565 510 1
b_2	8.522 400 629 2	7.301 381 021 9
c_2	3.075 247 472 4	2.909 046 517 8
EHEAT[a] (kcal mol^{-1})	928.9	1031.2
AMS (amu)	138.9055	174.9670

[a] Heats of formation of the La(III) and Lu(III) ions in Sparkle/PM3 obtained by adding the first three ionization potentials to the heat of atomization of each respective lanthanide.

of these UME values should be random around a mean whose value can be used as a measure of the accuracy of the model. Because the UME values are positive, defined in the domain (0, ∞), they should follow a gamma distribution that has the probability density function $\gamma(x;k,\theta)$, where $x > 0$ represents the UME values, $k > 0$ is the shape parameter, and $\theta > 0$ is the scale parameter of the gamma distribution. The expected value of the gamma distribution is simply $k\theta$.

To obtain the gamma distribution fit of the UME data, we estimated the shape and scale parameters with the method of maximum likelihood, and to assess the quality of the gamma distribution fit, we used the one-sample nonparametric Kolmogorov−Smirnov[28] test. This test allows us to verify statistically whether the distribution of the UME values is actually a gamma distribution indexed by the estimated parameters. In this case, the null hypothesis is that the UME values do follow the given gamma distribution. For the null hypothesis not to be rejected at the usual level of 5%, the p value of the test must thus be larger than 0.05. In addition, the higher the p value, whose maximum possible value is 1, the higher the probability that the UME deviations of the mean are random, and the more justifiable the use of the statistical tools employed here and, by extension, of the Sparkle model itself. If the p value is indeed larger than 0.05, the gamma distribution fit informs the probability of a UME, for an arbitrary complex, to belong to a given interval.

We now examine results for the Sparkle/PM3 model being presented in this article for La(III) and Lu(III), as well as the already published Sparkle/AM1 model for the same lanthanide ions.[3]

Figure 1a presents a gamma distribution fit of the UME$_{(\mathrm{La-L})}$ data for the Sparkle/PM3 model for La(III). As indicated in the figure, the p value is 0.655, thus indicating that the UME values are indeed significantly randomly distributed around the mean through a gamma distribution. To give a simple pictorial idea of this concept and also of where and how the actual UME values occurred, Figure 1 also presents superimposed histograms of the data with the number of bars chosen to best adjust the histogram to the curve obtained from the gamma distribution fit. For comparison purposes, Figure 1b presents the same gamma distribution fit for the UME$_{(\mathrm{La-L})}$ data for the already published Sparkle/AM1 model for La(III),[3] which gives a p value of 0.905.

Figure 2 presents gamma distribution fits of the UME$_{(\mathrm{Ln-L})}$ data for lutetium for both the present Sparkle/PM3 and the previously published Sparkle/AM1 model.[3] In both cases, the respective p values were well above the critical value of 0.05, thus validating the use of both Lu(III) Sparkle models.

