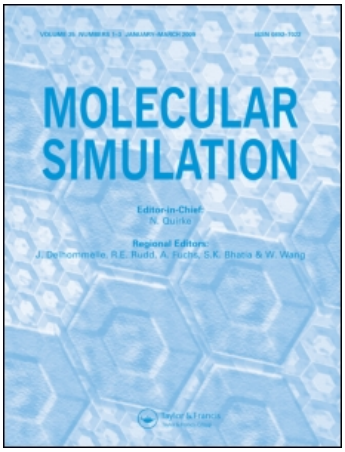

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

A point-charge model for electrostatic potentials based on a local projection of multipole moments

A. T. B. Gilbert[a]; P. M. W. Gill[a]
[a] Research School of Chemistry, Australian National University, Australia

To cite this Article Gilbert, A. T. B. and Gill, P. M. W.(2006) 'A point-charge model for electrostatic potentials based on a local projection of multipole moments', Molecular Simulation, 32: 15, 1249 — 1253
To link to this Article: DOI: 10.1080/08927020601052880
URL: http://dx.doi.org/10.1080/08927020601052880

Molecular Simulation, Vol. 32, No. 15, 30 December 2006, 1249–1253

Taylor & Francis
Taylor & Francis Group

A point-charge model for electrostatic potentials based on a local projection of multipole moments

A. T. B. GILBERT* and P. M.W. GILL

Research School of Chemistry, Australian National University, ACT 0200, Australia

(Received July 2006; in final form October 2006)

We introduce a method for obtaining atomic point-charges that yield accurate representations of the electrostatic potentials (ESP) of large systems. The method relies on a decomposition of the density and subsequent projection of the multipole moments of the density components onto neighbouring atomic sites. The resulting local multipole-derived charges (LMDCs) are well-defined, do not require sampling of the ESP at grid points around the molecule and provide a good description of the electrostatic potential. This local approach circumvents the numerical problems that arose in our original method which was designed to find the optimal atomic charge representation of the ESP of a system outside the electron density.

Keywords: Point-charge model; Electrostatic potentials; Multipole moments; Atomic point-charges

1. Introduction

Electrostatic forces play a dominant role in intermolecular interactions and an accurate, but simple, representation of the electrostatic potential is essential for the efficient simulation of molecular dynamics. Atomic point-charge models, adorned with various bond parameters, form the basis of molecular mechanics methods and are a proven compromise between cost and accuracy in this respect. However, the wide use of atomic charges belies their lack of a unique definition, a deficiency that has lead to a plethora of different approaches being advocated in the literature [1–13]. The relative usefulness of these methods is ultimately determined by the intended use of the resulting atomic charges.

If intermolecular interactions are of interest, then it makes sense to target the modelling of the electrostatic potential. Methods involving the direct fitting of point-charges to electrostatic potential data are well established [2,4,6–11]. Nevertheless, these methods are cumbersome in that the charges obtained depend on the choice of points around the molecule at which the ESP is evaluated, and in practice only a small subspace around the molecule is considered.

Recently, we proposed an atomic point-charge model that yields the best possible reproduction of the electrostatic potential outside the electron density [13]. The point-charges were determined by fitting to as many molecular multipole moments as possible rather than explicitly fitting to data obtained by evaluating the ESP at points in space. It was shown that this approach is well-defined, efficient and leads to excellent agreement between the true ESP and that obtained from the point-charges.

Our approach, however, is not without its limitations. Numerical problems can arise when calculating the atomic charges in large systems. These problems can be traced to the very large high-order moments which result in the fitting equations becoming ill-conditioned. Similar ill-conditioned equations arise when fitting point-charges to ESP data [14]. Additionally, charges associated with atoms "buried" within the molecule can have extremely large magnitudes, and although these charges reproduce the ESP well, they do so through a delicate mutual cancellation.

