Supporting Information: Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model

Tommaso Giovannini,* Alessandra Puglisi, Matteo Ambrosetti, and Chiara Cappelli*

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.

E-mail: tommaso.giovannini@sns.it; chiara.cappelli@sns.it

S1 FQFμ model without charge transfer between MM molecules

If each molecule is constrained to assume a fixed, total charge Q_α, the energy functional F (eq. 14) can be written by exploiting some Lagrangian multipliers (λ_α), whose number is equal to the total number of molecules in the MM portion.

$$
\begin{aligned}
F\left(\mathbf{q}, \boldsymbol{\mu}, \lambda_\alpha\right) &= E\left(\mathbf{r}, \mathbf{q}, \boldsymbol{\mu}\right) + \sum_\alpha \left[\lambda_\alpha \sum_i \left(q_{\alpha i}\right) - Q_\alpha\right] = \\
&= \frac{1}{2}\sum_{i\alpha}\sum_{j\beta} q_{i\alpha}\mathrm{T}^{qq}_{i\alpha,j\beta}q_{j\beta} + \frac{1}{2}\sum_i\sum_j \boldsymbol{\mu}^\dagger_{i\alpha}\mathbf{T}^{\mu\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} + \sum_i\sum_j q_{i\alpha}\mathbf{T}^{q\mu}_{i\alpha,j\beta}\boldsymbol{\mu}^\dagger_{j\beta} + \\
&+ \sum_{i\alpha} q_{i\alpha}\chi_{i\alpha} + \sum_\alpha \lambda_\alpha \left[\sum_i q_{\alpha i} - Q_\alpha\right] = \\
&= \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \boldsymbol{\chi}^\dagger\mathbf{q} + \boldsymbol{\lambda}^\dagger\mathbf{q}
\end{aligned}
\tag{S1}
$$

where α and β run over the molecules and λ_α are meant to preserve the total charge Q_α of each molecule. Therefore, the constrained minimum is found by imposing the derivatives of F with respect all the variables to be zero, thus resulting in the following linear problem:

$$
\begin{cases}
\sum_{j\beta}\mathbf{T}^{qq}_{i\alpha,j\beta}q_{j\beta} + \lambda_\alpha + \sum_{j\beta}\mathbf{T}^{q\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} = -\chi_{i\alpha} \\
\sum_{j\beta}\mathbf{T}^{\mu\mu}_{i\alpha,j\beta}\boldsymbol{\mu}_{j\beta} + \sum_{j\beta}\mathbf{T}^{q\mu}_{i\alpha,j\beta}q_{j\beta} = 0 \\
\sum_{i\alpha} q_{\alpha i} = Q_\alpha
\end{cases}
\tag{S2}
$$

The whole system can be recast in a more compact form as:[1,2]

$$
\begin{pmatrix}
\mathbf{T}^{qq} & \mathbf{1}_{\boldsymbol{\lambda}} & \mathbf{T}^{q\mu} \\
\mathbf{1}^\dagger_{\boldsymbol{\lambda}} & \mathbf{0} & \mathbf{0} \\
-\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu}
\end{pmatrix}
\begin{pmatrix}
\mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu}
\end{pmatrix}
=
\begin{pmatrix}
-\boldsymbol{\chi} \\ \mathbf{Q} \\ \mathbf{0}
\end{pmatrix}
\qquad \Rightarrow \qquad \mathbf{D}\mathbf{Q}_\lambda = -\mathbf{C}_Q
\tag{S3}
$$

where $\mathbf{1}_{\boldsymbol{\lambda}}$ is a rectangular matrix which accounts for the Lagrangians. $\mathbf{C}_Q$ is a vector containing atomic electronegativities and total charge constraints, whereas $\mathbf{Q}_\lambda$ is a vector containing charges, dipoles and Lagrange multipliers.

S2 The QM/PQEq approach

PQEq force field combines the Charge Equilibration model (QEq)[3] and the Drude Oscillator approach.[4] Each atom is considered to be composed of core and shell, on which Gaussian charge distributions are placed. In particular, the core is constituted by a fluctuating (q) and a fixed charge ($+Z$), the latter being connected trough an isotropic harmonic spring to the shell fixed charge ($-Z$), thus allowing displacement of the electron charge distributions. The formal equations defining such a force field can be found in Ref.[5] In this section, we will strictly follow the notation of that paper.

The coupling of the PQEq force field with a QM density can be expressed through Coulomb interaction:

$$E_{\mathrm{QM/PQEq}} = \sum_i \phi[\rho_{\mathrm{QM}}](\mathbf{r}_{ic})q_i + \sum_i \phi[\rho_{\mathrm{QM}}](\mathbf{r}_{ic})Z_i - \sum_i \phi[\rho_{\mathrm{QM}}](\mathbf{r}_{is})Z_i \tag{S4}$$

where $\phi[\rho_{\mathrm{QM}}](\mathbf{r}_i)$ is the electrostatic potential due to the QM charge density, calculated at PQEq charge positions. $\mathbf{r}_{ic}$ and $\mathbf{r}_{is}$ indicate core and shell positions. The QM potential can be divided into a nuclear and an electronic contribution, as expressed in Eq. 18.

Thus, Eq. S4 can be rewritten as:

$$E_{\mathrm{QM/PQEq}} = \sum_i q_i V(\mathbf{P})_{ic} + \sum_i Z_i \left(V(\mathbf{P})_{ic} - V(\mathbf{P})_{is}\right)$$

where $V(\mathbf{P})_{ik}$ is the potential calculated at $\mathbf{r}_{ik}$ (where $k = c$ or $k = s$).

The energy of the whole QM/PQEq system is the sum of three terms, $E = E_{\mathrm{QM}} +$

$E_{\mathrm{PQEq}} + E_{\mathrm{QM/PQEq}}$. Thus, the complete energy functional $\mathcal{F}$ reads:

$$\begin{aligned}\mathcal{F}(\mathbf{P},\mathbf{q},\lambda,\mathbf{r}_{is}) &= tr\mathbf{hP} + \frac{1}{2}tr\mathbf{PG}(\mathbf{P}) + \mathbf{q}^{\dagger}\boldsymbol{\chi} + \frac{1}{2}\sum_i \left[\eta_{ii}q_i^2 + K_s\left|\mathbf{r}_{ic}-\mathbf{r}_{is}\right|^2\right] + \\ &+ \lambda\mathbf{q} + \sum_{i>j}\left[C(\mathbf{r}_{ic,jc})q_{ic}q_{jc} - C(\mathbf{r}_{ic,js})q_{ic}Z_j - C(\mathbf{r}_{is,jc})q_{jc}Z_i + C(\mathbf{r}_{is,js})Z_iZ_j\right] + \\ &+ \sum_i q_iV(\mathbf{P})_{ic} + \sum_i Z_i\left(V(\mathbf{P})_{ic} - V(\mathbf{P})_{is}\right)\end{aligned} \tag{S5}$$

Where the expression for the MM portion has been taken from ref.[5] $C(\mathbf{r}_{ik,jk}), k = c, s$ represents the interaction between core and shell charges. The Lagrangian multiplier λ is imposed so that the charge of the whole MM portion is fixed to Q_{tot}. The contribution to the Fock matrix is obtained by differentiating Eq. S5 with respect to the density matrix elements:

$$\tilde{F}_{\mu\nu} = \frac{\partial\mathcal{E}}{\partial P_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{P}) + \mathbf{V}^{\dagger}_{\mu\nu,c}\mathbf{q} + \mathbf{V}^{\dagger}_{\mu\nu,c}\mathbf{Z} - \mathbf{V}^{\dagger}_{\mu\nu,s}\mathbf{Z} \tag{S6}$$

where, $\mathbf{q}$ and $\mathbf{Z}$ are the vectors containing fluctuating and fixed charges, respectively, whereas c and s indicate core and shell positions, where the QM potential is calculated.

