Controlled Diffusion of Photoswitchable Receptors by Binding Anti-Electrostatic Hydrogen-Bonded Phosphate Oligomers

Thomas S. C. MacDonald, Ben L. Feringa, William S. Price, Sander J. Wezenberg, and Jonathon E. Beves

Table of Contents

S1. General Experimental

Reagents were purchased from Combi-Blocks or Sigma-Aldrich and used as received. Deuterated NMR solvents were purchased from Cambridge Isotope Laboratories and used as received unless otherwise noted.

NMR experiments were conducted using a Bruker Avance III 500 MHz spectrometer equipped with a TBI (^{31}P) probe fitted with standard high-resolution gradient coils capable of generating gradient pulses up to 53.45 G cm^{-1}. Diffusion experiments were prepared and executed through the Bruker **diff5** utility and used the **diffSte** stimulated-echo pulse sequence. Typical parameters for ^{1}H diffusion experiments were Δ = 50 ms and δ = 4 ms. Typical parameters for ^{31}P diffusion experiments were Δ = 100 ms and δ = 7 ms. Gradient pulses were sine-shaped, and gradient lists typically consisted of 12 gradient strengths linearly increasing from 0 – 53.45 g cm^{-1}.

Spectra were automatically phased and baseline-corrected with MestReNova software 12.0, with peak integrals obtained using the built-in data analysis routines and exported as csv data for further processing.

High resolution mass spectra were collected on a Thermo Scientific LTQ-FT-ICR Mass Spectrometer equipped with a nanoESI source and Thermo Xcalibur software used for analysis.

S2. Speciation of acetate and dihydrogen phosphate: literature data

Given the focus placed upon the dihydrogen phosphate anion, it is important to consider potential protonation-deprotonation that may occur under the conditions of wet DMSO. Acid-base properties in non-aqueous solutions are notoriously uncharacterised, but in this case we have access to a carefully-conducted recent study of the acid-base properties of $H_2PO_4^-$ and $CH_3CO_2^-$ in DMSO.[1] That study also measured an association constant for anti-electrostatic $H_2PO_4^-$ dimerization of 180 M^{-1} in DMSO, but did not consider indefinite association to form higher oligomers as is presented here.

It has previously been shown that the deprotonation of $H_2PO_4^-$ in DMSO is highly unfavorable.[1] As such, we limit our consideration to the possibility of forming of phosphoric acid, H_3PO_4, by protonation of $H_2PO_4^-$ with water. Literature data for K^H, the acid association constant of a species, are given in Table S1.

Table S1. Literature acid-base association constants for phosphate,[1] acetate,[1] and water[2] in DMSO.

$$K^{\mathrm{H}} = \frac{[\mathrm{AH}]}{[\mathrm{H}][\mathrm{A}]}$$

Species	log K^H (DMSO)
H_3PO_4[1]	10.8
H_3CO_2H[1]	12.8
H_2O[2]	28

S2.1 Speciation of phosphate in DMSO

Proton transfer from water to dihydrogen phosphate would occur as:

$$\mathrm{H_2PO_4^-} + \mathrm{H_2O} \leftrightarrow \mathrm{H_3PO_4} + \mathrm{OH^-}$$

With the acid-base relationship and using the values in Table S1:

$$\frac{[\mathrm{OH^-}][\mathrm{H_3PO_4}]}{[\mathrm{H_2PO_4^-}][\mathrm{H_2O}]} = \frac{10^{-28}}{10^{-10.8}} = 10^{-17.2}$$

At 0.5% v/v H_2O, $[H_2O]_0$ = 280 mM. Considering the case for 50 mM $[H_2PO_4^-]$ and setting x as the concentration $[OH^-]$ and $[H_3PO_4]$, we can restate this as:

$$\frac{[\mathrm{OH^-}][\mathrm{H_3PO_4}]}{[\mathrm{H_2PO_4^-}][\mathrm{H_2O}]} = \frac{x^2}{([\mathrm{H_2PO_4^-}]_0 - x)([\mathrm{H_2O}]_0 - x)} = 10^{-17.2}$$

And then:

$$\frac{x^2}{(50 \times 10^{-3} - x)(280 \times 10^{-3} - x)} = 10^{-17.2}$$

$$x^2 = \left((50 \times 10^{-3})(280 \times 10^{-3}) - x(280 \times 10^{-3} + 50 \times 10^{-3}) + x^2\right)10^{-17.2}$$

$$x = 9.4 \times 10^{-9}\ \mathrm{M}^{-1}$$

$[H_3PO_4]$ is thus expected to be vanishingly small, and is not expected to play any part in supramolecular assembly.

S2.2 Speciation of acetate in DMSO

Following the same steps as above:

$$\frac{[\mathrm{OH^-}][\mathrm{AcOH}]}{[\mathrm{AcO^-}][\mathrm{H_2O}]} = \frac{10^{-28}}{10^{-12.8}} = 10^{-15.2}$$

And solving for the extent of proton transfer:

$$\frac{x^2}{(50 \times 10^{-3} - x)(280 \times 10^{-3} - x)} = 10^{-15.2}$$

$$x^2 = \left((50 \times 10^{-3})(280 \times 10^{-3}) - x(280 \times 10^{-3} + 50 \times 10^{-3}) + x^2\right)10^{-17.2}$$

$$x = 9.4 \times 10^{-8}\ \mathrm{M^{-1}}$$

While the formation of CH_3CO_2H is slightly more favorable than that of H_3PO_4, it is still expected to only be present at insignificant (~nM) concentration.

