Electronic Supplementary Information on
Solvation Energies of Ions with Ensemble Cluster–Continuum Approach

*Lukáš Tomaník, Eva Muchová, Petr Slavíček**

University of Chemistry and Technology. Department of Physical Chemistry, Technická 5, 16628
Prague 6, Czech Republic

*Corresponding author e-mail: petr.slavicek@vscht.cz

The adiabatic ionization energy of Cl^- in the aqueous phase

Fig. S1 shows in detail the calculation results for the adiabatic ionization energy of Cl^- in the aqueous phase as used in the EnCC cycle. It demonstrates that the Marcus theory can provide adiabatic ionization energies with reasonable accuracy. Note also, that the convergence with the increasing number of explicit solvent molecules provides a solid basis for the application in the proposed EnCC approach.

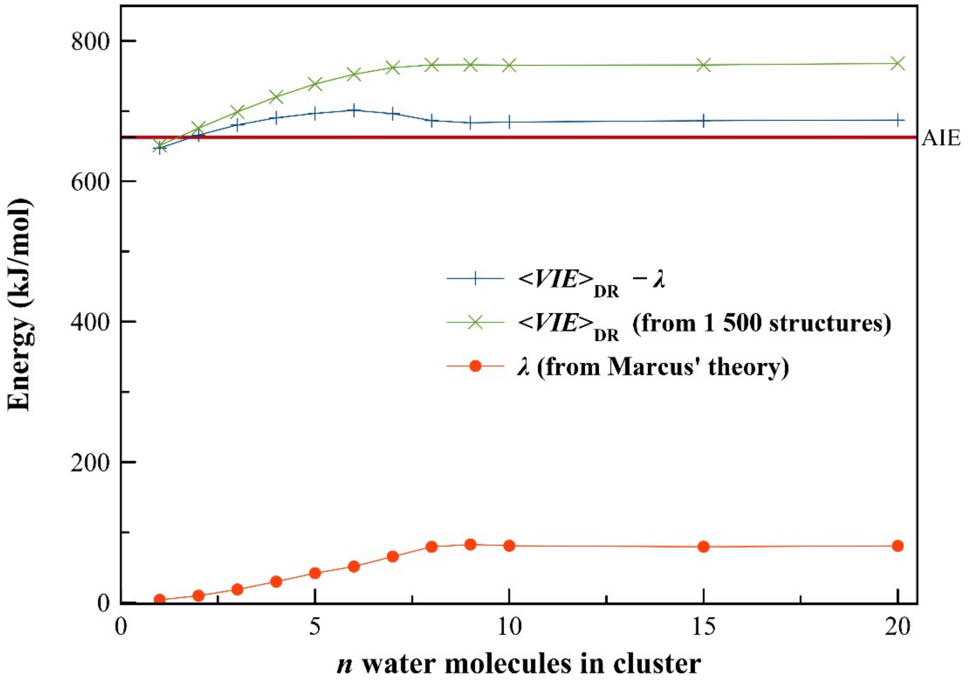

Fig. S1 Ionization energy (kJ/mol) of Cl^- in the aqueous phase obtained from the Marcus theory. Values for different numbers of explicit water molecules included in the calculation are shown. $\langle VIE \rangle_{DR}$ denotes the value when only surrounding SMD was relaxed averaged from 1 500 sampled structures. Parameter λ is derived from the variance of results using the Marcus theory. $\langle VIE \rangle_{DR} - \lambda$ is the final value used in our thermodynamic cycle. The calculations were performed at the LC-ωPBE/6-31+g* level of theory.

Solvation energies of neutral species calculated *via* cluster–continuum model

We compared the performance of the dielectric models to the cluster–continuum approach within the monomer cycle. The comparison is presented in Table S1. As can be seen, the results from PCM or SMD without additional water molecules are reliable and they typically differ from the cluster–continuum values by less than 7 kJ/mol. In the majority of cases, the optimal number of explicit water molecules which minimizes the solvation free energies is one. This represents the original cluster–continuum approach. However, the solvation free energies dramatically raise as two or three water molecules are included. Table S2 presents the results for the cluster cycle. As expected, it shows better stability with the increasing number of water molecules. Still, to fully grasp the predictive ability of the cluster cycle, one has to inspect the convergence in solvation free energy while the number of water molecules is increasing to more than three.

Table S1 Calculated solvation free energies (kJ/mol) of different neutral species using cluster–continuum model with the monomer cycle and HF/6-31+g* method with PCM or SMD solvation model. Results for 0 water molecules in cluster represent the performance of unaided PCM or SMD model as used to get ΔG^{*}_{solv} of neutral species within our approach.