The derivative with respect to the fluctuating charges q_i results in a linear equation, similar to what has been reported for QM/FQ:[6]

$$\begin{pmatrix}\mathbf{H} & \mathbf{1}\\ \mathbf{1}^{\dagger} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{q}\\ \lambda\end{pmatrix} = \begin{pmatrix}-\mathbf{A}\\ Q\end{pmatrix} + \begin{pmatrix}-\mathbf{V}\\ 0\end{pmatrix} = \begin{pmatrix}-\mathbf{B}\\ Q\end{pmatrix} \tag{S7}$$

that is:

$$\mathbf{Dq}_{\lambda} = \mathbf{K} \tag{S8}$$

where, $\mathbf{H}$ and $\mathbf{A}$ are defined as:

$$H_{ij} = \eta_{ii}\delta_{ij} + (1-\delta_{ij})C(\mathbf{r}_{ic,jc}) \tag{S9}$$

$$A_i = \chi_i + \sum_{i>j} \left[C(\mathbf{r}_{ic,jc}) - C(\mathbf{r}_{ic,js}) \right] Z_j \tag{S10}$$

Notice that the equations are identical to those applying to the QM/FQ approach, with the only difference being the definition of $\mathbf{A}$. In fact, in QM/FQ only the electronegativity χ enters the definition of $\mathbf{A}$, whereas in QM/PQEq additional polarization terms arise from the other MM charges. By inverting $\mathbf{D}$, both the fluctuating charges and the Lagrangian multipliers are calculated.

The condition on the derivative with respect to the shell positions needs to be also satisfied to minimise the energy functional $\mathcal{F}$. This results in imposing the forces acting on each shell mobile charge to be equal to zero. The forces acting on each shell charges can be divided in three main contributions: $\mathbf{F}_{\text{intra}}$, the spring force, $\mathbf{F}_{\text{inter}}$, the electrostatic forces due to the other MM charges, and $\mathbf{F}_{\text{QM}}$, which is the QM force, i.e. the QM electric field.

$$\mathbf{F}_{\text{intra}}(\mathbf{r}_{is}) + \mathbf{F}_{\text{inter}}(\mathbf{r}_{is}) + \mathbf{F}_{\text{QM}}(\mathbf{r}_{is}) = 0 \tag{S11}$$

where

$$\mathbf{F}_{\text{QM}} = -\boldsymbol{\nabla}_{r_{is}} \sum_i V(\mathbf{P})_{is}$$

The electric field produced by the QM density acting on the PQEq charges can be obtained as:

$$V_i^{(\xi)}(\mathbf{P}) = \sum_{\zeta=1}^{N_n} \frac{Z_\zeta(\mathbf{r}_i - \mathbf{R}_\zeta)}{|\mathbf{r}_i - \mathbf{R}_\zeta|^3} + \sum_{\mu\nu} P_{\mu\nu} \langle \chi_\mu \chi_\nu | \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} | \delta(\mathbf{r}' - \mathbf{r}_i) \rangle = \mathbf{E}_i(\mathbf{P}) \tag{S12}$$

S3 Parametrization

Table S1: QM/FQFμ electronegativities (χ), chemical hardnesses (η) and polarizabilities (α) for Oxygen and Hydrogen atoms in the water molecule.

Atom	χ	η	α
O	0.321226	0.718474	1.645295
H	0.166401	0.921014	0.424619

Table S2: Comparison of calculated QM/FQFμ and reference KM-EDA/6-31+G* electrostatic+polarization energies (a.u.). HF/6-31+G* was used in QM/FQFμ calculations. $\overset{ele}{\text{QM/FQF}\mu}|_q$ and $\text{E}^{ele}_{\text{QM/FQF}\mu}|_\mu$ indicate charge and dipole contributions to E_{ele}, respectively.

Structure	$\text{E}^{ele}_{\text{QM/FQF}\mu}\|_q$	$\text{E}^{ele}_{\text{QM/FQF}\mu}\|_\mu$	$\text{E}^{ele}_{\text{QM/FQF}\mu}$	E_{KM-EDA}	Error %
Dimer 1_{HDon}	-0.011826 (70%)	-0.004980 (30%)	-0.016806	-0.017678	4.93
Dimer 1_{HAcc}	-0.013376 (72%)	-0.005139 (28%)	-0.018515	-0.017678	4.73
Dimer 2_{HDon}	-0.009051 (72%)	-0.003443 (28%)	-0.012494	-0.012329	1.34
Dimer 2_{HAcc}	-0.009052 (72%)	-0.003445 (28%)	-0.012496	-0.012329	1.36
Dimer 3_{HDon}	-0.007275 (71%)	-0.002986 (29%)	-0.010260	-0.010862	5.54
Dimer 3_{HAcc}	-0.007626 (70%)	-0.003218 (30%)	-0.010844	-0.010862	0.17

S4 Dependence on the QM method

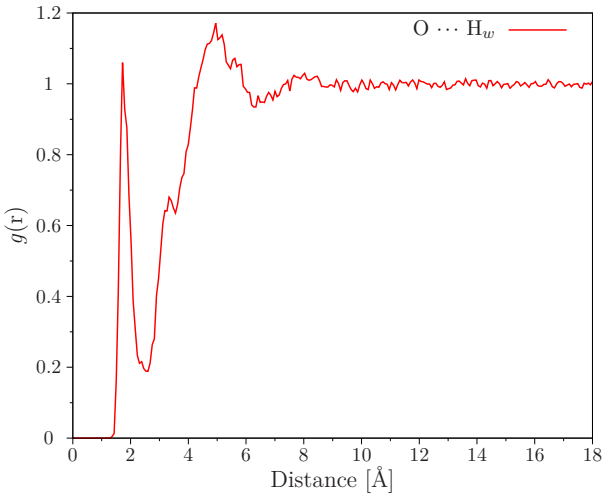
Table S3: Comparison of calculated QM/FQFμ and reference KM-EDA/6-31+G* electrostatic+polarization energies (a.u.). HF/6-31+G* was used in QM/MM calculations.

	6-31G	6-31+G	6-31G*	6-31+G*	6-311++G**	cc-pVDZ	cc-pVTZ	cc-pVQZ	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ
HF	-12.54	-13.16	-10.42	-10.91	-10.47	-9.53	-9.79	-9.93	-9.70	-10.28	-10.45
LDA	-12.29	-12.97	-10.09	-10.65	-10.25	-9.14	-9.50	-9.61	-8.83	-9.80	-10.07
PBE	-11.78	-12.44	-9.70	-10.25	-9.83	-8.77	-9.21	-9.35	-8.60	-9.49	-9.76
B97D	-11.78	-12.44	-9.74	-10.29	-9.86	-8.77	-9.22	-9.39	-8.70	-9.58	-9.84
RevTPSS	-11.65	-12.34	-9.71	-10.28	-9.85	-8.76	-9.23	-9.40	-8.77	-9.64	-9.88
BLYP	-11.59	-12.29	-9.59	-10.18	-9.80	-8.70	-9.23	-9.39	-8.53	-9.50	-9.78
M06	-12.25	-12.91	-10.02	-10.57	-10.18	-9.08	-9.35	-9.52	-9.05	-9.93	-10.06
PBE0	-12.10	-12.75	-9.97	-10.51	-10.07	-9.04	-9.39	-9.53	-8.98	-9.76	-9.99
B3LYP	-11.90	-12.59	-9.85	-10.42	-10.03	-8.95	-9.40	-9.54	-8.86	-9.74	-9.99
M062X	-12.28	-12.95	-10.10	-10.64	-10.19	-9.13	-9.53	-9.65	-9.08	-9.91	-10.16
sogga11x	-12.12	-12.80	-10.01	-10.56	-10.12	-9.08	-9.34	-9.52	-9.09	-9.85	-10.05
mPW1PW91	-12.10	-12.76	-9.98	-10.53	-10.09	-9.04	-9.39	-9.54	-9.00	-9.78	-10.01
CAM-B3LYP	-12.07	-12.79	-9.96	-10.55	-10.18	-9.08	-9.52	-9.67	-8.98	-9.85	-10.09