S2.3 Literature association data[1]

In a previous study, acetate was proposed to be incapable of forming an anti-electrostatic dimer, while a dimerization constant of K = 180 M^{-1} was reported for dihydrogen phosphate.[1] No indefinite association models were considered in that work.

These anions are able to associate with their respective conjugate acid ($K_{1,1}$), with a high (~10^4 M^{-1}) association constant found for phosphate.

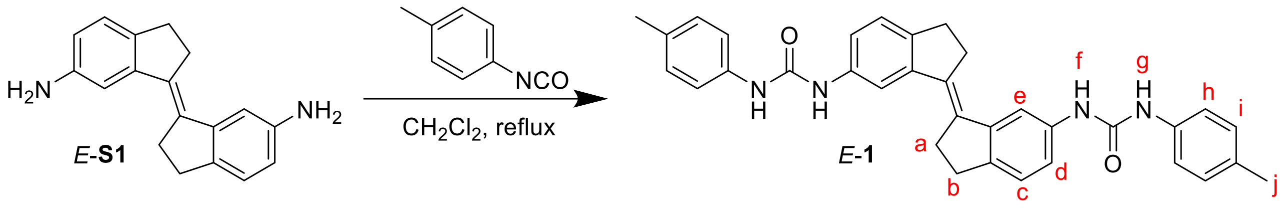
Table S2. Literature association values for dimerization of phosphate and acetate species, with method of determination noted in parentheses.[1] H_2PO_4 was found to form anti-electrostatic dimers with K = 180 M^{-1}, similar to the 120 M^{-1} isodesmic K_i found here using an indefinite self-association model.

Species	**log $K_{0,2}$ $[B^- + B^-] \rightarrow [B_2]^{2-}$**	**$K_{0,2}$ / M^{-1}**	**log $K_{1,1}$ $[HB + B^-] \rightarrow [HB_2]^-$**	**$K_{1,1}$ / M^{-1}**	**log $K_{1,2}$ $[HB\text{-}B]^- + B^- \rightarrow [HB_2]^{2-}$**	**$K_{1,2}$ / M^{-1}**
$H_2PO_4^-$[1]	2.26(6) (ITC)	1.8×10^2	4.23(3) (NMR)	1.7×10^4	2.92 (ITC)	8.3×10^2
OAc^-[1]	-	-	2.45(1) (ITC)	2.8×10^2	-	

S3. Synthesis

The *E* and *Z*-stiff-stilbene diamine **S1** was prepared as previously reported.[3]

S3.1 Preparation of bis-urea hosts *E*-1 and *Z*-1



Compound *E*-**S1** (100 mg, 0.38 mmol) was suspended in 50 mL dry CH_2Cl_2, and *p*-tolylisocyanate (101 mg, 95 µL, 0.76 mmol) was added by syringe. The reaction mixture was stirred at reflux overnight, cooled to rt, and the precipitate was filtered off to obtain crude *E*-**1**. The crude was washed with CH_2Cl_2 and aq. HCl (1M), suspended in water and sonicated, and the precipitate was collected by filtration to give *E*-**1** as a fine beige solid (167 mg, 83%).

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.70 (s, 2H, H^f), 8.57 (s, 2H, H^g), 7.89 (d, J = 1.9 Hz, 2H, H^e), 7.35 (d, J = 8.1 Hz, 4H, H^h), 7.26 – 7.20 (m, 4H, H^{c,d}), 7.09 (d, J = 8.1 Hz, 4H, H^i), 3.11 (m, 4H, H^a), 3.04 (m, 4H, H^b), 2.25 (s, 6H, H^j). ^{13}C NMR (151 MHz, DMSO-d_6) δ 152.80, 140.12, 138.37, 137.24, 135.08, 130.49, 129.15, 128.48, 124.82, 118.31, 117.58, 114.27, 31.66, 29.82, 20.35. HR-ESI-MS found: 529.25972; $[C_{34}H_{32}N_4O_2+H]^+$ requires 529.25980 *m*/z.

