Empirical Charge Scheme for Transition Metals and Lanthanoids: Development and Applications

Bodo R. Martin, Peter Comba
Inorganic Chemistry
Heidelberg University
Heidelberg, Germany
bodo.martin@aci.uni-heidelberg.de

Peter Comba
IWR
Heidelberg University
Heidelberg, Germany
peter.comba@aci.uni-heidelberg.de

Parameters for the fluctuating charge model are derived for In(III), La(III), Lu(III) and Bi(III), based on x-ray structural data and DFT single point calculations. A number of density partitioning schemes is tested.

atomic charges; transition metals; lanthanoids; force fields; fluctuating charge method;

I. INTRODUCTION

Atomic charges are a useful concept in computational chemistry. They are widely used to model electrostatic interactions in force field software. [1, 2] Even though polarization is known to be necessary for an adequate description of electrostatic interactions, few mayor software codes to date make use of these modified potentials [3]. Moreover, most force field programs are only parameterized for organic compounds and have poor or non-existent support for transition metal or lanthanoide complexes.

The aim of this work is to develop an accurate and fast model for atom charges, for elements up to and including actinides, implemented in the molecular mechanics program Momec [4]. This classical molecular mechanics software, developed in our group, is especially tailored for the calculation of coordination compounds. We have previously developed a fluctuating charge model, i.e. a geometry-dependent atomic charge model, for complexes containing copper and iron [5], and extend this here to complexes containing Indium, Lanthanum, Lutetium and Bismuth in the oxidation state III: In(III), La(III), Lu(III) and Bi(III). In a subsequent step, the atomic charges obtained for these elements can be the base for descriptors in a QSPR prediction of logD values, e.g. for complexes for PET imaging and radionuclide treatment [6]. Our focus here will be the determination of atomic charges from DFT calculations, based on approx. 200 compounds from the CSD [7].

II. THE FLUCTUATING CHARGE MODEL

In the charge equilibration (QEq) method, the charge distribution is calculated via (1)

$$\chi_{eq} = \chi_i^0 + 2J_{ii}^0 q_i + \sum_{j>i} q_j J_{ij} \tag{1}$$

where χ_i^0 is the electronegativity and J_{ii} is the idempotential of atom i. The interaction term J_{ij} is calculated as a two-center Coulomb integral [8]. The valence state ionization energies of an element vs. its ionization can be approximated via a quadratic fit as in (2)

$$E^a(q) = E^a(q=0) + A^a q + B^a q^2 \tag{2}$$

Where the superscript "*a*" denotes parameters for the isolated atom. Assuming a similar relationship for atoms as part of a molecule leads to the equivalent equation (3)

$$E^m(q) = E^m(q=0) + A^m q + B^m q^2 \tag{3}$$

The expressions for the parameters A and B in the molecule are now described as the atomic parameters and linear increments (4)

$$A^m = A^a + \Delta A \quad B^m = B^a + \Delta B \tag{4}$$

First and second derivatives of (3) gives

$$\frac{dE^m(q)}{dq} = A^m + 2B^m q \quad \frac{d^2 E^m(q)}{dq^2} = 2B^m$$

In a molecule, the total electrostatic energy can be expressed as a function of partial atomic charges, (5)

$$E(q_1 \ldots q_N) = \sum_{i=1}^{N} \left(\begin{array}{c} E(0)_i^{\mathrm{m}} + \left(\dfrac{\partial E_i^{\mathrm{m}}}{\partial q} \right)_{q_i} q_i \\ + \dfrac{1}{2} \left(\dfrac{\partial^2 E_i^{\mathrm{m}}}{\partial q^2} \right)_{q_i} q_i^2 \end{array} \right) \quad (5)$$

$$+ \sum_{i=1}^{N} \sum_{j>i} q_i q_j D_{ij}$$

Substituting we get (6)

$$E(q_1 \ldots q_N) = \sum_{i=1}^{N} \left(E(0)_i^{\mathrm{m}} + A_i^{\mathrm{m}} q_i + 3 B_i^{\mathrm{m}} q_i^2 \right)$$

$$+ \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i} q_i q_j D_{ij} \quad (6)$$