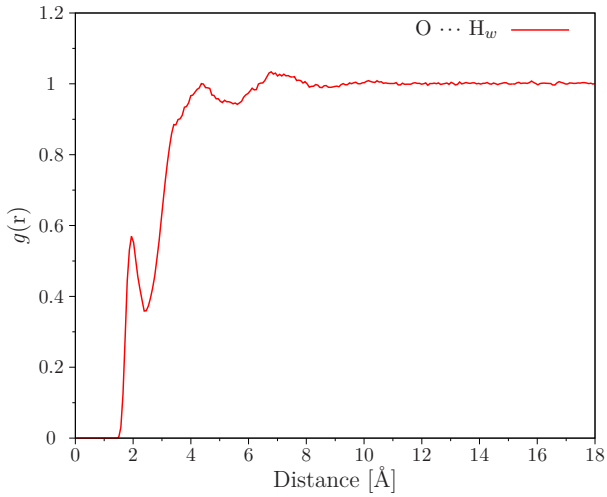
S5 Molecules in aqueous solution

S5.1 Classical MD simulations

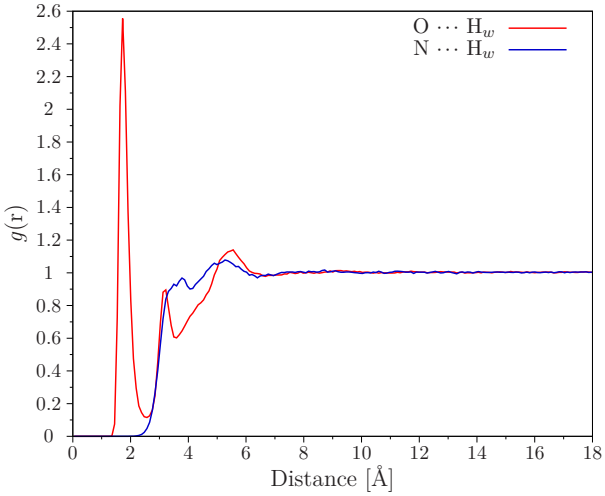
The geometries of (R)-methyloxirane (MOXY), Acrolein (ACRO), N-Methylacetamide (NMA) and Methanol (MeOH) were optimized at the B3LYP/ aug-cc-pVDZ level and the C-PCM[7] to represent the aqueous environment ($\varepsilon = 78.3553$). Following what has already been reported by some of the present authors,[8] in order to obtain a representative conformational sampling of solvated solutes, 10 ns MD simulations were performed in a pre-equilibrated box of 2865 SPC (Single Point Charge) water molecules in the NPT (isothermal-isobaric) ensemble using GROMACS.[9–12] All bonds were kept rigid using the Settle algorithm[13] for water; the geometry of the solute was kept rigid during the simulation. Electrostatic interactions were considered through the the Particle Mesh Ewald summation method.[14] The pressure was stabilized at 1 bar using the weak-coupling scheme with a coupling constant of 10 ps and an isotherm compressibility of $5 \cdot 10^{-5}\text{bar}^{-1}$. Each component of the system (both methyloxyrane and water) was coupled separately to a temperature bath at 300 K using the Berendsen thermostat[15] with a coupling constant of 0.5 ps. The all-atoms OPLS-AA (Optimized Potentials for Liquid Simulations - All Atoms) force field[16] was used for the solute