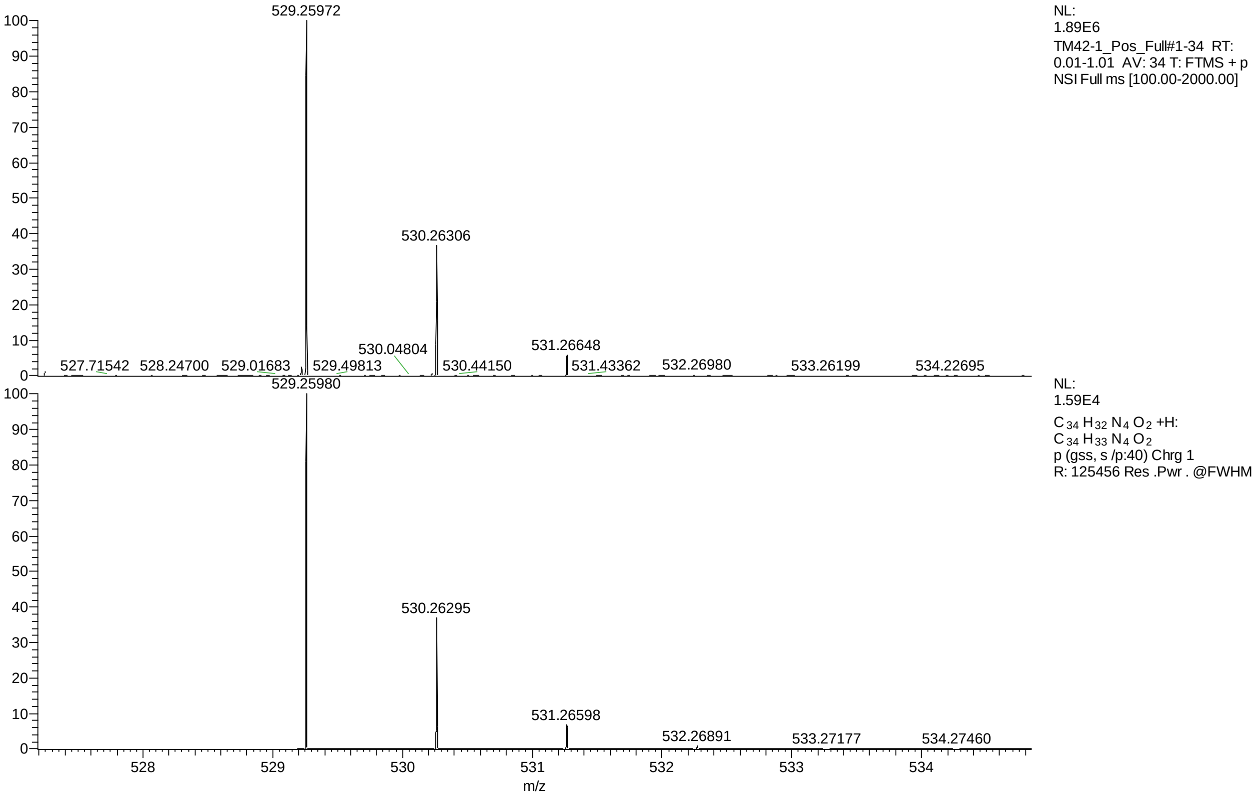
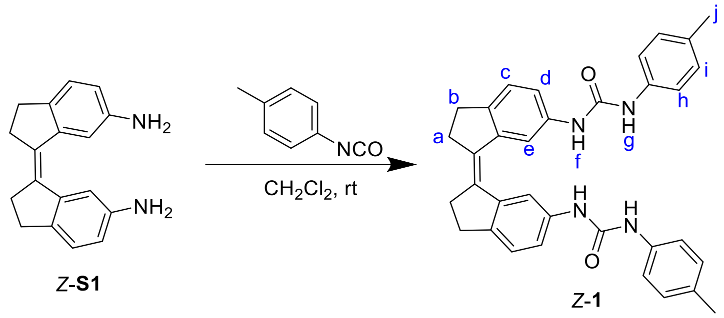


Figure S1.High resolution mass spectrum of $[E\text{-}\mathbf{1}+H]^+$



Compound Z-**S1** (100 mg, 0.38 mmol) was dissolved in 5 mL of dry CH_2Cl_2, and *p*-tolylisocyanate (101 mg, 95 µL, 0.76 mmol) was added by syringe. The reaction mixture was stirred at rt overnight and the precipitate was filtered off to obtain crude *Z*-**1**. The crude was washed with CH_2Cl_2 and aq. HCl (1M), suspended in water and sonicated, and the precipitate was collected by filtration to give *Z*-**1** as a fine beige solid (152 mg, 75%).

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.59 (s, 2H, H^g), 8.43 (s, 2H, H^f), 8.11 (s, 2H, H^e), 7.42 (dd, J = 8.2, 1.9 Hz, 2H, H^d), 7.29 (d, J = 8.0 Hz, 4H, H^h), 7.22 (d, J = 8.2 Hz, 2H, H^c), 6.88 (d, J = 8.0 Hz, 4H, H^i), 2.89 (m, 4H, H^b), 2.77 (m, 4H, H^a), 2.11 (s, 6H, H^j). ^{13}C NMR (151 MHz, DMSO-d_6) δ 152.71, 141.27, 140.05, 137.59, 137.28, 134.86, 130.20, 129.04, 125.10, 118.14, 117.71, 113.62, 34.45, 29.39, 20.22.

