Supporting information for:

On the brink of self-hydration: the water heptadecamer

B. Hartke

Institute for Physical Chemistry,
Kiel University,
Olshausenstr. 40,
24098 Kiel, GERMANY

∗ *e-mail: hartke@pctc.uni-kiel.de*

1 All (H₂O)₁₇ structures of the main paper

In the main paper, four different classes of $(H_2O)_{17}$ low-energy structures were shown with one representative structure. Here, all twelve structures are displayed. Their atomic coordinates are provided in separate xyz-files. All of them were locally optimized at the B2PYLP-D4/def2-TZVPPD level, with dense grids and tight convergence criteria.

The "prism" class features two fused pentagonal prisms, with one water dimer attached across a square side face of one of these prisms. As displayed in Fig. S1, the structures p1, p2, p3 and p4 have (qualitatively) the same O-atom cage but partly different hydrogen-bond arrangements. p1 is the global minimum candidate at half of the ab-initio/DFT levels employed here, and also for MB-pol2023.
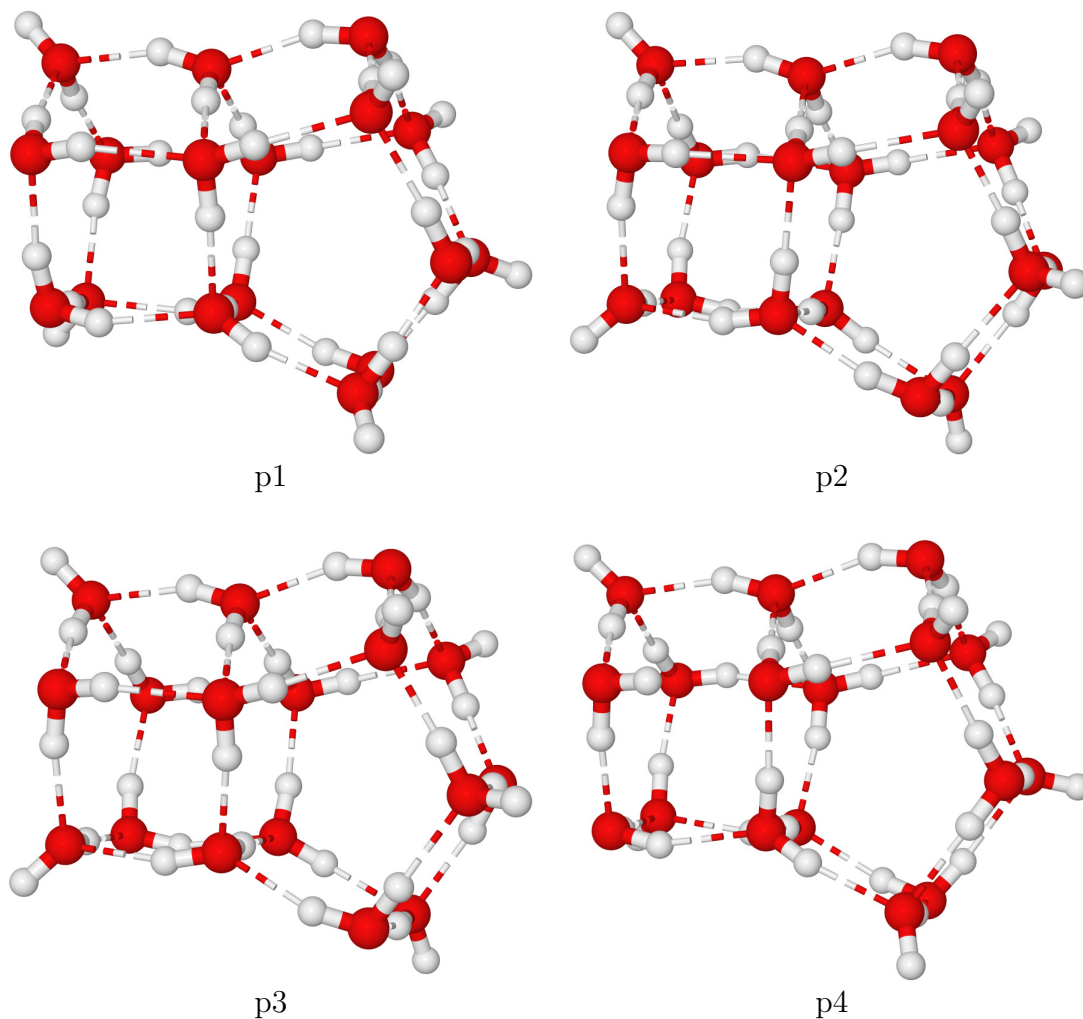