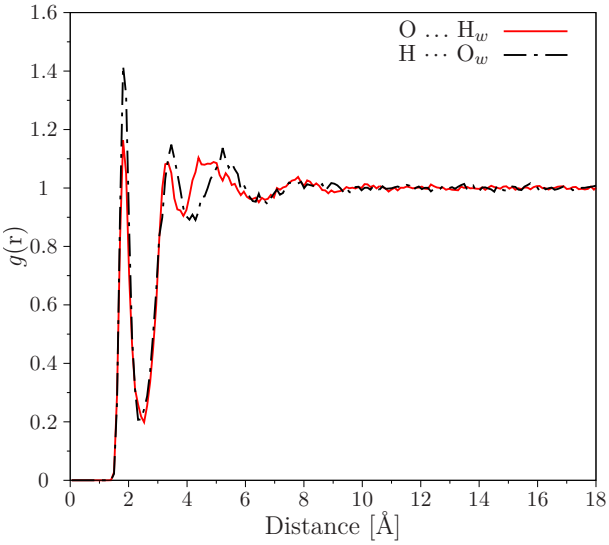
MOXY



ACRO



NMA



MeOH

S5.2 Structures of solute-water clusters

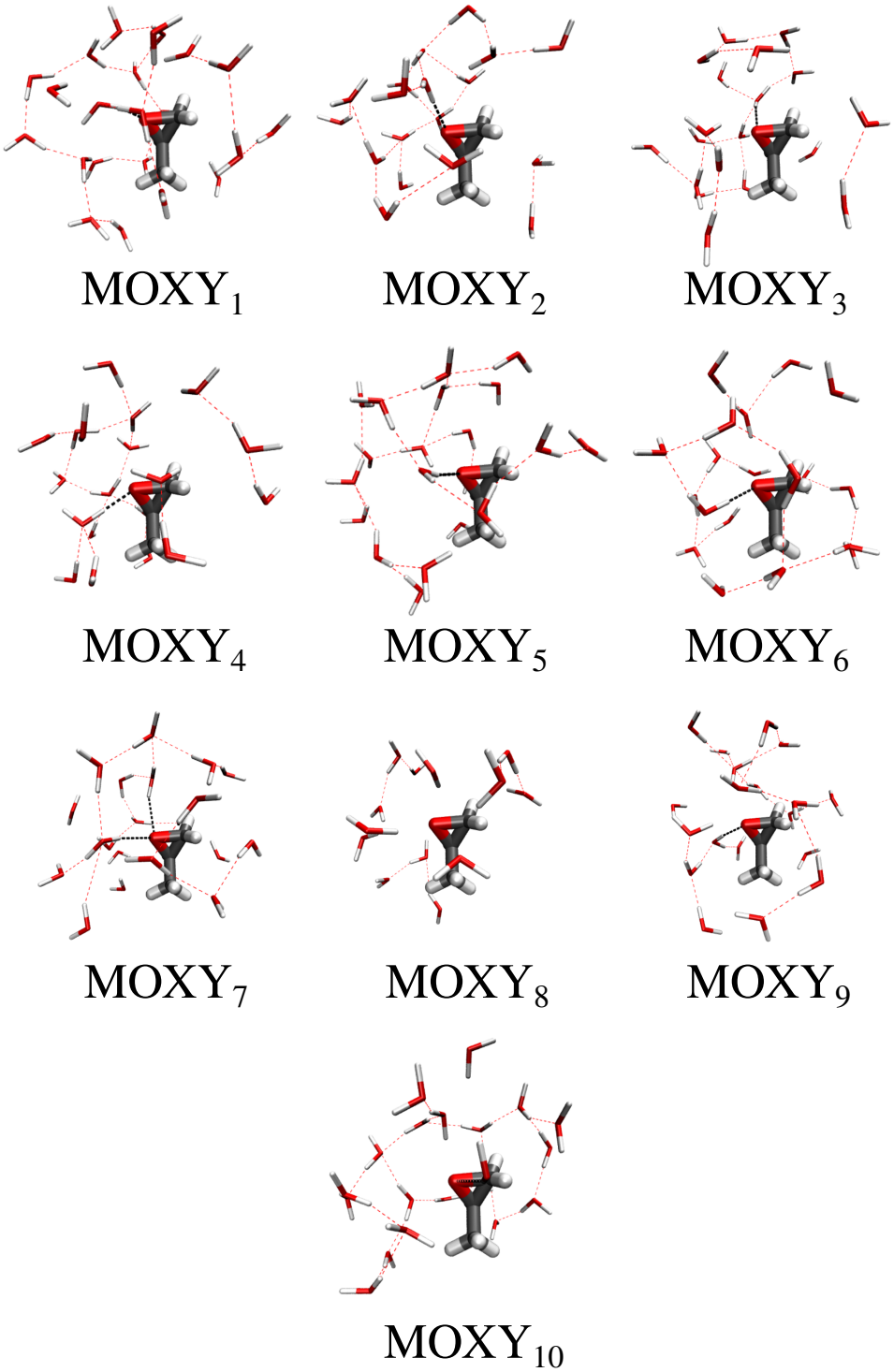


Figure S1: MOXY structures

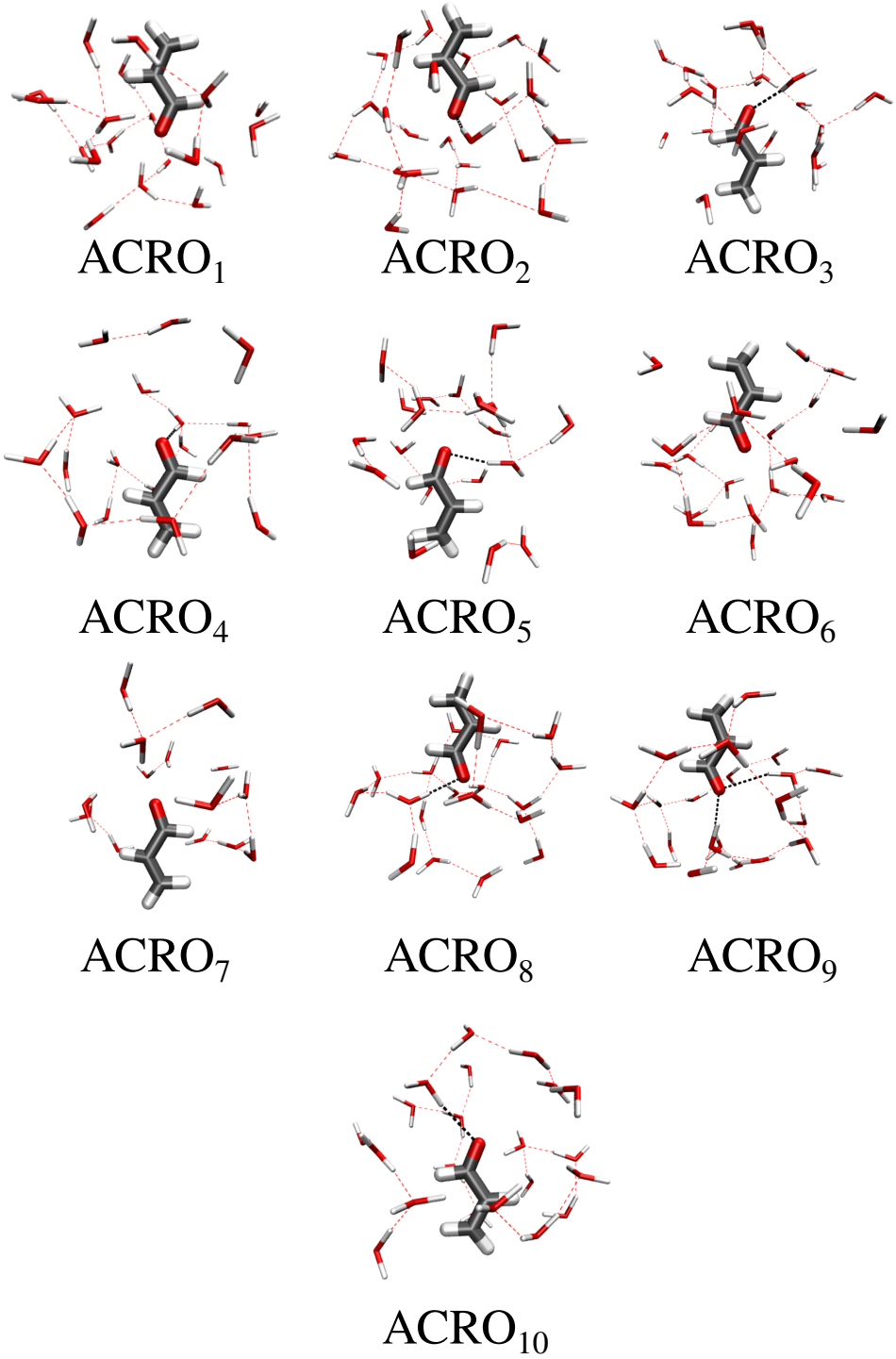


Figure S2: ACRO structures

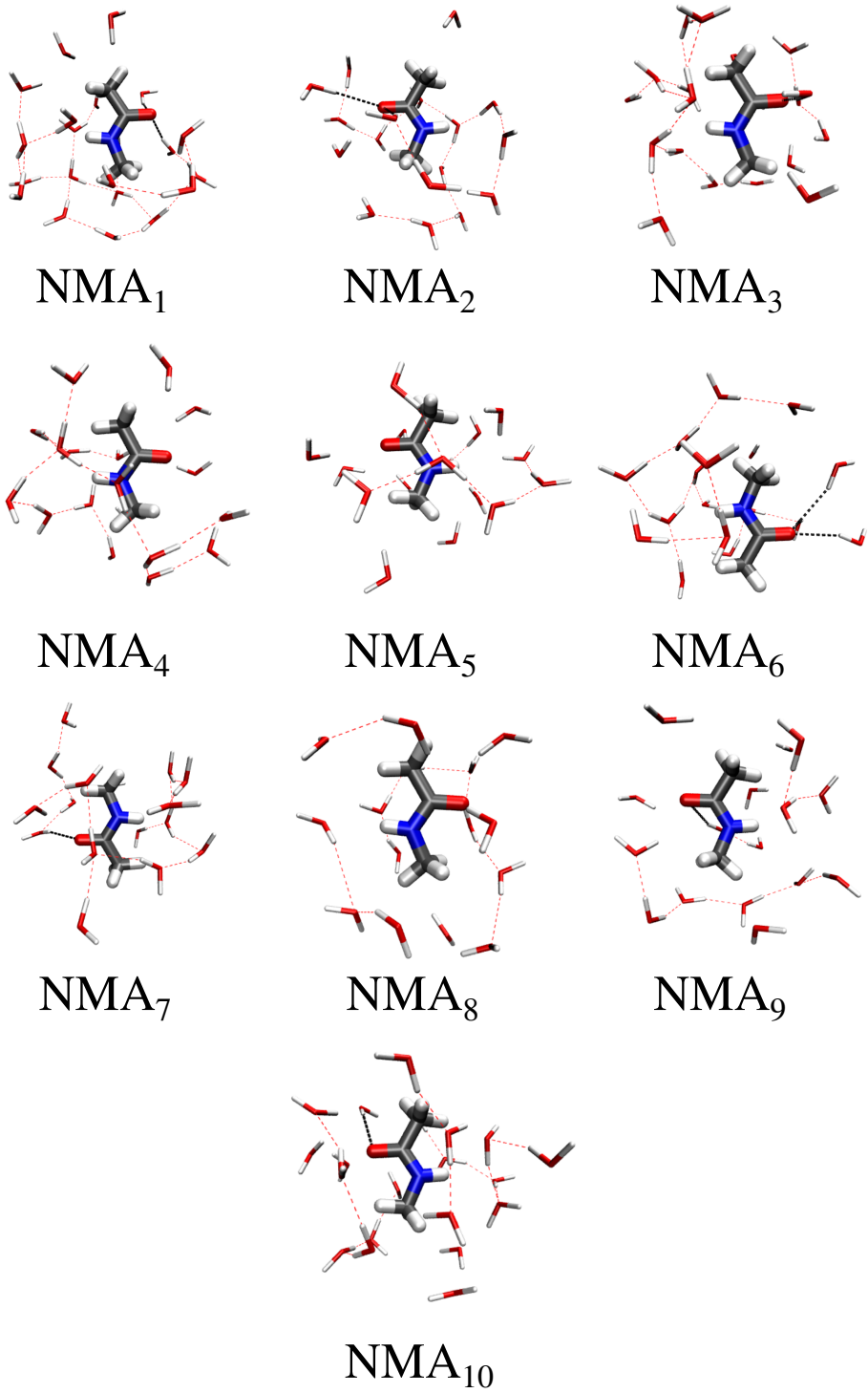


Figure S3: NMA structures

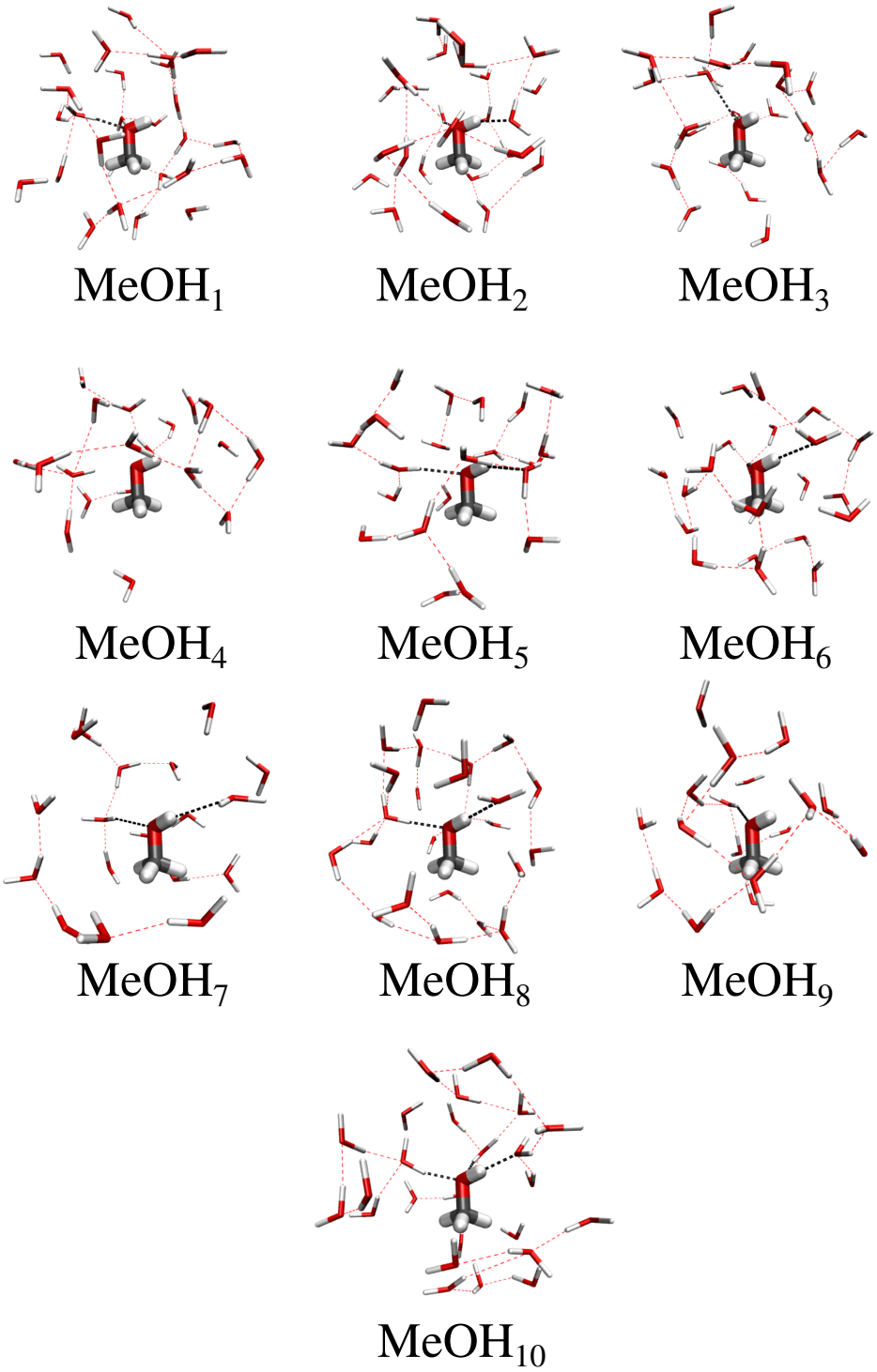


Figure S4: MeOH structures

S5.3 Electrostatic energy with CT between solvent molecules: 6-311++G** basis set

Table S4: Calculated electrostatic energies by exploiting QM/6-311++G**/FQFμ and SAPT0/6-311++G**. QM/FQF$\mu|_q$ represents the electrostatic contribution arising from the charges, whereas QM/FQF$\mu|_\mu$ that arising from the dipoles.

Structure	QM/FQF$\mu\|_q$	QM/FQF$\mu\|_\mu$	QM/FQF$\mu\|_{\text{tot}}$	SAPT0
$MOXY_1$	-21.95 (79%)	-5.67 (21%)	-27.62	-27.45
$MOXY_2$	-19.04 (75%)	-6.37 (25%)	-25.40	-24.33
$MOXY_3$	-21.62 (77%)	-6.44 (23%)	-28.05	-30.69
$MOXY_4$	-17.90 (74%)	-6.26 (26%)	-24.16	-25.73
$MOXY_5$	-22.74 (77%)	-6.65 (23%)	-29.40	-29.02
$MOXY_6$	-19.72 (76%)	-6.38 (24%)	-26.11	-26.35
$MOXY_7$	-15.73 (76%)	-4.99 (24%)	-20.72	-19.93
$MOXY_8$	-11.79 (75%)	-3.86 (25%)	-15.65	-17.95
$MOXY_9$	-14.15 (78%)	-4.06 (22%)	-18.22	-20.43
$MOXY_{10}$	-9.61 (72%)	-3.66 (28%)	-13.27	-16.20
$ACRO_1$	-16.97 (74%)	-5.88 (26%)	-22.86	-23.84
$ACRO_2$	-14.98 (76%)	-4.66 (24%)	-19.63	-18.95
$ACRO_3$	-5.67 (81%)	-1.35 (19%)	-7.02	-7.02
$ACRO_4$	-11.97 (75%)	-3.98 (25%)	-15.95	-19.36
$ACRO_5$	-15.11 (72%)	-5.78 (28%)	-20.89	-19.04
$ACRO_6$	-14.93 (75%)	-5.07 (25%)	-20.00	-19.85
