

Supporting Information for

Molecular Recognition at Mineral Interfaces: Implications for the Beneficiation of Rare Earth Ores

Jonathan E. Sutton,^{1#} Santanu Roy,¹ Azhad U. Chowdhury,¹ Lili Wu,² Anna K. Wanhala,¹ Nuwan De Silva,^{3##} Santa Jansone-Popova,¹ Benjamin P. Hay,⁴ Michael C. Cheshire,¹ Theresa L. Windus,² Andrew G. Stack,¹ Alexandra Navrotsky,^{2###} Bruce A. Moyer,¹ Benjamin Doughty,¹ and Vyacheslav S. Bryantsev^{1*}

1. Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6119, USA
2. Peter A. Rock Thermochemistry Laboratory and NEAT ORU, University of California Davis, 1 Shields Avenue, Davis, CA 95616, USA.
3. Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50011-3111, USA.
4. Supramolecular Design Institute, 127 Chestnut Hill Rd., Oak Ridge, Tennessee 37830-7185, USA

Current Address: Eastman Chemical Company, Kingsport, TN 37660, USA

Current Address: Department of Physical & Biological Sciences, Western New England University, Springfield, MA 01119, USA

School of Molecular Sciences, The College of Liberal Arts and Sciences, and the School for Engineering of Matter, Transport and Energy, Arizona State University, Arizona, USA

Email: bryantsev@ornl.gov, **phone:** (865) 576-4272, **fax:** (865) 576-7956

Supporting Information

The electronic supporting information includes additional computational details, experimental data, and discussion. This material is available free of charge at <http://pubs.acs.org>.

1. Binding modes of DMP on Ce-bastnäsite: The effect of Na⁺ counterion and supercell size

An important consideration for modeling the real aqueous environment is that ligands are typically added in a salt form (and often in the presence of a buffer), rather than in their acid form. To investigate this, we performed additional calculations at nine monolayers water (Ce-bastnäsite only) where we substituted the acidic proton for a Na⁺ counterion, placed in the outer bulk-like water layer. Binding energies are given in Table S1. We find that replacing a surface proton with the Na⁺ cation somewhat weakens the binding of DMP to the surface, but the two binding modes remain energetically very similar.

Table S1. Binding energies of DMP on Ce-bastnäsite for the acid and Na salt forms of the ligand at 9 ML coverage of water.

Counter Ion	Binding Energy (eV)	
	Monodentate	Bidentate
H ⁺	-1.66	-1.77
Na ⁺	-1.40	-1.39

We have also investigated some larger supercells of Ce-bastnäsite in order to search for adsorbate configurations that cannot be obtained simply by repeating the unit cell. For example, we have located stable configurations of two monodentate DMP ligands in 1x2 and 2x1 supercells. Images of these configurations in the absence of water are shown in Figure S1. In the 1x2 supercell (Figure S1a), the DMP ligands bind to alternating Ce atoms in a staggered or zig-zag fashion. In the 2x1 supercell (Figure S1b and S1c, which are the same structure only shifted by half the supercell), the two ligands take on more of a mirror image configuration. Of note is that the basic binding geometry of the ligands is essentially unchanged; only the position and identity of the binding O are altered (in the latter case leading to essentially a difference in chirality).

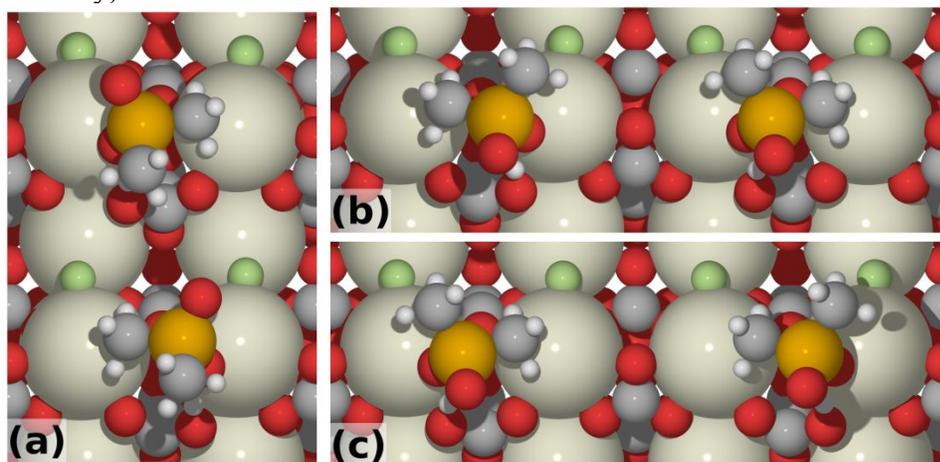


Figure S1. Images of non-trivial DMP configurations at 1/2 ML DMP coverage in 1x2 (a) and 2x1 (b, c) supercells. The images in the 2x1 supercell (b, c) are the same structure; (c) is merely translated in the x direction by half of the unit cell length to aid visualization.

We are also interested in the stability of these alternate configurations compared to the geometries of the 1x1 structures, both on the bare surfaces and in the presence of water. The ligand geometries in the presence of water are very similar to the structures in the absence of

water, with the exception of the usual rotation of the methyl groups about the P atom. The binding energies for the supercell configurations are given in Table S2. In all cases, the binding energies are similar to each other. In the absence of water, the binding energies are nearly identical to the corresponding monodentate structure in the 1x1 cell. In the presence of water, there is a 0.4 eV decrease in the binding energy in the supercell configurations compared to the 1x1 cell. We believe that these structures are simply local minima in the water network and that more stable water configurations are possible. The initial structures for the supercell calculations were obtained by extending the 1x1 cell in the appropriate direction and then rearranging the second ligand and proton and a minimal number of water molecules. Subsequent to the initial geometry optimization, we made no attempt (owing to the high computational expense of the larger supercells) to further optimize the water structure via short MD runs, as was done with the 1x1 cell calculations. In the course of our MD runs for the 1x1 cells, we found that the binding energy could decrease by this much due to favorable optimizations in the water network while leaving the ligand structure essentially unchanged. Were we to carry out MD runs, we believe the binding energies would be in better agreement with the results from the 1x1 cell.

Table S2. Average binding energies (eV) of DMP structures depicted in Figure S1 (0 ML) and the corresponding structures in the presence of a large amount of water (9 ML).

Supercell	Binding Energy (eV)	
	0 ML	9 ML
1x1	-1.31	-1.66
1x2	-1.34	-1.24
2x1	-1.27	-1.26

2. The effect of water on internal geometry of DMP

Before we could carry out a search for molecular fragments to link neighboring ligands, we needed to (1) account for the effect that the solvent has on ligand geometry and (2) quantify how much deviation from the baseline (unlinked) ligand geometry is permitted in the candidate ligands. As we have shown in the section 1 of the SI, the structures of individual ligands in larger supercells are very similar to single ligands in a 1x1 cell; therefore, we employed the smaller cell size.

Effect of Water on DMP Geometry

Before discussing our results on the detailed scans of the potential energy surface we first show that the presence of water has very little effect on the important degrees of freedom that we needed to consider in the linker search. We compare bond lengths for both monodentate and bidentate binding in the presence and absence of water in Figure S2. We do not show C-H bond lengths as these are wholly unaffected by water, as expected based on chemical intuition. The major difference in the monodentate binding is that the presence of water increases the Ce-O bond length; there are only minor perturbations to the other non-C-H bond lengths. The picture is roughly the same for bidentate DMP: adding water increases the Ce-O bond lengths slightly, though not as dramatically as for the monodentate complex, whereas the other bond lengths remain nearly identical.

Besides the bond lengths, we also consider the bond angles. These are compared in Figure S3. As before, we neglect any angles involving only C and H as these are internal to the methyl groups and are entirely unaffected by the solvent. For monodentate DMP, we observed two major changes: the P-O-Ce bond angle increases upon adding water, which is related to the lengthened Ce-O bond, and the angles between the other non-H atoms (shown in Figure S3) change by $\sim 5^\circ$. These changes are due to the water-induced rotation of the methyl groups and surface-unbound O about the P atom. The geometry of bidentate DMP is much less affected by the addition of water. The major change is in one of the Ce-O-P bonds, and this is due to the bound O atom being displaced by a water molecule competing with it for a surface Ce atom. Overall, the presence of water does not have any effects on the internal geometry of bound DMP that cannot be captured via perturbations to the structure on the bare surface.

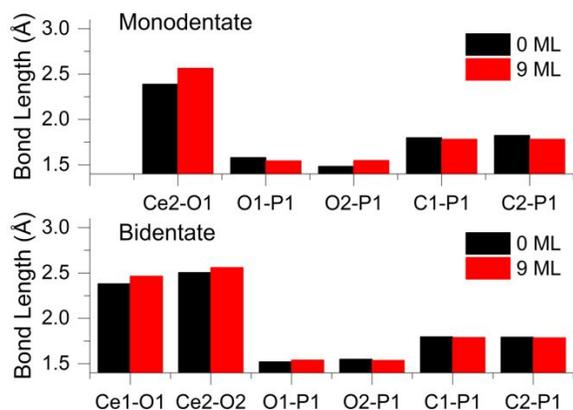


Figure S2. Effect of water coverage on select DMP bond lengths for monodentate and bidentate modes on Ce-bastnäsite. C—H bond lengths are excluded as they do not change appreciably.

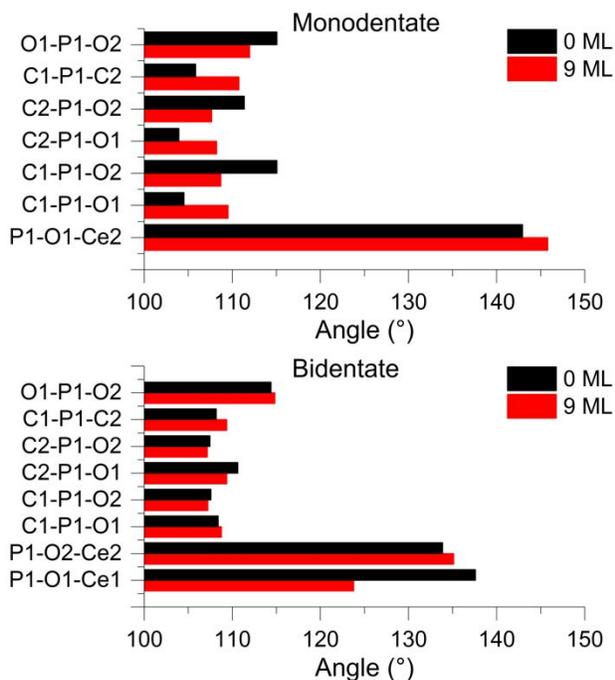


Figure S3. Effect of water coverage on select DMP bond angles for monodentate and bidentate modes on Ce-bastnäsite. Bond angles internal to the methyl groups (i.e., combinations of H-C-H) are excluded as they show no appreciable difference.