Taking the first derivative of (6) with respect to partial charge q_i gives the Iczkowski-Margrave electronegativity of atom i (7)

$$\chi(q_i) = \left(\frac{\partial E}{\partial q} \right)_{q_i} = A_i^{\mathrm{m}} + 6 B_i^{\mathrm{m}} q_i + \frac{1}{2} \sum_{j \neq i} q_j D_{ij} \quad (7)$$

With Sanderson`s electronegativity equalization principle (8),

$$\chi_1 = \chi_2 = \cdots \chi_N = \chi_{\mathrm{eq}} \quad (8)$$

conservation of the total charge Q_{tot} and by dropping the "m" superscripts we arrive at the matrix equation:

$$\begin{pmatrix} 6B_1 & \frac{1}{2}D_{12} & \cdots & \frac{1}{2}D_{1N} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{2}D_{N1} & \frac{1}{2}D_{N2} & \cdots & 6B_N & -1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} q_i \\ q_2 \\ \vdots \\ q_N \\ \chi_{\mathrm{eq}} \end{pmatrix} = \begin{pmatrix} -A_1 \\ -A_2 \\ \vdots \\ -A_N \\ Q_{\mathrm{tot}} \end{pmatrix}$$

The interatomic term D_{ij} is expressed as (9) [5]

$$D_{ij} = \left(\sqrt{r_{ij}^2 + \left(\frac{1}{2\sqrt{B_i B_j}} \right)^2} \right)^{-1} \quad (9)$$

The atomic parameter fits A^a and B^a can easily be calculated from ionization data; one example for Lu is shown in Fig. 1. The linear corrections for the atoms in a molecule, ΔA and ΔB, are the two parameters for the fluctuating charge model.
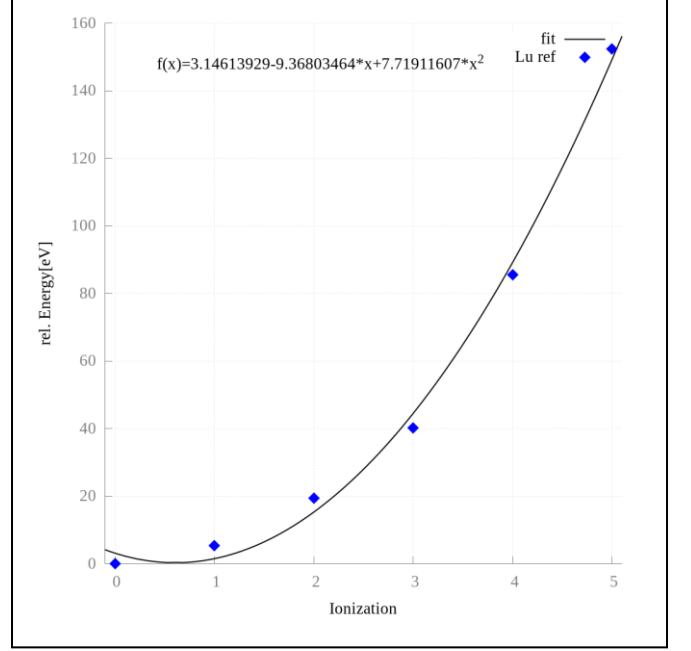
(5)

Fig. 1.   Lu energy vs. ionization fit to experimental data [9]

III.   METHODOLOGY

Starting from crystal structures, e.g. from structures obtained from the CSD [7], single point quantum mechanical calculations are followed by an iterative Hirshfeld treatment with Horton [10], or Chargemol [12], to obtain reference atomic charges. The reference charges are then approximated with our 2-parameter fluctuating charge model via a simplex minimization. Additionally, the geometries of the reference data set are reproduced by fitting the force field parameters of the Momec molecular mechanics program. The combined force field thus allows for a fast geometry optimization and charge prediction for complexes with elements up to Ac. The fluctuating charge method is of complexity $O(n^2)$; it is planned to reduce the computational cost by a divide and conquer algorithm. From an MM-calculated structure with atomic charges, one can then derive descriptors for e.g. logD or NMR isotropic shift predictions. A single point calculation with Gaussian typically takes a few hours on 16 cpu cores; Horton calculations take a similar time on one cpu core, but, dependent on the accuracy required, may require up to 100 GB of memory.
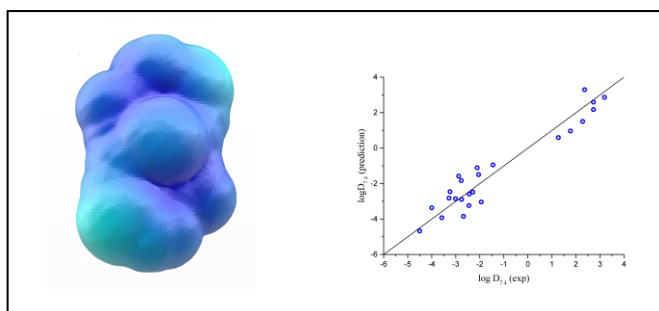
Fig. 2. SAS surface charge example (CSD code AGOVIB, left) and Cu complexes logD fit similar to [6].

