QCTFF: On the Construction of a Novel Protein Force Field

Paul L. A. Popelier\textsuperscript{[a,b]}

In this perspective, we explain the strategy behind QCTFF, the current name for a novel atomistic protein force field. The atoms are constructed using Quantum Chemical Topology (QCT). These topological atoms determine how a system’s energy is partitioned. We give a brief account of the results hitherto obtained, and a glimpse of unpublished results. Combining this QCT partitioning with the universal quantum expression of energy, leads to four types of fundamental energy contributions. The first of these is intra-atomic and the remaining three interatomic: (i) atomic self-energy, (ii) Coulomb energy, (iii) exchange energy, and (iv) correlation energy. All structural and dynamic effects emerge from the interplay of these contributions. The machine learning method kriging captures well how they change in response to a change in nuclear configuration. Second, the Coulomb energy is represented by a converging multipolar series expansion when the nuclei are sufficiently far apart. © 2015 The Authors International Journal of Quantum Chemistry Published by Wiley Periodicals, Inc.

DOI: 10.1002/qua.24900

Introduction

A force field is a short-cut: when used in geometry optimization or molecular simulation, it bypasses the Schrödinger equation. A force field mimics the behavior of molecules (and molecular assemblies) governed by this powerful equation. Let us take a diatomic molecule as an example. If its bond is stretched beyond the equilibrium bond length then the molecular energy will increase. This effect can be accurately reproduced (or predicted) by fitting an appropriate function capturing the relationship between energy and geometry. This is what a force field achieves, and while doing so, it ignores the mechanism behind this energy change. A force field focuses only on the link between the input (geometry) and the output (energy). This strategy appears effective because it delivers what we want: a computationally cheap prediction of energy as a function of geometry. Predicting the molecule’s energy using a force field is easily several orders of magnitude faster than by an \textit{ab initio} calculation, even for a diatomic molecule.

This advantage comes with major concerns. Can this force field be safely generalized to large molecules? Can we still sample enough molecular configurations? Will accuracy suffer? It is well-known that the number of configurations blows up as the number of degrees of freedom increases. In a diatomic, there is only one degree of freedom and it is perfectly feasible to sample the energy at say 10 different bond lengths. However, given that there are $3N-6$ degrees of freedom for a general nonlinear polyatomic, can one feasibly sample $10^{3N-6}$ configurations when $N$ is about 30, as is the case for a typical peptide-capped amino acid? A second major concern is whether a force field should necessarily be ignorant about the atomic and quantum mechanical mechanisms that make a molecule’s energy respond to a change in its geometry. In other words, should the relationship between a molecule’s energy and its geometry remain a black box? or should one probe into the reason why a molecule prefers a certain geometry, and then model the cause for an energy rise?

If one ignores the two concerns discussed above, one will soon be stuck when setting up a force field for a particular protein in the same way as for a given diatomic. Fragmentation is the way forward. Here, one takes advantage of the fact that quantum mechanics is kind enough to create localizable properties. Indeed, functional groups emerge naturally in the electron density: an imidazole group in an isolated histidine is very similar to one occurring in histidine in the protein cocrystal, for example. This similarity underpins the concept of transferability, the first cornerstone of any force field. Transferability can be quantified and systematically investigated. However, in spite of being so pivotal, this concept remains remarkably under-researched. There is a need to demonstrate quantitatively that the molecular fragments that are parameterized during the construction of a force field are indeed transferable units, within an accepted energy threshold.

Force fields such as AMBER and CHARMM continue to dominate the world of biomolecular simulation. The architecture of these force fields is very old and based on approximations that can be abandoned thanks to present-day computing power. A detailed motivation as to why this architecture should be overhauled in next-generation force fields has been made before (e.g., Refs. [1–5]) and cannot repeated here due to space restriction. One convincing argument in favor of this overhaul, which has again been made in detail\textsuperscript{[6]} is that point charges are inherently limited (except...
Paul Popelier educated in Flanders, he is currently Professor of Chemical Theory and Computation at the University of Manchester. At the age of 21, he rediscovered an algebraic expression for the exact roots of \( x^{17} - 1 = 0 \) (heptadecagon construction). His 185 publications include three books and 25 single-author items, the latter attracting more than 2850 citations. He plays modern jazz piano and composes. He has written poems, one-page images, an incomplete novel and funny pieces, all unpublished; perhaps one day

at very long range) and that multipolar electrostatics should, therefore, be standard.