3. Molecular Mechanics MM3 Force Field Parameters

Phosphinate parameters were developed by extending phosphine oxide parameters taken from De Silva, N.; Federico, Z.; Hay, B. P.; Gordon, M. S.; and Windus, T. L. "Conformations of Phosphine Oxides" *J. Phys. Chem. A* **2015**, 119, 8765-8773 and McCann, B. W.; De Silva, N.; Windus, T. L.; Gordon, M. S.; Moyer, B. A.; Bryantsev, V. S.; Hay, B.P "Computer-Aided Molecular Design of Bis-phosphine Oxide Lanthanide Extractants" *Inorg. Chem.* **2016**, 55, 5787-5803. These parameters were further refined by fitting to MP2/cc-pVTZ potential energy surfaces for 6 additional bond rotations in selected phosphinate molecules. These potential surfaces and fits are given below, where blue E(ref) is the MP2/cc-pVTZ surface and red E(calc) is the surface calculated with the modified MM3 parameters given above.

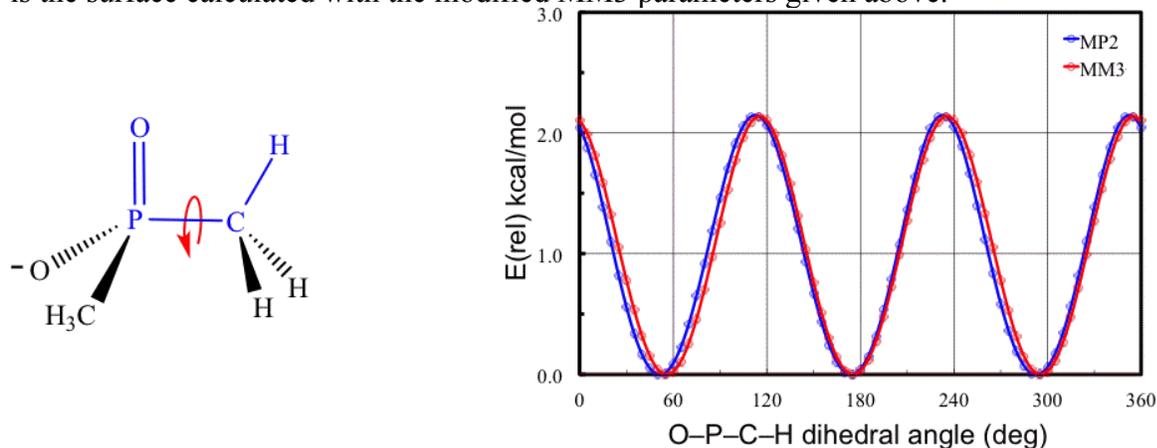


Figure S4. Potential energy surface for rotation about the O-P-C-H dihedral for a methyl substituent.

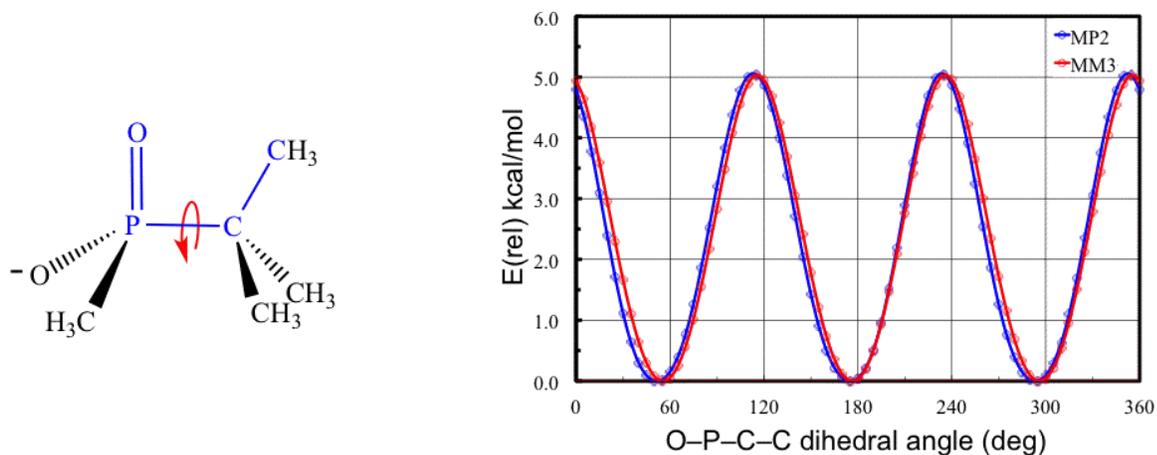


Figure S5. Potential energy surface for rotation about the O-P-C-C dihedral for a *tert*-butyl substituent.

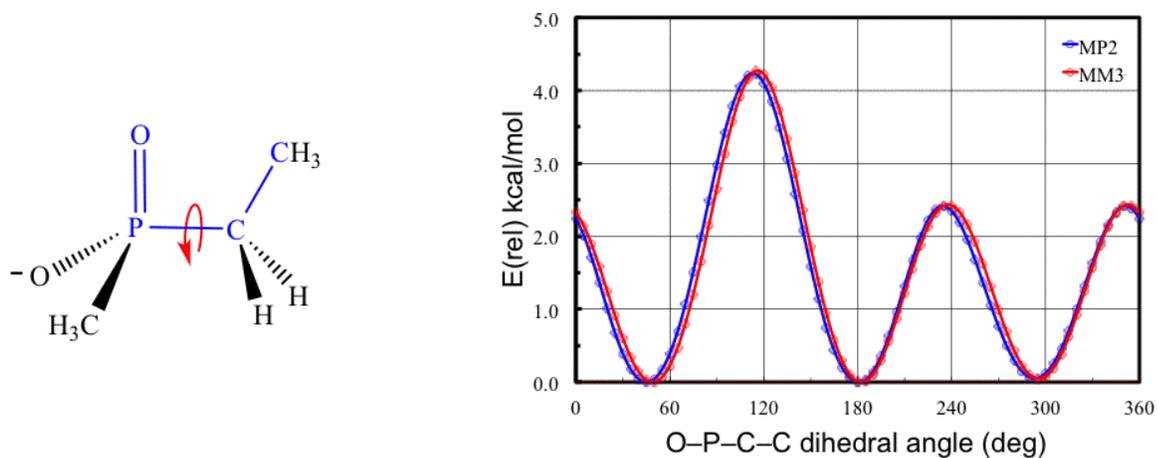


Figure S6. Potential energy surface for rotation about the O-P-C-C dihedral for an ethyl substituent.

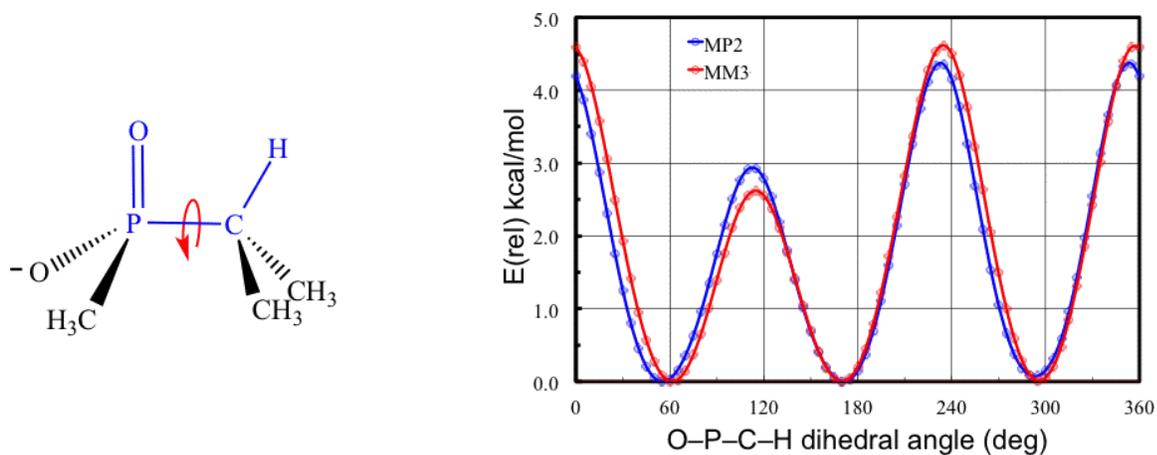


Figure S7. Potential energy surface for rotation about the O-P-C-H dihedral for an isopropyl substituent.

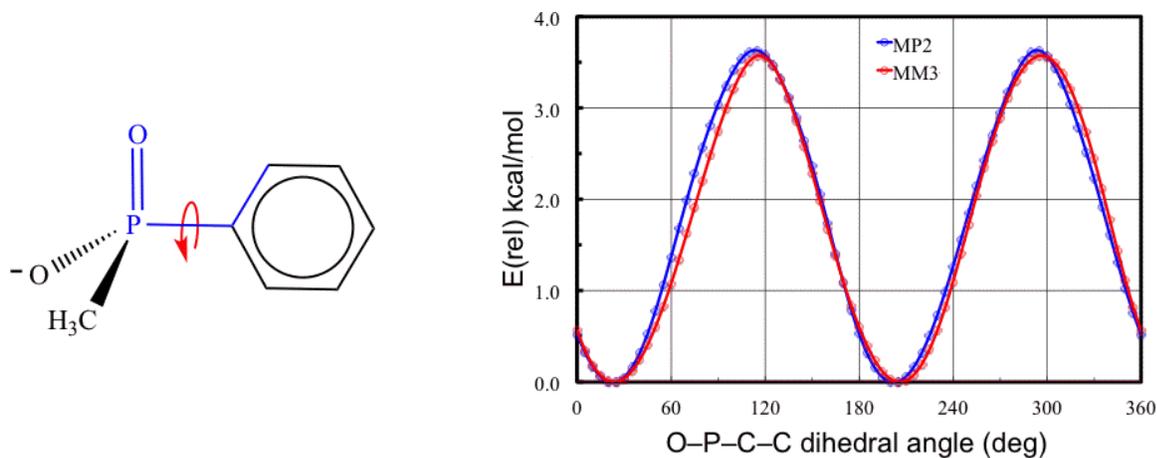


Figure S8. Potential energy surface for rotation about the O-P-C-C dihedral for a phenyl substituent.