IV. LOGD PREDICTION

We have previously demonstrated [6] that by using the above empirical charge scheme, it is possible to derive accurate predictions for logD values for a range of copper complexes, using a five-parameter model: the ovality O_v, the solvent accessible surface A, the local polarity π, the total variation from the mean positive and negative potentials σ_{tot}, and the polar surface area A_{pol}. With the parameterized charge model, we can now extend the logD prediction to all Transition Metal and Lanthanoide complexes, analogous to the surface charge distribution and logD fits from previous work [6] (cf. Fig. 2).
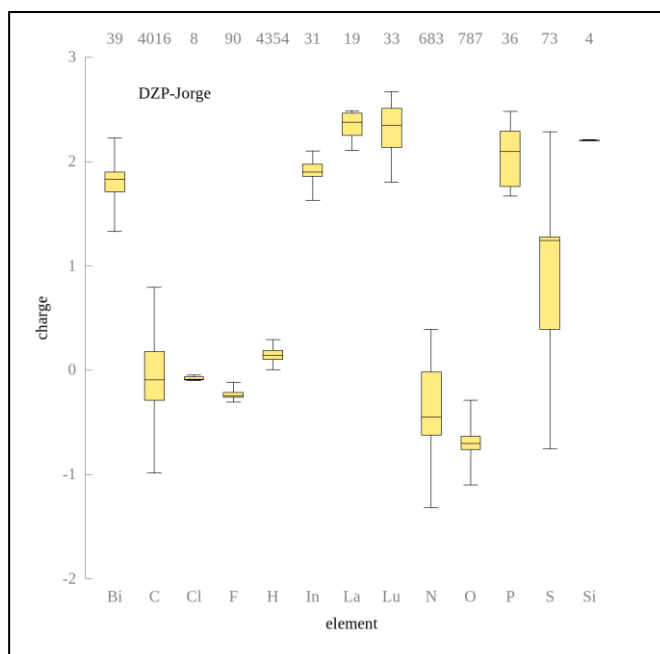

Fig. 3. Boxplots of reference atomic charges, calculated with Horton; Gaussian 09 with Jorge DZP [14] basis set. Numbers at top of figure are atom counts. Outliers (above/below 95%) not shown for clarity.

V. RESULTS

Fig. 3 shows iterative Hirshfeld atomic charges of a set of complexes of In(III), La(III), Lu(III) and Bi(III), for a range of coordination numbers, obtained via Horton 2.1.0b1, after calculating single points with B3LYP Gaussian 09 [11]; Fig. 4 shows DDEC6 [12] charges derived from the same basis set. Results for Orca 3.0.3 [13] with SVP/ZORA and TZP/ZORA basis sets are not shown for brevity.

The calculation of atomic charges depends on the DFT method, the basis set and the density partitioning scheme. As the data set for Fig 4 is larger than for Fig 3, the two cannot be directly compared; moreover, for the subsequent MM parameterization, consistency of the reference charges is more important than the absolute values, i.e. similar complexes should have similar atomic charges, and charge outliers should have a sufficiently different chemical environment.
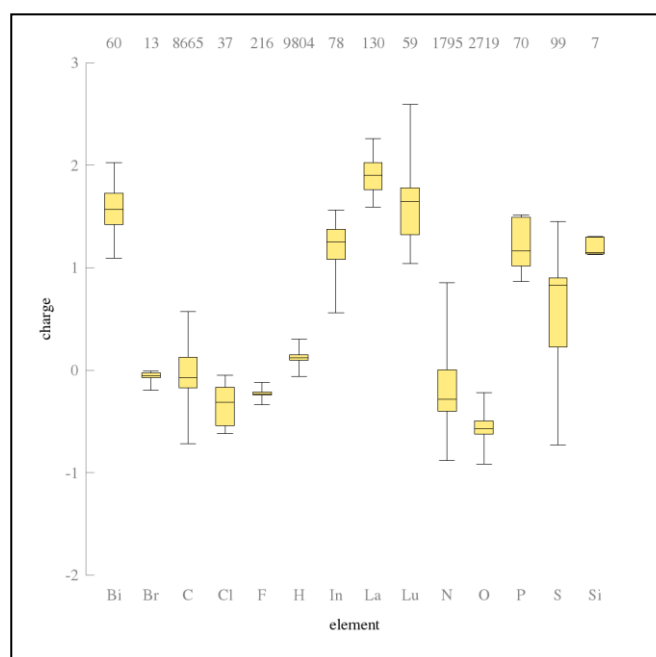