$ACRO_7$	-13.62 (74%)	-4.85 (26%)	-18.47	-18.75
$ACRO_8$	-11.10 (75%)	-3.73 (25%)	-14.83	-15.12
$ACRO_9$	-10.31 (73%)	-3.89 (27%)	-14.20	-14.56
$ACRO_{10}$	-8.38 (74%)	-2.89 (26%)	-11.26	-11.73
$MeOH_1$	-26.25 (76%)	-8.32 (24%)	-34.57	-37.27
$MeOH_2$	-28.50 (75%)	-9.64 (25%)	-38.14	-38.86
$MeOH_3$	-22.68 (75%)	-7.39 (25%)	-30.07	-29.62
$MeOH_4$	-10.69 (76%)	-3.35 (24%)	-14.04	-13.60
$MeOH_5$	-11.68 (76%)	-3.65 (24%)	-15.33	-16.40
$MeOH_6$	-23.26 (75%)	-7.59 (25%)	-30.85	-32.70
$MeOH_7$	-10.87 (75%)	-3.58 (25%)	-14.46	-15.80
$MeOH_8$	-29.55 (76%)	-9.31 (24%)	-38.85	-40.23
$MeOH_9$	-22.13 (75%)	-7.48 (25%)	-29.61	-31.68
$MeOH_{10}$	-19.25 (78%)	-5.55 (22%)	-24.80	-23.32
NMA_1	-20.73 (75%)	-6.89 (25%)	-27.62	-27.61
NMA_2	-22.68 (75%)	-7.51 (25%)	-30.19	-28.77
NMA_3	-26.09 (76%)	-8.18 (24%)	-34.27	-32.62
NMA_4	-23.94 (75%)	-8.08 (25%)	-32.02	-33.03
NMA_5	-13.59 (76%)	-4.37 (24%)	-17.96	-16.86
NMA_6	-26.59 (76%)	-8.37 (24%)	-34.95	-33.19
NMA_7	-19.94 (77%)	-5.95 (23%)	-25.89	-26.12
NMA_8	-15.52 (78%)	-4.44 (22%)	-19.97	-18.58
NMA_9	-18.55 (77%)	-5.53 (23%)	-24.07	-23.13
NMA_{10}	-24.41 (77%)	-7.19 (23%)	-31.60	-31.61

Table S5: Comparison between calculated QM/FQ[a], QM/FQ[b], QM/FQ[c], QM/FQFμ E_{ele} (HF/6-311++G** level for the QM portion) and SAPT0/6-311++G** data. In case of SAPT0 calculations electrostatic and induction energy contribution are summed up. All data are given in kcal/mol.
[a] FQ parametrization taken from Ref.[17]
[b] FQ parametrization taken from Ref.[18]
[c] FQ parametrization taken from Ref.[19]

