Clarification

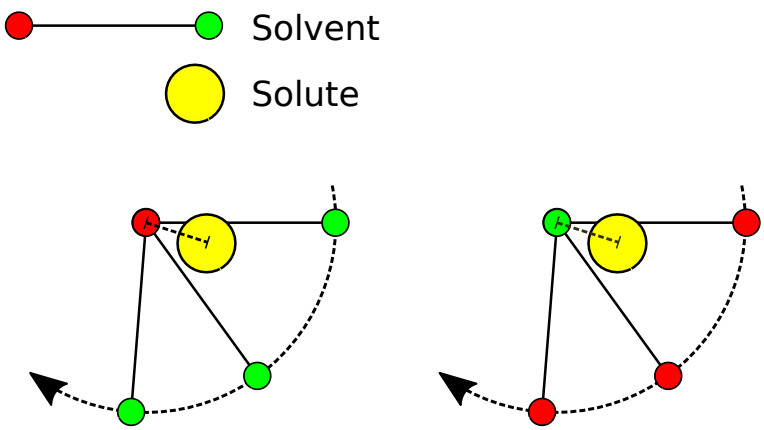
For: L. Martínez, S. Shimizu, **Molecular interpretation of preferential interactions in protein solvation: a solvent-shell perspective by means of minimum-distance distribution functions.** J. Chem. Theor. Comp. 13, 6358–6372, 2017

In the paragraph starting at the last phrase of page 6360 I affirm that

"Note that the minimum-distance density is in general larger than bulk density. In particular, if r is short enough, $n^*(r) = N\rho_s$ because $w_i^*(r) = 1$ for every i. In other words, the minimum-distance density at very short distances is the atomic density, not the molecular density of the solute."

These statements resulted to be imprecise and, although they do not affect any other conclusion or methodological detail in the manuscript, a clarification is needed.

Let us consider the simplest case in which the solute is monoatomic and the solvent molecule is diatomic, such that the MDDF differs from a standard radial distribution. This is illustrated in the figure below:



In the figure of the left, we show that the red atom of the solvent is the minimum-distance atom to the solute. Every rotation of the solvent molecule around the red atom preserves this fact. In the figure in the left I show a symmetric situation but in which the green atom of the solvent is the minimum-distance one.

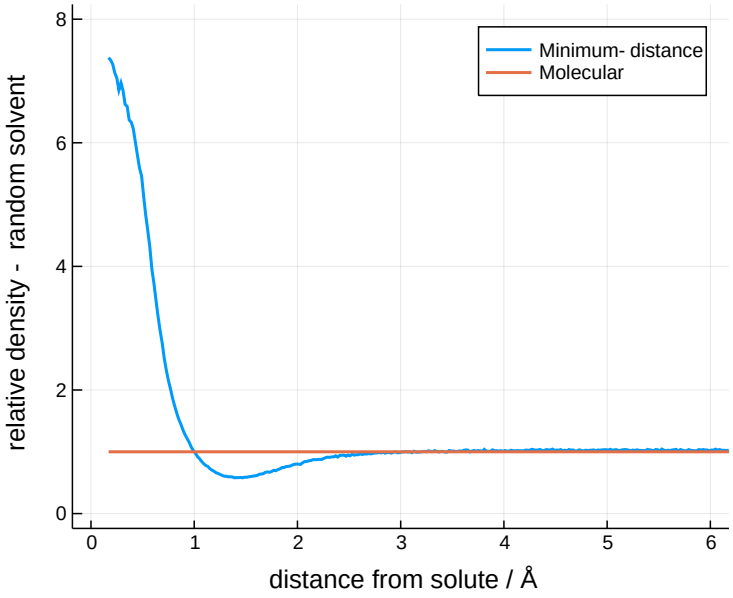
If one generates random positions (translations and rotations) for the solvent molecules around a non-interacting solute, the number density of red or green atoms at the vicinities of the distance indicated will be the same. Furthermore, these densities will be, for each atom, equal to the number density of the solvent molecule generated.

However, both the red and green atom contribute to the counting of minimum-distances associated to *the same distance*. Therefore, the minimum-distance density at that distance is twice the density of the solvent.

At the same time, the complementary atoms (green on the left and red on the right) will *not count* for the MDDF in these situations. Therefore, the distances at which they appear in these configurations will be associated with lower minimum-distance densities than that of the solvent.

If the solvent and solute molecules have many atoms and complex shapes, similar effects are present, but the geometrical analysis is not trivial. However, it is still true that for very short distances the minimum-distance count in a random distribution of the solvent will display larger densities than that of bulk solvent followed, still at short distances, by some density decrease.

We show this result in the case of a protein (RnaseT1) solvated by urea. The figure below displays the density of minimum-distances counted numerically by generating random urea configurations around the protein, normalized by the density of the urea molecules.



Note that the minimum-distance density, at short distances, approaches the atomic density of urea, which has 8 atoms. At large distances, the multiple count of atoms at the same minimum-distances is not possible (as explained in the Appendix A1 of the paper), and the minimum-distance density converges to the molecular density.

These particularities of the MDDF must be carefully taken into account when considering the normalization of the distribution functions. The normalization by a constant density is not appropriate and will most likely suggest an accumulation of molecules at the vicinity of the solvent which is an artifact of the minimum-distance counting. This is why the normalization must be performed by generating explicitly random solvent configurations around a non-interacting solute.

Leandro Martínez

August 14, 2019.