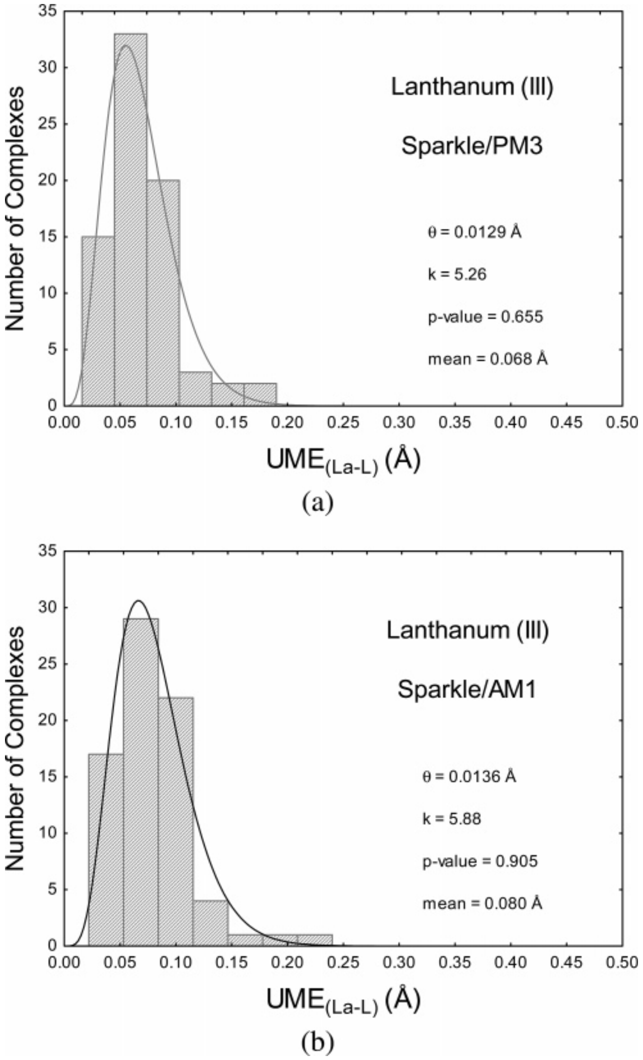

Figure 1. Probability densities of the gamma distribution fits of the UME$_{(\mathrm{Ln-L})}$ values for the La(III) Sparkle/PM3 and Sparkle/AM1 models, superimposed with histograms of the same data for all 75 La(III) complexes considered. k is the shape parameter and θ is the scale parameter of the gamma distribution; the p value is a measure of the significance of the gamma distribution fit; and mean is the expected value of the fitted gamma distribution, which was set to be equal to the arithmetic mean value of the 75 UME$_{(\mathrm{Ln-L})}$ values.

Table 2 presents unsigned mean errors for both Sparkle/PM3 and Sparkle/AM1 for various types of distances in the La(III) and Lu(III) complexes considered. The results indicate that there are no substantial differences between the two models for the two lanthanide ions considered. The distances between the lanthanide ion and its directly coordinated ligand atoms are predicted with significantly higher accuracy than the distances between atoms of the faces of the coordination polyhedron. This is fortunate because lanthanide ion−ligand atom distances are far more important for luminescent complex design.[29] For all complexes considered in the present work, the lanthanide−coordinated atom distances vary from 2.011 to 2.953 Å, whereas lanthanide−lanthanide distances range from 3.627 to 4.632 Å. Thus, the lanthanide−ligand atom average UME of 0.07 Å is 2.8% of the corresponding average distance of 2.50 Å. Likewise, the lanthanide−lanthanide average UME of 0.14 Å is 3.3% of the corresponding average distance of 4.19 Å. Therefore, the distances between two lanthanide ions, directly facing each other in a dilanthanide complex, present percentage errors that are comparable to those of the lanthanide−ligand atom distances.

