Supporting Information for

Quantum Chemistry for Solvated Molecules on Graphical Processing Units (GPUs) using Polarizable Continuum Models

Fang Liu,[1] Nathan Luehr,[1] Heather J. Kulik,[1,2] Todd J. Martínez[1,3]

[1]*Department of Chemistry and The PULSE Institute, Stanford University, Stanford, CA 94305*

[2]*Department of Chemical Engineering, Massachussetts Institute of Technology, Boston, MA*

[3]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025*

Contents

Analytical gradient of the total energy of the PCM solvated system

The total energy of the PCM solvated system is

$$E = E_0 + \Delta G_{els} \tag{S1}$$

where E_0 is the energy of the solute, and ΔG_{els} is the electrostatic solvation energy as defined in the main text Eq. (12). The total energy gradient with respect to the nuclear coordinates R_I of atom I is thus given by

$$\nabla_{R_I} E = \nabla_{R_I} E_0 + \nabla_{R_I} (\Delta G_{els}) \ , \tag{S2}$$

where

$$\nabla_{R_I} G_{els} = \sum_{\mu\nu} (\nabla_{R_I} P_{\mu\nu}) F^s_{\mu\nu} + \nabla^*_{R_I} (\Delta G_{els}) \ , \tag{S3}$$

and $F^s_{\mu\nu} = \sum_k q_k L^k_{\mu\nu}$ is the solvation contribution to Fock matrix. In Eq. (S3), the first term includes the derivative with respect to density matrix, whereas the other terms do not. Derivatives that do not include consideration of variation of the density matrix are denoted by the asterisk as $\nabla^*_{R_I}$. In the main text, we focus only on the calculation of $\nabla^*_{R_I}(\Delta G_{els})$, because the first term can be combined with the solute energy gradient term $\nabla_{R_I} E$ and readily obtained from the gas phase gradient subroutine as follows:

$$\begin{aligned} \nabla^*_{R_I}(E) &= \nabla_{R_I} E_0 + \sum_{\mu\nu} (\nabla_{R_I} P_{\mu\nu}) F^s_{\mu\nu} \\ &= \sum_{\mu,\nu} P_{\mu\nu} \nabla_{R_I} (H^0_{\mu\nu}) + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \nabla_{R_I} (\mu\lambda \| \nu\sigma) + \nabla_{R_I} (E_{nn}) \\ &+ \sum_{\mu,\nu} \nabla_{R_I} (P_{\mu\nu}) H^0_{\mu\nu} + \sum_{\mu\nu\lambda\sigma} (\nabla_{R_I} P_{\mu\nu}) P_{\lambda\sigma} (\mu\lambda \| \nu\sigma) + \sum_{\mu\nu} (\nabla_{R_I} P_{\mu\nu}) F^s_{\mu\nu} \end{aligned} \tag{S4}$$

In Eq.(S4), the last three terms can be combined as

$$\begin{aligned} & \sum\nolimits_{\mu\nu} (\nabla_{R_I} P_{\mu\nu}) \left\{ H^0_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} (\mu\lambda \| \nu\sigma) + F^s_{\mu\nu} \right\} \\ &= \sum_{\mu\nu} (\nabla_{R_I} P_{\mu\nu}) \left\{ H^0_{\mu\nu} + G^0_{\mu\nu} + F^s_{\mu\nu} \right\} \\ &= \sum_{\mu\nu} (\nabla_{R_I} P_{\mu\nu}) F_{\mu\nu} \end{aligned} \tag{S5}$$

Calculation of $\nabla_{R_I}(P_{\mu\nu})$ in the last three terms can be circumvented by introducing the energy weighted density matrix

$$W_{\mu\nu} = \sum_{i \in \text{occ.}} \varepsilon_i c_{\mu i} c_{\nu i} \tag{S6}$$

where ε_i denotes the orbital energy and $c_{\mu i}$ denotes the orbital coefficient of the i th spin-orbital, respectively. Then Eq. (S5) becomes

$$-\sum_{\mu\nu} W_{\mu\nu} (\nabla_{R_I} S_{\mu\nu}) \tag{S7}$$

Table S1. Geometries used to benchmark solvation energy errors versus different CG converge thresholds. Each system was treated at five geometries. Nuclear repulsion energies are given (in atomic units) to serve as a checksum. Geometry files in PDB format are listed separately in supplement information folder, with the same name as given in this table.

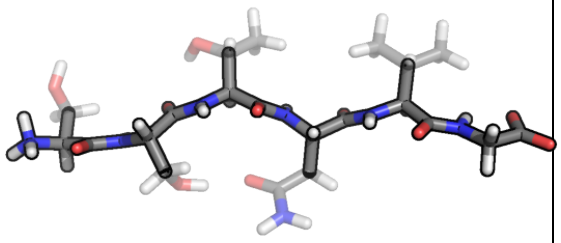
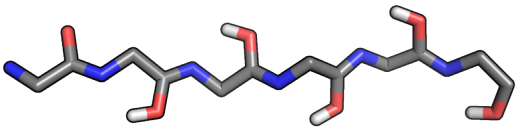
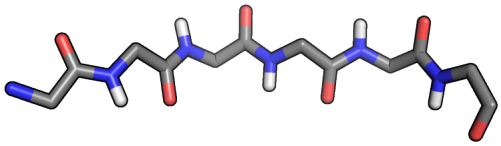
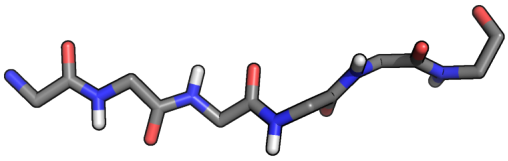
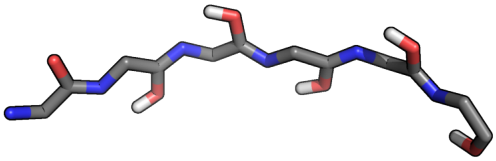
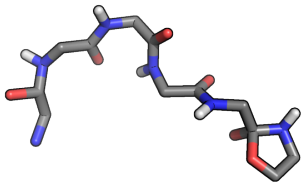
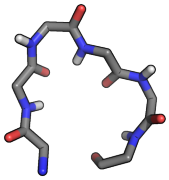
molecule	Atoms	Nuclear repulsion energy (au)	geometry name	prepation method
1LVR	158	12510.78558	frame0	COSMO RHF/MINI geometry optimization
		12834.1897	frame1	Amber forcefield Langevin dynamics at ~500K
		13321.28525	frame2	
		13152.51067	frame3	
		13075.35462	frame4	
2RMW	440	82588.05964	frame0	COSMO RHF/MINI geometry optimization
		82149.19194	frame1	Amber forcefield Langevin dynamics at ~500K
		79187.95346	frame2	
		78989.97469	frame3	
		79789.58322	frame4	
3FTR	76	4620.095683	minima0	COSMO RHF/STO-3G geometry optimization
		4523.668432	minima1	
		4575.992471	minima2	
		4942.330817	minima3	
		4627.493452	minima4	
Water13	39	992.1814668	frame0	Amber TIP3P Water pre-equilibrated box
		914.8135233	frame1	Amber forcefield Langevin dynamics at ~500K
		917.6179329	frame2	
		891.2139688	frame3	
		872.5772654	frame4	
water35	105	5676.191347	frame0	Amber TIP3P Water pre-equilibrated box
		5276.807228	frame1	Amber forcefield Langevin dynamics at ~500K
		5144.813421	frame2	
		5005.221816	frame3	
		4763.877215	frame4	
water121	363	47422.28785	frame0	Amber TIP3P Water pre-equilibrated box
		46149.85852	frame1	Amber forcefield Langevin dynamics at ~500K
		44546.25423	frame2	
		43389.74444	frame3	
		43570.17473	frame4	