p1   p2

p3   p4

Figure S1: The four p-structures.

In the "twisted-cubes" class, there are two slightly deformed cubes, not along a common axis and also rotated relative to each other. On one side these cubes are joined directly, on the other side via additional bridging water molecules (cf. Fig. S2). One version of this class, t2, is the TIP4P global minimum discovered by Wales and Hodges [1], which coincides with the "L-shape" structure presented by Xantheas et al. [2]. t2 has a rather high energy at the ab-initio/DFT levels used here. However, t1, a slight structural variation of t2 (not just a different H-bond pattern), has a very low energy

at the ab-initio/DFT levels employed here and is the global minimum candidate for q-AQUA-pol.
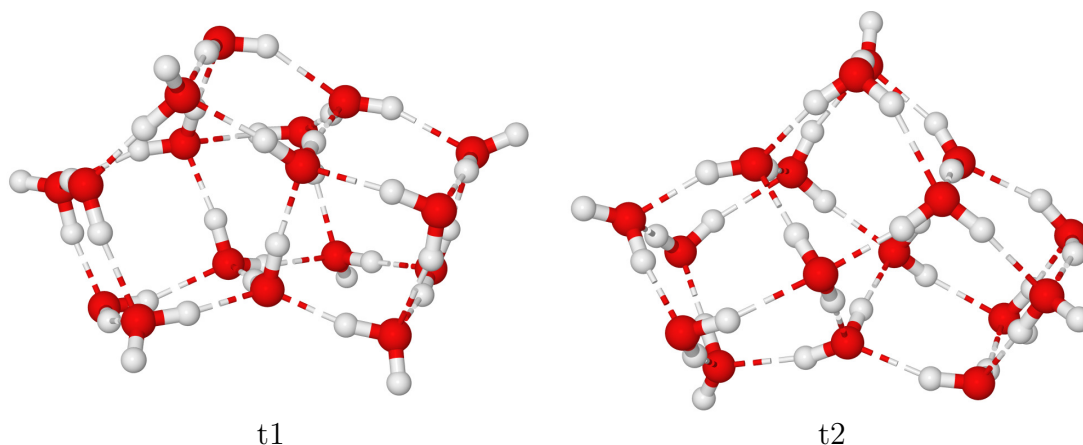

Figure S2: The two t-structures.

The "water-centered" class is the only one with actual water self-hydration, i.e., one water molecule is in or near the cluster center, surrounded by a hydration shell formed by water four-, five- and six-rings. w1 is the global minimum candidate at half of the ab-initio/DFT levels used here, and also for q-AQUA and TTM2-F/TTM3-F. When one of its six-rings is at the bottom (it does not matter which one), the ring sequence around its equator is 5-5-4-4-5-4, capped at the top by 4-5-6. w2, w3 and w4 are different H-bond patterns of a different ring arrangement: When a six-ring is at the bottom again, the ring sequence around the equator is 4-5-4-5-4-5, and the cap consists of three five-rings (cf. Fig. S3).