In this paper we introduce a modification of our original approach which effectively avoids these problems. This new approach shares some similarities to the distributed multipole analysis (DMA) [15] and Mulliken population analysis [1] in that it relies on a decomposition of the density in terms of basis function pairs. However, unlike the DMA, our approach restricts the representation to atom-centred point-charges and uses multiple centres to model the multipoles of each basis function pair. These

*Corresponding author. Email: Andrew.gilbert@anu.edu.au

Molecular Simulation
ISSN 0892-7022 print/ISSN 1029-0435 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/08927020601052880

restrictions ensure that the computational cost of using the model remains low, even in applications to large systems.

2. Theory

Many quantum chemical methods rely on the use of nuclear-centred Gaussian basis functions. Within these methods the electron density can be written

$$\rho(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) \tag{1}$$

where P is the density matrix and ϕ_μ are the Gaussian basis functions. The density is, therefore, naturally decomposed into density elements $\rho_{\mu\nu} = P_{\mu\nu}\phi_\mu\phi_\nu$ that are distributed within the molecule, either at nuclear centres (if ϕ_μ and ϕ_μ are concentric) or on internuclear axes (otherwise).

Using the Legendre expansion of the Coulomb operator, the potential of the density element $\rho_{\mu\nu}$ located at $\mathbf{r}_{\mu\nu}$ can be written

$$\tilde{V}_{\mu\nu}(\mathbf{r}) \simeq \sum_{\ell,m} \|\mathbf{r} - \mathbf{r}_{\mu\nu}\|^{-2\ell-1} R_\ell^m(\mathbf{r} - \mathbf{r}_{\mu\nu})\left\langle R_\ell^m\right\rangle_{\mu\nu} \tag{2}$$

where R_ℓ^m are the real, regular solid harmonics given by

$$R_\ell^0(\mathbf{r}) = \sqrt{\frac{4\pi}{2\ell+1}} r^\ell Y_\ell^0(\mathbf{r}/r)$$

$$R_\ell^{\pm m}(\mathbf{r}) = \sqrt{\frac{\pm 2\pi}{2\ell+1}} r^\ell \left[Y_\ell^{|m|}(\mathbf{r}/r) \pm Y_\ell^{|m|}(\mathbf{r}/r)\right] \tag{3}$$

where Y_ℓ^m are the usual spherical harmonics without the Condon–Shortley phase [16]. $\langle R_\ell^m\rangle_{\mu\nu}$ are the multipole moments of the density element, about its origin, given by

$$\left\langle R_\ell^m\right\rangle_{\mu\nu} = \int \rho_{\mu\nu}(\mathbf{r} - \mathbf{r}_{\mu\nu}) R_\ell^m(\mathbf{r})\mathrm{d}\mathbf{r} \tag{4}$$

Using the linearity of the potential operator the total potential of the molecule can be written

$$\tilde{V}(\mathbf{r}) = \sum_{\mu\nu} \tilde{V}_{\mu\nu}(\mathbf{r}) \tag{5}$$

where the tilde is to remind us that the expression is strictly valid only outside the electron density.

We observe that the potential in equation (2) is dominated by the low-order multipole moments and, in the spirit of our original multipole derived charges (MDCs) method [13], we will target these for each $\rho_{\mu\nu}$. The essence of the local multipole-derived charges (LMDC) method is that we stipulate that, to some order $L_{\max}$, the $\langle R_\ell^m\rangle_{\mu\nu}$ (for $\ell \le L_{\max}, |m| \le \ell$) must be reproduced by partial point charges located on the nuclei *closest* to $\mathbf{r}_{\mu\nu}$. By reproducing the multipole moments rather than directly fitting charges to ESP data on a grid, we avoid the problems associated with the grid-base approaches such as rotational dependence.

If there is no local symmetry about $\mathbf{r}_{\mu\nu}$, then the number of point charges required is given by $(L_{\max}+1)^2$ and the partial point-charges, $q_{\mu\nu}^i$, satisfy the following

$$\sum_i q_{\mu\nu}^i R_\ell^m(\mathbf{A}_i) = \left\langle R_\ell^m\right\rangle_{\mu\nu} \tag{6}$$

where $\mathbf{A}_i$ is the location of the ith atom and the the sum is over all atomic centres closest to $\mathbf{r}_{\mu\nu}$. The LMDC charge associated with atom i is given by $q_i = \sum_{\mu\nu} q_{\mu\nu}^i$. In the original MDC method [13] the highest order of multipole required to uniquely determine the atomic charges was fixed by the number of atoms in the system. For large systems this entailed using high-order multipoles which caused the associated matrix equations to become poorly conditioned. In the LMDC approach we are free to restrict the highest order of multipole used to lower values, and yet still have uniquely determined charges on the atoms.