Table S2. Characteristics of protein data set to benchmark PCM performance. The experimental method is solution NMR in aqueous solution unless otherwise noted.

					secondary structure		
PDB ID	Type	Res.	Atoms	Charge	%helical	%strand	%other
1LVR	signaling protein	9	158	1	0	0	100
1O53	Transferase	15	249	1	53	0	47
1ODP	lipid transport	20	330	0	75	0	25
1PJD[a]	membrane protein	18	250	0	50	0	50
1T2Y	metal binding protein	25	271	0	0	0	100
1Y03	antifreeze protein	35	437	2	74	0	26
1Y49[b]	neuropeptide	9	122	0	0	0	100
1YT6	de novo protein	10	134	-1	0	0	100
2CEH	coagulation protein	19	288	1	0	0	100
2FBU	antimicrobial protein	12	216	2	42	0	58
2FXZ	hydrolase	13	209	1	46	0	54
2I9M	de novo protein	17	246	1	71	0	29
2JOF	de novo protein	20	284	0	60	0	40
2JTA[c]	signaling protein	10	157	-2	0	0	100
2JXF	viral protein	30	503	1	83	0	17
2K59	transport protein	28	461	1	68	0	32
2KJM	RNA binding protein	30	516	1	57	0	43
2NX6	structural protein	27	349	0	26	0	74
2RLJ	viral protein	16	224	0	19	0	81
2RMW	protein fibril	26	440	-1	12	0	88

[a]Solid state NMR; [b]Solution NMR carried out in DMSO. [c]Solution NMR carried out in methanol.

Table S3. Characteristics of different types of 3FTR local minima

Name	structure[a]	features		
		end-to-end distance(Å)	protonation score	twisted backbone
min1n		>17	>2	no
min1u		>17	<1	no
min2n		14-17	>1.5	no
min2t		>14	>2	yes
min2u		14-17	<1.5	no
min3		8-14	>1.5	yes
min 4		<6	>1.5	no

[a] For clarity, side chain is only displayed for min1n, as transparent sticks.

Table S4. All stationary point structures found for 3FTR. Structures printed in **bold** are featured[a] minima further compared in Table S5 and in the main text.

method	basis	solvent	starting structure ID[b]	energy (au)	end-to-end distance (Å)	protonation score	minima category
RHF-D	STO-3G	COSMO	0	-1996.023821	11.68	1.62	3
			floppy[c] 0	**-1996.016285**	**14.28**	**1.58**	**2n**
			fixed[d] 0	**-1995.894082**	**17.08**	**2.09**	**1n**
			1	**-1996.054302**	**17.46**	**0.56**	**1u**
			2	**-1995.989764**	**14.46**	**2.28**	**2t**
			3	**-1996.177951**	**9.23**	**2.26**	**3**
			4	-1996.025622	10.50	1.64	3
			5	**-1996.065386**	**14.79**	**0.56**	**2u**
			6	**-1996.141393**	**4.67**	**2.95**	**4**
RHF-D	STO-3G	gas	0	-1996.028934	15.01	0.56	2u
			floppy 0	-1996.024596	15.38	0.56	2u
			fixed 0	**-1995.737411**	**17.08**	**2.09**	**1n**
			1	**-1996.021762**	**17.34**	**0.56**	**1u**
			2	-1996.135005	8.81	2.34	3
			3	**-1996.136989**	**9.19**	**2.28**	**3**
			4	-1996.106863	5.82	1.66	4
			5	-1996.02758	14.88	0.56	2u
			6	-1996.094373	5.07	2.50	4
			refine min2t	**-1995.834803**	**14.46**	**2.28**	**2t**
			floppy refine min2n	**-1995.904184**	**14.28**	**1.58**	**2n**
			refine min2u	**-1996.027583**	**14.87**	**0.56**	**2u**
			refine min4	**-1996.106458**	**4.82**	**2.96**	**4**
ωPBEh-D	STO-3G	COSMO	0	-2005.140981	14.86	0.87	2u