Finally, in the "cubes" class there are three fused cubes, aligned along a common central axis and not rotated relative to each other. The remaining 17th water molecule can now be added at different locations, as shown in Fig. S4: Adding it to one edge of the middle cube results in c1, which is lower in energy in most cases. Adding it to an end-edge of an end-cube produces c2. Neither c1 nor c2 are among the best structures at the ab-initio/DFT levels employed here, but c1 is the global minimum candidate for MB-pol2016.
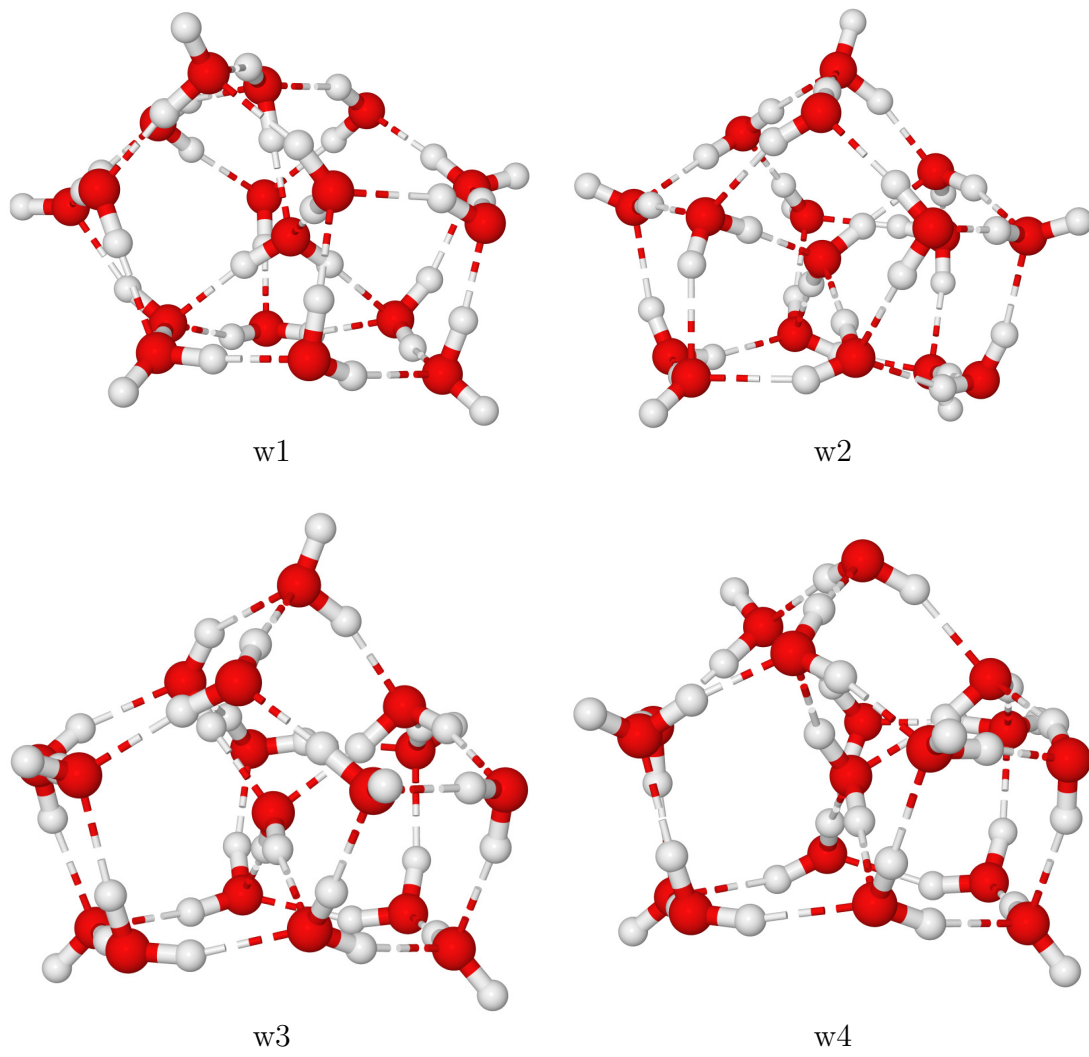
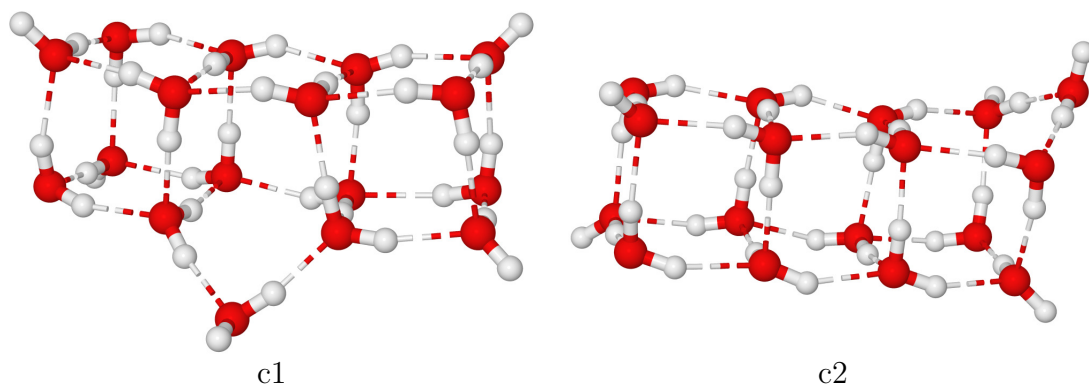Figure S3: The four w-structures.