S3.2 ^{1}H NMR spectra of *E*-1

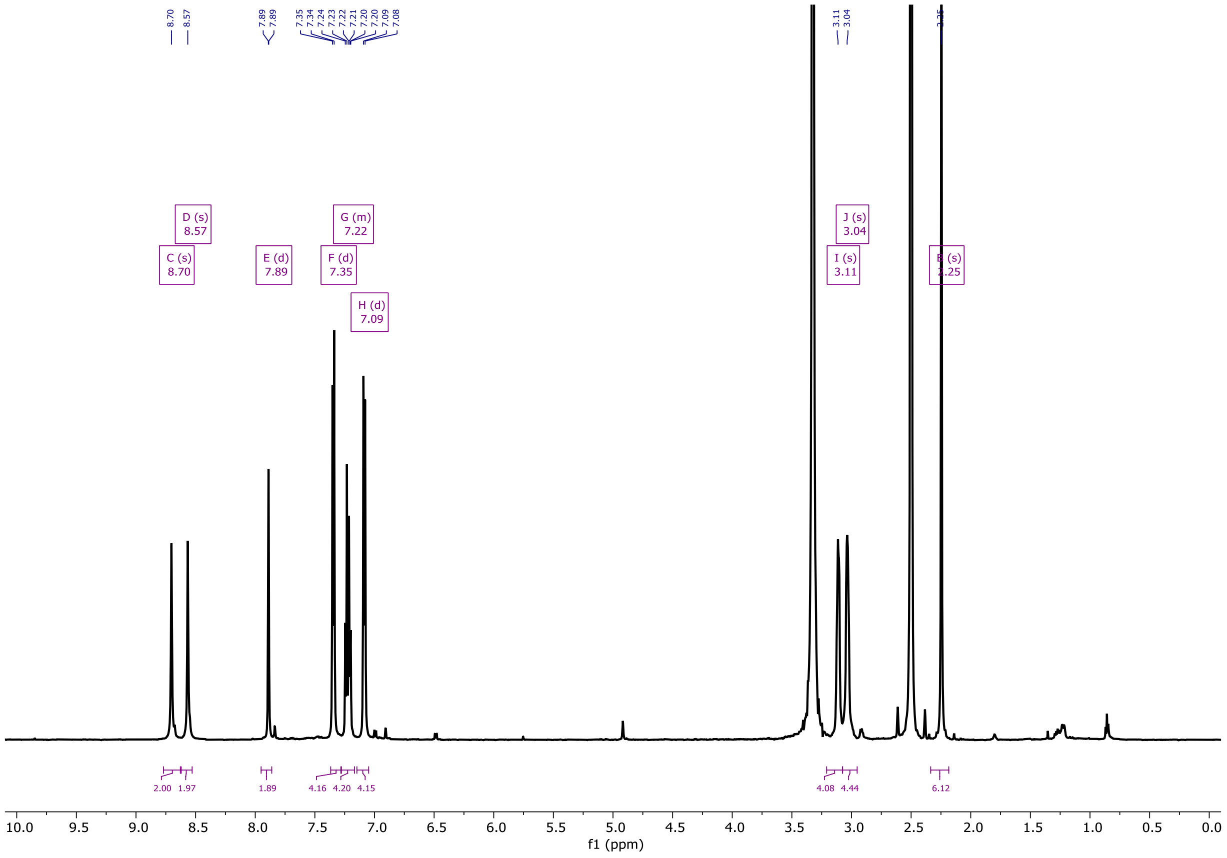


Figure S2. ^{1}H NMR spectrum of *E*-**1** (600 MHz, DMSO-d_6).