			floppy 0	-2005.11663	14.99	1.27	2u
			fixed 0	**-2004.858548**	**17.08**	**2.09**	**1n**
			1	**-2005.119089**	**17.51**	**0.86**	**1u**
			2	-2005.118946	17.51	0.86	1u
			3	**-2005.148226**	**9.44**	**2.38**	**3**
			4	-2005.140925	14.58	0.87	2u
			5	**-2005.140982**	**14.89**	**0.87**	**2u**
			6	**-2005.158206**	**4.88**	**2.14**	**4**
			0	-2005.092825	14.63	0.77	2u
			floppy 0	-2005.07033	15.25	0.76	2u
			fixed 0	**-2004.715571**	**17.08**	**2.09**	**1n**
			1	-2005.060985	17.58	0.77	1u
			2	SCF not converged			
ωPBEh-D	STO-3G	gas	3	-2005.10509	9.13	2.37	3
			4	-2005.092775	14.69	0.77	2u
			5	**-2005.092924**	**14.71**	**0.77**	**2u**
			6	**-2005.122296**	**4.83**	**1.90**	**4**
			refine min1u	**-2005.063855**	**17.43**	**0.77**	**1u**
			refine min3	**-2005.105096**	**9.15**	**2.37**	**3**
			0	**-1995.452803**	**14.51**	**1.68**	**2n**
			floppy 0	-1995.451064	14.84	1.65	2n
			fixed 0	**-1995.324353**	**17.08**	**2.09**	**1n**
			1	**-1995.492696**	**17.46**	**0.53**	**1u**
RHF	STO-3G	COSMO	**2**	**-1995.425339**	**16.04**	**2.49**	**2t**
			3	**-1995.588206**	**10.93**	**2.30**	**3**
			4	-1995.452802	14.51	1.68	2n
			5	**-1995.495978**	**15.69**	**0.54**	**2u**
			6	**-1995.545522**	**5.29**	**2.62**	**4**

RHF	STO-3G	gas	0	-1995.458018	16.24	0.54	2u
			floppy 0	-1995.456587	16.40	0.54	2u
			fixed 0	**-1995.167682**	**17.08**	**2.09**	**1n**
			1	-1995.453415	17.51	0.53	1u
			2	-1995.546805	10.58	2.32	3
			3	-1995.549134	10.24	2.31	3
			4	-1995.508679	15.46	0.88	2u
			5	-1995.457352	16.15	0.53	2u
			6	**-1995.508223**	**5.24**	**2.60**	**4**
			floppy refine min2t	**-1995.249979**	**16.04**	**2.49**	**2t**
			flppy refine min2n	**-1995.335582**	**14.51**	**1.68**	**2n**
			refine min1u	**-1995.453415**	**17.51**	**0.53**	**1u**
			refine min2u	**-1995.457376**	**16.19**	**0.53**	**2u**
			refine min3	**-1995.546791**	**10.58**	**2.32**	**3**
RHF	3-21G	COSMO	**0**	**-2010.367513**	**15.98**	**2.15**	**2n**
			fixed 0	**-2010.349997**	**17.08**	**2.09**	**1n**
			1	**-2010.228138**	**17.28**	**0.52**	**1u**
			2	-2010.362634	16.94	2.53	2n
			3	**-2010.302464**	**11.20**	**2.62**	**3**
			4	-2010.317177	16.12	1.64	2n
			5	-2010.287374	15.52	1.31	2u
			6	**-2010.401482**	**5.01**	**2.59**	**4**
			refine min2u	**-2010.235612**	**16.09**	**0.52**	**2u**
RHF	3-21G	gas	0	-2010.19325	13.67	1.95	3
			fixed 0	**-2010.169513**	**17.08**	**2.09**	**1n**
			1	-2010.161459	16.40	0.52	2u
			2	-2010.33004	5.08	2.90	4

			3	-2010.236178	10.22	2.67	3
			4	-2010.226089	8.12	1.79	3
			5	**-2010.162165**	**15.24**	**0.52**	**2u**
			6	-2010.314254	4.83	2.64	4
			floppy refine min1u	**-2010.147849**	**17.28**	**0.52**	**1u**
			flppy refine min2n	**-2010.171494**	**15.98**	**2.15**	**2n**
			refine min3	**-2010.236177**	**10.22**	**2.66**	**3**
			refine min4	**-2010.317156**	**5.05**	**2.66**	**4**
RHF	6-31G	COSMO	**0**	**-2020.739656**	**15.67**	**2.38**	**2n**
			fixed 0	**-2020.718568**	**17.08**	**2.09**	**1n**
			1	**-2020.574736**	**17.08**	**0.46**	**1u**
			2	-2020.732825	16.18	2.69	2n
			3	**-2020.624036**	**12.21**	**2.85**	**3**
			4	**-2020.67684**	**15.77**	**1.77**	**2n**
			5	-2020.639608	15.67	1.38	2u
			6	**-2020.741282**	**5.16**	**2.59**	**4**
			refine min2u	**-2020.577799**	**15.94**	**0.46**	**2u**
RHF	6-31G	gas	0	-2020.629089	4.74	2.50	4
			fixed 0	**-2020.523222**	**17.08**	**2.09**	**1n**
			1	-2020.490039	16.34	0.47	2u
			2	-2020.627914	4.52	2.86	4
			3	-2020.531861	11.19	2.70	3
			4	-2020.533863	11.90	1.67	3
			5	**-2020.490379**	**16.16**	**0.47**	**2u**
			6	-2020.629087	4.74	2.50	4
			floppy refine min1u	**-2020.476689**	**17.08**	**0.46**	**1u**
			floppy refine	**-2020.510262**	**15.67**	**2.38**	**2n**

			min2n				
			refine min3	**-2020.53022**	**11.15**	**2.86**	**3**
			refine min4	**-2020.629092**	**4.74**	**2.51**	**4**
ωPBEh	STO-3G	COSMO	0	-2005.140975	14.88	0.87	2u
			fixed 0	**-2004.858548**	**17.08**	**2.09**	**1n**
			floppy 0	-2005.11823	14.76	1.28	2u
			1	**-2005.119093**	**17.51**	**0.86**	**1u**
			2	-2005.119104	17.25	0.87	1u
			3	**-2005.148228**	**9.44**	**2.38**	**3**
			4	-2005.140929	14.57	0.87	2u
			5	**-2005.140982**	**14.89**	**0.87**	**2u**
			6	**-2005.158045**	**4.91**	**2.14**	**4**
ωPBEh	STO-3G	gas	0	-2005.092821	14.60	0.77	2u
			fixed 0	**-2004.715571**	**17.08**	**2.09**	**1n**
			floppy 0	-2005.063364	16.41	0.76	2u
			1	-2005.060985	17.58	0.77	1u
			2	SCF not converged			
			3	-2005.105096	9.14	2.37	3
			4	-2005.092775	14.69	0.77	2u
			5	**-2005.092923**	**14.71**	**0.77**	**2u**
			6	**-2005.122296**	**4.83**	**1.90**	**4**
			refine min1u	**-2005.063947**	**17.30**	**0.77**	**1u**
			refine min3	**-2005.105094**	**9.15**	**2.37**	**3**
ωPBEh	3-21G	COSMO	**0**	**-2020.307318**	**16.66**	**1.89**	**2n**
			fixed 0	**-2020.262178**	**17.08**	**2.09**	**1n**
			1	**-2020.300192**	**16.29**	**2.07**	**2t**
			2	-2020.293059	16.36	2.38	2t
			3	**-2020.246215**	**11.08**	**2.61**	**3**
			4	-2020.306403	14.96	1.98	2n