Figure S4: The two c-structures.

2 Tabulated energy values

Complementing the energy figures in the main paper, this section collects all numerical energy values depicted there in tabular form. Each theory level has its own table, for all 12 cluster structures, listing both the "absolute" energy value in Hartree (total electronic energy for the ab-initio/DFT methods, and typically with respect to dissocation into water monomers for the water model potentials) and the relative energy differences between the cluster structures in kJ/mol (relative to the lowest absolute energy, at this level of theory).

Structure labels are the same as in the main paper, and as explained and depicted in section 1 of this Supplementary Information.

e34	DLPNO-CCSD(T)/CBS	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-1298.683\,307$	0.00
p2	$-1298.682\,909$	1.04
p3	$-1298.682\,367$	2.46
p4	$-1298.682\,237$	2.80
t1	$-1298.683\,041$	0.69
t2	$-1298.681\,062$	5.89
w1	$-1298.682\,739$	1.49
w2	$-1298.682\,154$	3.02
w3	$-1298.681\,289$	5.29
w4	$-1298.681\,285$	5.31
c1	$-1298.682\,373$	2.45
c2	$-1298.681\,623$	4.42

ccsdt/f12	DLPNO-CCSD(T)/cc-pVQZ-F12	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-1298.607\,198$	0.23
p2	$-1298.606\,874$	1.08
p3	$-1298.606\,338$	2.49
p4	$-1298.606\,223$	2.79
t1	$-1298.607\,044$	0.63
t2	$-1298.605\,400$	4.95
w1	$-1298.607\,286$	0.00
w2	$-1298.606\,725$	1.47
w3	$-1298.605\,779$	3.95
w4	$-1298.605\,762$	4.00
c1	$-1298.606\,578$	1.85
c2	$-1298.605\,828$	3.82

ccsdt/qz	DLPNO-CCSD(T)/def2-QZVPP	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-1298.420\,935$	0.00
p2	$-1298.420\,592$	0.90
p3	$-1298.420\,053$	2.31
p4	$-1298.419\,945$	2.59
t1	$-1298.420\,849$	0.22
t2	$-1298.418\,925$	5.27
w1	$-1298.420\,883$	0.13
w2	$-1298.420\,379$	1.46
w3	$-1298.419\,453$	3.89
w4	$-1298.419\,462$	3.86
c1	$-1298.419\,979$	2.51
c2	$-1298.419\,128$	4.74

