

# Supporting Information for: Inclusion complexation of organic micropollutants with $\beta$ -Cyclodextrin

Máté Erdős, Remco Hartkamp, Thijs J. H. Vlugt, and Othonas A. Moulτος\*

*Engineering Thermodynamics, Process & Energy Department, Faculty of Mechanical,  
Maritime and Materials Engineering, Delft University of Technology, Leeghwaterstraat 39,  
2628CB Delft, The Netherlands*

E-mail: o.moultos@tudelft.nl

## Simulation details

In this section the simulation details of the Attach-Pull-Release (APR) and solvent balance method are presented.

### Attach-Pull-Release (APR) method

To calculate the binding free energy of the selected OMPs with  $\beta$ CD the Attach-Pull-Release (APR) method<sup>1</sup> was used. In this method, the calculation of the binding free energy is divided into four parts. In the first part (attachment phase), a thermodynamic integration is performed to measure the amount of work needed for transferring the host-guest complex to a reference state from its equilibrium configuration. In the second part (pulling phase), the reference state is taken as the initial configuration, and the guest molecule is pulled out of the host to a distance at which the host-guest interaction vanishes. In the third part (detachment

phase), a thermodynamic integration is performed, starting from the last configuration of the second part, to measure the amount of work which is needed to remove the restraints imposed on the host molecule. In the final part, the rest of the imposed restraints are accounted for using a semi-analytical expression and the amount of work needed to transfer the guest to standard conditions is calculated. The binding free energy can be calculated by combining all of the work components obtained from the four parts. The detailed description of the APR method can be found in the study by Henriksen et al.<sup>1</sup> The same terminology is used in this work as in the original publication of the method. In the attachment phase the distance/angle restraints between the OMP and the anchor ( $\lambda_p$ ), and the dihedral restraints of the  $\beta$ CD ( $\lambda_c$ ) are gradually turned on over 53 simulation windows from  $\lambda_p, \lambda_c = 0$  to  $\lambda_p, \lambda_c = 1$ . The  $\lambda$  values are uniformly distributed over the 53 simulation windows. The final force constant for the distance restraint at  $\lambda_p = 1$  is  $k_{\text{dist-OMP}} = 8368 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and for the angle restraints the final force constant is  $k_{\text{ang-OMP}} = 836.8 \text{ kJ mol}^{-1} \text{ rad}^{-2}$ . The force constant of the dihedral restraints at  $\lambda_c = 1$  is  $k_{\text{ang-}\beta\text{CD}} = 50.316 \text{ kJ mol}^{-1} \text{ rad}^{-2}$ . The free energy difference is calculated using the built-in BAR module in GROMACS which is based on the Bennett’s acceptance ratio method.

In the pulling phase, all of the restraints are fully turned on and the equilibrium distance between the OMP and the anchor is varied from 6 Å to 24 Å over 73 simulation windows which entails a 0.25 Å difference between two adjacent simulation windows. The free energy difference of the pulling phase is calculated using the built-in WHAM module in GROMACS which is based on the weighted histogram analysis method. The errors are calculated using the bootstrapping method built-in to the WHAM module.

In the detachment phase the dihedral restraints of the  $\beta$ CD are turned off using the same methodology as in the attachment. This stage consisted of 53 simulation windows with a uniformly distributed  $\lambda_c$  from 1 to 0. The free energy difference is calculated using the built-in BAR module in GROMACS which is based on the Bennett’s acceptance ratio method.

In all simulations the system consists of a  $\beta$ CD, an OMP, 3 anchor atoms, 6000 water molecules, 6 pairs of sodium chloride and the necessary amount of sodium to make the system neutral.

## Solvent balance method

The binding enthalpy of the OMPs with  $\beta$ CD was calculated using the solvent-balance method.<sup>2</sup> In this method four separate simulations are carried out and the binding enthalpy is calculated based on the potential energies of the four systems. To obtain the binding enthalpy, it is necessary to simulate the following distinct systems: (1) the host-guest complex in the solvent, (2) the free host in the solvent, (3) the free guest in the solvent, and (4) the pure solvent. The trajectories produced in the simulations (1), (2), (3), and (4) are referred to as (SOLV), (HOST), (GUEST), and (COMP) trajectories. In these simulations no restraints are used. In all simulations 3000 water molecules, 3 pairs of ions (sodium chloride) and the necessary amount of sodium ions to make the system neutral are used. All simulations are performed for 1.5  $\mu$ s and the averages are calculated from that trajectory. For more details and discussion of the theoretical background of this method the reader is referred to the study by Fenley et al..<sup>2</sup>