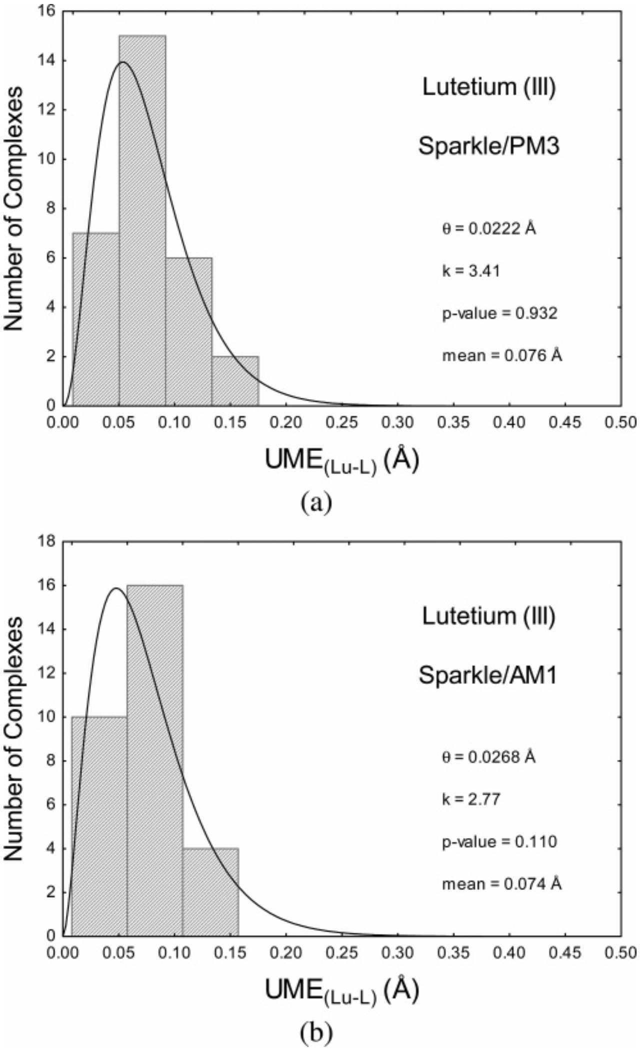

Figure 2. Probability densities of the gamma distribution fits of the $\mathrm{UME}_{(\mathrm{Ln-L})}$ values for the Lu(III) Sparkle/PM3 and Sparkle/AM1 models, superimposed with histograms of the same data for all 30 Lu(III) complexes considered. *k* is the shape parameter and θ is the scale parameter of the gamma distribution; the *p* value is a measure of the significance of the gamma distribution fit; and mean is the expected value of the fitted gamma distribution, which was set to be equal to the arithmetic mean value of the 30 $\mathrm{UME}_{(\mathrm{Ln-L})}$ values.

TABLE 2: Sparkle/PM3 and Sparkle/AM1 Unsigned Mean Errors (Å) for All Distances Involving the Central Lanthanide Ion, Ln, and the Ligand Atoms of the Coordination Polyhedron, L, for 75 La(III) Complexes and 30 Lu(III) Complexes

	type of distance					
model	Ln−Ln	Ln−O	Ln−N	L−L	Ln−L and Ln−Ln	Ln−L, Ln−Ln, and L−L
	Lanthanum (III)					
Sparkle/PM3	0.104	0.061	0.083	0.179	0.068	0.158
Sparkle/AM1	0.213	0.087	0.048	0.208	0.078	0.183
	Lutetium(III)					
Sparkle/PM3	0.176	0.083	0.054	0.145	0.076	0.130
Sparkle/AM1	0.222	0.084	0.047	0.170	0.075	0.150

Recently, a thorough study was carried out on the accuracies of predicting coordination polyhedron geometries using ab initio effective-core-potential (ab initio/ECP) calculations.[2] Contrary to what would normally be expected, it was reported that either an increase in the basis set or inclusion of electron correlation, or both, consistently enlarged the deviations and worsened the quality of the predicted coordination polyhedron geometries.[2] The conclusion was that RHF/STO-3G/ECP appears to be the most efficient model chemistry in terms of predicting coordination polyhedron crystallographic geometries from isolated lanthanide complex ion calculations.[2] However, because the study did not report calculations on La(III) and Lu(III) complexes, we chose to fully optimize the geometries of seven La(III) and five Lu(III) complexes with the model chemistry RHF/STO-3G/ECP. These complexes were selected to be representative of the various classes of ligands (β-diketones, nitrates, monodentates, bidentates, tridentates, polydentates, and dilanthanides) present in the validation set. At the time, we could not find any β-diketone and tridentate lutetium complexes in the CSD.