The extension\(^{[7]}\) of CHARMM to handle drug-like molecules is a straightforward example of how old approximations prevent accuracy enhancement. Here, charge optimization was carried out at the original HF/6–31G(d) level of theory for sake of consistency while it is well-known that higher levels of theory are mandatory for an accurate treatment of hydrogen bonding. Another recent example is the extension\(^{[8]}\) of OPLS to deal with halogen bonding, beyond the model of one charge for each nucleus. Here, the addition of extra sites (X-sites) is chosen as a route forward to model the inherent electrostatic features of chlorides, bromides, and iodides. However, the newly developed OPLS-AAx force field cannot be used in conjunction with the TIP5P water model\(^{[9]}\) because the X-sites cannot be handled by this model of one charge for each nucleus.

Fortunately, several research groups work with multipole moments, which avoid such problems. Many of these groups\(^{[10–20]}\) have made tremendous progress in going beyond the paradigm of current force field used in standard molecular dynamics simulations, by advancing polarization methodologies and by improving the electrostatics beyond the point charge framework. In particular, Kriging, which is identical to “Gaussian progress regression” was also used early on in the construction of potentials.\(^{[18]}\) The force field AMOEBA,\(^{[21]}\) both as approach and software product, is a well developed culmination of such progress beyond the traditional AMBER/CHARMM methodology. With the development of Quantum Chemical Topology Force Field (QCTFF) we joined this journey but started from a framework that differs dramatically from even AMOEBA, let alone AMBER/CHARMM. In the next section, we explain the strategy behind QCTFF and highlight the choices made about its core architecture.

**The Strategy Behind QCTFF**

**Generalities**

Instead of reporting the chronological development of QCTFF, we start with an outline of the complete strategy including the parts that are currently still being implemented. We only mention that QCTFF started with the interatomic electrostatic interaction, which is hence the best understood part. It is possible to combine this part, the details of which are unique to QCTFF, with the traditional expressions for the non-electrostatic part (e.g., Lennard-Jones and Hooke-like or Morse-like covalent bond potentials). However, this path is not followed.

A future-proof strategy defines all energy contributions from scratch and develops them together. As a result, one accurate energy term cannot be “swamped” by another less accurate term.

QCTFF is built on a constellation of five resolutions, which we each discuss in subsequent subsections below. Each of these resolutions is a research challenge by itself and hence not necessarily closed at this stage. First, we choose the so-called topological atom (or QCT atom) as the carrier of chemical information. More precisely, the system’s total energy can be predicted from the information that these atoms are endowed with. An important question that accompanies this decision is the size of the environment this atom needs to be aware of, which is again the question of transferability. Second, the total energy is partitioned into four types of energy contributions, each of which has a chemical meaning: intra-atomic self-energy, and interatomic Coulomb, exchange and correlation energy. Third, any \(1/r\) type of interaction can be expanded by means of the well-known Laplace expansion, which introduces spherical harmonic functions. This expansion generates the familiar multipole moments, which express the way electron density is distributed, in this case within a (topological) atom. This is of great practical use for the Coulomb energy provided the multipole expansion converges. Fourth, a variation in nuclear configuration causes a change in a given atom’s energies and multipole moments. The machine learning method Kriging the capacity to capture the link between the coordinates of the atom’s surrounding nuclei (input) and this atom’s energies and multipole moments (output).

**Quantum chemical topology**

The “Quantum Theory of Atoms in Molecules (QTAIM),”\(^{[22]}\) which is embedded in the more modern extension to Quantum Chemical Topology (QCT),\(^{[23]}\) rigorously defines atoms by means of the gradient vector field of the electron density. In other words, QTAIM is a special case of QCT, where the central idea of partitioning by gradient is applied to electron density or to its Laplacian. The gradient also partitions other\(^{[24]}\) three-dimensional functions, such as the Electron Localization Function, the electrostatic potential, the virial field, the nuclear potential, the magnetically induced molecular current distribution, the Ehrenfest force field, and the intracule density. There are five conceptual and technical advantages\(^{[25]}\) associated with a topological partitioning of the electron density.
energy of $\Omega$. To complete the intra-atomic energy one must add the intra-atomic potential energy. Our lab was early to compute\textsuperscript{127} the latter quantity, which needs the second-order reduced density-matrix $\rho_{ij}(r_{1},r_{2})$. We can now define the electron–electron repulsion energy between two atoms $\Omega_{i}$ and $\Omega_{j}$,

$$V_{\text{ee}}^{\Omega_{i}\Omega_{j}} = \int d\vec{r}_{1} d\vec{r}_{2} \frac{\rho_{ij}(\vec{r}_{1},\vec{r}_{2})}{r_{12}}$$  \hspace{1cm} (1)$$