	QM/FQ[a]	QM/FQ[b]	QM/FQ[c]	QM/FQFμ	SAPT0
$MOXY_1$	-20.24	-23.33	-41.06	-27.62	-27.45
$MOXY_2$	-13.24	-16.41	-28.75	-25.40	-24.33
$MOXY_3$	-16.70	-20.32	-33.79	-28.05	-30.69
$MOXY_4$	-13.21	-16.25	-27.53	-24.16	-25.73
$MOXY_5$	-19.51	-22.78	-36.40	-29.40	-29.02
$MOXY_6$	-15.15	-18.21	-29.57	-26.11	-26.35
$MOXY_7$	-12.74	-15.80	-25.46	-20.72	-19.93
$MOXY_8$	-8.72	-10.92	-17.54	-15.65	-17.95
$MOXY_9$	-11.46	-13.85	-22.35	-18.22	-20.43
$MOXY_{10}$	-6.80	-7.71	-11.56	-13.27	-16.20
$ACRO_1$	-12.63	-15.42	-24.42	-22.86	-23.84
$ACRO_2$	-11.75	-14.15	-23.09	-19.63	-18.95
$ACRO_3$	-6.65	-7.30	-10.97	-7.02	-7.02
$ACRO_4$	-9.21	-11.27	-18.38	-15.95	-19.36
$ACRO_5$	-10.49	-12.84	-20.19	-20.89	-19.04
$ACRO_6$	-11.63	-14.10	-21.54	-20.00	-19.85
$ACRO_7$	-9.45	-11.90	-18.06	-18.47	-18.75
$ACRO_8$	-9.01	-10.88	-16.93	-14.83	-15.12
$ACRO_9$	-7.64	-9.64	-15.56	-14.20	-14.56
$ACRO_{10}$	-6.99	-8.39	-12.91	-11.26	-11.73
$MeOH_1$	-22.01	-27.10	-43.98	-34.57	-37.27
$MeOH_2$	-16.39	-20.02	-30.83	-38.14	-38.86
$MeOH_3$	-19.46	-23.84	-39.02	-30.07	-29.62
$MeOH_4$	-8.91	-10.95	-17.19	-14.04	-13.60
$MeOH_5$	-9.78	-11.72	-17.91	-15.33	-16.40
$MeOH_6$	-18.63	-22.48	-36.19	-30.85	-32.70
$MeOH_7$	-8.82	-10.77	-16.88	-14.46	-15.80
$MeOH_8$	-24.34	-29.07	-46.23	-38.85	-40.23
$MeOH_9$	-16.39	-20.02	-30.83	-29.61	-31.68
$MeOH_{10}$	-16.88	-19.93	-31.05	-24.80	-23.32
NMA_1	-14.71	-17.56	-26.94	-27.62	-27.61
NMA_2	-15.70	-18.61	-27.14	-30.19	-28.77
NMA_3	-20.83	-24.02	-35.90	-34.27	-32.62
NMA_4	-15.74	-19.02	-28.52	-32.02	-33.03
NMA_5	-9.61	-11.11	-16.02	-17.96	-16.86
NMA_6	-18.32	-22.16	-33.63	-34.95	-33.19
NMA_7	-15.20	-17.61	-25.81	-25.89	-26.12
NMA_8	-13.99	-16.35	-25.44	-19.97	-18.58
NMA_9	-14.15	-16.56	-25.09	-24.07	-23.13
NMA_{10}	-19.10	-22.45	-33.79	-31.60	-31.61

S5.4 Electrostatic energies without CT between solvent molecules: 6-311++G** basis set

Table S6: Comparison between calculated QM/FQ[a], QM/FQ[b], QM/FQ[c], QM/FQFμ E_{ele} (HF/6-311++G** level for the QM portion) and SAPT0/6-311++G** data. In case of SAPT0 calculations electrostatic and induction energy contribution are summed up. All data are given in kcal/mol.
[a] FQ parametrization taken from Ref.[17]
[b] FQ parametrization taken from Ref.[18]
[c] FQ parametrization taken from Ref.[19]

	QM/FQ[a]	QM/FQ[b]	QM/FQ[c]	QM/FQFμ	SAPT0
$MOXY_1$	-10.73	-13.10	-23.81	-18.48	-27.45
$MOXY_2$	-12.24	-16.63	-32.94	-22.89	-24.33
$MOXY_3$	-13.60	-17.59	-32.65	-24.37	-30.69
$MOXY_4$	-12.44	-16.67	-30.87	-22.69	-25.73
$MOXY_5$	-13.46	-16.67	-28.85	-22.97	-29.02
$MOXY_6$	-12.86	-16.76	-30.61	-22.38	-26.35
$MOXY_7$	-9.31	-12.05	-20.49	-16.78	-19.93
$MOXY_8$	-8.00	-10.86	-19.25	-14.21	-17.95
$MOXY_9$	-6.71	-8.25	-13.49	-12.90	-20.43
$MOXY_{10}$	-6.36	-8.49	-16.04	-12.27	-16.20
$ACRO_1$	-11.73	-15.67	-28.38	-21.31	-23.84
$ACRO_2$	-9.28	-12.16	-22.92	-16.02	-18.95
$ACRO_3$	-0.80	-0.58	-1.03	-2.18	-7.02
$ACRO_4$	-9.21	-12.34	-22.38	-15.50	-19.36
$ACRO_5$	-9.25	-12.28	-21.84	-18.01	-19.04
$ACRO_6$	-8.56	-11.33	-19.60	-15.86	-19.85
$ACRO_7$	-10.68	-14.47	-24.58	-19.01	-18.75
$ACRO_8$	-6.16	-7.63	-12.37	-11.52	-15.12
$ACRO_9$	-7.58	-10.32	-18.50	-13.69	-14.56
$ACRO_{10}$	-4.99	-6.61	-12.12	-8.93	-11.73
$MeOH_1$	-17.11	-21.91	-37.83	-30.15	-37.27
$MeOH_2$	-18.45	-24.14	-41.54	-33.81	-38.86
$MeOH_3$	-11.93	-15.07	-24.69	-23.03	-29.62
$MeOH_4$	-6.69	-8.70	-14.26	-11.69	-13.60
$MeOH_5$	-7.11	-9.13	-14.84	-12.70	-16.40
$MeOH_6$	-16.47	-21.14	-36.06	-28.20	-32.70
$MeOH_7$	-7.89	-10.29	-17.69	-13.37	-15.80
$MeOH_8$	-18.64	-23.62	-39.95	-33.05	-40.23
$MeOH_9$	-17.56	-23.30	-40.68	-29.84	-31.68
$MeOH_{10}$	-9.80	-12.11	-19.46	-17.53	-23.32
NMA_1	-13.17	-17.70	-30.48	-22.17	-27.61
NMA_2	-16.78	-22.57	-39.52	-25.07	-28.77
NMA_3	-12.27	-16.10	-28.32	-23.76	-32.62
NMA_4	-13.89	-19.41	-34.36	-27.82	-33.03
NMA_5	-7.78	-10.68	-19.15	-15.15	-16.86
NMA_6	-16.78	-22.57	-39.52	-31.15	-33.19
NMA_7	-12.27	-16.10	-28.32	-21.55	-26.12
NMA_8	-5.62	-7.49	-13.36	-12.10	-18.58
NMA_9	-9.17	-12.30	-21.76	-17.86	-23.13
NMA_{10}	-12.39	-15.94	-26.81	-22.79	-31.61

S5.5 Electrostatic energy with CT between solvent molecules: 6-31+G* basis set

Table S7: Calculated electrostatic energies by exploiting QM/6-31+G*/FQFμ and SAPT0/6-31+G*. QM/FQF$\mu|_q$ represents the electrostatic contribution arising from the charges, whereas QM/FQF$\mu|_\mu$ that arising from the dipoles.