mp2	RI-SCS-MP2/def2-TZVPPD	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-1298.126\,007$	1.60
p2	$-1298.125\,687$	2.43
p3	$-1298.125\,201$	3.71
p4	$-1298.124\,957$	4.35
t1	$-1298.125\,899$	1.88
t2	$-1298.123\,355$	8.56
w1	$-1298.126\,616$	0.00
w2	$-1298.126\,275$	0.89
w3	$-1298.125\,072$	4.05
w4	$-1298.125\,237$	3.62
c1	$-1298.124\,396$	5.83
c2	$-1298.123\,574$	7.98

b2plyp	B2PLYP-D4/def2-TZVPPD	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-1299.366\,933$	0.00
p2	$-1299.366\,525$	1.07
p3	$-1299.366\,022$	2.39
p4	$-1299.365\,961$	2.55
t1	$-1299.366\,684$	0.65
t2	$-1299.364\,053$	7.56
w1	$-1299.366\,612$	0.84
w2	$-1299.366\,352$	1.52
w3	$-1299.365\,243$	4.43
w4	$-1299.365\,154$	4.67
c1	$-1299.365\,344$	4.17
c2	$-1299.364\,355$	6.76

wb97	ω**B97M-D4/def2-TZVPPD**	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-1300.422\,829$	2.29
p2	$-1300.422\,494$	3.17
p3	$-1300.421\,941$	4.62
p4	$-1300.421\,832$	4.91
t1	$-1300.422\,798$	2.37
t2	$-1300.421\,255$	6.42
w1	$-1300.423\,703$	0.00
w2	$-1300.423\,101$	1.58
w3	$-1300.422\,013$	4.43
w4	$-1300.422\,232$	3.86
c1	$-1300.422\,653$	2.75
c2	$-1300.421\,790$	5.02

mbpol2016	**MB-pol2016**	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.274\,008$	0.77
p2	$-0.272\,592$	4.49
p3	$-0.272\,038$	5.94
p4	$-0.272\,626$	4.40
t1	$-0.273\,327$	2.56
t2	$-0.271\,935$	6.22
w1	$-0.273\,706$	1.56
w2	$-0.272\,680$	4.26
w3	$-0.271\,919$	6.26
w4	$-0.271\,613$	7.06
c1	$-0.274\,304$	0.00
c2	$-0.273\,688$	1.61

mbpol2023	**MB-pol2023**	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.277\,555$	0.00
p2	$-0.275\,663$	4.96
p3	$-0.275\,084$	6.48
p4	$-0.275\,773$	4.67
t1	$-0.276\,705$	2.23
t2	$-0.274\,098$	9.07
w1	$-0.276\,777$	2.04
w2	$-0.275\,796$	4.61
w3	$-0.275\,300$	5.92
w4	$-0.273\,834$	9.76
c1	$-0.276\,861$	1.82
c2	$-0.276\,361$	3.13

qaqua	q-AQUA	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.281\,220$	4.56
p2	$-0.281\,651$	3.43
p3	$-0.280\,649$	6.06
p4	$-0.280\,528$	6.38
t1	$-0.281\,699$	3.30
t2	$-0.277\,311$	14.82
w1	$-0.282\,959$	0.00
w2	$-0.282\,638$	0.84
w3	$-0.281\,140$	4.77
w4	$-0.278\,712$	11.14
c1	$-0.277\,228$	15.04
c2	$-0.275\,978$	18.32

qaquapol	q-AQUA-pol	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.275\,637$	2.52
p2	$-0.275\,179$	3.72
p3	$-0.275\,817$	2.04
p4	$-0.276\,113$	1.27
t1	$-0.276\,598$	0.00
t2	$-0.273\,634$	7.78
w1	$-0.274\,806$	4.70
w2	$-0.275\,177$	3.73
w3	$-0.275\,195$	3.68
w4	$-0.271\,764$	12.69
c1	$-0.274\,764$	4.81
c2	$-0.273\,753$	7.46