This equation covers both the intra-atomic ($\Omega_{i} = \Omega_{j} = A$) and interatomic case ($\Omega_{i} = A \neq \Omega_{j} = B$), which features in Interatomic electrostatic energy. Only the electron-nucleus attraction energy is needed to fully complete $E_{\text{intra}}$ of atom $A$,

$$V_{\text{en}}^{AA} = -Z_{A} \int d\vec{r} \frac{\rho(\vec{r})}{r_{1A}}$$ \hspace{1cm} (2)

where $r_{1A}$ is the distance between an electron and the nucleus of atom $A$, so that we can finally write

$$E_{\text{intra}}^{A} = T_{A} + V_{\text{ee}}^{AA} + V_{\text{en}}^{AA}$$ \hspace{1cm} (3)$$

This intra-atomic energy $E_{\text{intra}}^{A}$ is the energy that a single atom possesses inside a system, regardless of whether this system is a single molecule or a cluster of molecules (including even ions). Unpublished work from our lab shows that an oxygen (or nitrogen atom) has the same energy, within maximum 0.8 kJ mol\textsuperscript{-1} only, when appearing in a tripeptide compared with appearing in a pentapeptide, with their common nuclear skeleton in exactly the same configuration. This energetic transferability is based on observations made in at least four test cases, that is, the homo-oligopeptides of Ala, Ser, Gly, and Leu. The generation of energy-transferability data on nonprotein systems (e.g., water clusters, ions in water, and carbohydrates) is planned but we already know such systems can be successfully kriged. This very high degree of transferability of QCT has also been detected and conceded\textsuperscript{230} in terms of atomic charges by those who develop alternative partitioning schemes (such as the Hirshfeld partitioning\textsuperscript{238} scheme and all its variants). Such favorable assessment makes us hypothesize that high transferability is also a hallmark of topological atoms in nonprotein matter, including drugs and natural ligands.

### Interatomic electrostatic energy

The second energy type is interatomic in nature and builds on (but is not equal to) the solely electronic electron–electron Coulomb repulsion energy $V_{\text{ee}}^{AB}$ between two different atoms or $V_{\text{en}}^{AA}$. This quantity can only be calculated from Eq. (1) provided the pure Coulomb part is filtered out of $\rho_{ij}(r_{1},r_{2})$. This can be done by knowing the fine-structure of $\rho_{ij}$,

$$\rho_{ij}(r_{1},r_{2}) = \rho_{ij}^{\text{coul}} + \rho_{ij}^{\text{exch}} + \rho_{ij}^{\text{corr}} = \rho(r_{1})\rho(r_{2}) - \rho_{1}(r_{1})\rho_{j}(r_{2}) + \rho_{i}(r_{1})\rho_{j}(r_{2})$$  \hspace{1cm} (4)$$

where the first term refers to the quantum-mechanically uncorrelated Coulomb-like pair density, the second term to

---

**Figure 1.** An example of a QCT partitioning of the electron density of porphyrin ring in cytochrome P450 BM3 (PDB code 2IJ2). A selection of topological atoms excluding the iron atom in the center of the porphyrin ring. At the inside of the molecule, a topological atom is sharply bounded by an interatomic surface. At the outside of the molecule, the atom is bounded by a constant electron density envelope (for practical reasons, in the gas phase). A bond critical point (purple) is a saddle point in the electron density, that is, a minimum on the line connecting the two respective nuclei, and a maximum in the plane orthogonal to this line. A ring critical point (pink) is saddle point dual to a bond critical point, that is, a minimum in the ring’s plane and a maximum along the line orthogonal to it. This picture has been generated by in-house software called IRIS, which incorporates a version of MORPHY\textsuperscript{231}.
the Fock–Dirac exchange (which is dominant by and associated with the Fermi hole), while the third term is at least an order of magnitude smaller than the second, and connected with the Coulomb hole. Each term in Eq. (4) is associated with a type of potential energy, so that the corresponding fine-structure of $V_{ee}^{AB}$ automatically follows,