			5	-2020.306418	15.45	1.98	2n
			6	**-2020.351833**	**5.06**	**2.54**	**4**
ωPBEh	3-21G	gas	0	-2020.172435	11.69	1.66	3
			fixed 0	**-2020.102764**	**17.08**	**2.09**	**1n**
			1	-2020.132055	14.89	0.71	2u
			2	-2020.285947	4.70	2.52	4
			3	-2020.184184	10.61	2.57	3
			4	-2020.172066	10.10	1.61	3
			5	-2020.18756	11.03	1.73	3
			6	-2020.279871	4.83	2.52	4
			flppy refine min2n	**-2020.146783**	**16.66**	**1.89**	**2n**
			refine min2t	**-2020.150712**	**15.43**	**1.86**	**2t**
			refine min3	**-2020.184181**	**10.61**	**2.57**	**3**
			refine min4	**-2020.279864**	**4.83**	**2.51**	**4**
ωPBEh	6-31G	COSMO	0	-2030.851837	16.78	2.14	2n
			fixed 0	**-2030.826738**	**17.08**	**2.09**	**1n**
			1	**-2030.697814**	**16.71**	**0.62**	**2u**
			2	**-2030.842599**	**17.15**	**2.46**	**2t**
			3	**-2030.747500**	**12.10**	**2.64**	**3**
			4	-2030.852535	15.67	2.18	2n
			5	**-2030.852783**	**16.00**	**2.19**	**2n**
			6	**-2030.871458**	**4.89**	**2.55**	**4**
ωPBEh	6-31G	gas	0	-2030.681419	14.43	1.83	2n
			fixed 0	**-2030.654648**	**17.08**	**2.09**	**1n**
			1	-2030.623102	16.47	0.57	2u
			2	-2030.680996	14.01	1.86	2n
			3	-2030.672887	11.40	2.59	3
			4	-2030.686729	8.73	1.73	3

5	-2030.732875	10.08	2.12	3
6	-2030.780522	4.69	2.42	4
floppy refine min2t	**-2030.624340**	**17.15**	**2.46**	**2t**
flppy refine min2n	**-2030.657123**	**16.00**	**2.19**	**2n**
refine min2u	**-2030.61617**	**16.66**	**0.58**	**2u**
refine min3	**-2030.672877**	**11.40**	**2.59**	**3**
refine min4	**-2030.779536**	**4.92**	**2.52**	**4**

[a] Featured structure: At the same level of theory, many minima may fall into the same category although they are minimized from different starting structures. At such circumstances, the only one chosen for discussion in the main text is referred here as the "featured" structure. The criteria for choosing the featured structure are as follows: in COSMO, the minimum should be the lowest in energy among other minima belonging to the same category. In gas phase, the minima has the smallest deviation in end-to-end distance and protonation score from the featured COSMO minimum belonging to the same category.

[b] Some of the stationary point structures are obtained from geometry minimization from the eight starting structures. Others are obtained in the refinement process: to better compare the gas phase and COSMO structures, a gas phase minimization is carried out from a COSMO minimum and vice versa. For structures obtained in this way, its starting structure is labeled as "refine min".

[c] In some cases a strict stationary point cannot be obtained. Then a looser convergence threshold is used for the minimization. The obtained structure has only a few imaginary frequencies with small magnitude. Such structures are labeled with "floppy". Detailed information about the imaginary frequencies of "floppy" minima is given in Table S-7.

[d] The featured minima, min1n (starting structure 0), is not a stationary point for many methods, but it is an important structure as it is closest to the crystal structure. The single point energy of min1n is evaluated with different methods without minimization, labeled as "fixed 0".

Table S5. Relative energy of featured 3FTR local minima ordered by energy

method	basis	solvent	minima category (energy[a]: kcal/mol)					
RHF-D	STO-3G	COSMO	3	4	2u	1u	2n	2t
			0.00	22.94	70.63	77.59	101.45	118.09
		gas	3	4	2u	1u	2n	2t
			0.00	19.16	68.65	72.31	146.09	189.62
RHF	STO-3G	COSMO	3	4	2u	1u	2n	2t
			0.00	26.78	57.87	59.93	84.97	102.20
		gas	3	4	2u	1u	2n	2t
			0.00	24.20	56.11	58.59	132.53	186.25
	3-21G	COSMO	4	2n	1n	3	2u	1u
			0.00	21.32	32.31	62.13	104.08	108.77
		gas	4	3	2n	1n	2u	1u
			0.00	50.81	91.40	92.65	97.26	106.24
	6-31G	COSMO	4	2n	1n	3	2u	1u
			0.00	1.02	14.25	73.57	102.60	104.51
		gas	4	3	1n	2n	2u	1u
			0.00	62.04	66.43	74.57	87.00	95.63
ωPBEh-D	STO-3G	COSMO	4	3	2u	1u		
			0.00	6.26	10.81	24.55		
		gas	4	3	2u	1u		
			0.00	10.79	18.43	36.67		
ωPBEh	STO-3G	COSMO	4	3	2u	1u		
			0.00	6.16	10.71	24.44		
		gas	4	3	2u	1u		
			0.00	10.79	18.43	36.61		
	3-21G	COSMO	4	2n	2t	1n	3	
			0.00	27.93	32.40	56.26	66.28	
		gas	4	3	2t	2n	1n	
			0.00	60.04	81.04	83.51	111.13	
	6-31G	COSMO	4	2n	2t	1n	3	2u
			0.00	11.72	18.11	28.06	77.78	108.96
		gas	4	3	2n	1n	2t	2u
			0.00	66.93	76.81	78.37	97.39	102.51

[a] The energy listed for each geometry is the energy relative to the most stable geometry obtained at the specific level of theory. Each row is sorted by the energies in ascending order.