ttm3f	TTM3-F	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.276\,652$	6.90
p2	$-0.277\,107$	5.70
p3	$-0.276\,111$	8.32
p4	$-0.275\,200$	10.71
t1	$-0.276\,070$	8.43
t2	$-0.277\,264$	5.29
w1	$-0.279\,281$	0.00
w2	$-0.277\,455$	4.79
w3	$-0.276\,924$	6.18
w4	$-0.278\,846$	1.14
c1	$-0.277\,955$	3.48
c2	$-0.278\,420$	2.25

amoeba+	AMOEBA+	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.274\,308$	3.88
p2	$-0.273\,985$	4.73
p3	$-0.273\,180$	6.84
p4	$-0.273\,171$	6.86
t1	$-0.272\,872$	7.65
t2	$-0.273\,689$	5.50
w1	$-0.275\,785$	0.00
w2	$-0.274\,295$	3.91
w3	$-0.274\,094$	4.44
w4	$-0.274\,489$	3.40
c1	$-0.275\,520$	0.69
c2	$-0.274\,908$	2.30

amoeba14	AMOEBA-2014	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.271\,896$	5.44
p2	$-0.271\,985$	5.20
p3	$-0.270\,965$	7.88
p4	$-0.270\,730$	8.50
t1	$-0.271\,025$	7.72
t2	$-0.271\,928$	5.35
w1	$-0.273\,967$	0.00
w2	$-0.272\,673$	3.40
w3	$-0.272\,309$	4.35
w4	$-0.272\,800$	3.06
c1	$-0.273\,125$	2.21
c2	$-0.272\,895$	2.81

tip4p2005	TIP4P/2005	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-0.302\,483$	4.37
p2	$-0.302\,732$	3.71
p3	$-0.302\,134$	5.28
p4	$-0.302\,015$	5.59
t1	$-0.302\,574$	4.13
t2	$-0.304\,145$	0.00
w1	$-0.303\,314$	2.18
w2	$-0.302\,800$	3.53
w3	$-0.302\,370$	4.66
w4	$-0.303\,407$	1.94
c1	$-0.302\,821$	3.48
c2	$-0.303\,027$	2.94

tip4p	TIP4P	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	−0.274 296	3.64
p2	−0.274 526	3.04
p3	−0.273 980	4.47
p4	−0.273 871	4.76
t1	−0.274 340	3.52
t2	−0.275 684	0.00
w1	−0.274 897	2.06
w2	−0.274 381	3.42
w3	−0.273 989	4.45
w4	−0.274 937	1.96
c1	−0.274 490	3.13
c2	−0.274 700	2.58

gfn2-xtb	GFN2-xTB	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	−86.479 330	7.41
p2	−86.479 382	7.28
p3	−86.478 902	8.54
p4	−86.478 527	9.52
t1	−86.479 299	7.50
t2	−86.481 660	1.30
w1	−86.482 156	0.00
w2	−86.481 114	2.73
w3	−86.481 518	1.67
w4	−86.481 499	1.72
c1	−86.481 428	1.91
c2	−86.481 625	1.39

tip3p	TIP3P	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	−0.269 678	5.57
p2	−0.270 697	2.90
p3	−0.269 411	6.27
p4	−0.267 847	10.38
t1	−0.268 016	9.93
t2	−0.266 757	13.24
w1	−0.271 802	0.00
w2	−0.268 872	7.69
w3	−0.266 597	13.66
w4	−0.270 062	4.56
c1	−0.266 512	13.88
c2	−0.268 489	8.69

gfn-ff	GFN-FF	
structure	absolute energy / Hartree	relative energy / kJ/mol
p1	$-5.836\,569$	2.30
p2	$-5.837\,448$	0.00
p3	$-5.835\,457$	5.22
p4	$-5.833\,811$	9.54
t1	$-5.834\,756$	7.06
t2	$-5.829\,560$	20.71
w1	$-5.836\,606$	2.21
w2	$-5.833\,809$	9.55
w3	$-5.831\,423$	15.81
w4	$-5.833\,225$	11.08
c1	$-5.828\,489$	23.52
c2	$-5.830\,776$	17.51