Our algorithm batches over all shell-pairs arising from the same pair of nuclear centres and also only considers those shell-pairs that are significant, as determined by the two-electron integral cut-off threshold. The number of fitting equations that have to be solved therefore scales as $\mathcal{O}(N_{\text{atoms}})$. In practice, the cost of determining the LMDCs is a small fraction of the cost of the self-consistent field calculation required to obtain the density.

3. Results and discussion

We have implemented the above algorithm within a development version of the QCHEM [17] package and have applied it to the base-pair system adenine–thymine. The structure was obtained from the SPARTAN [18] package and was capped with hydrogen atoms to give a system with a total of 62 atoms. The HF/6-31G(d) density was computed and the corresponding shell-pair decomposition of the density was used to compute the LMDCs with the maximum order of multipole, $L_{\max}$, set to 5. This ensures all molecular multipoles up to and including 32-poles are reproduced, and requires projection of the $\langle R_\ell^m\rangle_{\mu\nu}$ onto the nearest 36 nuclei. For comparison the Mulliken and CHELPG [7] charges were also computed using the same density. The CHELPG charges were obtained from the GAMESS [19] package.

Figure 1 shows the error in the ESP made by each point-charge approximation. The cross section is taken through the xy plane of the molecule which approximately corresponds to the plane of the hydrogen bonds. It is well-known that the Mulliken charges yield a poor representation of the ESP and this is illustrated by the large spacings between the contours in the error plot. CHELPG, on the other hand, is designed to provide a good model for the ESP just outside the van der Waals (vdW) surface and consequently yields a good overall potential in the region shown. The error in the LMDC potential is much larger close to the vdW surface and this is a reflection of the fact that the multipole expansion, upon