Ion / Method	Number of water molecules in a cluster			
	0	1	2	3
CH_3O / SMD	−7.1	−1.7	22.6	38.9
CH_3O / PCM	−9.2	−4.2	11.3	16.3
SCN / SMD	0.4	−0.8	6.7	43.1
SCN / PCM	−13.8	−17.2	−18.4	6.7
S / SMD	−5.0	−5.4	15.1	35.6
S / PCM	−4.6	−7.5	5.9	15.1
Na / SMD	−2.9	−13.8	−0.4	26.4
Na / PCM	−27.2	−25.5	−9.2	13.4
Ca / SMD	−100.4	−85.4	−54.8	−34.3
Ca / PCM	−52.7	−45.2	−23.8	−10.5
Li / SMD	−8.8	−43.1	−43.1	−21.3
Li / PCM	−51.5	−69.9	−65.7	−61.5

Table S2 Calculated solvation free energies (kJ/mol) of different neutral species using cluster–continuum model with the cluster cycle and HF/6-31+g* method with PCM or SMD solvation model. Results for 0 water molecules in cluster represent the performance of unaided PCM or SMD model as used to get ΔG^{*}_{solv} of neutral species within our approach.

Ion / Method	Number of water molecules in a cluster			
	0	1	2	3
CH_3O / SMD	−7.1	−1.7	15.5	2.9
CH_3O / PCM	−9.2	−4.2	11.3	−0.4
SCN / SMD	0.4	−0.8	−0.4	7.1
SCN / PCM	−13.8	−17.2	−18.4	−10.0
S / SMD	−5.0	−5.4	7.9	−0.8
S / PCM	−4.6	−7.5	5.9	−1.7
Na / SMD	−2.9	−13.8	−7.5	−9.6
Na / PCM	−27.2	−25.5	−9.6	−3.3
Ca / SMD	−100.4	−85.4	−61.9	−70.3
Ca / PCM	−52.7	−45.2	−23.8	−27.2
Li / SMD	−8.8	−43.1	−50.2	−57.3
Li / PCM	−51.5	−69.9	−65.7	−78.2

The effect of MD sampling

The resulting solvation free energies of Cl⁻ from EnCC when different molecular dynamics simulations ate used to sample the configurational space are demonstrated in Fig. S2. Those results only differ in one component of the EnCC cycle – the ionization energy in the aqueous phase. The source of this inconsistency is a different level of theory applied to the molecular dynamics, where the ensembles of structures come from. This represents an additional aspect which affects the accuracy of calculated solvation free energies.