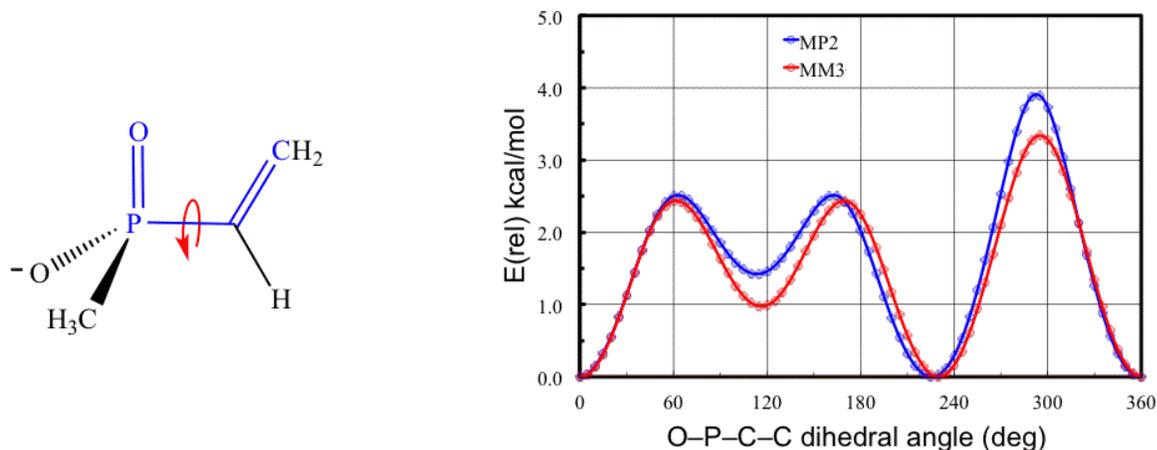


Figure S9. Potential energy surface for rotation about the O-P-C-C dihedral for a vinyl substituent.

Comparison of geometric parameters for phosphinate-lanthanide complexes with those for lanthanide-phosphine oxide complexes in the Cambridge Crystal Structure Database revealed indistinguishable mean values for Ln-O bond distances, Ln-O-P bond angles, and Ln-O-P-C torsion angles. Thus, parameters for phosphinate-Ce(III) interactions were assigned the same values used for phosphine oxide-Ce(III) interactions in McCann, B. W.; De Silva, N.; Windus, T. L.; Gordon, M. S.; Moyer, B. A.; Bryantsev, V. S.; Hay, B.P “Computer-Aided Molecular Design of Bis-phosphine Oxide Lanthanide Extractants” *Inorg. Chem.* **2016**, *55*, 5787-5803. In order to restrict movement of bastnäsité surface atoms during geometry optimization, initial lengths and angles taken from the DFT model of the surface and large stretching and bending force constants were assigned to all bonds and angles involving two or more surface atoms. Torsion barriers for all dihedral angles involving two or more surface atoms were set to zero. In this way, the geometry of the surface did not change during the optimization.

4. Host Designer sample input files (for the bidentate case)

Control file

```
OVER hosta=bidentate drivea numview=3000 and
nochiral noprochiral
```

Input fragment file

```
CC1 Bidentate
46
C 1 -3.80971 2.82391 -0.41612 1 4 38 39 40
O 2 -2.99719 0.85874 -2.07327 47 4
O 3 -4.21665 0.20094 0.10588 147 4
P 4 -3.20177 1.14275 -0.59290 153 1 2 3 36
41
C 5 -5.72728 -1.82765 -2.25179 215 14 15
Ce 6 -4.10019 -2.14971 0.95914 339 8 9 11 13
15 28
F 7 -0.41221 -0.85329 -2.77400 211 46 9 10 21
```

F	8	-2.08483	-1.86166	1.99510	210	6	9	11	22
O	9	-2.76885	-2.03462	-1.16575	219	6	7	8	
O	10	-2.18199	-0.02322	-5.79352	212	46	7	12	
O	11	-5.45162	-2.06304	3.14773	218	6	8	13	
O	12	-3.95082	-1.35465	-5.54083	213	46	10	14	
O	13	-6.46645	-2.87388	1.34182	217	6	11	15	
O	14	-5.09762	-1.16642	-3.11629	214	5	46	12	
O	15	-5.81724	-1.37067	-0.98824	216	5	6	13	
C	16	0.84237	-1.27512	0.27607	215	23	24		
Ce	17	3.96819	-0.48374	-0.93556	339	19	21	23	26
29									
Ce	18	2.46940	-1.59866	3.48732	339	20	22	24	25
27	29								
F	19	6.15678	-0.30177	-0.24703	211	17	25	26	
F	20	4.48402	-1.31012	4.52224	210	18	25	27	
O	21	2.61786	-0.80236	-3.01328	213	7	17	23	26
O	22	0.10306	-2.32162	3.86978	217	8	18	24	27
O	23	1.47184	-0.61545	-0.58803	214	16	17	21	
O	24	0.75126	-0.81828	1.53926	216	16	18	22	
O	25	3.80032	-1.48288	1.36309	219	18	19	20	
O	26	4.38669	0.52904	-3.26637	212	17	19	21	
O	27	1.11807	-1.51164	5.67518	218	18	20	22	
Du	28	-3.35031	-1.59288	-1.25213	200	46	6	36	
Du	29	3.21865	-1.04119	1.27585	200	17	18	35	
Du	30	3.29336	0.32631	1.60569	200	35			
C	31	4.91981	1.56489	2.82317	1	34	42	43	44
O	32	3.57265	1.41022	0.45400	47	34			
O	33	2.35208	0.75239	2.63300	147	34			
P	34	3.36784	1.69370	1.93519	153	31	32	33	35
45									
Du	35	2.96247	1.08130	1.54347	200	29	30	34	
Du	36	-3.60709	0.52991	-0.98364	200	4	28	37	
Du	37	-3.27604	-0.22506	-0.92252	200	36			
H	38	-4.74859	2.91330	-0.98650	5	1			
H	39	-4.00384	3.05564	0.64303	5	1			
H	40	-3.07399	3.53381	-0.82733	5	1			
H	41	-1.64888	1.01251	0.29563	5	4			
H	42	5.11475	0.49503	3.01400	5	31			
H	43	4.86014	2.07157	3.80017	5	31			
H	44	5.74300	1.98060	2.22128	5	31			
H	45	2.75905	3.37620	2.11155	5	34			
Ce	46	-2.60036	-1.03593	-3.46319	339	7	10	12	14
28									
G	46								
H	2								
41	X	0							
45	X	0							

```

D    4
T    I   -15.00  15.00   5.00   28   36           //twist on z-axis
T    I   -15.00  15.00   5.00   29   35           //twist on z-axis
A    I   -20.00  40.00   5.00   29   35   30       //rotate on O--O
axis
A    I   -20.00  40.00   5.00   28   36   37       //rotate on O--O
axis

```

HostDesigner Input Structures

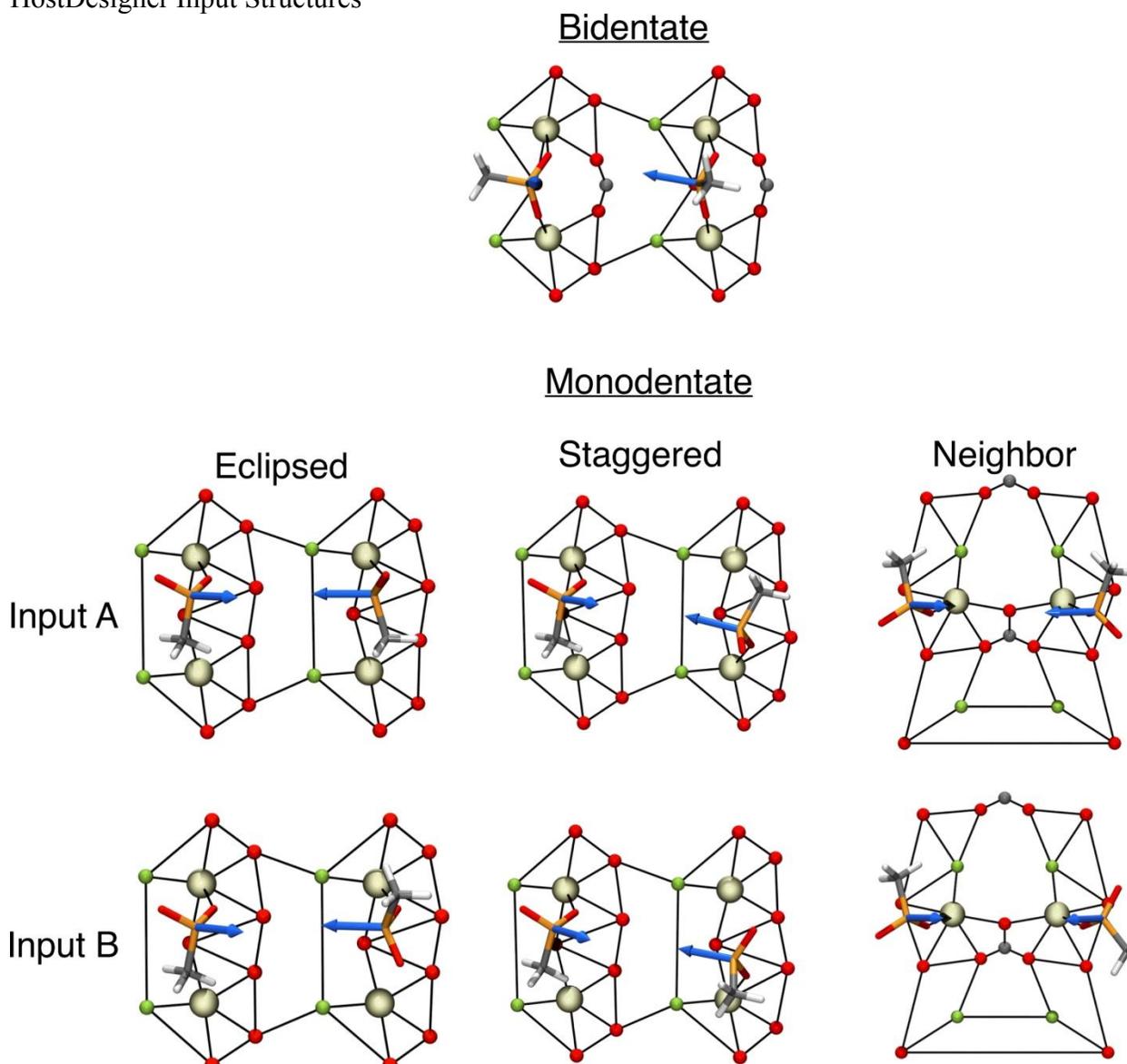


Figure S10. HostDesigner input fragments.