The enthalpy decomposition followed a similar methodology as presented by Tang et al.<sup>3</sup> The decomposed enthalpy terms are defined as follows:

$$\Delta H_{\text{host-guest}} = \langle E_{\text{host-guest}} \rangle_{\text{COMP}} - \langle E_{\text{host}} \rangle_{\text{HOST}} - \langle E_{\text{guest}} \rangle_{\text{GUEST}} \quad (\text{S1})$$

$$\Delta H_{\text{host configuration}} = \langle E_{\text{host}} \rangle_{\text{COMP}} - \langle E_{\text{host}} \rangle_{\text{HOST}} \quad (\text{S2})$$

$$\Delta H_{\text{guest configuration}} = \langle E_{\text{guest}} \rangle_{\text{COMP}} - \langle E_{\text{guest}} \rangle_{\text{GUEST}} \quad (\text{S3})$$

$$\Delta H_{\text{solute intermolecular}} = \langle E_{\text{host-guest}} \rangle_{\text{COMP}} - \langle E_{\text{host}} \rangle_{\text{COMP}} - \langle E_{\text{guest}} \rangle_{\text{COMP}} \quad (\text{S4})$$

$$\Delta H_{\text{desolvation}} = \Delta H_{\text{host-solvent}} + \Delta H_{\text{guest-solvent}} + \Delta H_{\text{solvent-solvent}} \quad (\text{S5})$$

$$\begin{aligned} \Delta H_{\text{host-solvent}} = & (\langle E_{\text{solvent-host}} \rangle_{\text{COMP}} - \langle E_{\text{solvent}} \rangle_{\text{COMP}} - \langle E_{\text{host}} \rangle_{\text{COMP}}) \\ & - (\langle E_{\text{solvent-host}} \rangle_{\text{HOST}} - \langle E_{\text{solvent}} \rangle_{\text{HOST}} - \langle E_{\text{host}} \rangle_{\text{HOST}}) \end{aligned} \quad (\text{S6})$$

$$\begin{aligned} \Delta H_{\text{guest-solvent}} = & (\langle E_{\text{solvent-guest}} \rangle_{\text{COMP}} - \langle E_{\text{solvent}} \rangle_{\text{COMP}} - \langle E_{\text{guest}} \rangle_{\text{COMP}}) \\ & - (\langle E_{\text{solvent-guest}} \rangle_{\text{GUEST}} - \langle E_{\text{solvent}} \rangle_{\text{GUEST}} - \langle E_{\text{guest}} \rangle_{\text{GUEST}}) \end{aligned} \quad (\text{S7})$$

$$\Delta H_{\text{solvent-solvent}} = \langle E_{\text{solvent}} \rangle_{\text{COMP}} + \langle E_{\text{solvent}} \rangle_{\text{solvent}} - \langle E_{\text{solvent}} \rangle_{\text{HOST}} - \langle E_{\text{solvent}} \rangle_{\text{GUEST}} \quad (\text{S8})$$

,where the  $\langle E_{\text{solvent}} \rangle_{\text{COMP}}$ ,  $\langle E_{\text{host}} \rangle_{\text{COMP}}$ ,  $\langle E_{\text{guest}} \rangle_{\text{COMP}}$ ,  $\langle E_{\text{host-guest}} \rangle_{\text{COMP}}$  are the average potential energies of the solvent, the  $\beta$ CD, the OMP and the  $\beta$ CD-OMP complex in the (COMP) simulation trajectory, respectively. The  $\langle E_{\text{solvent}} \rangle_{\text{HOST}}$ ,  $\langle E_{\text{host}} \rangle_{\text{HOST}}$ ,  $\langle E_{\text{solvent-host}} \rangle_{\text{HOST}}$  terms are representing the average potential energies of the solvent, the  $\beta$ CD, and the host+solvent together in the (HOST) trajectory. The  $\langle E_{\text{solvent}} \rangle_{\text{GUEST}}$ ,  $\langle E_{\text{guest}} \rangle_{\text{GUEST}}$ ,  $\langle E_{\text{solvent-guest}} \rangle_{\text{GUEST}}$  terms are representing the average potential energies of the solvent, the OMP, and the guest+solvent together in the (GUEST) trajectory.

