Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.


# SUPPORTING INFORMATION

for

## Assessment of the DNA Partial Specific Volume and Hydration Layer Properties from the CHARMM Drude Polarizable and Additive MD Simulations

Alexey Savelyev

*Department of Chemistry and Biochemistry, University of Montana, Missoula, MT 59812, USA.*
*Tel: 406 243 4022; E-mail: alexsav.science@gmail.com*

## S1. Definition and Computation of the Volume Jacobian

As explained in the main text (Theory and Analysis Methodologies→ Practical Implementation), and also in the caption to Fig. 1, the volume Jacobian, *J(r)*, is defined as the volume of the shell equidistant from the DNA and located at the distance *r* from the DNA surface. These shells are build out of the 3D elementary grid volumes (of linear size 0.25Å) the entire simulation box volume is divided into. Similarly to the water-solute distance histograms, *N(r),* Eq. (3), the volume Jacobian is also a function of the distance *r* defined as the *closest* separation between a given grid elementary volume and the DNA. Thus, equidistant shells are formed out of the grid elementary volumes grouped according to their separations from the DNA. Figs. 1A & B schematically illustrate the equidistant shells (in blue). The entire volume enclosing all the space up to the distance *R* from the DNA is then obtained by integration of the *J(r)* over *r*, with the upper limit being *R*.

Take, for example, Eq. (4) of the main text, representing an expression of the total volume enclosing the DNA molecule and its hydration layer: the *R* corresponds to what is defined as "bulk", the onset of the transition from the structured water to the bulk unstructured region (as judged based on the RDF properties, Fig. 1D); the *dr*, the infinitesimal interval over integration variable, *r*, is nothing else but the linear size of the elementary grid volume, 0.25Å. It important to remember that in practice everything is discrete, so the integrals should be understood as summations of the elementary grid volumes grouped into shells (i.e., into *J(r)*s at different *r*).

Ideally, the above procedure should be repeated every simulation step (each snapshot). However, due to being computationally intensive, the numerical computaiton was performed every 1ns of the respective simulation trajectory. As mentioned in the text, for the case of the DNA this is well justified since DNA is relatively rigid macromolecule, not fluctuating to an extent that volume Jacobian would differ significantly from snapshot to snapshot, so closer spacing along the trajectory would have been necessary. The resulting *J(r)* curves (e.g. Fig. 1C) are then obtained by averaging over Jacobians numerically computed from these every-1ns-snapshots (i. e. averaging over the simulations trajectory).

For more information, we refer interested readers to other computational studies [1-4] (also cited in the main text) employing the closest-separation concept and the volume Jacobians (similarly defined).

## S2. Minimum Water-DNA Separation Distances for Intrinsic DNA Volume Calculation

To avoid possible confusion, we reiterate that these minimum separation definitions are required *only* for our proposed method to estimate DNA's intrinsic volume from MD simulations, in order to estimate the water density in the first solvation shell (as an alternative to other approximate fixed-structure methods, the 'probe-on-a-surface' approach and the approach using standard DNA volume definitions, see the text). These minimum separation definitions are *not* required for PSV calculations since we avoid the volume separation scheme relying on knowledge of both the solute and the hydration layer volumes (the novelty and the strength of the present approach).

As described in the Theory and Analysis Methodologies→Assessment of the Hydration Layer Density, we suggest to use the minimum possible separations among the water molecules and selected DNA atoms in order to compute a modified volume Jacobian, *J(r)_wat*, whose computation procedure is analogous to the procedure described above (section S1), but whose definition is based on use of the elementary grid volumes accessible to the water molecules only in the course of simulation (so, computed volumes based on integration of the *J(r)_wat* will not include the DNA's volume, i. e. the volume inaccessible to water).

The definition of the minimum water–DNA-atoms separations are still based on the closest-separation concept used in computing the distance histograms, *N(r)*; as described in the mains text, when the elementary grid volumes are considered for inclusion into *J(r)_wat*, this concept allows for unique association of the grid volumes with the specific DNA atom; same is true for the water molecules, so the minimum possible distances at which water molecules are located with respect to specific DNA atoms in the simulation course are used in judging if the elementary grid volume belongs to the DNA intrinsic volume (grid volume–DNA-atom separation is less than corresponding water-DNA-atom minimum distance), or to the outer region, the hydration layer (otherwise).