3 Deviation measures

To support the $(H_2O)_{17}$ energy graphs in the main paper, the following table list mean energy differences and standard deviations of these differences, between each method used and a reference. As shown in the main paper as one of the central points, there are small but significant differences between the results of the different ab-initio methods used. However, the differences between those ab-initio results on the one hand and the water model results on the other hand are larger still. Therefore, here it does not matter much which of the ab-initio methods are chosen as reference. Hence, we take as reference "e34", i.e., DLPNO-CCSD(T) CBS-extrapolated from triple- and quadruple-zeta basis sets. All data are based on the "relative energy" numbers in kJ/mol listed in the tables in the previous section and shown in the graphs in the main paper.

differences to e34		
method	mean / kJ/mol	standard deviation / kJ/mol
ccsdt/f12	0.680	0.617
ccsdt/qz	0.646	0.611
mp2	1.93	0.837
b2plyp	0.824	0.817
wb97m	1.42	0.686
mbpol2023	1.97	1.55
mbpol2016	1.73	1.13
quaquapol	2.29	1.85
quaqua	5.19	4.38
gfn2-xtb	4.19	2.52
gfn-ff	7.65	6.20
ttm3f	3.77	2.78
amoeba+	2.69	1.92
amoeba14	2.92	2.46
tip4p2005	2.48	1.66
tip4p	2.17	1.58
tip3p	5.54	3.35

In the present case, these statistical measures are less meaningful than one could hope. For example, for q-AQUA-pol, much of the total deviation is due to one single value (w4, easily visible in Fig.4 of the main paper). With this value deleted, the mean value of the differences is somewhat smaller (1.82 kJ/mol) while the standard deviation is cut in half (0.98 kJ/mol, even less than for both MB-pol versions).

However, this focus on one big deviation is not true for the other methods. In fact, deviation distributions turn out to be highly method/model-dependent, which tends to be hidden by the integrative statistical indicators shown in the above table. Hence, it is more informative to inspect the 12 actual values for each method, as shown in the main paper.

4 Further water model versions

To avoid cluttering the figures in the main paper, only one version of AMOEBA and TIP4P, respectively, is shown there. Fig. S5 demonstrates that for the $(H_2O)_{17}$ cluster structures under study, TIP4P [3] and TIP4P/2005 [4] data are very similar, as well as the AMOEBA2014 [5] and AMOEBA+ [6,7] results, respectively.
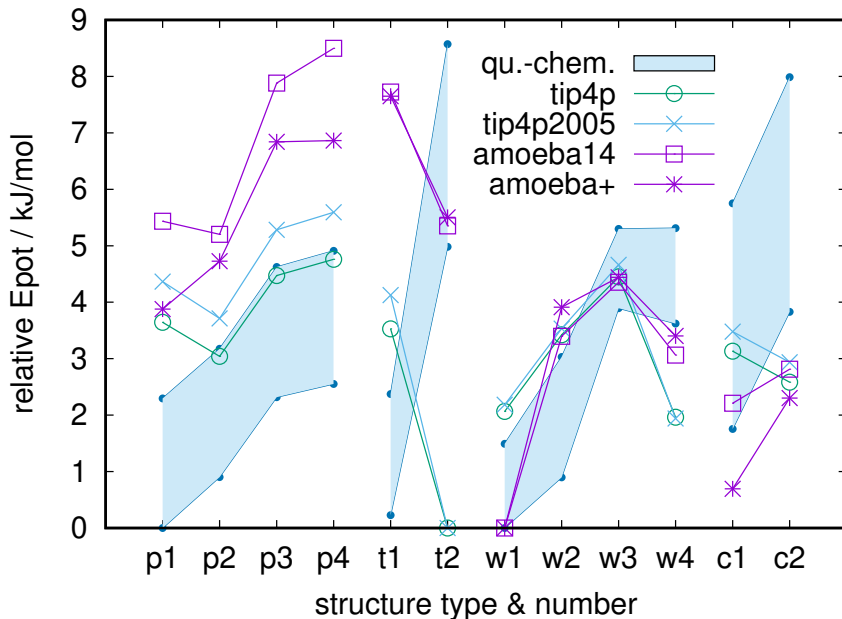

