## **Supplementary Information for:**

3D-RISM-KH Molecular Theory of Solvation and Density Functional Theory Investigation of the Role of Water in the Aggregation of Model Asphaltenes

Leonardo M. da Costa,<sup>a,b</sup> Seigo Hayaki,<sup>a,c</sup> Stanislav R. Stoyanov,<sup>a</sup> Sergey Gusarov,<sup>a</sup> Xiaoli Tan,<sup>d</sup> Murray R. Gray,<sup>d</sup> Jeffrey M. Stryker,<sup>e</sup> Rik Tykwinski,<sup>f</sup> J. Walkimar M. Carneiro,<sup>b</sup> Hirofumi Sato,<sup>c</sup> Peter R. Seidl,<sup>g</sup> and Andriy Kovalenko<sup>a,h,\*</sup>

<sup>a</sup> National Institute for Nanotechnology, National Research Council of Canada, 11421 Saskatchewan Drive, Edmonton, Alberta, T6G 2M9, Canada; <sup>b</sup> Institute of Chemistry, Universidade Federal Fluminense, Niteroi, Rio de Janeiro, Brazil; <sup>c</sup> Department of Molecular Engineering, Kyoto University, Kyoto Daigaku Katsura, Kyoto, Japan; <sup>d</sup>Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, Canada; <sup>e</sup> Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada; <sup>f</sup> Department of Chemistry and Pharmacy, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; <sup>g</sup> Department of Organic Processes, School of Chemistry, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; <sup>h</sup> Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada

\* Corresponding author. E-mail: <u>andriy.kovalenko@nrc-cnrc.gc.ca</u>

.

Table S1: DFT method exploration: Gas phase <sup>1</sup>H NMR shifts for benzene, pyridine and pyrene relative to TMS in CHCl<sub>3</sub>

Basis	Proto	n EXP	B3LYP-D	Error (%)	WB97XD	Error (%)	B97D	Error (%)
	Benzene							
6-31G(d)	Н	7.34	6.91	-5.86	7.12	-3.00	6.75	-8.04
6-31G(d,p)	Н		7.12	-3.00	7.30	-0.54	6.99	-4.77
6-31+G(d)	Н		7.09	-3.41	7.29	-0.68	6.89	-6.13
6-31+G(d,p)	Н		7.08	-3.54	7.45	1.50	7.14	-2.72
					Pyridine			
6-31G(d)	$\mathrm{H}_{\mathrm{l}}$	7.23	6.69	-7.47	6.89	-4.70	6.53	-9.68
	$\mathrm{H}_{2}$	8.59	8.34	-2.91	8.52	-0.81	8.17	-4.89
	$H_3$	7.62	7.07	-7.22	7.31	-4.07	6.88	-9.71
6-31G(d,p)	$\mathrm{H}_{1}$		6.93	-4.15	7.10	-1.80	6.79	-6.09
	$\mathrm{H}_{2}$		8.52	-0.81	8.67	0.93	8.40	-2.21
	$H_3$		7.28	-4.46	7.49	-1.71	7.12	-6.56
6-31+G(d)	$\mathrm{H}_{1}$		6.87	-4.98	7.04	-2.63	6.67	-7.75
	$\mathrm{H}_{2}$		8.50	-1.05	8.62	0.35	8.31	-3.26
	$H_3$		7.28	-4.46	7.48	-1.84	7.05	-7.48
6-31+G(d,p)	$\mathrm{H}_{1}$		7.10	-1.80	7.23	-0.01	6.93	-4.15
	$\mathrm{H}_{2}$		8.64	0.58	8.75	1.86	8.51	-0.93
	$\mathrm{H}_3$		7.48	-1.84	7.65	0.39	7.23	-5.12
					Pyrene			
6-31G(d)	$\mathrm{H}_{1}$	7.99	7.60	-4.88	7.74	-3.13	7.46	-6.63
	$\mathrm{H}_{2}$	8.16	7.72	-5.39	7.91	-3.06	7.55	-7.48
	${\rm H}_3$	8.05	7.71	-4.22	7.89	-1.99	7.55	-6.21
6-31G(d,p)	$\mathrm{H}_{1}$		7.79	-2.50	7.92	-0.88	7.66	-4.13
	$\mathrm{H}_{2}$		7.91	-3.06	8.09	-0.86	7.77	-4.78
	${\rm H}_3$		7.89	-1.99	8.06	0.12	7.75	-3.73
6-31+G(d)	$\mathrm{H}_{1}$		7.87	-1.50	7.79	-2.50	7.62	-4.63
	$\mathrm{H}_{2}$		7.93	-2.82	8.60	5.39	7.69	-5.76
	${\rm H}_3$		7.89	-1.99	7.94	-1.37	7.63	-5.22
6-31+G(d,p)	$\mathrm{H}_{1}$		8.04	0.63	7.91	-1.00	7.90	-1.13
	$\mathrm{H}_{2}$		8.14	-0.25	8.64	5.88	7.94	-2.70
	$H_3$		8.06	0.12	8.11	0.75	7.91	-1.74