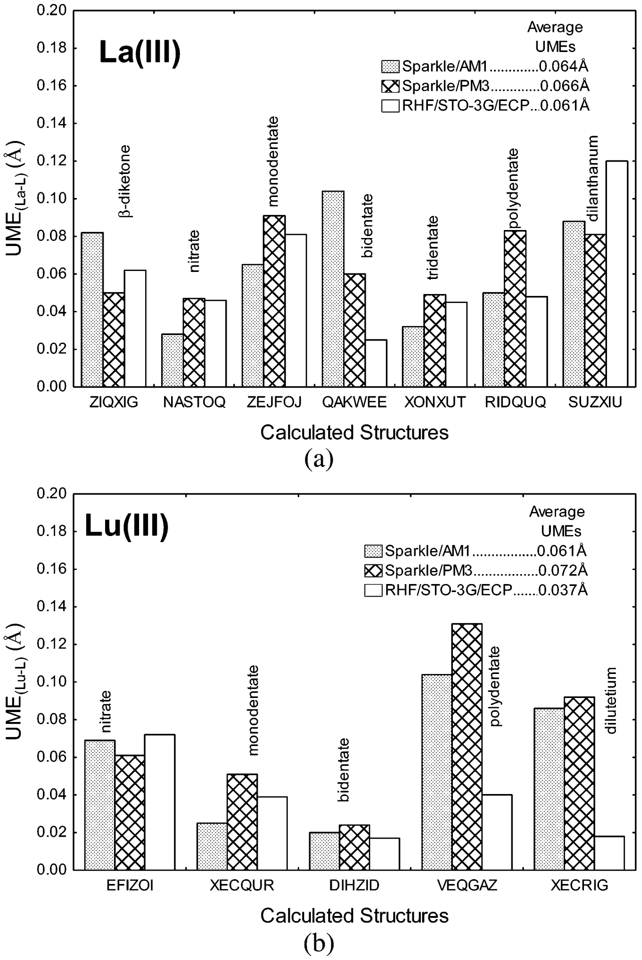

Figure 3. Unsigned mean errors, $\mathrm{UME}_{(\mathrm{Ln-L})}$ values (in Å), between the lanthanide central ion and the atoms of the coordination polyhedron obtained from Sparkle/AM1, Sparkle/PM3, and ab initio RHF/STO-3G/ECP calculations of the ground-state geometries for each of the representative La(III) and Lu(III) complexes, identified by their respective Cambridge Structural Database 2004 codes.

As indicated in Figure 3, the RHF/STO-3G/ECP average $\mathrm{UME}_{(\mathrm{Ln-L})}$ and UME values for the six La(III) complexes were calculated to be 0.061 and 0.143 Å, which are comparable to the respective Sparkle/PM3 values of 0.066 and 0.117 Å and to the corresponding Sparkle/AM1 values of 0.064 and 0.135 Å.

Likewise, from Figure 3, the RHF/STO-3G/ECP average $\mathrm{UME}_{(\mathrm{Ln-L})}$ and UME values for the five Lu(III) complexes calculated were 0.037 and 0.086 Å, which are comparable to the respective Sparkle/AM1 values of 0.061 and 0.112 Å and to the corresponding Sparkle/PM3 values of 0.072 and 0.106 Å.

Conclusion

The conclusions drawn for Sparkle/AM1, for both La(III) and Lu(III) complexes,[3] can safely be extended to Sparkle/PM3. Both models exhibit prediction accuracies for coordination polyhedron geometries that are competitive with those of present-day full geometry optimization ab initio/ECP calculations on lanthanide complexes, while being hundreds of times faster. Indeed, Sparkle/PM3 might prove useful for ligand design, where many different combinations of ligands and substituents must be evaluated in a combinatorial way. The choice of which model to use rests with the user who must evaluate the impact of using either PM3 or AM1 on the description of the properties of interest, as they depend on the ligands.