Fig. 4. Boxplot of reference atomic DDEC6 [12] charges; Gaussian 09 with Jorge DZP [14] basis set. Numbers at top of figure are atom counts. Outliers (above/below 95%) not shown for clarity.

VI. CONCLUSION/OUTLOOK

Calculation of reference atomic charges for elements in the 5th row and above is not straightforward; only a few general basis sets exist, and relativistic effects have to be accounted for either implicitly or via ZORA or DKH. Here we present two methods for obtaining charges, after post-processing via Horton [10] and Chargemol [12], to obtain iterative Hirshfeld and DDEC6 charges. We are still in the process of collecting data to decide on the best set of reference charges for our fluctuating charge model, which we have shown by a model parameterization and previous work to be able to reproduce atomic charges accurately. In a next step, we will extend the

logD prediction, based on descriptors obtained with this charge model. The NMR prediction based on charges may require higher moments; use of higher charge moments is currently under development.

REFERENCES

[1] D.A. Case, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman (2017), AMBER 2017, University of California, San Francisco.

[2] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers", SoftwareX, vol. 1-2, pp. 19-25, 2015.

[3] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J.W. Ponder, P. Ren, "The Polarizable Atomic Multipole-based AMOEBA Force Field for Proteins," J. Chem. Theory Comput. Vol 9(9), pp. 4046-4063, 2013.

[4] P. Comba, T. W. Hambley, B. Martin, "Molecular modeling of inorganic compounds", Wiley-VCH, Weinheim 2009.

[5] P. Comba, B. Martin, A. Sanyal, "An efficient fluctuating charge model for transition metal complexes," J. Comput. Chem. vol. 34, pp. 1598-1608, 2013.

[6] P. Comba, B. Martin, A. Sanyal, H. Stephan, "The computation of lipophilicities of Cu PET systems based on a novel approach for fluctuating charges," Dalton Trans. vol. 42(31), pp. 11066-11073, 2013.

[7] C.R. Groom, I.J. Bruno, M.P. Lightfoot, S. C. Ward, "The Cambridge Structural Database" Acta Cryst. vol. B72, pp. 171-179, 2016.

[8] A. K. Rappe, W. A. Goddard III, "Charge Equilibration for Molecular Dynamics Simulations," J. Phys. Chem. vol. 95, pp.3358-3363, 1991.

[9] A. Kramida, Yu. Ralchenko, J. Reader, NIST ASD Team [2017, August 23], "NIST Atomic Spectra Database (ver. 5.2)" https://physics.nist.gov/asd , 2014.

[10] T. Verstraelen, P. Tecmer, F. Heidar-Zadeh, E. González-Espinoza, C. M. Chan, T. D. Kim, K. Boguslawski, S. Fias, S. Vandenbrande, D. Berrocal, P.W. Ayers, "HORTON 2.1.0", http://theochem.github.com/horton/

[11] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, and G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, and J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian 09 , Revision D.01.

[12] T. A. Manz, N. G. Limas, "Introducing DDEC6 atomic population analysis: part 1. Charge partitioning theory and methodology," RSC Adv., vol. 6(53), pp. 47771-47801, 2016.

[13] F. Neese, "The ORCA program system", Rev.: Comput. Mol. Sci., vol. 2, pp. 73-78. 2002.

[14] F. E. Jorge, L. S. C. Martins, M. L. Franco, "All-electron double zeta basis sets for the lanthanides: Application in atomic and molecular property calculations," Chem. Phys. Lett. , vol 642, pp. 84-88, 2016.