5. Molecular Mechanics Binding Energies

Table S3. Candidate ligand scoring based on relative free energy estimates, ΔG_{tot} (in kcal/mol), computed using molecular mechanics.

Structure	Bidentate	Neighboring	Repeated	Staggered	N_{rot}
1		1.16		0.90	2
2		1.31	0.96	0.92	2
3		1.54		0.92	2
4	1.84	1.37	2.07	1.07	3
5	4.43	1.74	1.64	1.31	3
6		1.58		2.49	2
7		1.59			2
8			1.64		2
9		2.79	2.99	1.84	5
10		2.47	2.00	2.45	4
11	3.54	2.95	2.24	2.39	6
12			2.31		4
13		2.33	3.09	3.67	4
14	4.99	2.97	3.42	2.45	7
15		4.29			2
16		2.95			4
17		3.56		4.12	6
18		4.19		3.76	4
19		4.12	3.93		2
20				3.96	2
21			4.49		6

Table S4. Bidentate ligand scoring based on relative free energy estimates, ΔG_{tot} (kcal/mol), computed with molecular mechanics. Contributions from the difference between the binding and the lowest energy conformations, ΔU_1 , strain energy ΔU_2 , and the number of rotatable bonds that are restricted in the ligand-metal complex, N_{rot} are shown.

Structure	ΔU_1	ΔU_2	N_{rot}	ΔG_{tot}
4	0.19	0.72	3	1.84
5	2.34	1.16	3	4.43
11	1.18	0.50	6	3.54
14	2.24	0.58	7	4.99

6. Images of Strongest Binding Ligands

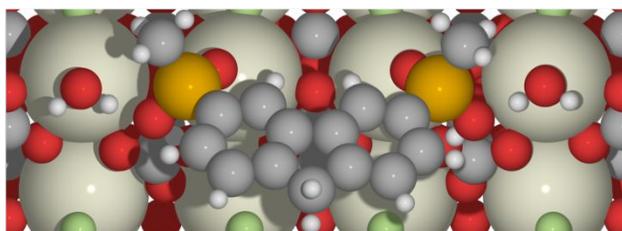


Figure S11. Image of ligand **2** bound to Ce-bastnäsite.

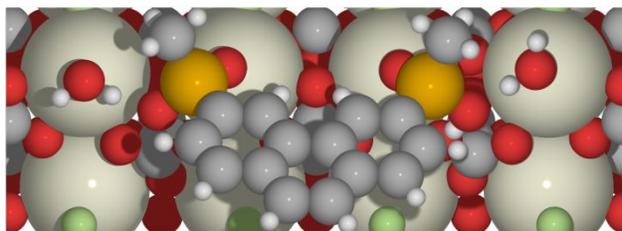


Figure S12. Image of ligand **6** bound to Ce-bastnäsite.

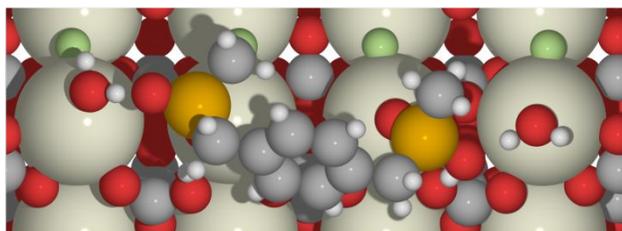


Figure S13. Image of ligand **10** bound to Ce-bastnäsite.

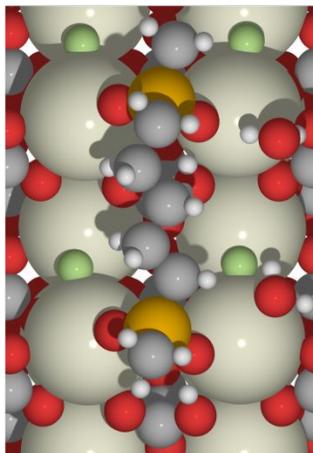


Figure S14. Image of ligand **11** bound to Ce-bastnäsite.

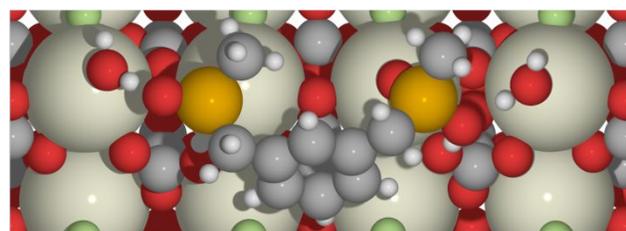


Figure S15. Image of ligand **13** bound to Ce-bastnäsite.

7. Additional MD Simulations Results

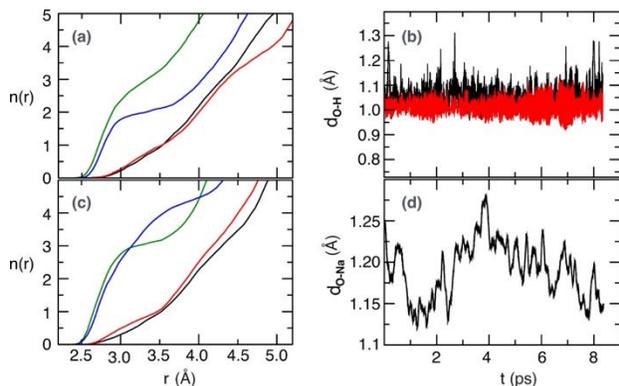


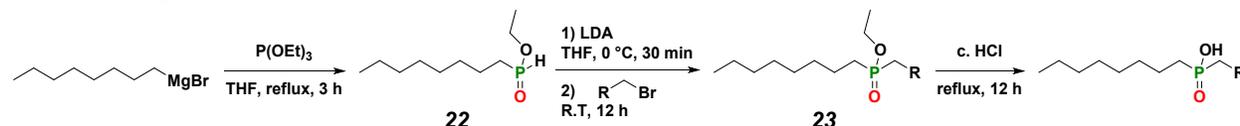
Figure S16. (a and c) Number of water molecules that are bound to the phosphinate oxygens of **13b**; black and red indicate partial water-exposure as these oxygens bind to the mineral surface Ce atoms, and green and blue indicate complete water-exposure. The cases with protonated and deprotonated carbonates (with Na^+ counterions) are represented by the top and bottom rows, respectively. (b) The time evolution of the distances between carbonate oxygens and the protons, indicating that protons remain bound to surface carbonates throughout the simulations. (d) The time evolution of one of the distances between carbonate oxygen and Na^+ , showing that Na^+ ions prefer to stay in the bulk water and away from the surface.

8. Additional Experimental Details

General

The ^1H and ^{13}C NMR spectra were recorded on an AvanceIII-400 MHz NMR spectrometer (Bruker Company) equipped with a 5 mm PABBO probe. Residual solvent peak was used as an internal reference (CDCl_3 : 7.26 ppm for ^1H NMR and 77.16 ppm for ^{13}C NMR). $\text{P}(\text{OEt})_3$ was used as an internal reference (137.5 ppm) for ^{31}P NMR. HRMS analyses were performed under contract by UT Austin mass spectrometric facility using positive mode ESI method. Commercially available compounds were purchased from Aldrich Chemical Co., Acros Organics, Alfa Aesar or TCI America and were used without further purification.

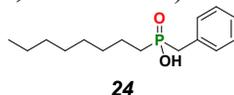
General Synthesis Scheme



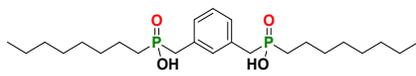
Scheme S1. Three-step synthesis of bis-phosphinic acids.

Synthesis

Ethyl octylphosphinate (**22**) was prepared according to a literature procedure by Petneházy et al. (*Synthetic Communication*, **2003**, 33, 10, 1665–1674).

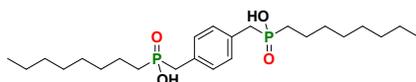


Benzyl(octyl)phosphinic acid (24): A solution of ethyl octylphosphinate **22** (5.0 g, 0.024 mol) in anhydrous THF (110 mL) was cooled in ice/water bath under inert atmosphere, to this solution 2 M lithium diisopropylamide solution in THF (12.0 mL, 0.024 mol) was added dropwise. After 30 minutes, benzyl bromide (2.6 mL, 0.022 mol) was added to this solution and the reaction mixture was allowed to warm to room temperature and stir at this temperature for 12 hours. DI water was then added to the reaction mixture and the product was extracted with diethyl ether (3x), organic phases were combined, washed with DI water, then brine, and dried over anhydrous MgSO_4 . After filtration, solvent was evaporated under reduced pressure on a rotary evaporator. The crude material was purified on CombiFlash R_f automated flash chromatography system using normal phase silica gel as a stationary phase and isocratic 3% MeOH in CH_2Cl_2 as an eluent system. The product **23** ($\text{R} = \text{Ph}$) was obtained as a yellow, viscous oil (1.5 g, 23% yield). In a round bottom flask equipped with a magnetic stirrer and reflux condenser, compound **23** ($\text{R} = \text{Ph}$) (1.0 g, 3.37 mmol) was mixed with 12 M HCl solution (5 mL) and the mixture was refluxed for 12 hours. Afterwards, the reaction mixture was allowed to cool to room temperature, before being diluted with DI water (~20 mL). The precipitate was filtered, washed with DI water, dried, and then washed with Et_2O . The product **24** was obtained as white powder (0.45 g, 50% yield). ^1H NMR (400 MHz, CDCl_3) 7.38-7.19 (m, 5H), 3.19 (d, $J = 15.6$ Hz, 2H), 1.82-1.69 (m, 2H), 1.65-1.51 (m, 2H), 1.41-1.19 (m, 10H), 0.93-0.83 (m, 6H). ^{13}C NMR (100.66 MHz, CDCl_3) 130.0, 129.6, 129.0, 127.7, 36.4, 31.9, 30.6 ($J = 15.3$ Hz), 29.1, 22.7, 21.0 ($J = 5.6$ Hz), 14.2. ^{31}P NMR (400 MHz, CDCl_3) 85.9. HRMS m/z : $[\text{M}+\text{Na}]^+$, calculated for $\text{C}_{12}\text{H}_{25}\text{O}_2\text{P}$, 291.1484; found 291.1478.



