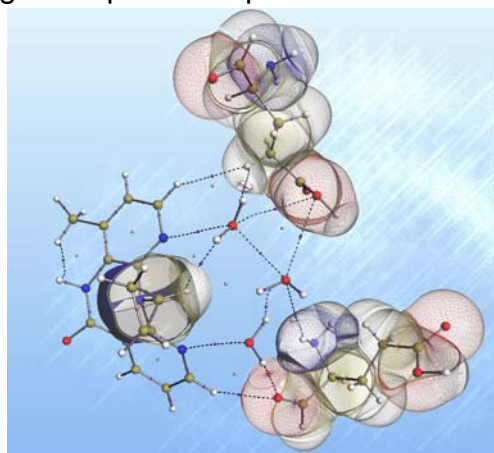


Paul Popelier – Force field design with malleable atoms

Computational chemistry is now recognised as a popular and important tool by many (bio)chemists and materials scientists. It is used alongside spectroscopic methods and crystallography. Computational chemistry aims both at making independent predictions of molecular properties and reactions, and providing atomistic insight. We work in the area of Quantum Chemical Topology (QCT), which uses the language of dynamical systems (e.g. critical point, basin, gradient path, attractor, catastrophe theory,...) to partition quantum systems and extract chemical insight from them. However, more recently, QCT is also being investigated in its potential to act as a basic framework for a new force field. The current highlight focuses on this aspect of our research. The picture on the right illustrates how QCT reveals, from a quantum mechanical wave function, the atoms and their bonding network in selected amino acids of nevirapine. Note that the atoms have a finite volume, touch each other without overlap and have a particular shape depending on their position in the molecular system.



Despite the success of computational chemistry, there is still a need to make its predictions more reliable, especially for large systems. Research in computational chemistry mainly aims at striking a favourable balance between computational speed and accuracy. This is the case in the area of force field development. A force field is a formula designed to deliver structural and dynamical properties of molecular systems very quickly. This formula contains many parameters specific to the systems at hand. As computers are steadfastly becoming ever more powerful, it is vital to support research into more realistic force fields.

Force fields are typically several orders of magnitude faster than algorithms to solve the Schrödinger equation. This equation acts as the ultimate source of molecular information. Nowadays force field parameters are determined from data provided by the Schrödinger equation. Because of its computational demands this is only possible for small systems. For example, protein force fields acquire their parameters from single peptides and, more recently, from oligopeptides. The true predictive power of a force field depends on how reliably information of small molecules (or their assemblies) can be transferred to large ones. Only if transferability is high, reliable predictions can be made outside the force field's "training ground". This is where we propose an ambitious and novel approach. Our main idea is to construct "knowledgeable" atoms, drawn from small molecules and made to interact in order to predict properties of large molecules. These atoms are three-dimensional fragments of electron density that possess a finite volume. They have sharp (i.e. non-fuzzy) boundaries, which endows them with an almost "malleable" character. Their shape responds to the immediate environment of the large molecule they are part of. Machine learning techniques (e.g. neural networks or Kriging) can capture how these atoms change their properties in response to the positions of their neighbours. We have successfully reached the proof-of-concept stage of this idea for glycine¹ and water clusters². This is how we handle polarization, the change of an atom's electron density in response to its environment. We note that the so-called charge transfer term is just the monopole response and does not need separate treatment. Polarization is a very important issue in force field design (see recent dedicated issue of the *Journal of Chemical Theory and Computation*). Our approach operates in the context of accurate electrostatic

interaction, delivered by multipole moments. They secure a more realistic description than point charges. Encouraged by a consistent body of published results we are excited to tackle the two remaining terms of interatomic interaction. They are dispersion and exchange-repulsion. The former is pivotal in pi-pi stacking, which features in DNA base pairing and between aromatic rings in proteins. The latter is responsible for both covalent bonding and the repulsive interaction between (closed-shell) molecules. Dispersion should be treatable along the lines of polarization (inspired by the Casimir-Polder equation). The exchange-repulsion derives from the second-order reduced density matrix, which can also be partitioned in atomic contributions. Perhaps "chemical hardness" can be reformulated in this context, using the finite atomic boundaries. All four interactions will be consolidated into a well tested and user-friendly computer package. The approach is generic, although developments so far have focused on peptides in aqueous solution. Preliminary work on ions in water is most promising and will lead to improve ion channel modeling. Novel scoring functions for docking will also be investigated.

1. Darley, M. G.; Handley, C. M.; Popelier, P. L. A., Beyond Point Charges: Dynamic Polarization from Neural Net Predicted Multipole Moments. *J.Chem.Theory & Comp.* **2008**, 4 (9), 1435-1448.
2. Handley, C. M.; Hawe, G. I.; Kell, D. B.; Popelier, P. L. A., Optimal Construction of a Fast and Accurate Polarisable Water Potential based on Multipole Moments trained by Machine Learning. *Phys.Chem.Chem.Phys.* **2009**, 11, 6365–6376.