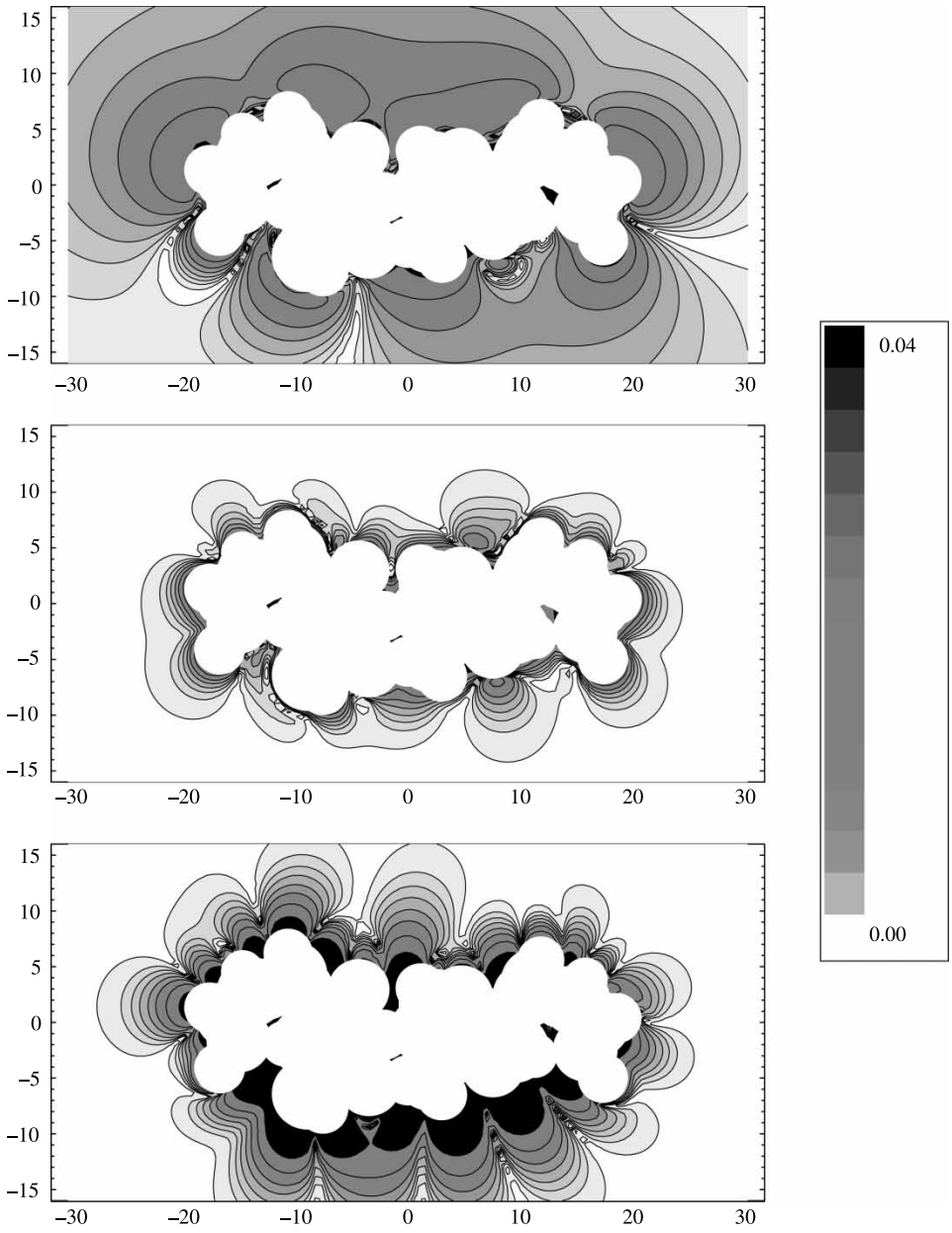

Figure 1. Contour plots showing the absolute error between the HF/6-31G(d) ESP for the AT base pair and that obtained using the Mulliken (top) CHELPG (middle) and LMDCs (bottom). The white area in the centre indicates the vdW volume. All values are in atomic units.

which the LMDC approach is based, is an asymptotic series and the error is therefore much larger in regions close to the electron density.

Further from the vdW surface the LMDC potential becomes more competitive. This is shown more clearly in figures 2 and 3 which show log–log plots of the absolute error in the potential along the positive regions of the *x* and *y* axes. The singularities in these plots correspond to points where the potentials cross the true potential and hence the error goes to zero. The major axis of the molecule is aligned along the *x* axis and figure 2 shows that along this axis the ESP of the LMDCs is superior to both of the other charge models. Along the *y* axis (figure 3) the reproduction of the potential is less satisfactory close to the vdW surface. The reason for this is that the expansion of the potential in terms of molecular multipole moments is only valid outside a sphere with a radius large enough to envelope all the electron density of the molecule. In an anisotropic system such as this, the radius is determined by the extent of the density along the principle axis, in this case the *x* axis, and we therefore, expect the LMDCs to perform better in this direction. In other directions, along the *y* axis for example, there will be regions of space where the density is effectively zero, but yet are still within the sphere where the multipole expansion is not valid.

By construction, the LMDCs accurately reproduce all the molecular multipole moments of this system up to 5th order. No attempt is made to model the 6th order moments

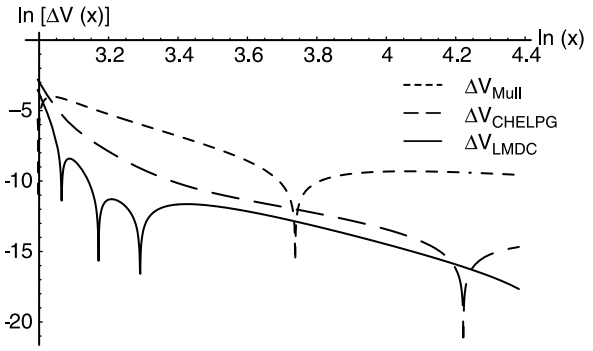

Figure 2. Plot showing detail of the error between the HF/6-31G(d) ESP and the model potentials along the x axis. The region plotted corresponds to that outside the vdW surface. All values are in atomic units.