13b

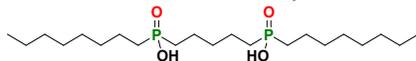
(1,3-Phenylenebis(methylene))bis(octylphosphinic acid) (**13b**): A solution of ethyl octylphosphinate **22** (2.6 g, 12.6 mmol) in anhydrous THF (20 mL) was cooled in ice/water bath under inert atmosphere, to this solution 2 M lithium diisopropylamide solution in THF (6.2 mL, 12.6 mmol) was added dropwise. After 30 minutes, 1,3-bis(bromomethyl)benzene (1.5 g, 5.68 mmol) dissolved in 8 mL of anhydrous THF was added to this solution. The reaction mixture was allowed to warm to room temperature and stir at this temperature for 12 hours. DI water was then added to the reaction mixture and the product was extracted with diethyl ether (3x), organic phases were combined, washed with DI water, then brine, and dried over anhydrous MgSO₄. After filtration, solvent was evaporated under reduced pressure on a rotary evaporator. The crude material was purified on CombiFlash R_f automated flash chromatography system using normal phase silica gel as a stationary phase and isocratic 4% MeOH in CH₂Cl₂ as an eluent system. The product **23** was obtained as a yellow, viscous oil (1.31 g, 45% yield). In a round bottom flask equipped with a magnetic stirrer and reflux condenser compound **23** (1.0 g, 1.94 mmol) was mixed with 12 M HCl solution (5 mL) and the mixture was refluxed for 12 hours. Afterwards, the reaction mixture was allowed to cool to room temperature, before being diluted with DI water (~20 mL). The precipitate was filtered, washed with DI water, dried, and then washed with Et₂O. The product **13b** was obtained as white powder (0.72 g, 81% yield). ¹H NMR (400 MHz, CDCl₃) 7.30-7.22 (m, 1H), 7.17-7.08 (m, 3H), 3.05 (d, *J* = 14.4 Hz, 4H), 1.78-1.65 (m, 4H), 1.62-1.49 (m, 4H), 1.42-1.19 (m, 20H), 0.94-0.83 (m, 6H). ¹³C NMR (100.66 MHz, CDCl₃) 131.7, 130.8, 130.7, 129.2, 36.2 (*J* = 82.7 Hz), 31.9, 30.8 (*J* = 16.3 Hz), 29.2, 28.0 (*J* = 93.1 Hz), 22.8, 21.3, 14.2. ³¹P NMR (400 MHz, CDCl₃) 89.1. HRMS *m/z*: [M+Na]⁺, calculated for C₂₄H₄₄O₄P₂, 481.2607; found 481.2608.



10b

(1,4-Phenylenebis(methylene))bis(octylphosphinic acid) (**10b**): A solution of ethyl octylphosphinate **22** (2.6 g, 12.6 mmol) in anhydrous THF (20 mL) was cooled in ice/water bath under inert atmosphere, to this solution 2 M lithium diisopropylamide solution in THF (6.2 mL, 12.6 mmol) was added dropwise. After 30 minutes, 1,4-bis(bromomethyl)benzene (1.5 g, 5.68 mmol) was added to this solution. The reaction mixture was allowed to warm to room temperature and stir at this temperature for 12 hours. DI water was then added to the reaction mixture and the product was extracted with diethyl ether (3x), organic phases were combined, washed with DI water, then brine, and dried over anhydrous MgSO₄. After filtration, solvent was evaporated under reduced pressure on a rotary evaporator. The crude material was purified on CombiFlash R_f automated flash chromatography system using normal phase silica gel as a stationary phase and isocratic 2.5% MeOH in CH₂Cl₂ as an eluent system. The product **23** was obtained as a viscous oil (1.2 g, 41% yield). In a round bottom flask, equipped with a magnetic stirrer and reflux condenser compound **23** (0.97 g, 1.88 mmol) was mixed with 12 M HCl solution (5 mL) and the mixture was refluxed for 12 hours. Afterwards, the reaction mixture was allowed to cool to room temperature, before being diluted with DI water (~20 mL). The precipitate was filtered, washed with DI water, dried, and then washed with Et₂O. The product **10b** was obtained as white powder (0.70 g, 87% yield). ¹H NMR (400 MHz, CDCl₃) 7.17 (s, 4H), 2.96 (d, *J* = 13.8 Hz, 4H), 1.81-1.66 (m, 4H), 1.64-1.50 (m, 4H), 1.41-1.18 (m, 20H), 0.93-0.83 (m, 6H). ¹³C NMR (100.66 MHz,

CDCl₃) 130.5, 129.4, 31.9, 30.7 ($J = 13.8$ Hz), 29.1, 22.7, 21.2, 14.2. ³¹P NMR (400 MHz, CDCl₃) 82.4. HRMS m/z : [M+Na]⁺, calculated for C₂₄H₄₄O₄P₂, 481.2607; found 481.2614.



11b

Pentane-1,5-diylbis(octylphosphinic acid) (11b): A solution of ethyl octylphosphinate **22** (5.0 g, 0.024 mol) in anhydrous THF (120 mL) was cooled in ice/water bath under inert atmosphere, to this solution 2 M lithium diisopropylamide solution in THF (12.2 mL, 0.024 mol) was added dropwise. After 30 minutes, 1,5-dibromopentane (1.5 mL, 11.01 mmol) was added to the solution. The reaction mixture was allowed to warm to room temperature and stir at this temperature for 12 hours. DI water was then added to the reaction mixture and the product was extracted with diethyl ether (3x), organic phases were combined, washed with DI water, then brine, and dried over anhydrous MgSO₄. After filtration, solvent was evaporated under reduced pressure on a rotary evaporator. The crude material was purified on CombiFlash R_f automated flash chromatography system using normal phase silica gel as a stationary phase and isocratic 3% MeOH in CH₂Cl₂ as an eluent system. The product **23** was obtained as a viscous oil (2.5 g, 47% yield). In a round bottom flask equipped with a magnetic stirrer and reflux condenser compound **23** (1.0 g, 2.08 mmol) was mixed with 12 M HCl solution (5 mL) and the mixture was refluxed for 12 hours. Afterwards, the reaction mixture was allowed to cool to room temperature, before being diluted with DI water (~20 mL). The precipitate was filtered, washed with DI water, dried, and then washed with Et₂O. The product **11b** was obtained as white powder (0.78 g, 88% yield). ¹H NMR (400 MHz, CDCl₃) 1.86-1.72 (m, 8H), 1.72-1.46 (m, 10H), 1.44-1.34 (m, 4H), 1.34-1.19 (m, 16H), 0.95-0.82 (m, 6H). ¹³C NMR (100.66 MHz, CDCl₃) 31.9, 30.7 ($J = 15.4$ Hz), 29.2, 29.1, 28.6, 28.3, 27.7, 22.7, 21.2 ($J = 4.3$ Hz), 20.8 ($J = 5.5$ Hz), 14.2. ³¹P NMR (400 MHz, CDCl₃) 76.2. HRMS m/z : [M+Na]⁺, calculated for C₂₁H₄₆O₄P₂, 447.2764; found 447.2766.

1 mM sample preparation for vSFG measurements

Each bis-phosphinic acid (0.01 mmol) was combined with 9.8 mL of DI water and 0.2 mL of 0.1 M NaOH (2 equiv.) in 15 mL centrifuge tube. The mono-phosphinic acid (0.01 mmol) was combined with 9.9 mL of DI water and 0.1 mL of 0.1 M NaOH (1 equiv.). The four tubes were rotated on the wheel at 24 °C for 14 hours. Each solution was then filtered through a 0.45 μm Nylon filter. pH of each solution was measured using the universal pH-indicator strip to be ~5.

ATR-FTIR spectra

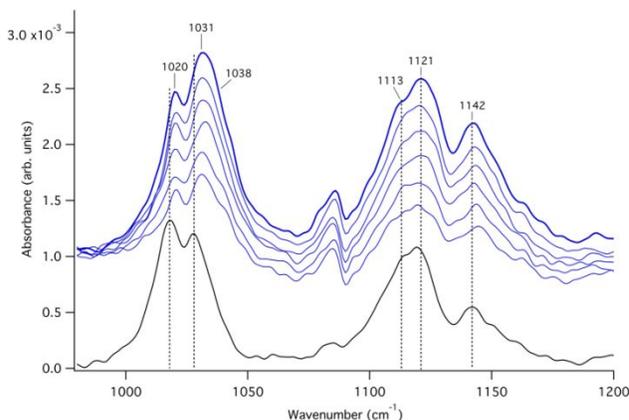


Figure S17. FTIR spectra of **10b** (0.4 mM) adsorbed onto synthetic Ce-bastnäsité. The spectrum of the ligand in aqueous solution (4 mM) is shown in black at the bottom of the plot. The colored traces are the spectra taken at (bottom to top) 10, 20, 30,

40, 50, and 60 minutes of the adsorbed ligand *in situ*, with the bastnäsité particle film and bulk water spectra subtracted (adsorbed spectra are vertically offset from aqueous spectrum for clarity). The unlabeled peaks at 1090 cm^{-1} in each adsorbed spectrum are a residual peak from the bastnäsité background.