$$
V_{ee}^{AB} = V_{ee,coul}^{AB} + V_{ee,exch}^{AB} + V_{ee,corr}^{AB} = \int_{\Omega_A} dr_1 \int_{\Omega_B} dr_2 \frac{\rho_1(r_1) \rho_2(r_2)}{r_{12}} - \int_{\Omega_A} dr_1 \int_{\Omega_B} dr_2 \frac{\rho_1(r_1) \rho_2(r_2)}{r_{12}} + \int_{\Omega_A} dr_1 \int_{\Omega_B} dr_2 \frac{\rho_1(r_1) \rho_2(r_2)}{r_{12}}
$$

(5)

The first term in Eq. (5) is what we are after. Note that it is useful to complete the electron–electron Coulomb energy $V_{ee,coul}^{AB}$ with the three extra terms that cover the nuclear contributions: the electron-nucleus attraction (potential) energy, denoted $V_{en}^{AB}$, which is the interatomic generalization of Eq. (2), its dual denoted $V_{en,r}$, and the classical nucleus–nucleus repulsion, $V_{nn}^{AB}$. When added, these terms lead to the full electrostatic interaction between two atoms A and B, $V_{elec}^{AB}$, or

$$
V_{elec}^{AB} = V_{ee,coul}^{AB} + V_{en}^{AB} + V_{en,r}^{AB} + V_{nn}^{AB}
$$

(6)

This fundamental energy type accounts for the ubiquitous electrostatics, including charge transfer and polarization effects. Exchange-delocalization energy $V_x$: a measure of covalency. The second term in Eq. (5) represents the exchange de-localization energy, $V_{ee,exch}^{AB}$, which is (already) present at Hartree–Fock level. This term teases out the interaction that keeps bonded atoms together.

To what extent atoms are bonded can be estimated by a nonenergy measure, which is typically a quantum-mechanical bond order. QCT offers such a measure. However, it was shown by our lab that this bond order is only the first term of the multipolar expansion of $V_x$. Hence, the latter contains more information than a bond order.

A recent article[33] studied the behavior of $V_x$ with increasing distance for a sizeable number of bond types. It is clear that $V_x$ is always stabilizing and, therefore, always negative. This is why one invariably works with its negative value, to represent it conveniently by a logarithmic scale. The larger the absolute value of $V_x$, or $|V_x|$, the stronger the bond it is associated with. Figure 2 illustrates how $V_x$ decreases exponentially with increasing internuclear distance (yellow band) in the water dimer. The four covalent O–H bonds (marked as 1,2 in Fig. 2) found in both water monomers are the strongest bonds in the dimer, as expected. Their $|V_x|$ values are of the order of 400 kJ mol$^{-1}$ (red disk in Fig. 2), which is in line with the listed mean bond enthalpy of 463 kJ mol$^{-1}$ for OH. The $|V_x|$ values of the 1.3 H...H interactions (yellow disk) are about two orders of magnitude smaller, while the $|V_x|$ value of 1.5 H...H interaction is negligible. Interestingly, the H...O hydrogen bond (bottom purple disk) is anomalously strong, compared with the yellow band marking overall behavior. The same is true for O...O (upper purple disk), which suggests that hydrogen bonding is best localized by considering a subset of three atoms, rather than two, in this case O–H...O (pink elliptical disk).

The $V_x$ energy measure is a novel and exciting gauge to detect unusual interactions, in an orbital- and parameter-free manner. In the long term, it will be instrumental in designing a QCTFF version for chemical reactions, given its ability to quantum mechanically quantify bond-breaking and bond-making. This type of energy governs chemical bonding and (hyper)conjugation effects. Indeed, the latter are easily revealed from a graph of $V_x$ versus distance, again as deviations from the general “yellow band” trend. A fine example of this can be found in Ref. [33] where $V_x$ values for (C)H...H(C) interactions are always anomalously large if the hydrogens are in trans position with respect to each other. One does not have to explain this effect (or trend or even rule) by hyperconjugation but can instead simply name it the “trans effect” and use it to make predictions independently from hyperconjugation. What matters is that the effect is consistent and clearly spotted, in a good number of saturated hydrocarbons, and that it emerges from the relationship between two well-defined quantities: $V_x$ and internuclear distance.