and these are in error by as much as 100%. The error introduced by this neglect of the 64-poles decays as r^{-7} and corresponds to a slope of -7 on the log–log plots. The slope of the best fit line to the LMDC curve in figure 2 for x values between 40 and 80 bohr is -7.01 indicating that this asymptotic limit has been obtained. The corresponding slopes for the other two curves are harder to estimate due to the singularities in the curves, but they are clearly not as steep and this is consistent with both the Mulliken and CHELPG charges making errors in the lower-order moments. In figure 3 the corresponding LMDC slope is -5.7 and therefore the asymptotic limit has not yet been reached on the length scale shown.

Figure 3 also clearly shows the cross-over point at $y \approx e^{3.25} \approx 25.8$ bohr where the LMDC potential becomes better than the CHELPG potential. The decision as to which of these two model potentials to use will clearly depend on the length-scales that are important for the particular application, with long-range effects favouring the LMDCs. An accurate representation of the long range potential such as that provided by the LMDCs is important for modelling some properties (see for example reference [20]).

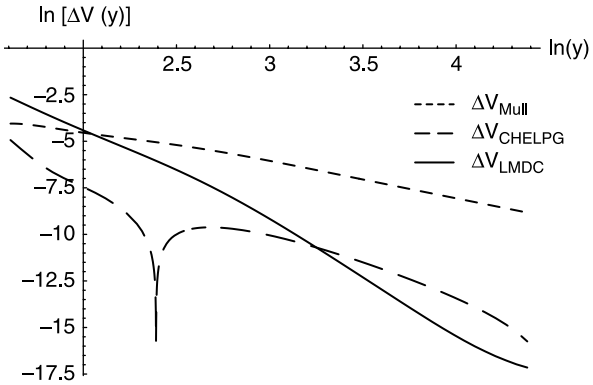

Figure 3. Plot showing detail of the error between the HF/6-31G(d) ESP and the model potentials along the y axis. The region plotted corresponds to that outside the vdW surface. All values are in atomic units.

For this system, the original MDC approach also attempts to fit the 64-poles and least-squares fits the 128-poles to determine the remaining 13 charges. The resulting matrix equation has a condition number of 1.03×10^{28} and it is not possible to solve this using our Fortran implementation. In contrast, the condition numbers for the fitting equations in the LMDC approach do not exceed 9.53×10^{8} and do not present any numerical difficulties.

4. Conclusions

We have presented a method for obtaining atomic charges for large molecules that yields an accurate description of the electrostatic potential, especially in the asymptotic regions. By decomposing the density into shell-pair components, and projecting the multipoles of each of these components onto only neighbouring atom sites, greater control over the maximum order of multipole required is achieved, leading to matrix equations that are much better conditioned. The LMDC charges do not reproduce the maximum number of multipoles possible, and are therefore theoretically inferior to the original prescription, however they are are tractable for large systems and share many of the positive qualities of the MDC approach. For example, the method is efficient, yields an accurate representation of the ESP and does not require the ESP to be evaluated on grid points surrounding the molecule. Unlike the original MDC approach, LMDCs require a decomposition of the density and, therefore, cannot be determined from experimentally determined multipoles. In practice this is not a problem as the high-order moments required in either method are difficult to obtain empirically. The LMDC charges outperform others in regions far from the molecule and are ideally suited to applications involving long-range electrostatic forces.

Acknowledgements

ATBG would like to thank Dr C.Y. Lin for assistance running the GAMESS package.

References

[1] R.S. Mulliken. Electronic population analysis on LCAOMO molecular wave functions. I. *J. Chem. Phys.*, **23**, 1833 (1955).
[2] U.C. Singh, P.A. Kollman. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.*, **5**, 129 (1984).
[3] A.E. Reed, R.B. Weinstock, F.A. Weinhold. Natural population analysis. *J. Chem. Phys.*, **83**, 735 (1985).
[4] L.E. Chirlian, M.M. Francl. Atomic charges derived from electrostatic potentials—a detailed study. *J. Comput. Chem.*, **6**, 894 (1987).
[5] R.F.W. Bader. *Atoms in Molecules: A Quantum Theory*, Clarendon Press, Oxford (1990).
[6] B.H. Besler, K.M. Merz, P. Kollman. Atomic charges derived from semiempirical methods. *J. Comput. Chem.*, **11**, 431 (1990).
[7] C.M. Breneman, K.B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials—the need for high

sampling density in formamide conformational-analysis. *J. Comput. Chem.*, **11**, 361 (1990).