Proton EXP Basis B3LYP-D Error (%) WB97XD Error (%) **B97D** Error (%) Benzene 6-31G(d) Η 7.34 7.17 -2.32 7.25 -1.23 6.86 -6.54 6-31G(d,p)Η 7.35 0.14 7.44 1.36 7.10 -3.27 6-31+G(d)Η 7.37 0.41 7.43 1.23 7.02 -4.36 3.13 7.59 6-31+G(d,p)Η 7.57 3.41 7.27 -0.95 Pyridine 6-31G(d) 7.23 7.08 -2.07 -6.92  $H_1$ 7.10 -1.806.73  $H_2$ 8.59 8.47 -1.40 8.52 -0.81 8.16 -5.01 7.09 -6.96 H<sub>3</sub> 7.62 7.48 -1.84 7.55 -0.92 6-31G(d,p) $H_1$ 7.31 1.11 7.32 1.24 6.99 -3.32 0.58 0.93 -2.33  $H_2$ 8.64 8.67 8.39  $H_3$ 7.67 0.66 7.72 1.31 7.33 -3.81 7.29 6.90 6-31+G(d) $H_1$ 0.83 7.28 0.69 -4.56 0.47 8.62  $H_2$ 8.63 0.35 8.31 -3.26 7.72 7.74 -4.33  $H_3$ 1.31 1.57 7.29 7.52 4.01 6-31+G(d,p) $H_1$ 7.47 3.32 7.15 -1.11 8.78 2.21 8.75 1.86 8.50 -1.05  $H_2$  $H_3$ 7.92 3.94 7.91 3.81 7.53 -1.18 Pyrene 7.99 7.91 7.90 7.60 6-31G(d) -1.00 -1.13 -4.88  $H_1$ 8.16 8.02 -1.72 8.08 -0.98 7.70 -5.64  $H_2$ 8.01 -0.50 8.04 -4.47  $H_3$ 8.05 -0.12 7.69 6-31G(d,p)  $H_1$ 7.96 -0.38 8.08 1.13 7.81 -2.25 8.09 -0.86 8.26 1.23 7.92 -2.94  $H_2$ -0.05 8.10 -1.99  $H_3$ 8.05 0.62 7.89 8.06 0.88 7.96 -0.38 6-31+G(d) $H_1$ 7.79 -2.50  $H_2$ 8.12 -0.49 8.79 7.72 7.84 -3.92  $H_3$ 8.06 0.12 8.13 0.99 7.79 -3.23 6-31+G(d,p)8.19 2.50 1.00  $H_1$ 8.07 1.00 8.07 8.30 8.21 8.09 -0.86  $H_2$ 1.72 8.83  $H_3$ 8.19 1.74 8.29 2.98 8.08 0.37

Table S2: DFT method exploration: CPCM <sup>1</sup>H NMR shifts for benzene, pyridine and pyrene relative to TMS in CHCl<sub>3</sub>.