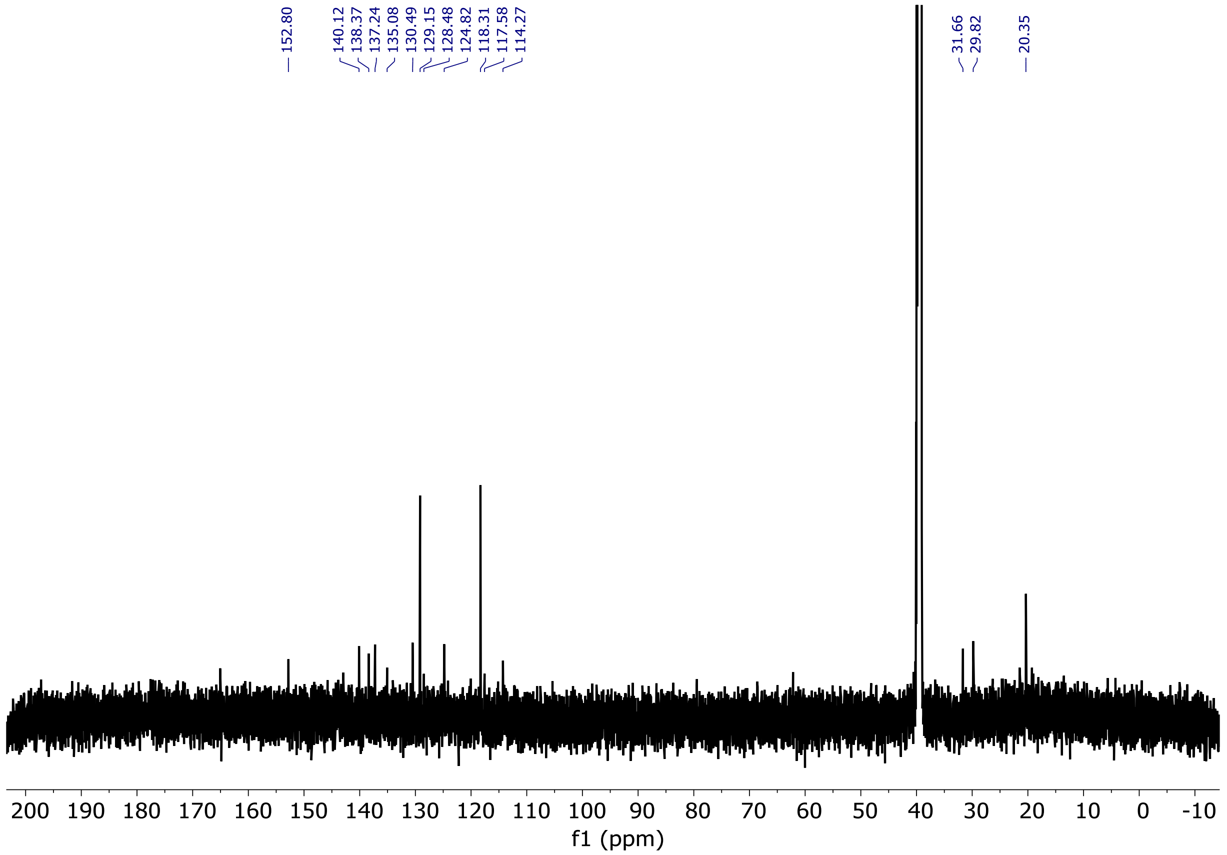


Figure S3. $^{13}C\{^1H\}$ NMR spectrum of *E*-**1** (125 MHz, DMSO-d_6). Low signal-to-noise due to poor solubility.

S3.3 ^{1}H NMR spectra of *Z*-1

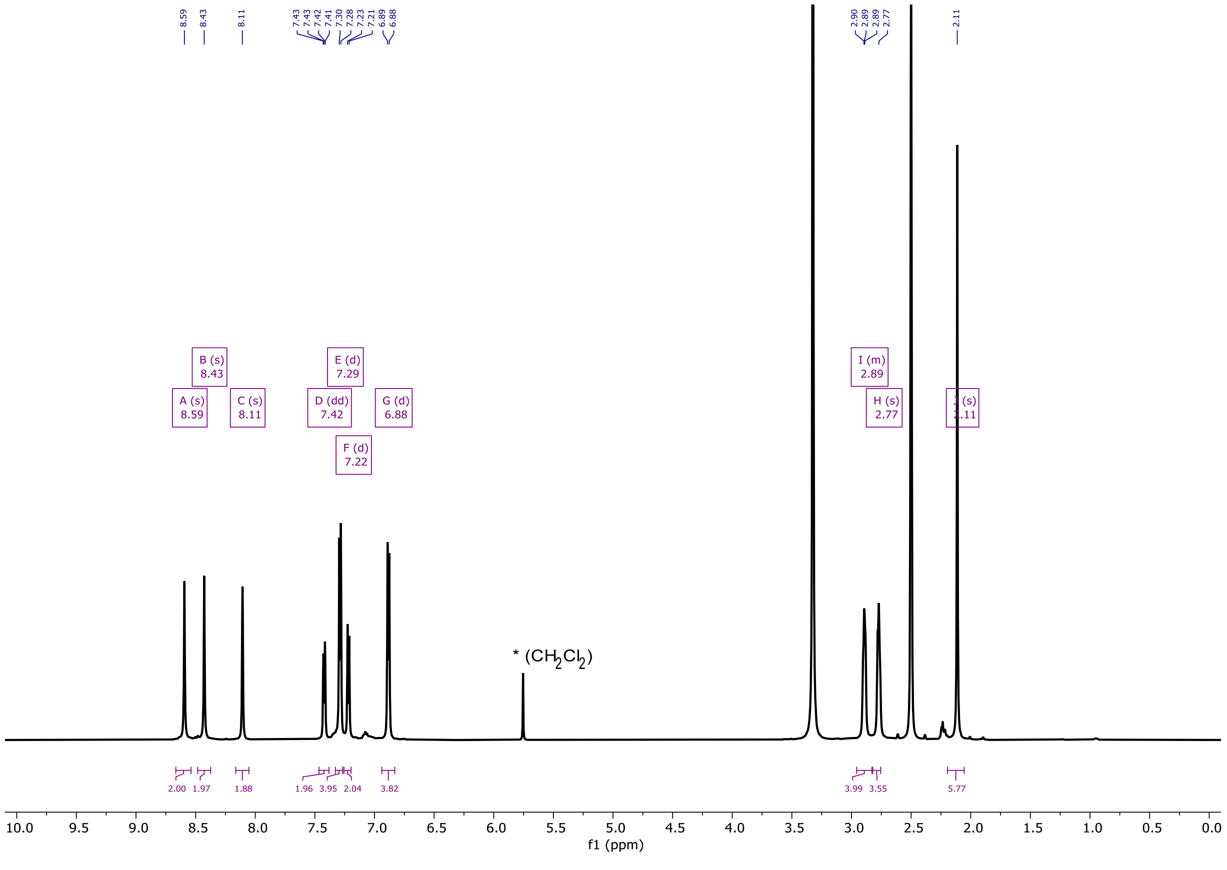


Figure S4. ^{1}H NMR spectrum of *Z*-**1** (600 MHz, DMSO-d_6).