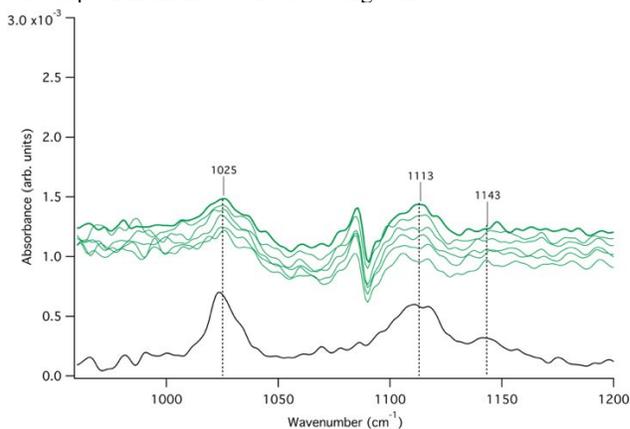


Figure S18. FTIR spectra of **24** (1 mM) adsorbed onto synthetic Ce-bastnäsité. The spectrum of the ligand in aqueous solution (4 mM) is shown in black at the bottom of the plot. The colored traces are the spectra taken at (bottom to top) 10, 20, 30, 40, 50, and 60 minutes of the adsorbed ligand *in situ*, with the bastnäsité particle film and bulk water spectra subtracted (adsorbed spectra are vertically offset from aqueous spectrum for clarity). The unlabeled peaks at 1090 cm^{-1} in each adsorbed spectrum are a residual peak from the bastnäsité background.

Spectra of **10b** show two components of the symmetric O-P-O stretch band, likely due to a slight difference between the two phosphinic functional groups. In the aqueous phase, these are centered at 1018 and 1028 cm^{-1} . When adsorbed to the bastnäsité film, these peaks shift slightly to 1020 and 1031 cm^{-1} . An additional component to this band appears as a shoulder at 1038 cm^{-1} . The asymmetric stretch band at 1142 cm^{-1} does not appear to shift, but increases in intensity relative to the neighboring aromatic ring band in the adsorbed state. The adsorption affinity of **10b** at the bastnäsité interface is apparent by the intensity of ligand bands in the *in-situ* spectra. Symmetric and asymmetric O-P-O stretching bands of **24** (1025 cm^{-1} and 1143 cm^{-1} , respectively) are very weak in the adsorbed spectra, and do not shift from aqueous positions. This indicates a weaker adsorption of **24**, compared to **10b** and **13b**.

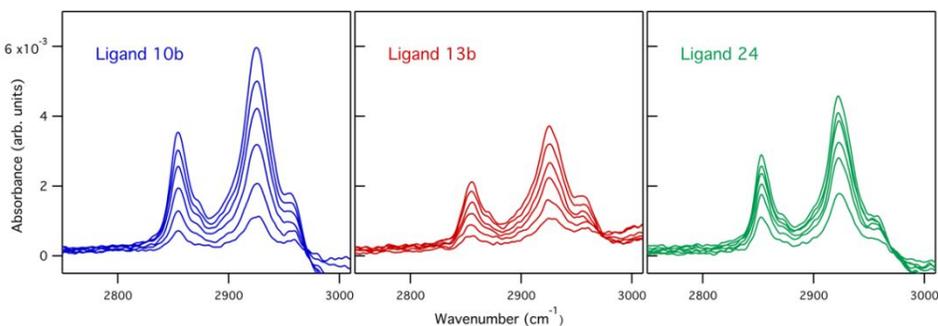


Figure S19. FTIR spectra of **10b**, **13b**, and **24** adsorbed to synthetic bastnäsité film after (bottom to top) 10, 20, 30, 40, 50, 60 minutes. Peaks in this region of the spectrum correspond to CH stretching.

The CH stretching vibrations appear in the FTIR in the same region as the SFG spectra, indicating the presence of the ligand at the surface (Figure S19). These characteristic bands do not change considerably with the adsorption mechanism. While these *in situ* experiments were not controlled to measure the quantity of adsorbed ligands directly, the trend in relative intensities of the CH stretching bands is consistent with the SFG results. **10b** has the most intense CH stretching peaks, followed by **24**. **13b** has relatively weaker CH stretching bands, despite the strong evidence of phosphinate chemisorption. In contrast, **24** shows very weak IR

absorption in the region of O-P-O stretching, but relatively strong CH stretching peaks. This supports the interpretation that **24** adsorbs at the surface, but not via a strong interaction with the phosphinate functional group.

Hydrogen bonding plays a significant role in adsorption, affecting ligand interaction with the solution, as well as the surface and other ligand molecules. The ligands are deprotonated under the conditions tested, forming hydrogen-bonds to water molecules. The broad IR absorbance of bulk water cannot be separated from ligand-specific OH stretches by simple background subtraction, preventing the inclusion of OH stretches in the *in-situ* FTIR analysis.

Background spectra from vSFG and ATR-FTIR experiments

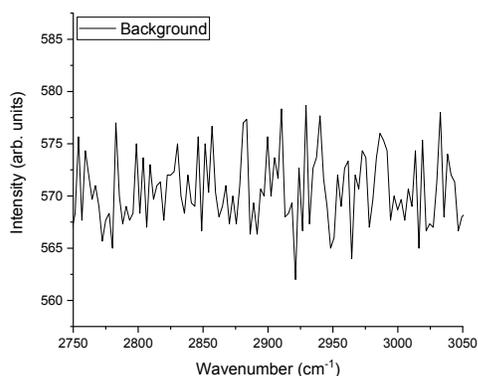


Figure S20. Representative background spectrum from vSFG measurements.

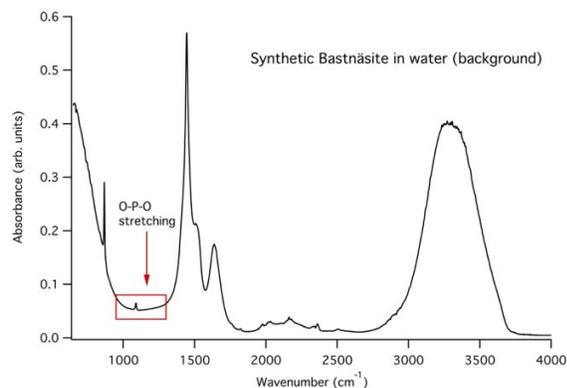


Figure S21. ATR-FTIR spectrum of the synthetic bastnäsite particle film in water, with the O-P-O stretching region (used for *in situ* spectra) in red. The subtracted background is mostly flat in the region of O-P-O stretching, except for a small peak at 1089 cm^{-1} , corresponding to CO_3 symmetric stretching in the bastnäsite solid.

Mineral surface characterization

Table S5. As-prepared Ce-bastnäsite and calcite surface areas (m^2/g) and water content (%).

Mineral	Surface area measured previously by BET*	Weight percentage of H_2O measured by TG-DSC
CeFCO_3	21.50	3.00%
CaCO_3	22.41	0.66%

*Srinivasan, S. G.; Shivaramaiah, R.; Kent, P. R. C.; Stack, A. G.; Riman, R.; Anderko, A.; Navrotsky, A.; Bryantsev, V. S. *Phys. Chem. Chem. Phys.* **2017**, *19*, 7820-7832.

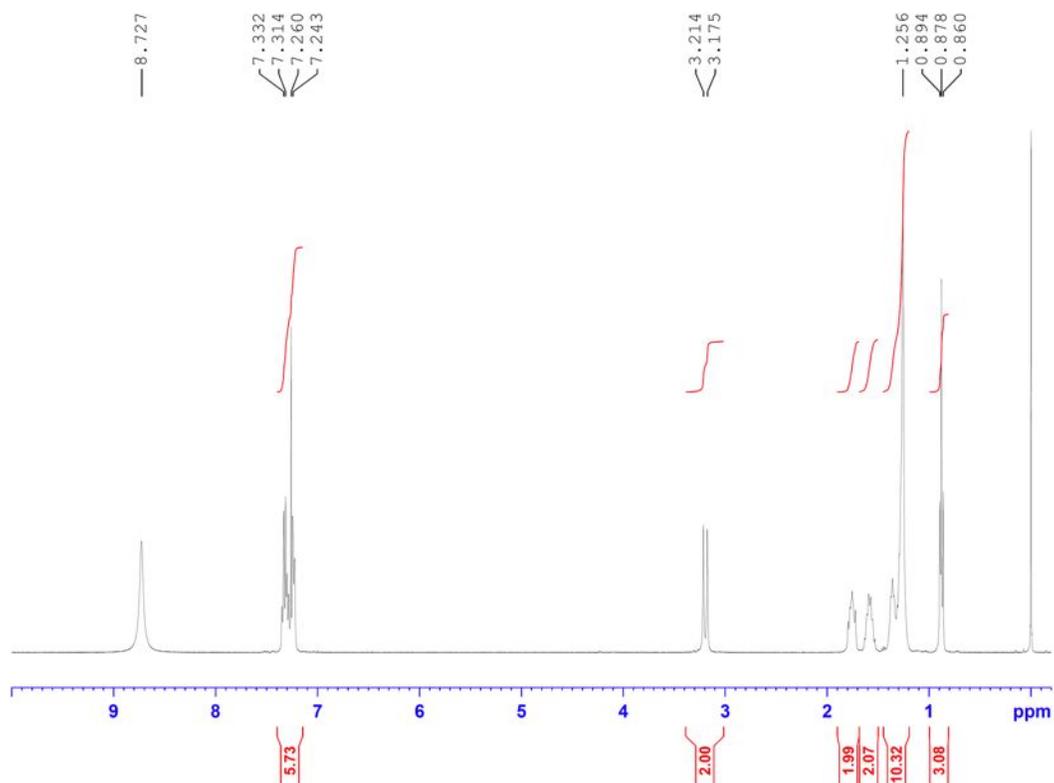


Figure S22. ^1H NMR spectrum of **24** in CDCl_3 .