Figure S5: Relative potential energies for the full set of 12 structures, comparing TIP4P, TIP4P2005, AMOEBA2014 and AMOEBA+. As in the main paper, the shaded regions represent the quantum-chemistry results.

5 Water hexamer

To help gauging the DFT and ab-initio approaches used in the main paper, the following table lists total electronic energies of the well-known low-energy $(H_2O)_6$ structures "cage", "book" and "6-ring", relative to the "prism" and in direct comparison to literature values [8–13]. Since there is some spread in these literature values, due to differences in basis sets, CBS-extrapolations and other methodical issues, they are summarized in

value intervals. However, MP2 and CCSD(T) from the literature are shown separately. All values are in kJ/mol; the method abbreviations from the main paper are used. As in the main paper for $(H_2O)_{17}$, local structure optimizations were performed for SCS-MP2 and B2PLYP, while all CCSD(T) results are single-point calculations for the B2PLYP-optimized $(H_2O)_6$ structures.

structure	b2plyp	mp2	ccsdt/qz	ccsdt/f12	e34	lit.mp2	lit.ccsdt
prism	0.0	0.0	0.0	0.0	0.0	0.0	0.0
cage	0.22	0.42	1.49	1.31	1.73	$0.38 - 0.39$	$0.84 - 1.26$
book	0.32	0.62	2.86	2.93	2.54	$1.13 - 1.54$	$2.64 - 5.02$
6ring	3.78	3.23	6.74	7.15	5.84	$4.27 - 5.31$	$6.44 - 9.12$

References

[1] D. J. Wales and M. P. Hodges, Chem. Phys. Lett. 286 (1998) 65.

[2] S. Yoo, E. Aprà, X. C. Zeng and S. Xantheas, J. Phys. Chem. Lett. 1 (2010) 3122.

[3] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, J. Chem. Phys. 79 (1983) 926.

[4] J. L. F. Abascal and C. Vega, J. Chem. Phys. 123 (2005) 234505.

[5] M. L. Laury, L.-P. Wang, V. S. Pande, T. Head-Gordon and J. W. Ponder, J. Phys. Chem. B 119 (2015) 9423.

[6] C. Liu, J.-P. Piquemal and P. Ren, J. Chem. Theory Comput. 15 (2019) 4122.

[7] C. Liu, J.-P. Piquemal and P. Ren, J. Phys. Chem. Lett. 11 (2020) 419.

[8] K. M. Herman and S. S. Xantheas, Phys. Chem. Chem. Phys. 25 (2023) 7120.

[9] B. Temelso, K. A. Archer and G. C. Shields, J. Phys. Chem. A 115 (2011) 12034.

[10] D. M. Bates and G. S. Tschumper, J. Phys. Chem. A 113 (2009) 3555.

[11] D. M. Bates, J. R. Smith, T. Janowski and G. S. Tschumper, J. Chem. Phys. 135 (2011) 044123.

[12] R. M. Olson, J. L. Bentz, R. A. Kendall, M. W. Schmidt and M. S. Gordon, J. Chem. Theory Comput. 3 (2007) 1312.

[13] E. Miliordos and S. S. Xantheas, J. Chem. Phys. 142 (2015) 234303.