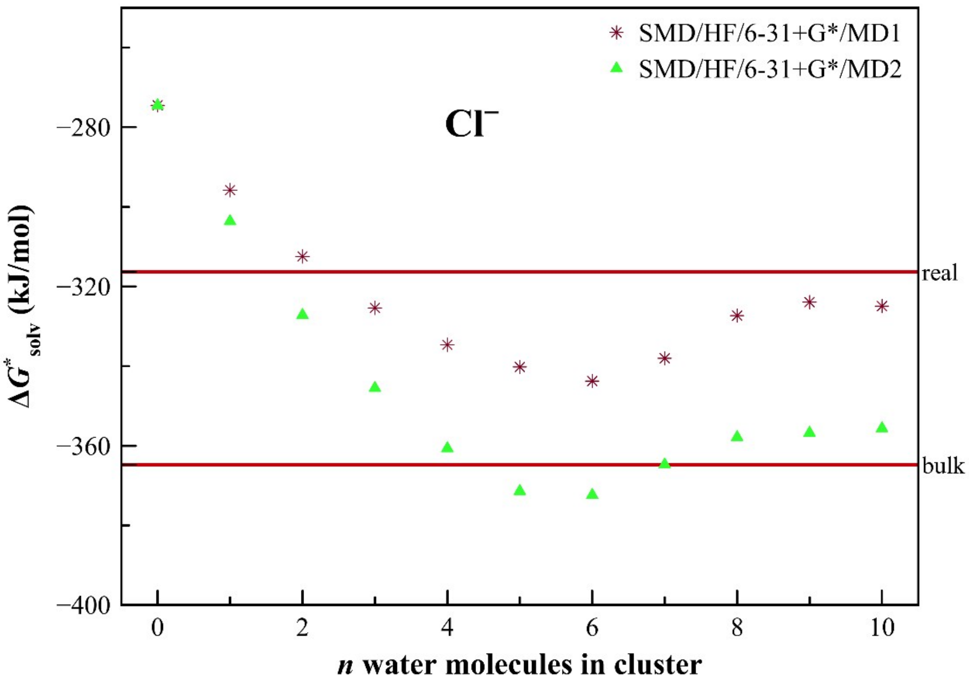

Fig. S2 Calculated solvation free energies (kJ/mol) of Cl⁻ using the EnCC approach based on the ion-water cluster geometries sampled from different dynamic simulations. MD1 refers to QM/MM molecular dynamics simulation as used in the main article. MD2 refers to a large molecular dynamics simulation with periodic boundary conditions and water molecules represented by the SPC/E model. Detailed description in the following text.

Molecular dynamics depicted as "MD1" is the same as used in the whole main article for all the tested ions. It is QM/MM molecular dynamics simulation using the B3LYP/6-31+g* potential energy surface and Grimme's dispersion correction D2. For the MM part, where all the water molecules were included, TIP3P model was used. (See Computational Details in the main article for more details.)

"MD2" was done as follows. Classical molecular dynamics simulation of a chloride anion in water was performed to sample the configurational space. The MD simulation was performed using the GROMACS 5.0.5 package. Water molecules were described with the SPC/E non-polarizable model. The Lennard-Jones parameters were taken from Joung and Cheatham[1] and are summarized in Table S3. The simulation box contained one chloride ion, one sodium cation (at an average distance of 50 Å from the chloride anion to keep the solution electroneutral) and 33151 water molecules in 99.8 Å × 99.8 Å × 99.8 Å box. The total length of the simulation was 200 ns, the time step for the propagation was set to 2 fs, and 3D periodic boundary conditions were employed. The simulation temperature was set to 300 K and was controlled by v-rescale thermostat with a time coupling set to 0.1 ps. The pressure of the system was set to 1 bar which was controlled by the Parrinello-Rahman barostat with a coupling constant of 2 ps. LINCS constrain algorithm of the fourth-order was applied to all bonds. The van der Waals interactions were truncated at 1.0 nm; the long-range electrostatic interactions were calculated by the particle mesh Ewald method. 1500 frames, equally distanced in time, were extracted from the simulation for further *ab initio* calculations.

Table S3 Lennard-Jones parameters used in the classical MD simulations for Cl^- ion, water was represented by the SPC/E model.

	$R_{min}/2$ [Å]	ε [kcal/mol]
Cl^-	2.919	0.0427845

Uncertainty of calculated values

To construct error bars, we seek to find the standard deviation of $\Delta G^{*}_{solv}(A^{-})$. The calculated value of the solvation free energy through the EnCC cycle is obtained as (see the main article):

$$\Delta G^{*}_{solv}(A^{-}) = IE\ (g) + \Delta G^{*}_{solv}(A^{\cdot}) - IE\ (aq)$$

Since $IE\ (g)$ and $\Delta G^{*}_{solv}(A^{\cdot})$ are single calculations for one fixed geometry, they represent constants. Therefore it is sufficient to figure out the standard deviation of $IE\ (aq)$ calculated from the ensemble of structures. $IE\ (aq)$ is calculated as

$$IE\ (aq) = \langle VIE \rangle_{\mathrm{DR}} - \lambda\ = \langle VIE \rangle_{\mathrm{DR}} - \frac{\sigma^{2}}{2k_{B}T}$$

Then the standard deviation of $IE\ (aq)$ (and of the $\Delta G^{*}_{solv}(A^{-})$ as well) is calculated as

$$s\ (IE\ (aq)) = s\left(\Delta G^{*}_{solv}(A^{-})\right) = \sqrt{s^{2}\left(\langle VIE \rangle_{\mathrm{DR}}\right) + s^{2}(\lambda)}$$

where

$$s\left(\langle VIE \rangle_{\mathrm{DR}}\right) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(VIE_{\mathrm{DR},i} - \langle VIE \rangle_{\mathrm{DR}}\right)^{2}}$$

Assuming that σ is now a variable depending on the set of geometries used for the calculations, the standard deviation of reorganization energy, $s(\lambda)$, can be obtained as follows:

$$s\ (\lambda) = \frac{1}{2k_{B}T}\, s\ (\sigma^{2})$$

where

$$s\ (\sigma^{2}) = \sigma^{2}\sqrt{\left(\frac{s\ (\sigma)}{\sigma}\right)^{2} \cdot \left(\frac{s\ (\sigma)}{\sigma}\right)^{2}} = \sqrt{2}\ \cdot\ \sigma\ \cdot\ s\ (\sigma)$$

and

$$s\ (\sigma) = \sigma\,\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{n-1}{2} - \left(\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right)^{2}}$$

Here n is the number of samples (specifically, 1 500 sampled geometries in this work) and $\Gamma\ (n)$ is the gamma function.