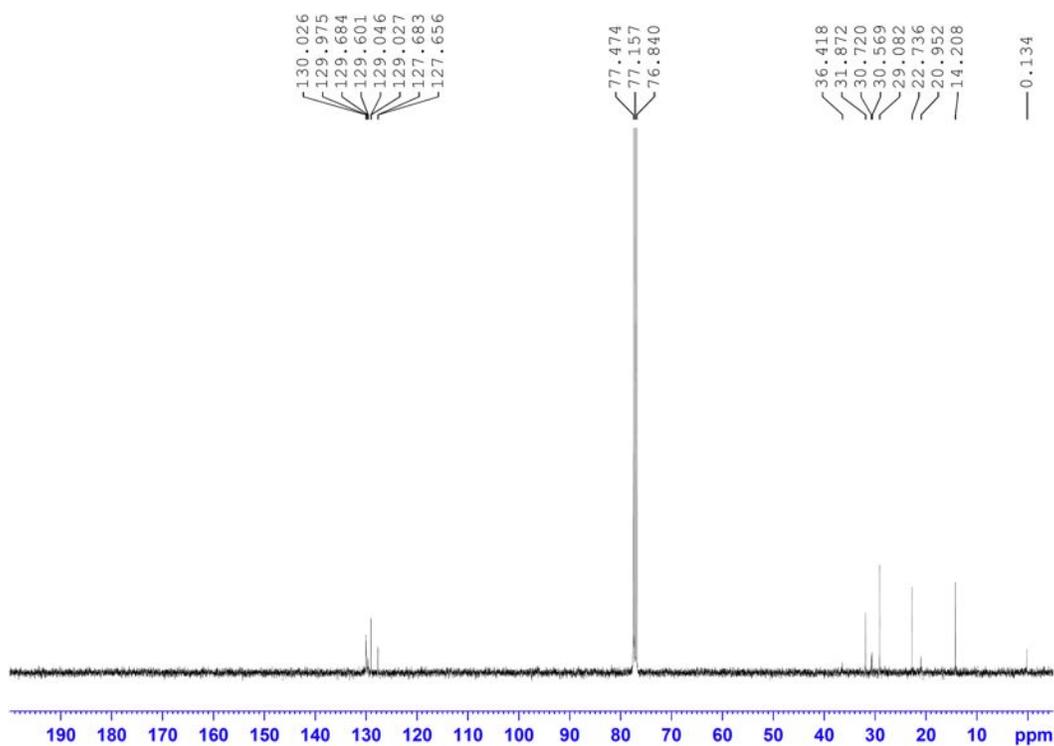


Figure S23. ^{13}C NMR spectrum of **24** in CDCl_3 .

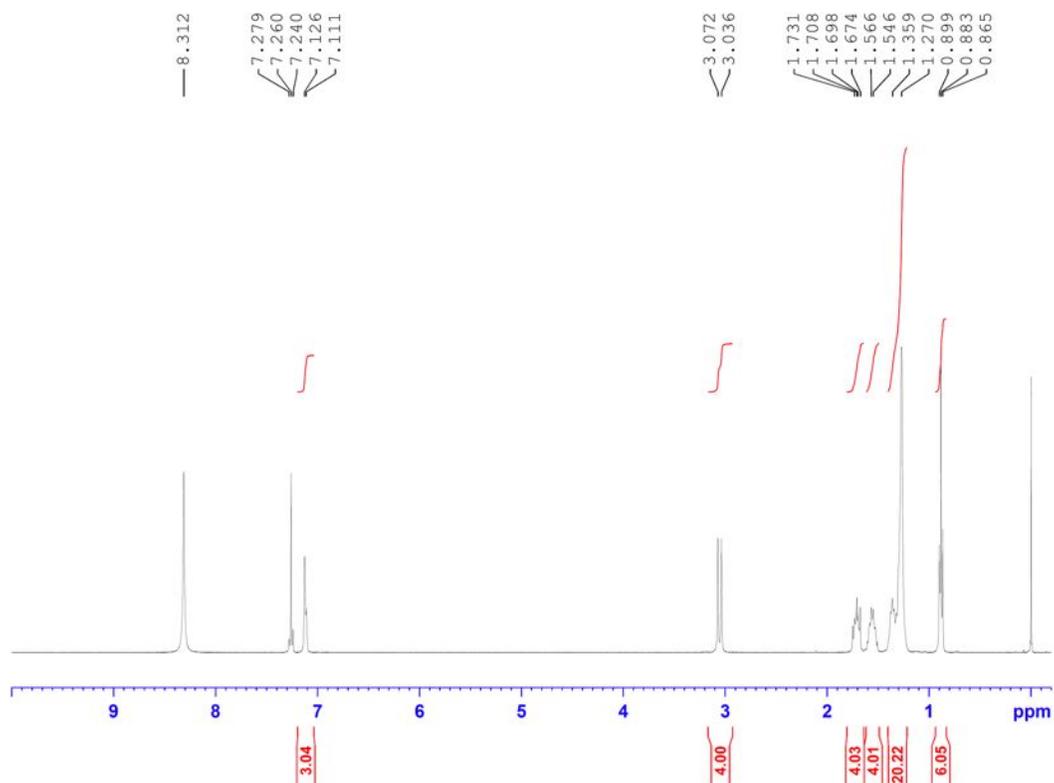


Figure S24. ^1H NMR spectrum of **13b** in CDCl_3 .

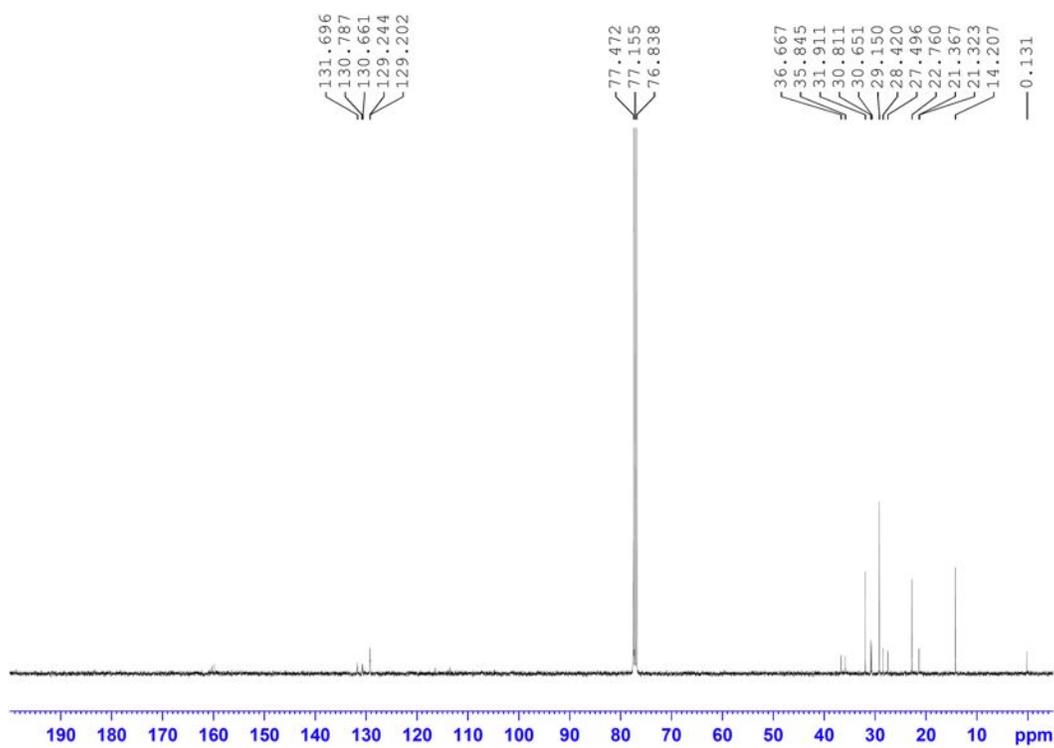


Figure S25. ^{13}C NMR spectrum of **13b** in CDCl_3 .

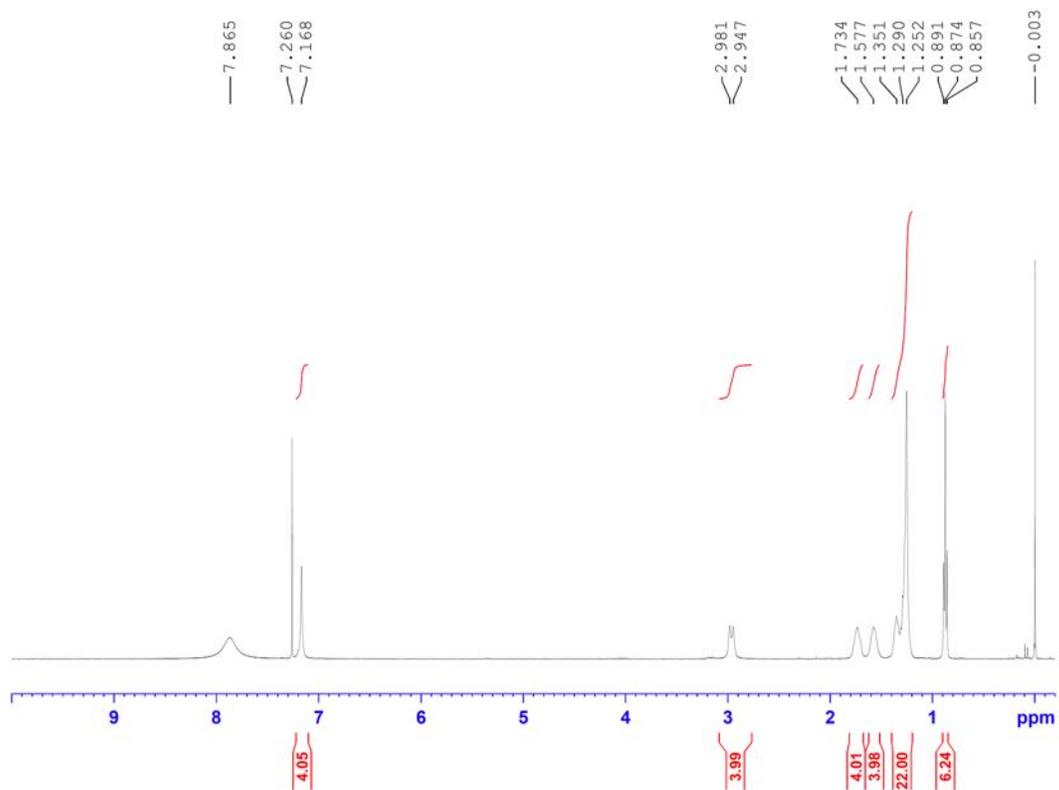


Figure S26. ^1H NMR spectrum of **10b** in CDCl_3 .