Basis	Proto	n EXP	B3LYP-D	Error (%)	WB97XD	Error (%)	B97D	Error (%)
	Benzene							
6-31G(d)	Н	7.34	7.14	-2.72	7.23	-1.50	6.85	-6.68
6-31+G(d)	Н		7.33	-0.14	7.41	0.95	7.03	-4.22
6-31+G(d,p)	Н		7.51	2.32	7.56	3.00	7.27	-0.95
					Pyridine			
6-31G(d)	$\mathrm{H}_{1}$	7.23	7.04	-2.63	7.08	-2.07	6.71	-7.19
	$H_2$	8.59	8.45	-1.63	8.51	-0.93	8.16	-5.01
	$H_3$	7.62	7.44	-2.36	7.52	-1.31	7.07	-7.22
6-31+G(d)	$\mathrm{H}_{\mathrm{1}}$		7.24	0.14	7.25	0.28	6.88	-4.84
	$H_2$		8.61	0.23	8.61	0.23	8.31	-3.26
	$H_3$		7.67	0.66	7.71	1.18	7.27	-4.59
6-31+G(d,p)	$\mathrm{H}_{1}$		7.47	3.32	7.44	2.90	7.13	-1.38
	$H_2$		8.75	1.86	8.74	1.75	8.50	-1.05
	$H_3$		7.87	3.28	7.88	3.41	7.51	-1.44
					Pyrene			
6-31G(d)	$\mathrm{H}_{1}$	7.99	7.75	-3.00	7.87	-1.50	7.58	-5.08
	$H_2$	8.16	7.86	-3.68	8.05	-1.35	7.68	-5.93
	$H_3$	8.05	7.86	-2.36	8.01	-0.50	7.67	-4.80
6-31+G(d)	$\mathrm{H}_{1}$		8.01	0.25	7.90	-1.13	7.78	-2.76
	$H_2$		8.06	-1.23	8.26	1.23	7.77	-4.79
	$H_3$		8.01	-0.50	8.09	0.50	7.82	-2.93
6-31+G(d,p)	$\mathrm{H}_{1}$		8.17	2.25	8.05	0.75	8.04	-0.69
	$\mathrm{H}_{2}$		8.27	1.35	8.30	1.72	8.07	-1.16
	$\mathrm{H}_3$		8.17	1.49	8.26	2.61	8.05	-0.03

Table S3: DFT method exploration: PCM <sup>1</sup>H NMR shifts for benzene, pyridine and pyrene relative to TMS in CHCl<sub>3</sub>.

Table S4: Interaction energies, enthalpies, and Gibbs free energies (kcal.mol<sup>-1</sup>) defined in eq. 6 for the dimers of 2,2'-bipyridine with alternative N atom configuration presented in Figure S2 calculated with the WB97Xd(B3LYP-DCP)/6-31G(d,p) methods and PCM solvation model.

Structure	ΔΕ	$\Delta H$	$\Delta G^{298}$	$d_{\mathrm{BIP}}$	$d_{PYR}$
4	-9.2(-8.3)	-8.7(-8.5)	1.7(4.7)	3.39(3.39)	-
5	-49.0(-41.6)	-51.5(-44.1)	2.4(10.0)	3.33(3.23)	-
6	-64.8(-55.2)	-68.2(-58.8)	3.8(15.1)	3.31(3.21)	-
8	-43.1(-37.5)	-43.4(-37.7)	-23.4(-14.1)	3.32(3.28)	3.60(3.49)
9	-67.63(56.9)	-69.24(-59.5)	-25.12(-10.0)	3.29(3.26)	3.62(3.52)
10	-88.2(-72.2)	-90.1(-75.7)	-26.7(-7.9)	3.33(3.35)	3.75(3.48)
11	-107.3(-84.8)	-112.4(-90.1)	-27.1(-0.6)	3.31(3.33)	3.70(3.69)
12	-129.1(-103.4)	-135.5(-109.1)	-26.8(3.1)	3.38(3.40)	3.31(3.35)

Table S5: Interaction energies, enthalpies, and Gibbs free energies (kcal.mol<sup>-1</sup>) defined in eq. 6 for the PBP dimers calculated with the WB97Xd(B3LYP-DCP)/6-31G(d,p) methods in gas phase.

Structure	ΔΕ	ΔΗ	$\Delta G^{298}$
8	-47.7(-39.0)	-47.7(-39.2)	-28.9(-15.2)
9	-76.5(-62.1)	-78.0(-64.0)	-36.8(-18.3)
10	-101.7(-83.7)	-104.7(-87.5)	-41.9(-21.1)
11	-125.9(-103.4)	-130.6(-108.4)	-45.7(-22.0)
12	-153.0(-125.8)	-159.8(-132.7)	-52.7(-25.0)

Figure S1: DFSE of PBP in water-saturated chloroform solvent (left) and water in water (right).





Figure S2: Optimized structures of aggregates, in an alternative N atom configuration, containing 2,2<sup>c</sup>-bipyridine, PBP, and water.