Structure	QM/FQF$\mu\|_q$	QM/FQF$\mu\|_\mu$	QM/FQF$\mu\|_{\text{tot}}$	SAPT0
$MOXY_1$	-21.98 (80%)	-5.66 (20%)	-27.64	-27.45
$MOXY_2$	-19.04 (75%)	-6.37 (25%)	-25.41	-24.33
$MOXY_3$	-21.55 (77%)	-6.43 (23%)	-27.97	-30.69
$MOXY_4$	-17.98 (74%)	-6.27 (26%)	-24.25	-25.73
$MOXY_5$	-22.76 (77%)	-6.65 (23%)	-29.41	-29.02
$MOXY_6$	-19.82 (76%)	-6.41 (24%)	-26.23	-26.35
$MOXY_7$	-15.75 (76%)	-5.00 (24%)	-20.75	-19.93
$MOXY_8$	-11.91 (75%)	-3.89 (25%)	-15.80	-17.95
$MOXY_9$	-14.38 (78%)	-4.13 (22%)	-18.52	-20.43
$MOXY_{10}$	-9.70 (72%)	3.69 (28%)	-13.39	-16.20
$ACRO_1$	-16.82 (74%)	-5.84 (26%)	-22.66	-23.84
$ACRO_2$	-15.01 (76%)	-4.67 (24%)	-19.69	-18.95
$ACRO_3$	-5.64 (81%)	1.34 (19%)	-6.99	-7.02
$ACRO_4$	-12.13 (75%)	-4.03 (25%)	-16.16	-19.36
$ACRO_5$	-15.12 (72%)	-5.80 (28%)	-20.92	-19.04
$ACRO_6$	-14.84 (75%)	-5.04 (25%)	-19.89	-19.85
$ACRO_7$	-13.60 (74%)	-4.85 (26%)	-18.45	-18.75
$ACRO_8$	-11.18 (75%)	-3.77 (25%)	-14.95	-15.12
$ACRO_9$	-10.24 (73%)	-3.87 (27%)	-14.11	-14.56
$ACRO_{10}$	-8.29 (74%)	2.86 (26%)	-11.14	-11.73
$MeOH_1$	-26.61 (76%)	-8.44 (24%)	-35.05	-37.27
$MeOH_2$	-28.82 (75%)	-9.74 (25%)	-38.57	-38.86
$MeOH_3$	-22.91 (75%)	-7.45 (25%)	-30.36	-29.62
$MeOH_4$	-10.89 (76%)	-3.42 (24%)	-14.31	-13.60
$MeOH_5$	-11.92 (76%)	-3.71 (24%)	-15.63	-16.40
$MeOH_6$	-23.89 (75%)	-7.79 (25%)	-31.68	-32.70
$MeOH_7$	-11.05 (75%)	-3.64 (25%)	-14.68	-15.80
$MeOH_8$	-30.17 (76%)	-9.50 (24%)	-39.66	-40.23
$MeOH_9$	-22.45 (75%)	-7.58 (25%)	-30.03	-31.68
$MeOH_{10}$	-19.59 (78%)	-5.64 (22%)	-25.23	-23.32
NMA_1	-20.76 (75%)	-6.89 (25%)	-27.66	-27.61
NMA_2	-22.61 (75%)	-7.48 (25%)	-30.09	-28.77
NMA_3	-26.03 (76%)	-8.17 (24%)	-34.20	-32.62
NMA_4	-23.77 (75%)	-8.04 (25%)	-31.81	-33.03
NMA_5	-13.50 (76%)	-4.36 (24%)	-17.86	-16.86
NMA_6	-26.43 (76%)	-8.32 (24%)	-34.75	-33.19
NMA_7	-19.79 (77%)	-5.90 (23%)	-25.70	-26.12
NMA_8	-15.43 (78%)	-4.40 (22%)	-19.83	-18.58
NMA_9	-18.40 (77%)	-5.49 (23%)	-23.89	-23.13
NMA_{10}	-24.30 (77%)	-7.14 (23%)	-31.44	-31.61

Table S8: Comparison between calculated QM/FQ[a], QM/FQ[b], QM/FQ[c], QM/FQFμ E_{ele} (HF/6-31+G* level for the QM portion) and SAPT0/6-31+G* data. In case of SAPT0 calculations electrostatic and induction energy contribution are summed up. All data are given in kcal/mol.
[a] FQ parametrization taken from Ref.[17]
[b] FQ parametrization taken from Ref.[18]
[c] FQ parametrization taken from Ref.[19]

	QM/FQ[a]	QM/FQ[b]	QM/FQ[c]	QM/FQFμ	SAPT0
$MOXY_1$	-20.27	-23.36	-40.86	-27.64	-27.45
$MOXY_2$	-13.35	-16.49	-28.70	-25.41	-24.33
$MOXY_3$	-16.80	-20.38	-33.60	-27.97	-30.69
$MOXY_4$	-13.37	-16.39	-27.56	-24.25	-25.73
$MOXY_5$	-19.58	-22.83	-36.36	-29.41	-29.02
$MOXY_6$	-15.32	-18.37	-29.65	-26.23	-26.35
$MOXY_7$	-12.81	-15.84	-25.39	-20.75	-19.93
$MOXY_8$	-8.87	-11.07	-17.66	-15.80	-17.95
$MOXY_9$	-11.70	-14.10	-22.60	-18.52	-20.43
$MOXY_{10}$	-6.95	-7.85	-11.66	-13.39	-16.20
$ACRO_1$	-12.53	-15.29	-24.18	-22.66	-23.84
$ACRO_2$	-11.79	-14.18	-23.16	-19.69	-18.95
$ACRO_3$	-6.62	-7.26	-10.89	-6.99	-7.02
$ACRO_4$	-9.36	-11.44	-18.66	-16.16	-19.36
$ACRO_5$	-10.53	-12.88	-20.21	-20.92	-19.04
$ACRO_6$	-11.58	-14.03	-21.38	-19.89	-19.85
$ACRO_7$	-9.47	-11.91	-18.02	-18.45	-18.75
$ACRO_8$	-9.09	-10.97	-17.04	-14.95	-15.12
$ACRO_9$	-7.61	-9.59	-15.44	-14.11	-14.56
$ACRO_{10}$	-6.97	-8.33	-12.72	-11.14	-11.73
$MeOH_1$	-21.21	-25.40	-40.26	-35.05	-37.27
$MeOH_2$	-22.47	-27.54	-44.33	-38.57	-38.86
$MeOH_3$	-25.68	-30.52	-48.37	-30.36	-29.62
$MeOH_4$	-9.14	-11.19	-17.42	-14.31	-13.60
$MeOH_5$	-10.05	-11.99	-18.20	-15.63	-16.40
$MeOH_6$	-19.28	-23.17	-37.17	-31.68	-32.70
$MeOH_7$	-9.02	-10.97	-17.09	-14.68	-15.80
$MeOH_8$	-25.05	-29.78	-47.03	-39.66	-40.23
$MeOH_9$	-16.75	-20.40	-31.24	-30.03	-31.68
$MeOH_{10}$	-17.29	-20.34	-31.49	-25.23	-23.32
NMA_1	-14.84	-17.67	-26.98	-27.66	-27.61
NMA_2	-15.74	-18.62	-27.03	-30.09	-28.77
NMA_3	-20.87	-24.01	-35.71	-34.20	-32.62
NMA_4	-15.71	-18.93	-28.25	-31.81	-33.03
NMA_5	-9.61	-11.09	-15.89	-17.86	-16.86
NMA_6	-18.29	-22.09	-33.42	-34.75	-33.19
NMA_7	-15.17	-17.53	-25.55	-25.70	-26.12
NMA_8	-14.03	-16.32	-25.20	-19.83	-18.58
NMA_9	-14.12	-16.47	-24.76	-23.89	-23.13
NMA_{10}	-19.10	-22.40	-33.57	-31.44	-31.61

Table S9: Root Mean Squared Deviation (RMSD), Maximum Absolut Error (MAE) and Relative Error (RE) of ten selected structures of MOXY, ACRO, MeOH and NMA in aqueous solution extracted from aqueous solution. SAPT0/6-31+G* E_{ele} values are taken as reference. TOT indicates statistical parameters calculated on all 40 structures extracted from MD runs. RMSD and MAE are given in kcal/mol.
[a] FQ parametrization taken from Ref.[17]
[b] FQ parametrization taken from Ref.[18]
[c] FQ parametrization taken from Ref.[19]