For every calculated value, the 95% confidence interval was constructed as

$$\Delta G^{*}_{solv}(A^{-}) \pm 2 \cdot s\left(\Delta G^{*}_{solv}(A^{-})\right)$$

The calculated solvation free energies together with error bars for Cl^-, CH_3O^- and Na^+ ions are shown in Figs. S3–S5.

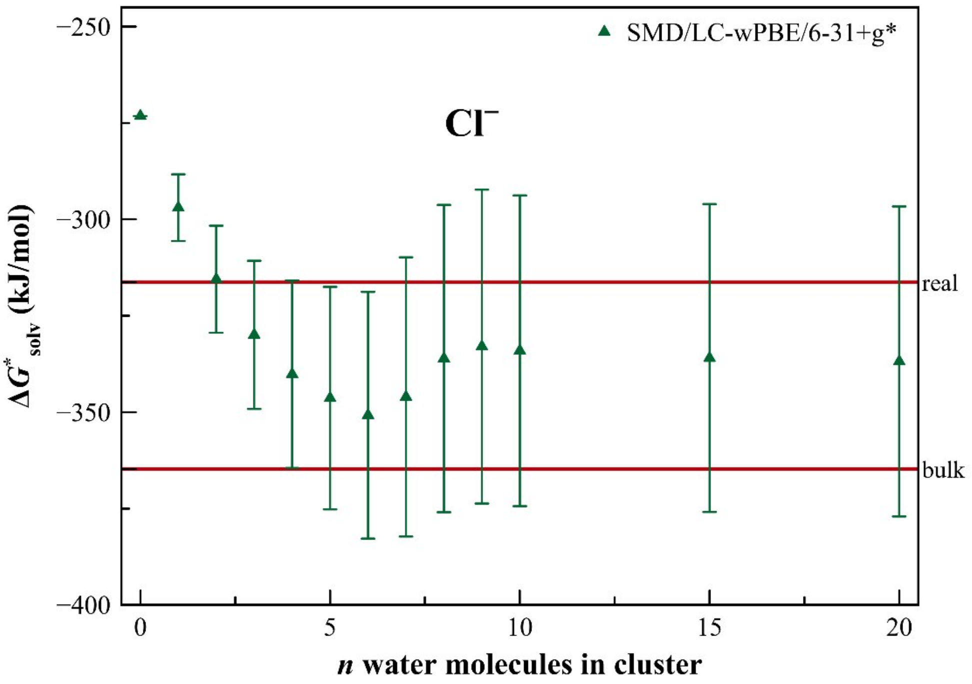

Fig. S3 Calculated solvation free energies (kJ/mol) of Cl^- together with error bars using the EnCC approach. Calculations were performed at the LC-ωPBE/6-31+g* level of theory.