Table S6. Mulliken charges for different minima calculated with gas phase RHF/6-31G.

Residue ID	Mulliken charges					
	min1n	min1u	min2n	min2u	min3	min4
1	1.03	0.46	1.05	0.43	0.63	0.99
2	0.05	0.04	0.06	0.05	0.08	-0.01
3	0.02	-0.02	-0.01	-0.01	-0.04	0.00
4	-0.03	-0.05	0.00	-0.02	-0.02	0.03
5	-0.07	0.04	-0.07	0.00	-0.09	-0.06
6	-1.01	-0.46	-1.04	-0.45	-0.55	-0.95

Table S7. Frequency analysis for 3FTR floppy minima

method	basis	solvent	minima	# imaginary freq.	max imaginary freq.	max imaginary force const	
						(mdyn/Å)	(Hartree/bohr^2)
RHF-D	sto-3G	cosmo	2t	12	-84.92	0.0104	6.66E-05
RHF-D	sto-3G	gas	2n	11	-153.65	0.03	1.92E-04
RHF-D	sto-3G	gas	2t	14	-188.45	0.049	3.14E-04
RHF	sto-3G	gas	2n	2	-19.36	0.0016	1.02E-05
RHF	sto-3G	gas	2t	4	-124.65	0.0004	2.56E-06
RHF	3-21G	gas	2n	1	-18.18	0.0008	5.12E-06
RHF	3-21G	gas	1u	1	-18.05	0.0009	5.76E-06
RHF	6-31G	gas	2n	1	-133.92	0.0118	7.55E-05
RHF	6-31G	gas	1u	2	-34.04	0.0028	1.79E-05
ωPBEh	3-21G	gas	2n	1	-16.09	0.0007	4.48E-06
ωPBEh	6-31G	gas	2n	3	-30.37	0.0027	1.73E-05

Table S8. Timings for COSMO (RHF/6-31G, 590 points/atom) of TeraChem and Q-Chem (no FMM[a]). TeraChem on GTX TITAN GPUs and Qchem on Intel Xeon X5690@3.47 GHz CPUs.[b]

	GPU /CPU cores	Total runtime			PCM gradient			Gas phase gradient			PCM SCF			Gas phase SCF		
		QC	TC	Speedup	QC	TC	Speedup	QC	TC	Speedup	QC	TC	Speedup	QC	TC	Speedup
1y49 (122,5922)	1	1702	115	15	88 (5%)	2 (2%)	44	364 (21%)	22 (19%)	17	307 (18%)	25 (22%)	12	944 (55%)	66 (57%)	14
	4	815	41	20	110 (13%)	2 (5%)	55	96 (12%)	6 (15%)	16	314 (39%)	10 (24%)	31	296 (36%)	23 (56%)	13
	8	947	31	31	209 (22%)	1 (3%)	209	55 (6%)	4 (13%)	14	439 (46%)	8 (27%)	55	244 (26%)	17 (57%)	14
2kjm (516,26025)	1	34916	1787	20	2324 (7%)	40 (2%)	58	6514 (19%)	417 (23%)	16	6126 (18%)	445 (25%)	14	19952 (57%)	885 (50%)	23
	4	17331	622	28	2731 (16%)	26 (4%)	105	1647 (10%)	116 (19%)	14	6421 (37%)	181 (29%)	35	6533 (38%)	299 (48%)	22
	8	22051	419	53	5078 (23%)	23 (5%)	221	1002 (5%)	59 (14%)	17	10534 (48%)	141 (34%)	75	5438 (25%)	196 (47%)	28

	GPU/ CPU cores	CG			Build c			Build Fs		
		QC	TC	speed-up	QC	TC	speed-up	QC	TC	speed-up
1y49 (122,5922)	1	10 (1%)	6 (5%)	2	157 (9%)	9 (8%)	17	140 (8%)	10 (9%)	14
	4	10 (1%)	4 (9%)	3	161 (20%)	3 (7%)	54	143 (18%)	3 (7%)	48
	8	10 (1%)	4 (12%)	3	229 (24%)	2 (7%)	115	200 (21%)	2 (7%)	100
2kjm (516,26025)	1	252 (1%)	124 (7%)	2	3145 (9%)	131 (7%)	24	2728 (8%)	176 (10%)	16
	4	253 (1%)	81 (16%)	3	3302 (19%)	39 (6%)	85	2864 (17%)	48 (8%)	60
	8	270 (1%)	81 (19%)	3	5507 (25%)	21 (5%)	262	4755 (22%)	25 (6%)	190

[a] For the no-FMM version of Q-Chem COSMO, where matrix is explicitly stored in memory, the CG method is faster. MPI is used for parallelization, as the CG does not work properly with OpenMP.

[b] TeraChem's CPU calculation is conducted on the same group of Intel Xeon X5690@3.47 GHz CPUs as Q-Chem.