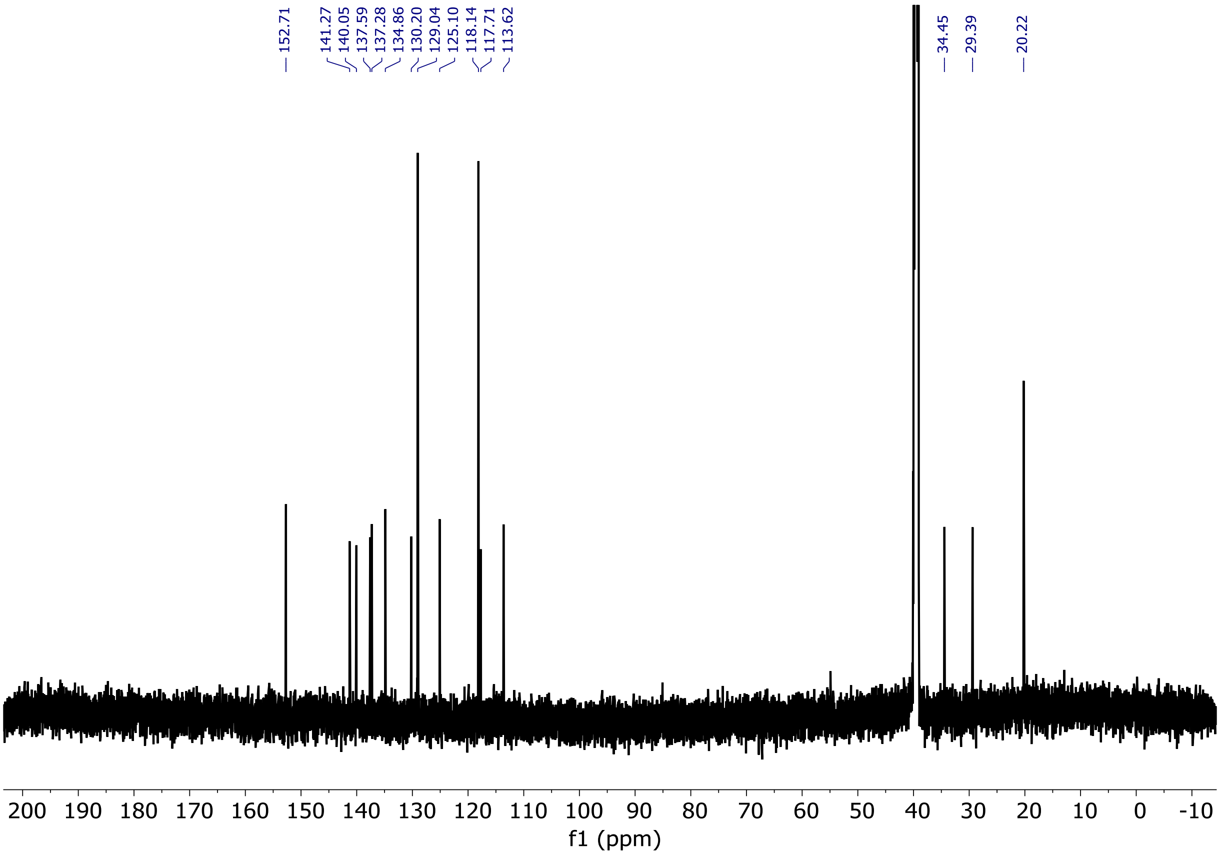


Figure S5. $^{13}C\{^1H\}$ NMR spectrum of *Z*-**1** (125 MHz, DMSO-*d*$_6$).

S4. Viscosity measurements of tetrabutylammonium dihydrogen phosphate solutions

Viscosities and densities were measured at 25 °C using a Lovis 2000M rolling ball viscometer (Ø1.59 mm capillary; 1 mm stainless steel ball), with results shown in Table S3.

Table S3. Viscosity measurements of solutions of $[NBu_4][H_2PO_4]$ in DMSO-*d*$_6$ with 0.5% v/v added water.

$[NBu_4][H_2PO_4]$ / mM	**Density / g·cm^{-3}**	**Temperature / °C/**	**Dyn. Viscosity / mPa·s**	**Kin. Viscosity / mm²·s^{-1}**	**Temperature / °C**	**Error / %**
0	1.1833	25.04	2.149	1.816	25.00	0.02
5	1.1833	25.00	2.210	1.868	25.00	0.08
20	1.1828	25.08	2.259	1.910	25.00	0.04
50	1.1820	25.04	2.334	1.974	24.99	0.03
100	1.1805	25.06	2.499	2.117	25.00	0.01
300	1.1752	25.04	3.251	2.767	25.00	0.02