[8] C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.*, **97**, 10269 (1993).

[9] J.G. Ángyán, C. Chipot. A comprehensive approach to molecular charge-density models—from distributed multipoles to fitted atomic charges. *Int. J. Quant. Chem.*, **52**, 17 (1994).

[10] K. Hinsen, B. Roux. A potential function for computer simulation studies of proton transfer in acetylacetone. *J. Comput. Chem.*, **18**, 368 (1997).

[11] E. Sigfridsson, U. Ryde. Comparison of methods for deriving atomic charges from the electrostatic potential and moments. *J. Comput. Chem.*, **19**, 377 (1998).

[12] M. Swart, P.T. van Duijnen, J.G. Snijders. A charge analysis derived from an atomic multipole expansion. *J. Comput. Chem.*, **22**, 79 (2001).

[13] A.C. Simmonett, A.T.B. Gilbert, P.M.W. Gill. An optimal point-charge model for molecular electrostatic potentials. *Mol. Phys.*, **103**, 2789 (2005).

[14] M.M. Francl, C. Carey, L.E. Chirlian, D.M. Gange. Charges fit to electrostatic potentials 2. Can atomic charges be unambiguously fit to electrostatic potentials? *J. Comput. Chem.*, **17**, 367 (1996).

[15] A.J. Stone. Distributed multipole analysis, or how to describe a molecular charge-distribution. *Chem. Phys. Lett.*, **83**, 233 (1981).

[16] E.W. "Spherical Harmonic." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/SphericalHarmonic.html Weisstein

[17] Y. Shao, L. Fusti–Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S.T. Brown, A.T.B. Gilbert, L.V. Slipchenko, S.V. Levchenko, D.P. O'Neill, R.A. DiStasio Jr., R.C. Lochan, T. Wang, G.J.O. Beran, N.A. Besley, J.M. Herbert, C.-Y. Lin, T. Van Voorhis, S.-H. Chien, A. Sodt, R.P. Steele, V.A. Rassolov, P.E. Maslen, P.P. Korambath, R.D. Adamson, B. Austin, J. Baker, E.F.C. Byrd, H. Daschel, R.J. Doerksen, A. Dreuw, B.D. Dunietz, A.D. Dutoi, T.R. Furlani, S.R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R.Z. Khalliulin, P. Klunzinger, A.M. Lee, M.S. Lee, W.Z. Liang, I. Lotan, N. Nair, B. Peters, E.I. Proynov, P.A. Pieniazek, Y.M. Rhee, J. Ritchie, E. Rosta, C.D. Sherrill, A.C. Simmonett, J.E. Subotnik, H.L. Woodcock III, W. Zhang, A.T. Bell, A.K. Chakraborty, D.M. Chipman, F.J. Keil, A. Warshel, W.J. Hehre, H.F. Schaefer III, J. Kong, A.I. Krylov, P.M.W. Gill, M. Head-Gordon. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.*, **8**, 3172 (2006).

[18] Spartan '04 Macintosh v1.0.1, Wavefunction, Inc. 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612 USA

[19] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery. General atomic and molecular electronic structure system. *J. Comput. Chem.*, **14**, 1347 'Advances in electronic structure theory: GAMESS a decade later' M.S. Gordon, M.W. Schmidt in "Theory and Applications of Computational Chemistry, the First Forty Years" C.E. Dykstra, G. Frenking, K.S. Kim, G.E. Scuseria, Elsevier, Amsterdam, (2005) (1993).

[20] A. Krushelnitsky. Intermolecular electrostatic interactions and brownian tumbling in protein solutions. *Phys. Chem. Chem. Phys.*, **8**, 2117 (2005).