## Combined binding free energy and enthalpy calculation

To compare the simulation results with the measurements the following equation is used to obtain the binding free energy based on the different binding orientations:<sup>1</sup>

$$\Delta G_{\text{combined}} = -RT \ln \sum_{n=1}^{N_m} e^{-\frac{\Delta G_n}{RT}} \quad (\text{S9})$$

where  $\Delta G_{\text{combined}}$  is the combined free energy of the different orientations,  $R$  is the universal gas constant,  $T$  is temperature,  $N_m$  is the total number of different binding orientation, and  $\Delta G_n$  is the binding free energy of binding orientation  $n$ . The  $\Delta$  indicates the difference between two states of the system, namely, the state in which the OMP is in the bulk phase and the state in which the OMP forms an inclusion complex with the  $\beta$ CD. Similarly, the combined binding enthalpy,  $\Delta H_{\text{combined}}$ , is calculated using:<sup>1</sup>

$$\Delta H_{\text{combined}} = \frac{\sum_{n=1}^{N_m} \Delta H_n e^{-\frac{\Delta G_n}{RT}}}{\sum_{n=1}^{N_m} e^{-\frac{\Delta G_n}{RT}}} \quad (\text{S10})$$

where  $\Delta H_n$  is the binding enthalpy of binding orientation  $n$ .

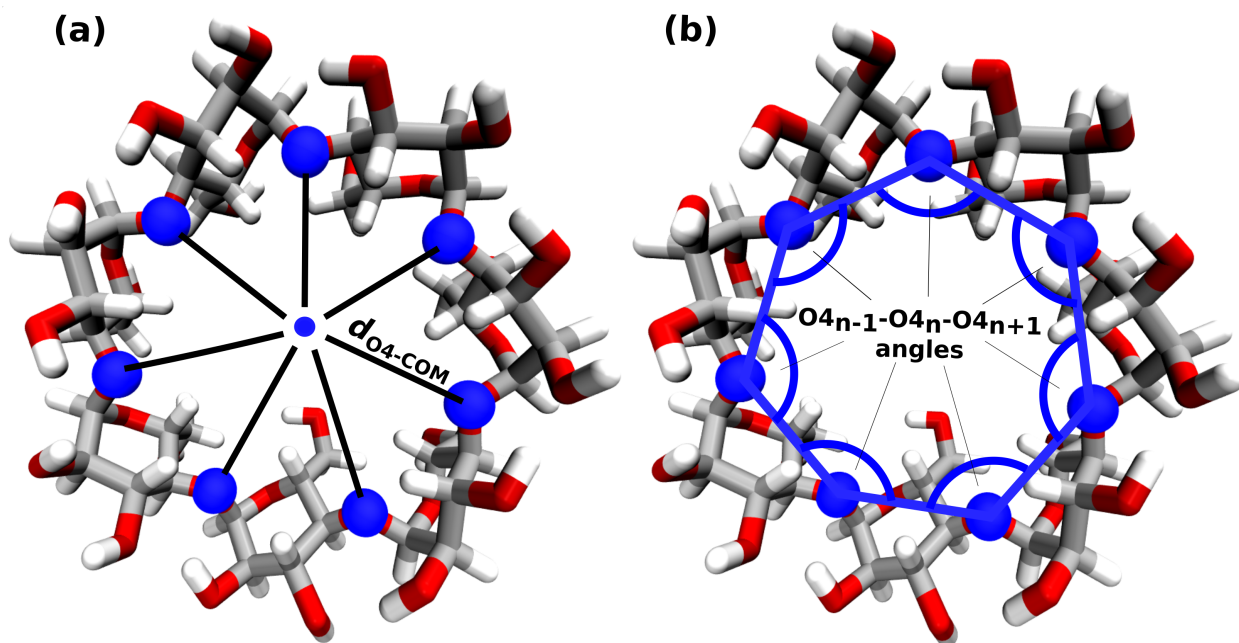


Figure S1: Schematic representation of the (a) distance of the O4 oxygens from the centre of mass of the seven O4 atoms in the  $\beta$ CD and the (b) angle between the O4 oxygens of the residues of the  $\beta$ CD. The O4 oxygens are shown in blue. In our simulations, the average of the  $d_{O4-COM}$  distances and  $O_{4n-1}-O_{4n}-O_{4n+1}$  angles are found to be  $4.93 \pm 0.3 \text{ \AA}$  and  $127.15 \pm 1.64$  degrees, respectively.

## References

- (1) Henriksen, N. M.; Fenley, A. T.; Gilson, M. K. Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics. *Journal of Chemical Theory and Computation* **2015**, *11*, 4377–4394.
- (2) Fenley, A. T.; Henriksen, N. M.; Muddana, H. S.; Gilson, M. K. Bridging Calorimetry and Simulation through Precise Calculations of Cucurbituril–Guest Binding Enthalpies. *Journal of Chemical Theory and Computation* **2014**, *10*, 4069–4078.
- (3) Tang, Z.; Chang, C.-e. A. Binding Thermodynamics and Kinetics Calculations Using Chemical Host and Guest: A Comprehensive Picture of Molecular Recognition. *Journal of Chemical Theory and Computation* **2018**, *14*, 303–318.