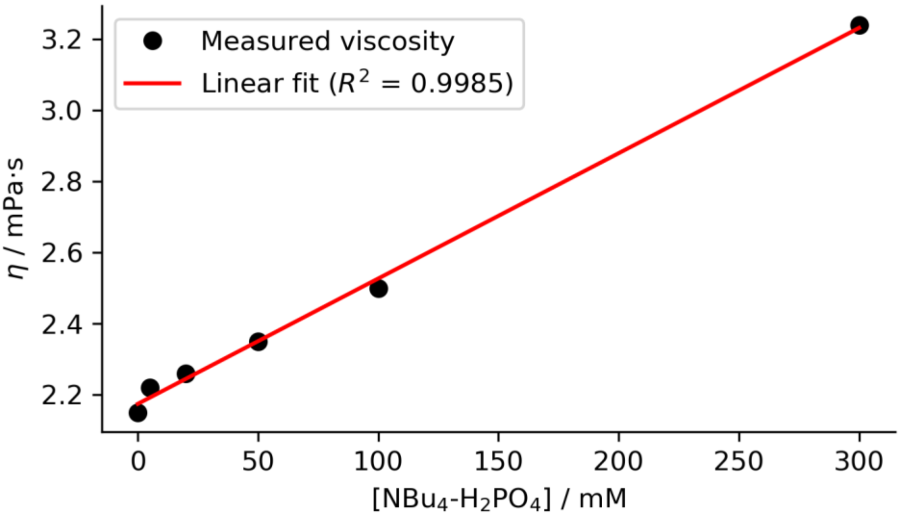


Figure S6. Concentration dependence of the measured viscosity of tetrabutylammonium dihydrogen phosphate solutions in DMSO-*d*$_6$ with 0.5% added water. Linear fit: $\eta = 2.17 + (3.53 \times 10^{-3})([NBu_4][H_2PO_4])$. For calibrating diffusion measurements, it is convenient to express this in relative form: $\eta/\eta_0 = 1 + 1.63 \times [NBu_4][H_2PO_4]$ / mM]

S5. Mathematical modelling of isodesmic self-association

The association of dihydrogen phosphate was fitted to an isodesmic model,[4] constructed as follows.

S5.1 Isodesmic model for self-association

Consider a monomer A capable of forming an oligomer of arbitrary length, where each association step $\mathrm{A}_{n-1} + \mathrm{A} \rightarrow \mathrm{A_n}$ occurs with the same (isodesmic) association constant K_i:

$$K_i = \frac{[\mathrm{A_n}]}{[\mathrm{A}][\mathrm{A_{n-1}}]} \tag{1}$$

And the total concentration of A is given as:

$$[\mathrm{A}]_0 = \sum_{n=1}^{\infty} n[\mathrm{A_n}] \tag{2}$$

It is convenient to define $z = [\mathrm{A}]K_i$. Then:

$$\begin{aligned}[\mathrm{A_n}] &= K_i[\mathrm{A}][\mathrm{A_{n-1}}] \\ &= z[\mathrm{A_{n-1}}]\end{aligned}$$

As $[\mathrm{A_1}] = [\mathrm{A}]$, by induction:

$$\begin{aligned}[\mathrm{A}_n] &= [\mathrm{A}]z^{n-1} \\ &= \frac{z^n}{K_i}\end{aligned} \tag{3}$$

We can combine (2) and (3) to build a mass balance equation for $[\mathrm{A}]_0$:

$$\begin{aligned}[\mathrm{A}]_0 &= \sum_{n=1}^{\infty} n[\mathrm{A_n}] \\ &= \frac{1}{K_i}\sum_{n=1}^{\infty} nz^n \\ &= \frac{z}{K_i(z-1)^2}\end{aligned}$$

Allowing us to express z as the root of a quadratic:

$$[\mathrm{A}]_0 K_i (z-1)^2 = z$$

$$[\mathrm{A}]_0 K_i z^2 - (2[\mathrm{A}]_0 K_i + 1)z + [\mathrm{A}]_0 K_i = 0$$

Defining $L = [\mathrm{A}]_0 K_i$ and assuming that z is the smallest solution to the quadratic:

$$z = \frac{2L + 1 - \sqrt{4L+1}}{2L} \tag{4}$$

From z we can obtain the concentration of monomeric (unbound) species $[\mathrm{A}] = \frac{z}{K_i}$, and by combining equations (4) and (3) we can calculate the size distribution of oligomers from the starting concentration $[\mathrm{A}]_0$ and association constant K_i.

One other relationship may be useful. If we assume that *only the terminal sites* of each oligomer are available for binding by a host (ie each monomeric or oligomeric guest species has two binding sites, regardless of size), then the guest concentration of guest binding sites in solution will be:

$$\begin{aligned}[\mathrm{A}_{eff}] &= \sum_{n=1}^{\infty} [\mathrm{A}_n] \\ &= \frac{1}{K_i}\sum_{n=1}^{\infty} z^n \\ &= \frac{z}{K_i(1-z)}\end{aligned} \tag{5}$$

S5.2 Isodesmic model for diffusion

The Stokes-Einstein-Sutherland equation for molecular diffusion states that for hard spherical particles[5]:

$$D = \frac{k_B T}{6\pi\eta R_H} \tag{6}$$

For the sake of simplicity, we can approximate each monomer as a hard sphere of volume V and each n-unit oligomer as a hard sphere of volume nV. This assumption of perfect packing of hard spheres to form a larger hard sphere is clearly unphysical, but the results of this approximation have previously been shown to be quite close to those obtained from more detailed treatments.[6]