Dispersion energy treated as dynamic correlation. Dispersion is an important effect even in biomolecular modeling but its accurate quantification continues to be a challenge. The four natural amino acids that contain an aromatic residue have already been successfully treated by our methodology in terms of pure multipolar electrostatics, but the next big challenge is to integrate dispersion in a rigorous and streamlined manner. The key to achieve this is being able to calculate the third term in Eq. (5) or $V_{elec,exch}^{AB}$, which embodies dynamic correlation. This is the general quantum mechanical effect that governs dispersion, and allows it to be included as an energy contribution without the context of Rayleigh–Schrödinger (long-range) perturbation theory. As proven a long time ago by London, one then recovers the familiar $C_d r^{-6}$ expression given in textbooks for dipolar dispersion, and higher powers ($C_d$, $C_p$, $C_6$, $C_{10}$, ...) for higher rank multipoles. This treatment is only valid for intermolecular interactions but if represented by the third term in Eq. (5), dispersion can also be included in
In 2014, we wrote a perspective[6] entitled “Multipolar Electrostatics,” which collects dozens of case studies from the literature, systematically proving the superior accuracy of multipolar moments over the traditional point charge approach (of AMBER, CHARMM, OPLS, etc.). Even David Case, a main developer of the point-charge based force field cluster and program AMBER, urged for further research into the accuracy of force-field potentials, particularly electrostatic terms.[37] Applying the Laplace multipole expansion leads to

$$V_{\text{elec}}^{AB} = \sum_{i_k, m_k, j_k} T_{i_k, m_k, j_k} Q_{i_k, m_k} Q_{j_k, m_k}$$

where $Q_i$ represent the atomic multipole moments of rank $\ell$ while $T$ is a purely geometrical interaction tensor. There are three conceptual and technical advantages associated with QCT multipole moments: (i) more compact than Cartesian (no redundancies), (ii) excellent convergence at short-range, and (iii) no penetration effects because of nonoverlapping atoms.

Machine learning: kriging tackles complexity of all energy terms responding to geometrical change

Our lab pioneered the use of kriging in the context of potential design.[38] Kriging[39] is excellent in handling the high-dimensional complexity of configurational change in condensed matter. Kriging learns the mapping between an atomic quantity (as output, e.g., multipole moment or self-energy) and the coordinates of the neighboring atoms (as input, called features $F = (f_i)$). Figure 3 illustrates this principle on a capped amino acid. In the formula of Figure 3, $\mu$ is the mean of the observed values, $N_{\text{tr}}$ is the number of training examples used to build the model (typically of the order of a thousand), $w_k$ is the kriging weight (computed from the correlation matrix), $d$ is the number of (geometric) features (or dimensionality of the feature space), and $N_{\text{feat}}$ is the number of features. By analytically maximizing the (statistical) likelihood function, the vector of parameters $\theta_k$ and $p_k$ are optimized. We then maximize the maximum likelihood in feature space by particle swarm optimization.[40] The division of the data set into training and test set is made less critical by k-fold cross validation.[41]

The four advantages of kriging are: (i) ranking of feature values according to importance: chemical insight, (ii) performance scales linearly with the dimension of feature space, (iii) trained kriging model is analytical and differentiable, so forces[42] can be computed quickly and accurately, and (iv) the knowledge of how an environment influences an atom is stored in kriging parameters $\theta_k$ and $p_k$. Hence, there is no need to perform iterative calculations to self-consistent energies during a molecular dynamics simulation.

In Figure 4, we give an impression of the success of kriging.[43] The sigmoidal curves show which percentage of an external test set of two representative doubly peptide-capped amino acids (isoleucine and cysteine) have all their atoms predicted within a certain error. For example, one can read off from the red curve in Figure 4 that 70% of the intramolecular electrostatic energies of isoleucine are within 4 kJ mol$^{-1}$ ($\sim$1 kcal mol$^{-1}$), an often quoted energy threshold of “chemical accuracy.”