DNA is built from only five chemical elements: P, O, C, N, and H, so, in principle, only five types of the minimum separation water–DNA-atom distances are required. However, such *averaged* (over the DNA molecule) definitions would lead to the low resolution in determining the resulting intrinsic DNA volume since, for example, oxygen atoms contribute to different chemical groups (backbone, minor and major grooves), and the interactions between water and the oxygen atoms contributing to these distinct chemical groups are different (so are the minimum separations observed from MD simulations driven by these distinct interactions). Therefore, to increase the resolution, but at the same time, not to make the procedure computationally too demanding, we grouped all DNA atoms into the following chemical groups: backbone (phosphate and sugar groups, and phosphodiester bonds), minor and major groove regions, and the rest of the DNA atoms which do not contribute to any of the above groups. Fig. 1S provides a visual representation of such atom categorizing, while Table S1 summarizes all the 13 types of the minimum separations extracted from MD simulations (4 for the backbone, 3 for the minor groove, 3 for the major groove, and 3 for the remaining atoms).

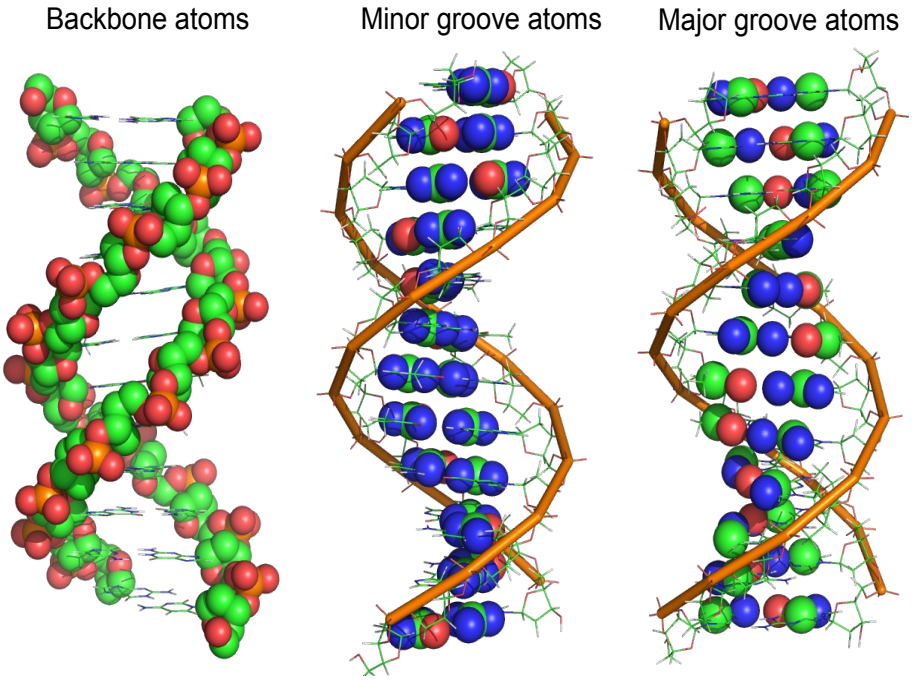


***Figure S1:*** *To compute the minimum water-DNA separation thresholds DNA atoms were grouped into the backbone, minor and major groove categories.*

| **Backbone atoms** | | **Minor groove atoms** | | **Major groove atoms** | | **Remaining atoms** | |
|---|---|---|---|---|---|---|---|
| *Name* | *Separation(Å)* | *Name* | *Separation(Å)* | *Name* | *Separation(Å)* | *Name* | *Separation(Å)* |
| **Additive C36 simulations** | | | | | | | |
| P | 2.65 | N1/N2/N3 | 2.15 | N4/N6/N7 | 2.15 | H | 1.45 |
| O1P/O2P | 1.95 | O2 | 1.95 | O4/O6 | 2.05 | N | 2.65 |
| O3’/O4’/O5’ | 2.05 | C2 | 2.35 | C5 | 2.55 | C | 2.35 |
| C1’/C2’/C3’/ C4’/C5’ | 2.35 | | | | | | |
| **Drude Polarizable simulations** | | | | | | | |
| P | 2.55 | N1/N2/N3 | 2.15 | N4/N6/N7 | 1.95 | H | 1.35 |
| O1P/O2P | 2.05 | O2 | 2.25 | O4/O6 | 2.35 | N | 2.85 |
| O3’/O4’/O5’ | 1.85 | C2 | 2.55 | C5 | 2.65 | C | 2.45 |
| C1’/C2’/C3’/ C4’/C5’ | 2.55 | | | | | | |

***Table S1:*** *Minimum separations between DNA atoms grouped into different categories and geometric centers of the water molecules observed in the course of MD simulations; these threshold distances were used in assessment of the DNA intrinsic volume from simulations, as described in the main text. CHARMM force-field notations are used for atomic names.*

## References


[1] V. Lounnas, B. Pettitt and G. Phillips, Biophysical Journal, 1994, 66, 601–614.
[2] F. Merzel and J. C. Smith, Proceedings of the National Academy of Sciences, 2002, 99, 5378–5383
[3] A. Savelyev and G. A. Papoian, J. Am. Chem. Soc., 2006, 128, 14506–14518
[4] M. Feig and B. M. Pettitt, Biophysical Journal, 1999, 77, 1769–1781