Table S9. Number of SCF iterations for COSMO of TeraChem (TC) and Q-Chem (QC) (corresponding to the timing tests shown in of **Table 1** and **Table 3**).

basis set, grid	molecule (#atoms, #MS points)	Total SCF iterations	
		QC	TC
6-31G, 110 points/atom	1y49 (122, 5922)	11	16
	2kjm (516, 26025)	10	15
6-31++G*, 590 points/atom	1y49 (122, 22430)	10	16
	2kjm (516, 97923)	7	12

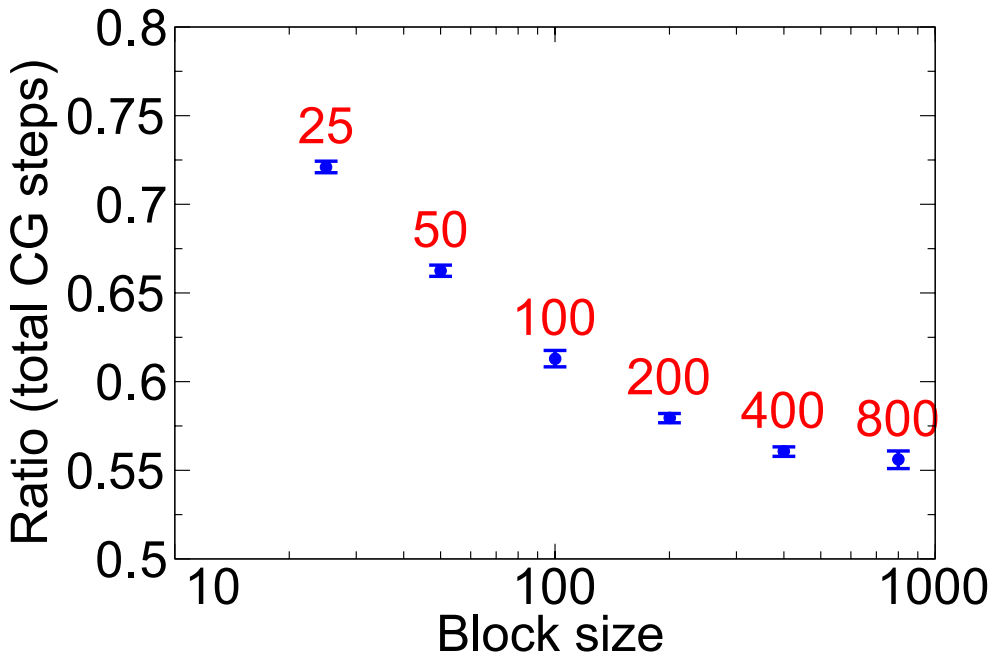


Figure S1. Dependence of CG steps (ratio for the steps taken by specific RBJ preconditioned CG versus Jacobi preconditioned CG) on RBJ preconditioner block size. The COSMO calculation is conducted for a protein (PDB ID: 2KJM, 516 atoms) at RHF/6-31G level. Red numbers labels the block size for each data point.

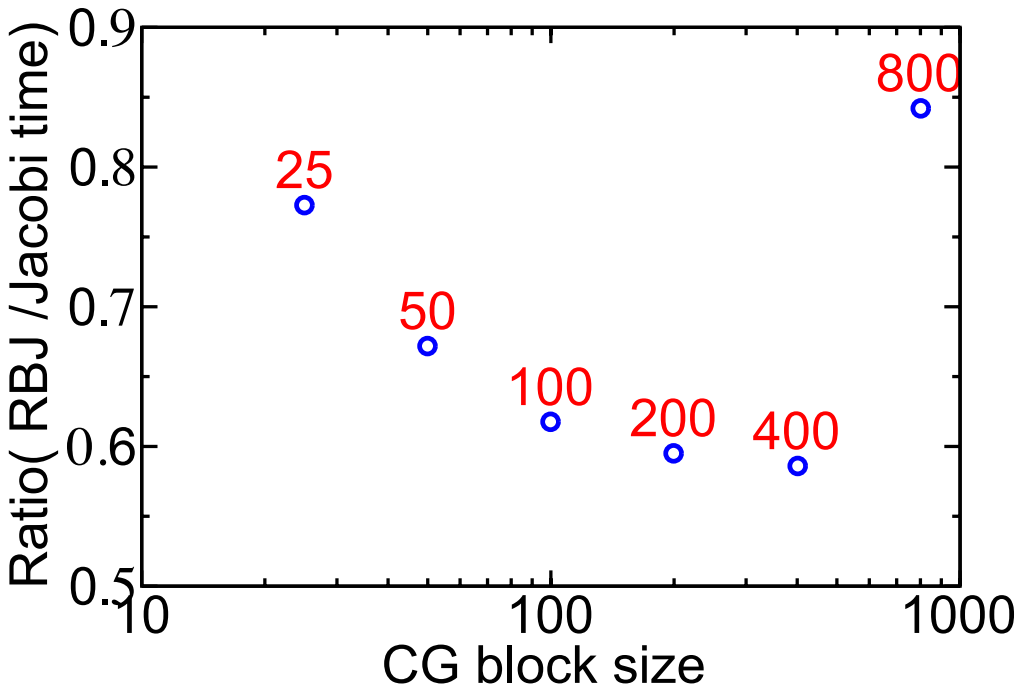


Figure S2. Dependence of CG time for the entire COSMO calculation (including preconditioner formation) on RBJ preconditioner block size. The COSMO calculation is conducted for a protein (PDB ID: 2KJM, 516 atoms) at RHF/6-31G level. Red numbers labels the block size for each data point.