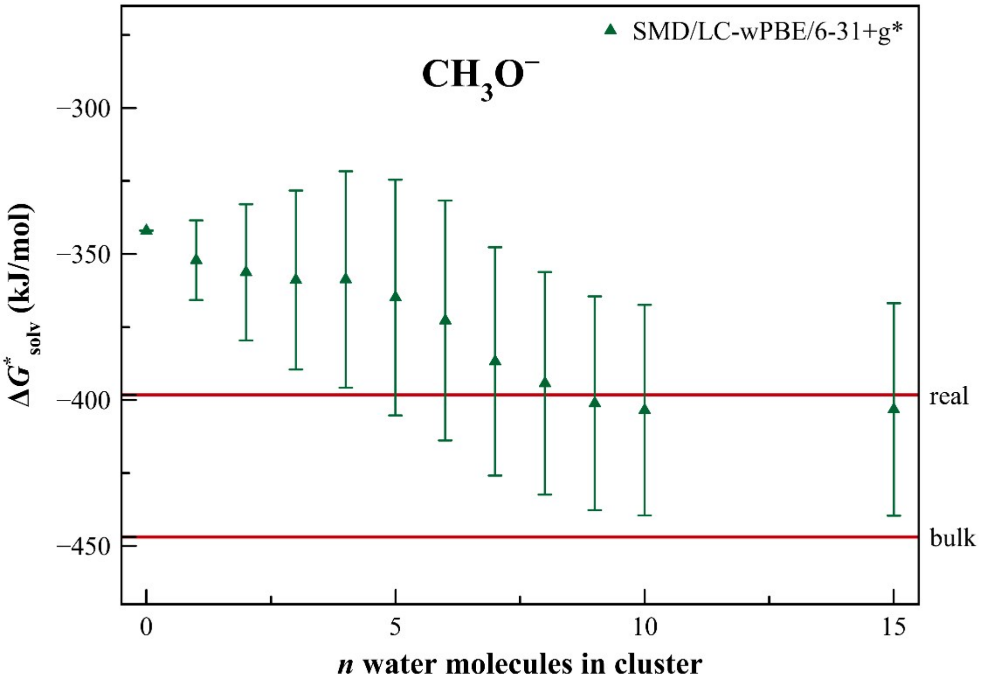

Fig. S4 Calculated solvation free energies (kJ/mol) of CH_3O^- together with error bars using the EnCC approach. Calculations were performed at the LC-ωPBE/6-31+g* level of theory.

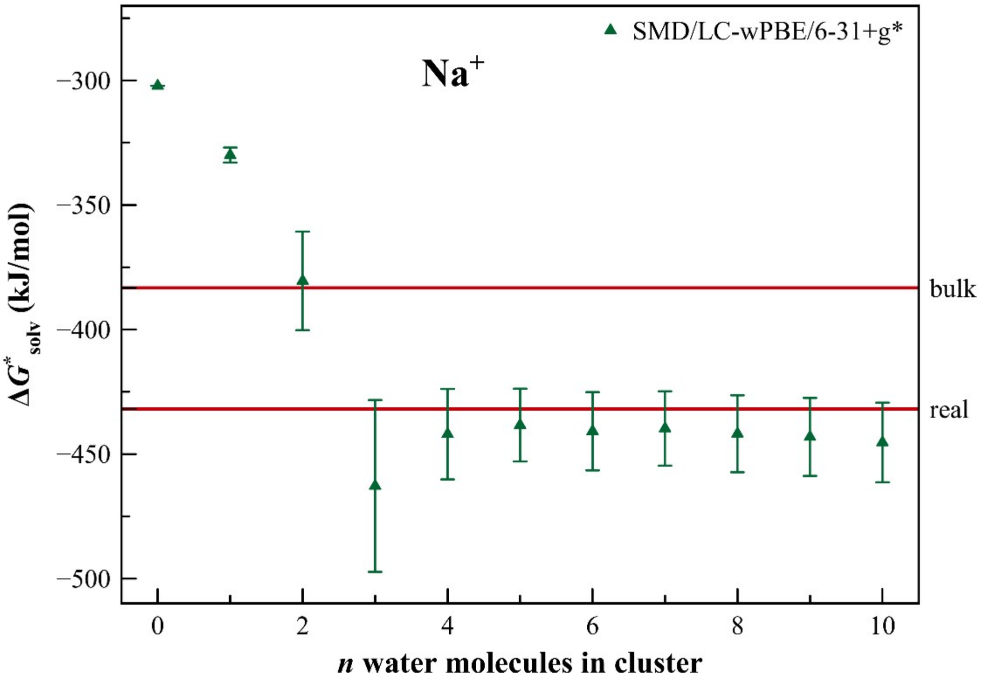

Fig. S5 Calculated solvation free energies (kJ/mol) of Na^+ together with error bars using the EnCC approach. Calculations were performed at the LC-ωPBE/6-31+g* level of theory.

The convergence of the calculated solvation free energy with the basis set

We investigated, how much the choice of the basis set can affect the resulting value of the solvation free energy calculated *via* the EnCC cycle. The results are shown in Table S4. The use of a more extensive basis set leads to rather minor changes in the resulting values of the solvation free energy. Therefore, the use of a relatively small basis set 6-31+g* does not affect the result significantly.

Table S4 Calculated solvation free energies (kJ/mol) of Cl^- using the EnCC approach and the number of explicit water molecules $n = 10$ with different basis sets applied. Calculations were performed at the LC-ωPBE level of theory.

Basis set	ΔG^*_{solv}
6-31+g*	−334.1
6-31++g**	−333.7
6-311+g*	−339.2
6-311++g**	−335.6
aug-cc-pVDZ	−331.9

References

1. I. S. Joung and T. E. Cheatham, *J. Phys. Chem. B*, 2008, **112**, 9020-9041.