The training set of molecules is obtained by distorting a particular energy minimum according to calculated normal modes. An alternative generator for a training set is sampling local environments from PDB files. Yet another generator can

![Electrostatic energy predictions.](image)

Figure 4. An illustration of the typical performance of kriging for single peptide-capped amino acids.
be a crude (point charge) molecular dynamics simulation in case the training focuses on water clusters or hydrated peptides. How to generate a training data set is still an active topic of research in our group, with the possibility of distorting around configurations that are not energy minima. More details will be published as applied to carbohydrates. All of the 20 natural amino acids have been successfully trained for by kriging, yielding comparable mean errors to the two amino acids illustrated in Figure 4.

Parameterization Procedure

Figure 5 shows a flow chart of the parameterization protocol for QCTFF, encompassing in-house FORTRAN90 programs, external programs and a number of scripts. A central program is TYCHE, which generates the training set. It was originally developed via the so-called “normal-mode distortion,” which needs local energy minima as a feed. The structure in Figure 5 has been implemented from scratch, streamlined, CPU-optimized, documented, and systematically tested. Finally, the training procedure is being turned user-friendly such that users can QCTFF train for their own systems of interest. The full details behind this flow chart cannot be given here due to space restriction but can be found in Ref. [35]. A more advanced training protocol is currently being investigated, with a deeper understanding of the performance of particle swarm optimization, dynamic training set construction, parallelization and computational efficiency, feature reduction, feature interpretation and scrutiny, and safe default settings.

The very first tests of geometry optimization, as implemented in the simulation package DL_POLY_4[44] using all fundamental energy types (except dynamic correlation) are carried out right now.

Conclusions

The driving force behind QCTFF’s development is a step-change enhancement of realism. This step-change should be achieved in a bottom-up fashion, from small to large molecules, and by systematic validation. Constrained by the journal imposed brevity of perspective we outlined the strategy of the design of QCTFF, a novel protein force field, based on malleable atoms that adjust their electronic structure to the coordinates of the neighboring atoms. Each atom is endowed with this knowledge in terms of four fundamental types of energy contributions (both intra- and interatomic), making the mimicking (through the traditional expressions) of their combined, overall effects obsolete.

Acknowledgments

The author is very grateful to Dr Tim Fletcher for providing the flow chart of Figure 5 and to Joseph Thacker for preparing Figure 1. The
author thank the EPSRC for the award of an Established Career Fellowship.

Keywords: quantum chemical topology · force field design · potential energy types · multipole moment · quantum theory of atoms in molecules · Kriging · polarization

How to cite this article: P. L. A., Popelier Int. J. Quantum Chem. 2015, 115, 1005–1011. DOI: 10.1002/qua.24900