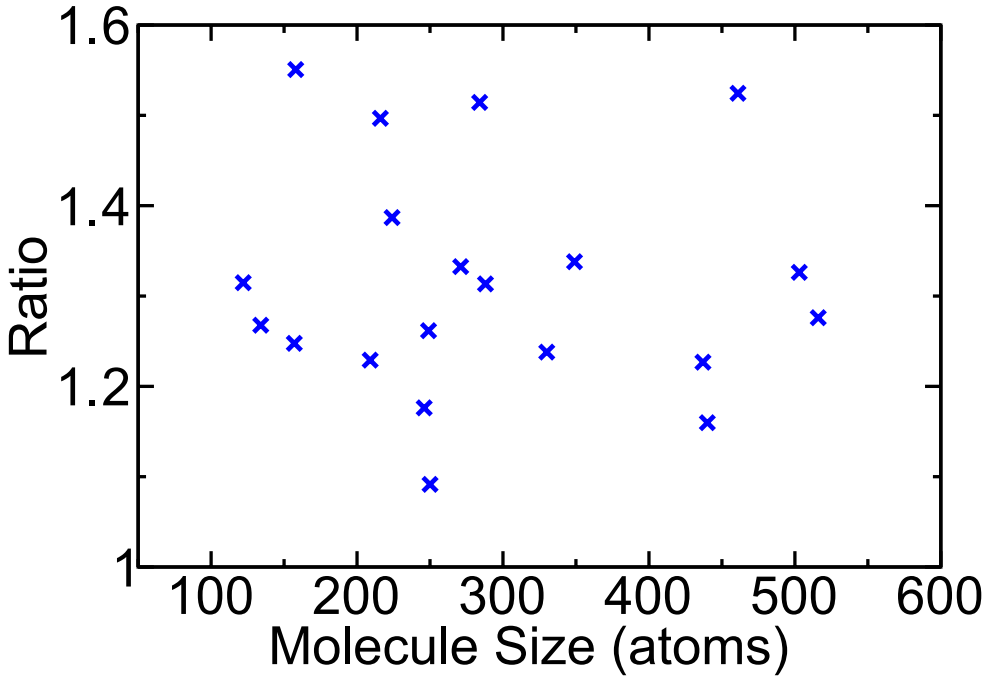


Figure S3. Ratio of time for COSMO versus gas phase gradient calculation for 20 small proteins at RHF/6-31G level. Dynamic precision for 2-electron integral is used; SCF convergence threshold for DIIS error: 1e-6 au; COSMO cavity radii: 1.2* Bondi. ISWIG discretization scheme is used with 110 Lebedev points/atom.

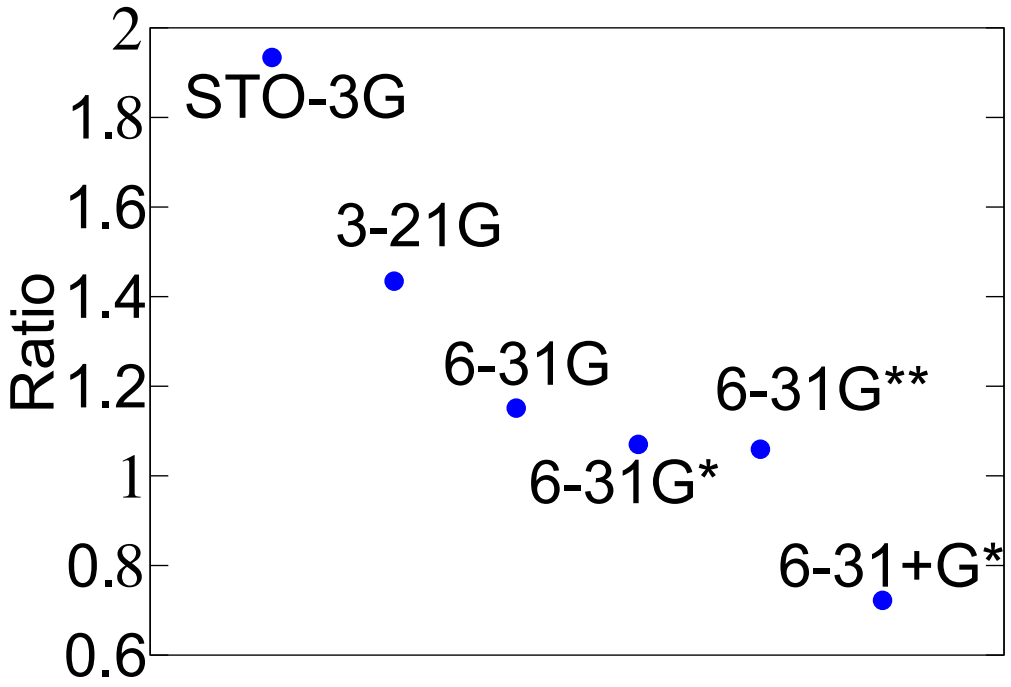


Figure S4. Dependence of COSMO calculation time on basis set for a protein (PDB ID: 1ODP, 330 atoms). Ratio in the graph is the time spent on COSMO calculation versus the time spent on gas phase calculation with the same basis set. Note that the COSMO calculation can be less expensive than the gas phase calculation for large basis sets (ratio < 1, e.g. for 6-31+G*). This occurs because SCF convergence is accelerated in the presence of the reaction field.

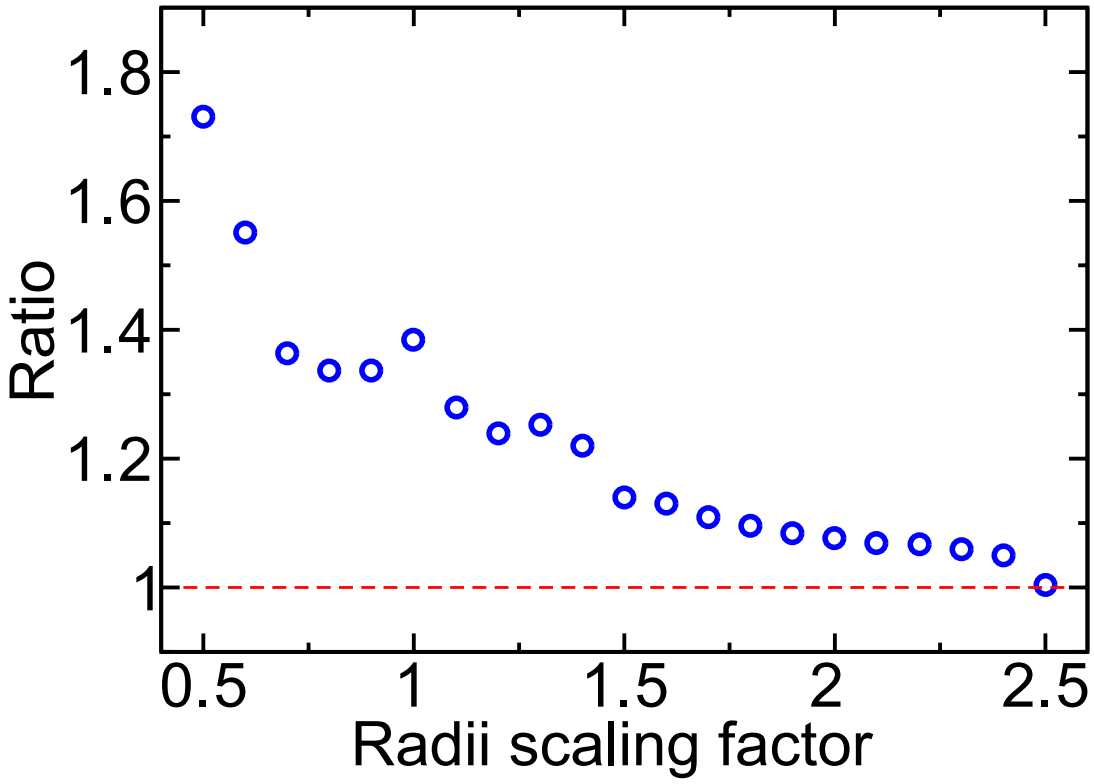


Figure S5. Dependence of COSMO calculation time on cavity radii for a protein (PDB ID: 1ODP, 330 atoms). Ratio in the graph is the time spent on COSMO calculation with specific cavity radii versus the time spent on gas phase calculation at RHF/6-31G level.