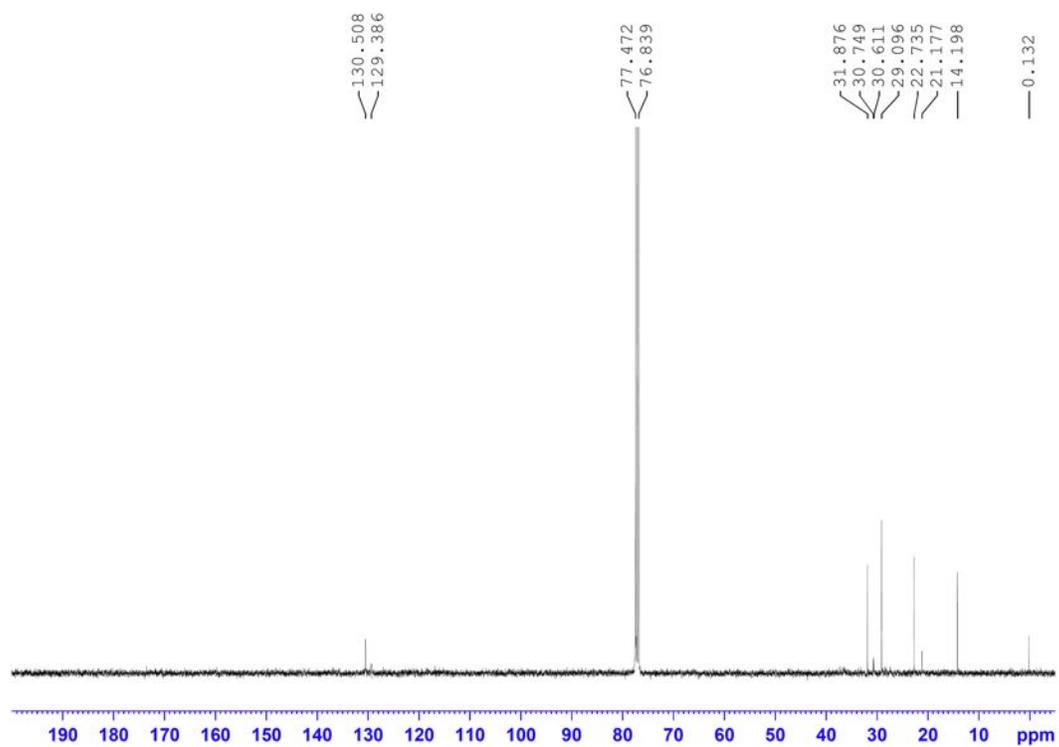


Figure S27. ^{13}C NMR spectrum of **10b** in CDCl_3 .

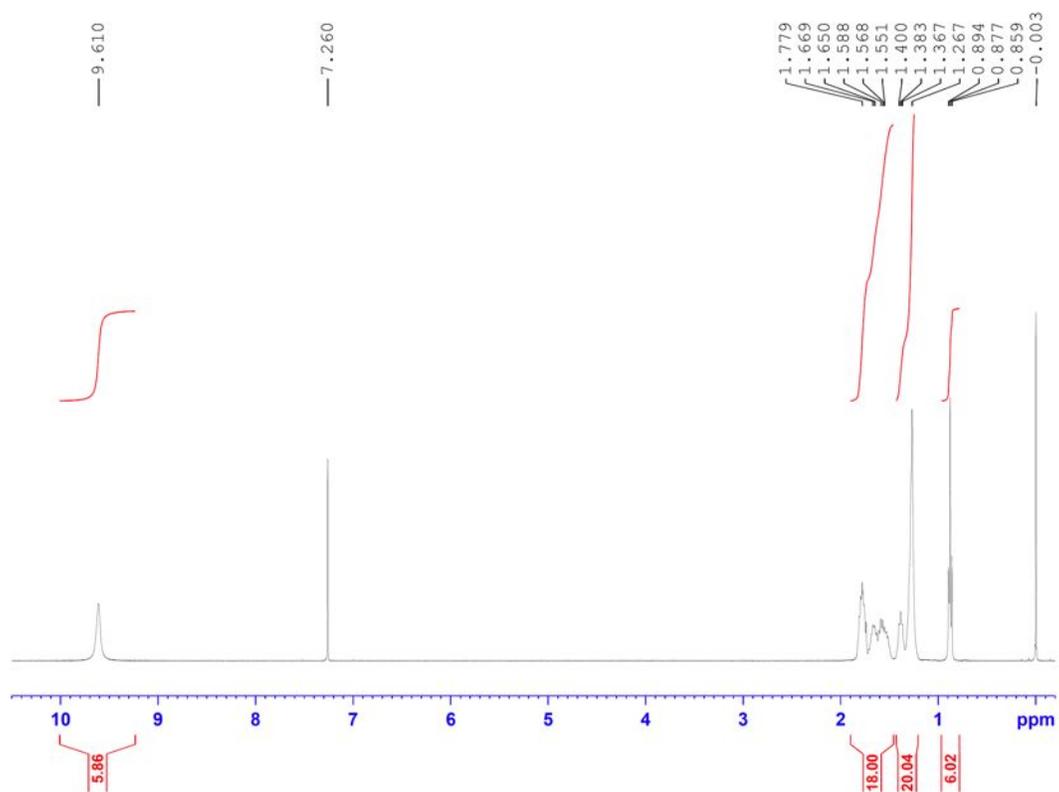


Figure S28. ^1H NMR spectrum of **11b** in CDCl_3 .

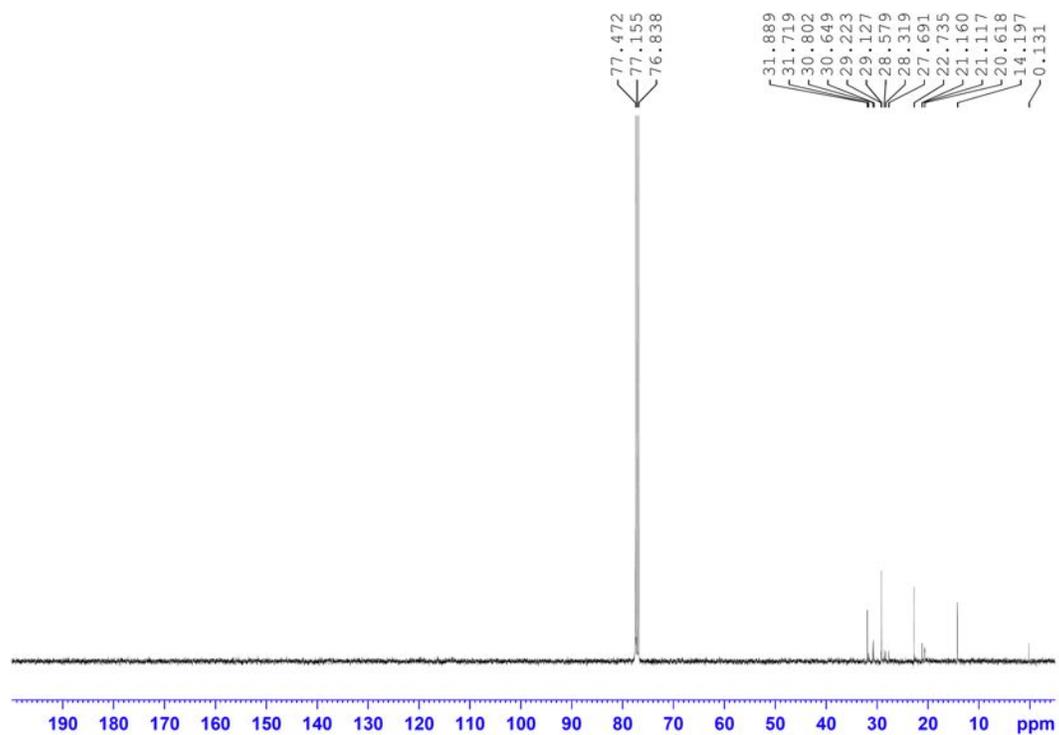


Figure S29. ^{13}C NMR spectrum of **11b** in CDCl_3 .

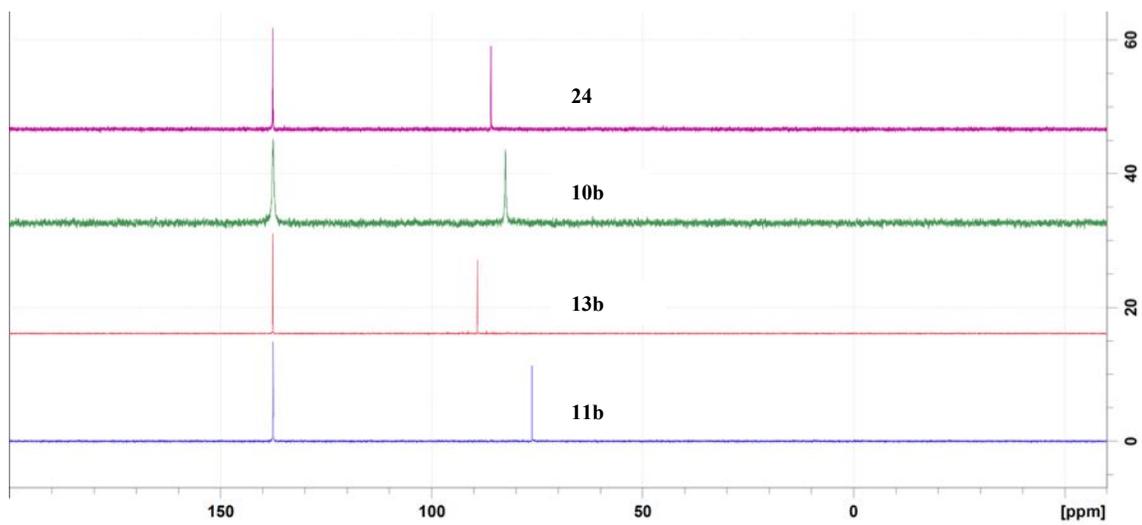


Figure S30. ^{31}P NMR spectra of **10b**, **11b**, **13b**, and **24** in CDCl_3 .