		MOXY	ACRO	NMA	MeOH	TOT
QM/FQ[a]	RMSD	10.12	7.83	11.93	11.59	10.49
	MAE	13.89	11.32	17.32	16.40	17.32
	RE	42.51%	40.66%	41.28%	37.28%	40.43%
QM/FQ[b]	RMSD	7.40	5.75	9.16	9.19	8.00
	MAE	10.31	8.55	14.10	14.09	14.10
	RE	31.00%	29.41%	31.10%	29.04%	30.14%
QM/FQ[c]	RMSD	5.75	2.08	2.96	7.37	4.52
	MAE	13.41	4.22	6.63	18.75	18.75
	RE	18.90%	12.75%	8.99%	19.96%	15.06%
QM/FQFμ	RMSD	1.67	1.29	1.07	1.25	1.34
	MAE	2.81	3.20	1.58	2.23	3.20
	RE	6.44%	4.67%	3.61%	4.41%	4.78%

S5.6 Electrostatic energies without CT between solvent molecules: 6-31+G* basis set

Table S10: Comparison between calculated QM/FQ[a], QM/FQ[b], QM/FQ[c], QM/FQFμ E_{ele} (HF/6-31+G* level for the QM portion) and SAPT0/6-31+G* data. In case of SAPT0 calculations electrostatic and induction energy contribution are summed up. All data are given in kcal/mol.
[a] FQ parametrization taken from Ref.[17]
[b] FQ parametrization taken from Ref.[18]
[c] FQ parametrization taken from Ref.[19]

	QM/FQ[a]	QM/FQ[b]	QM/FQ[c]	QM/FQFμ	SAPT0
$MOXY_1$	-10.73	-13.10	-23.74	-18.46	-27.45
$MOXY_2$	-12.27	-16.62	-32.73	-22.86	-24.33
$MOXY_3$	-13.65	-17.58	-32.29	-24.28	-30.69
$MOXY_4$	-12.53	-16.72	-30.78	-22.73	-25.73
$MOXY_5$	-13.45	-16.61	-28.71	-22.93	-29.02
$MOXY_6$	-13.01	-16.90	-30.69	-22.53	-26.35
$MOXY_7$	-9.33	-12.06	-20.40	-16.78	-19.93
$MOXY_8$	-8.11	-10.97	-19.31	-14.32	-17.95
$MOXY_9$	-6.85	-8.40	-13.66	-13.13	-20.43
$MOXY_{10}$	-6.43	-8.53	-15.98	-12.33	-16.20
$ACRO_1$	-11.64	-15.54	-28.13	-21.13	-23.84
$ACRO_2$	-9.29	-12.16	-22.94	-16.05	-18.95
$ACRO_3$	-0.79	-0.55	-0.99	-2.15	-7.02
$ACRO_4$	-9.30	-12.44	-22.52	-15.65	-19.36
$ACRO_5$	-9.30	-12.33	-21.86	-18.06	-19.04
$ACRO_6$	-8.52	-11.26	-19.41	-15.76	-19.85
$ACRO_7$	-10.69	-14.46	-24.49	-18.98	-18.75
$ACRO_8$	-6.24	-7.72	-12.48	-11.64	-15.12
$ACRO_9$	-7.56	-10.28	-18.37	-13.61	-14.56
$ACRO_{10}$	-4.94	-6.51	-11.85	-8.79	-11.73
$MeOH_1$	-17.44	-22.26	-38.12	-30.56	-37.27
$MeOH_2$	-18.79	-24.48	-41.83	-34.18	-38.86
$MeOH_3$	-19.97	-25.04	-41.79	-35.97	-29.62
$MeOH_4$	-6.81	-8.82	-14.35	-11.86	-13.60
$MeOH_5$	-7.25	-9.27	-14.98	-12.90	-16.40
$MeOH_6$	-16.94	-21.67	-36.80	-28.89	-32.70
$MeOH_7$	-8.04	-10.45	-17.87	-13.56	-15.80
$MeOH_8$	-19.10	-24.09	-40.46	-33.69	-40.23
$MeOH_9$	-17.85	-23.61	-40.98	-30.20	-31.68
$MeOH_{10}$	-10.01	-12.32	-19.68	-17.81	-23.32
NMA_1	-10.83	-14.59	-25.86	-22.18	-27.61
NMA_2	-13.21	-17.70	-30.34	-25.02	-28.77
NMA_3	-12.14	-15.73	-26.72	-23.71	-32.62
NMA_4	-13.82	-19.27	-33.98	-27.63	-33.03
NMA_5	-7.77	-10.64	-18.94	-15.08	-16.86
NMA_6	-16.74	-22.49	-39.25	-30.98	-33.19
NMA_7	-12.22	-15.98	-27.98	-21.38	-26.12
NMA_8	-5.56	-7.35	-13.01	-11.89	-18.58
NMA_9	-9.17	-12.25	-21.52	-17.75	-23.13
NMA_{10}	-12.38	-15.89	-26.58	-22.66	-31.61

S5.7 MOXY$_1$ with a cutting radius of 7 Å

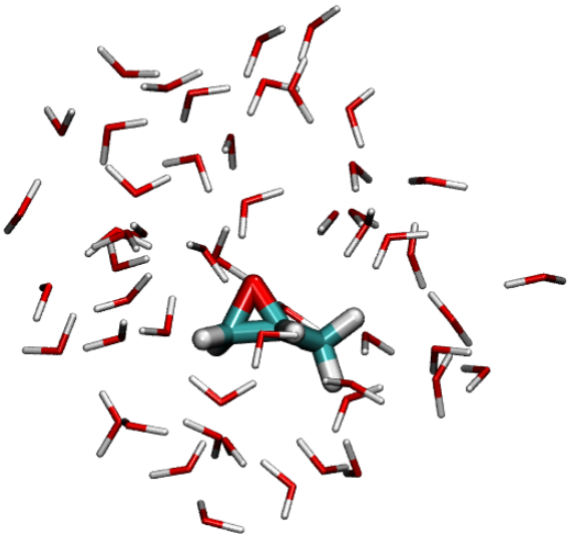


Figure S5: MOXY$_1$-cluster structure obtained by using a cutting radius of 7.0 Å. The number of water molecules included in the cluster is 53.

Table S11: Calculated electrostatic energies for MOXY$_1$ in Figure S5 obtained by exploiting QM/6-31G*/FQFμ and SAPT0/6-31G*. QM/FQF$\mu|_q$ represents the electrostatic contribution arising from the charges, whereas QM/FQF$\mu|_\mu$ that arising from the dipoles.

Structure	QM/FQF$\mu\|_q$	QM/FQF$\mu\|_\mu$	QM/FQF$\mu\|_{\text{tot}}$	SAPT0
MOXY$_1$	-22.56	-5.74	-28.31	-27.74

References

(1) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730–4737.

(2) Rinkevicius, Z.; Li, X.; Sandberg, J. A.; Mikkelsen, K. V.; Ågren, H. *J. Chem. Theory Comput.* **2014**, *10*, 989–1003.

(3) Rappe, A. K.; Goddard III, W. A. *J. Phys. Chem* **1991**, *95*, 3358–3363.

(4) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell Jr, A. D. *Chem. Rev.* **2016**, *116*, 4983–5013.

(5) Naserifar, S.; Brooks, D. J.; Goddard III, W. A.; Cvicek, V. *J. Chem. Phys.* **2017**, *146*, 124117.

(6) Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.

(7) Barone, V.; Cossi, M. *J. Phys. Chem A* **1998**, *102*, 1995–2001.

(8) Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2013**, *9*, 1880–1884.

(9) Berendsen, H.; van der Spoel, D.; van Drunen, R. *Comp. Phys. Comm.* **1995**, *91*, 43 – 56.

(10) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.

(11) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(12) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(13) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem* **1992**, *13*, 952–962.

(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(15) Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(16) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(17) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

(18) Carnimeo, I.; Cappelli, C.; Barone, V. *J. Comput. Chem.* **2015**, *36*, 2271–2290.

(19) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. *submitted*