Acknowledgment. We appreciate the financial support from CNPq (Brazilian agency), as well as grants from PADCT and the Instituto do Milênio de Materiais Complexos. 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 La(III) and Lu(III) 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 both La(III) and Lu(III). Examples of Mopac93r2 crystallographic geometry input (.dat) and optimized geometry summary output (.arc) files from Sparkle/AM1 calculations for the La(III) complex EGOBAD, for the dilanthanum complex HETALA11, for the Lu(III) complex XEPLUZ, and for the dilutetium complex XEWYIH. Additional figures showing schematic two-dimensional representations of the 15 complexes that constituted both parametrization training sets. Additional tables with Sparkle/AM1 unsigned mean errors for each of the 75 La(III) and 30 Lu(III) complex CSD[25−27] high-quality crystallographic structures. Chemical structures of the seven La(III) and five Lu(III) complexes that had their geometries fully optimized via RHF/STO-3G/ECP ab initio calculations and tables with their respective accuracies as compared to Sparkle/PM3 and Sparkle/AM1 results. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Freire, R. O.; Rocha, G. B.; Simas, A. M. *Inorg. Chem.* **2005**, *44*, 3299.

(2) Freire, R. O.; Rocha, G. B.; Simas, A. M. *J. Mol. Model.* **2006**, *12*, 373.

(3) Freire, R. O.; Costa, N. B., Jr.; Rocha, G. B.; Simas, A. M. *J. Phys. Chem. A* **2006**, *110*, 5897.

(4) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.

(5) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221−264.

(6) Dewar, M. J. S.; Zoebish, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

(7) Stewart, J. J. P. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley & Sons: New York, 1998.

(8) Stewart, J. J. P. *MOPAC 93.00 Manual*; Fujitsu Limited: Tokyo, Japan, 1993.

(9) 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.

(10) *HyperChem Professional*, version 7.51; Hypercube, Inc.: Gainesville, FL, 2002.

(11) *Spartan '04*; Wavefunction, Inc.: Irvine, CA, 2004.

(12) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California, San Francisco, CA, 2004.

(13) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21* (12), 1088.

(14) *VAMP*; Accelrys: San Diego, CA, 2002.

(15) Rowley, C.; Hassinen, T. *Ghemical for GNOME*, version 1.00; University of Iowa, Iowa City, IA, 2002.

(16) Jorgensen, W. L. BOSS—Biochemical and Organic Simulation System. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley & Sons: New York, 1998; Vol. 5, p 3281.

(17) *AMPAC 8.0 Manual*; Semichem, Inc.: Shawnee Mission, KS, 2004; see http://www.semichem.com.

(18) Katagi, T. *J. Mol. Struct. (THEOCHEM)* **2005**, *728* (1−3), 49.

(19) Li, F.-H.; Zhao, G.-H.; Wu, H.-X.; Lin, H.; Wu, X.-X.; Zhu, S.-R.; Lin, H.-K. *J. Inorg. Biochem.* **2006**, *100*, 36.

(20) Bhatt, A. I.; May, I.; Volkovich, V. A.; Collison, D.; Helliwell, M.; Polovov, I. B.; Lewin, R. G. *Inorg. Chem.* **2005**, *44* (14), 4934.

(21) Ramkumar, J. *Talanta* **2006**, *68* (3), 902.

(22) Verigos, K.; Stripp, D. C. H.; Mick, R.; Zhu, T. C.; Whittington, R.; Smith, D.; Dimofte, A.; Finlay, J.; Busch, T. M.; Tochner, Z. A.; Malkowicz, S. B.; Glatstein, E.; Hahn, S. M. *J. Environ. Pathol. Toxicol. Oncol.* **2006**, *25* (1−2), 373.

(23) Zhu, T. C.; Finlay, J. C.; Hahn, S. M. *J. Photochem. Photobiol. B* **2005**, *79* (3), 231.

(24) Gelis, V. M.; Chuveleva, E. A.; Firsova, L. A.; Kozlitin, E. A.; Barabanov, I. R. *Russ. J. App. Chem.* **2005**, *78* (9), 1420.

(25) Allen, F. H. *Acta Crystallogr. B* **2002**, *58*, 380.

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

(27) Allen, F. H.; Motherwell, W. D. S. *Acta Crystallogr. B* **2002**, *58*, 407.

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

(29) 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.