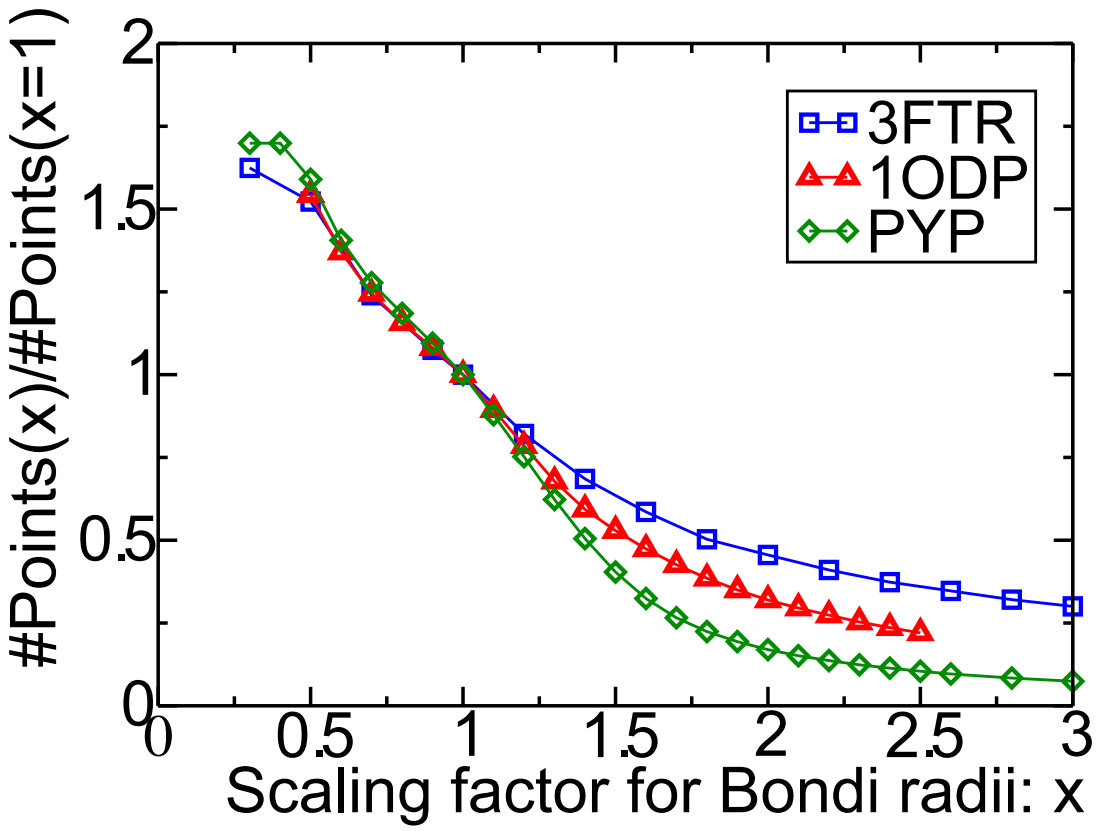


Figure S6. Dependence of number of surface points on cavity radii for three proteins(PDB ID: 3FTR, 76 atoms; PDB ID: 1ODP, 330 atoms; Photoactive Yellow Protein (PYP), PDB ID: 1XFN, 1573 atoms)

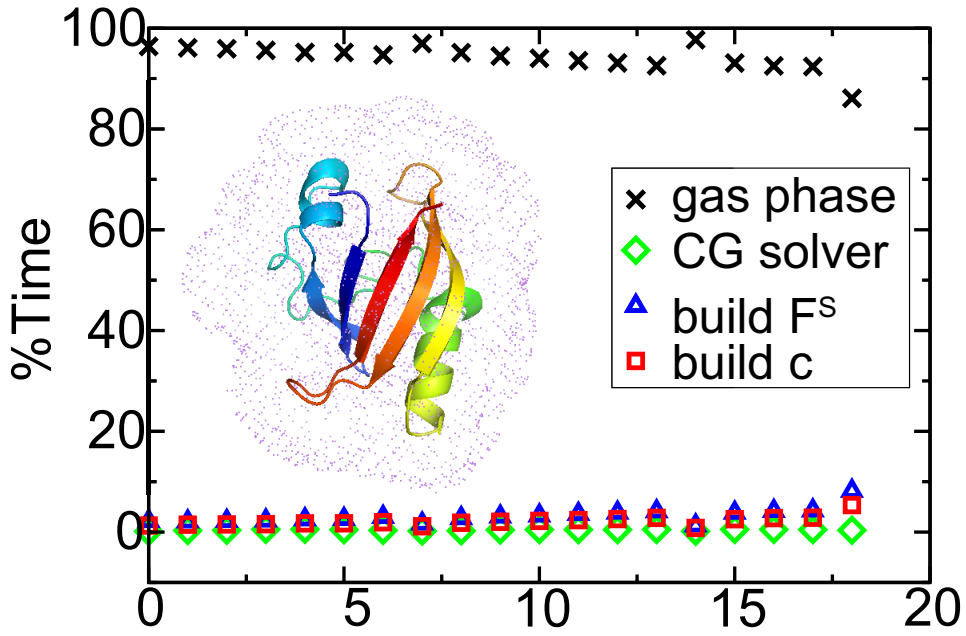


Figure S7. Timing of COSMO RHF/6-31G* for PYP with radii 2.0*Bondi.