[1] P. L. A. Popelier, In Structure and Bonding. Intermolecular Forces and Clusters, Vol. 715; D. J. Wales, Ed.; Springer: Heidelberg, Germany, 2005; p. 1.
[2] P. L. A. Popelier, In Modern Charge-Density Analysis, Vol. 14; C. Gatti, P. Macchi, Eds.; Springer: Germany, 2012; p. 505.
[3] P. L. A. Popelier, AIP Conf. Proc. 2012, 1456, 261.
[4] P. L. A. Popelier, Curr. Top. Med. Chem. 2012, 12, 1924.
[5] P. L. A. Popelier, In Drug Design Strategies: Computational Techniques and Applications, Vol. 20, Chapter 6; L. Banting, T. Clark, Eds.; Royal Society of Chemistry, Cambridge, Great Britain, 2012; p. 120.
[6] S. Cardamone, T. J. Hughes, P. L. A. Popelier, Phys. Chem. Chem. Phys. 2014, 16, 10367.
[7] K. Vanommeslaeghe, A. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, I. Lopes, I. Vorobyov, A. D. J. McKerrell, J. Comput. Chem. 2010, 31, 671.
[8] W. L. Jorgensen, P. Schyman, Phys. Chem. Chem. Phys. 2000, 12, 8910.
[9] F. Archambault, C. Chipot, I. Soteras, F. J. Luque, K. Schulten, F. Dehez, J. Chem. Theory Comput. 2009, 5, 3022.
[10] R. Chaudret, N. Gresh, O. Parisel, J. P. Piquemal, J. Comput. Chem. 2011, 32, 2949.
[11] A. Holt, J. Bostrom, G. Karlström, R. Lindh, J. Comput. Chem. 2010, 31, 1583.
[12] I. Soteras, C. Curutchet, A. Bidon-Chanal, F. Dehez, J. G. Angyán, M. Orozco, C. Chipot, F. J. Luque, J. Chem. Theory Comput. 2007, 3, 1901.
[13] M. Mezei, J. Chem. Theory Comput. 2007, 3, 2138.
[14] P. Y. Ren, C. J. Wu, J. W. Ponder, J. Chem. Theory Comput. 2011, 7, 3143.
[15] S. L. Price, Acc. Chem. Res. 2009, 42, 117.
[16] K. V. J. Jose, N. Artrith, J. Behler, J. Chem. Phys. 2012, 136, 194111.
[17] A. Bartok, M. C. Payne, R. Kondor, G. Csanyi, Phys. Rev. Lett. 2010, 104, 136403.
[18] C. M. Handley, R. J. Deeth, J. Chem. Theory Comput. 2012, 8, 194.
[19] T. Verstraalen, P. W. Ayers, V. Van Speybroeck, M. Waroquier, J. Chem. Theory Comput. 2013, 9, 2221.
[20] J. W. Ponder, C. Wu, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. J. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, T. Head-Gordon, J. Phys. Chem. B 2010, 114, 2549.
[21] R. F. W. Bader, Atoms in Molecules. A Quantum Theory; Oxford University Press: Oxford, Great Britain, 1990.
[22] P. L. A. Popelier, F. M. Aicken, Chem. Phys. Chem. 2003, 4, 824.
[23] P. L. A. Popelier, In The Nature of the Chemical Bond Revisited, Chapter 8; G. Frenking, S. Shaik, Eds.; Wiley-VCH: Weinheim Germany, 2014; p. 271.
[24] M. Rafat, P. L. A. Popelier, J. Comput. Chem. 2007, 28, 2602.
[25] T. L. Fletcher, S. M. Kandathil, P. L. A. Popelier, Theor. Chem. Acc. 2014, 133, 1499.
[26] P. L. A. Popelier, D. S. Kosov, J. Chem. Phys. 2001, 114, 6539.
[27] F. L. Hirshfeld, Theor. Chim. Acta 1977, 44, 129.
[28] A. M. Pendas, E. Francisco, M. A. Blanco, J. Phys. Chem. A 2006, 110, 12864.
[29] M. A. Blanco, A. M. Pendas, E. Francisco, J. Chem. Theory Comput. 2005, 1, 1096.
[30] J. G. Angyán, M. Loos, I. Mayer, J. Phys. Chem. 1994, 98, 5244.
[31] M. Rafat, P. L. A. Popelier, In Quantum Theory of Atoms in Molecules, Vol. 5; C. F. Matta, R. J. Boyd, Eds.; Wiley-VCH: Weinheim, Germany, 2007; p. 121.
[32] M. Garcia-Revilla, E. Francisco, P. L. A. Popelier, A. M. Martín-Pendas, Chem. Phys. Chem. 2013, 14, 1211.
[33] J. Klimeš, A. Michaelides, J. Chem. Phys. 2012, 137, 120901.
[34] T. Fletcher, S. J. Davie, P. L. A. Popelier, J. Chem. Theory Comput. 2014, 10, 3708.
[35] S. Grimm, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. 2010, 132, 154104.
[36] J. W. Ponder, D. A. Case, Adv. Protein Chem. 2003, 66, 27.
[37] M. Rafat, P. L. A. Popelier, Phys. Chem. Chem. Phys. 2009, 11, 6365.
[38] N. Cressie, Statistics for Spatial Data; Wiley: New York, USA, 1993.
[39] J. Kennedy, R. C. Eberhart, In Proceedings of the IEEE International Conference on Neural Networks, Vol. 4; 1995; p. 1942.
[40] D. J. Livingstone, In Data Analysis for Chemists, Oxford University Press: Oxford, Great Britain, 1995.
[41] M. J. L. Mills, P. L. A. Popelier, J. Chem. Theory Comput. 2014, 10, 3840.
[42] S. M. Kandathil, T. L. Fletcher, Y. Yuan, J. Knowles, P. L. A. Popelier, J. Comput. Chem. 2013, 34, 1850.
[43] I. T. Todorov, W. Smith, Phil. Trans. R. Soc. Lond. A 2004, 362, 1835.