The diffusion coefficient D_n of an n-unit oligomer can then be expressed as

$$D_n = n^{-\frac{1}{3}} D_0 \tag{7}$$

where D_0 is the diffusion coefficient of the monomeric species (*i.e.* at infinite dilution). For system of oligomers in fast exchange on the NMR timescale, the measured diffusion coefficient $\bar{D}$ is a concentration-weighted sum of the diffusion coefficients of all possible oligomers†:

$$\bar{D} = \sum_{n=1}^{\infty} x_n D_n \tag{8}$$

Where x_n is the molfraction of the oligomerizing species present in an oligomer of size n, ie $x_n = \frac{n[\mathrm{A_n}]}{[\mathrm{A}]_0}$

Using the previously obtained expressions (7) and (3) for $D_n$and $[\mathrm{A_n}]$ respectively, we can then say:

$$\bar{D} = \frac{D_0}{[\mathrm{A}]_0 K_i} \sum_{n=1}^{\infty} n^{\frac{2}{3}} z^n \tag{9}$$

There is no closed-form analytic solution to this infinite sum, but for numerical modelling and curve-fitting it is convenient to restate it as a polylogarithm. The polylogarithm is a special function defined as

$$\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}$$

allowing us to express $\bar{D}$ as:

$$\bar{D} = \frac{D_0}{[\mathrm{A}]_0 K_i} \mathrm{Li}_{-\frac{2}{3}}([\mathrm{A}] K_i) \tag{10}$$

Some results of this model for the concentration-dependent diffusion of isodesmically self-associating species are plotted in Figure S7.

---

† This treatment neglects potential differences in longitudinal relaxation times T_1 across oligomers of different sizes or for the different sites present within an oligomer.

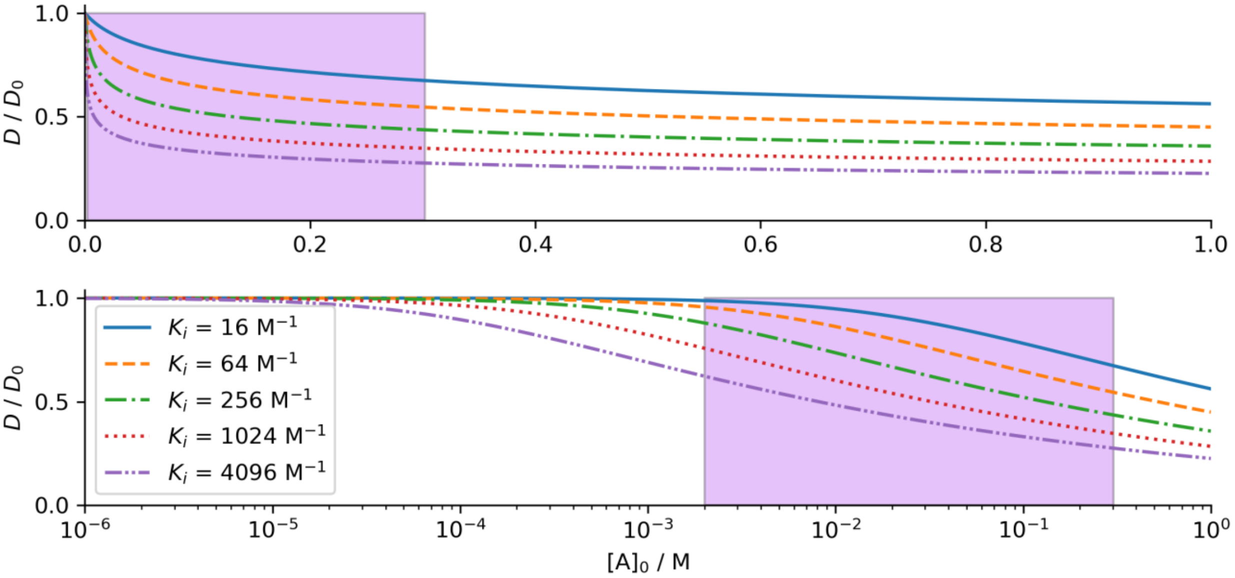


Figure S7. Modelled changes in relative diffusion coefficients *D/D_0* under conditions of isodesmic self-association (Equation 10), for five different association constants *K_i* over the concentration range 1 μM to 1 M. The 2 – 300 mM concentration range studied here has been shaded in purple. Both plots show the same curves, against a linear (top) and log (bottom) concentration scale.

S5.3 Estimated diffusion of host-guest complexes

Consider a supramolecular complex made up from *m* host molecules and *n* guest molecules, of the form $[H_mG_n]$, where the diffusion coefficients of free host and free guest have been obtained as D_H and D_G. Under the same approximation of perfectly packing hard spheres used previously, we can then approximate the diffusion coefficient of the complex as:

$$D(m,n) = (mD_h^{-3} + nD_G^{-3})^{-\frac{1}{3}} \tag{11}$$

As the hosts and guests will not pack perfectly into hard spherical complexes, the diffusion coefficients of complexes estimated by this method can be considered upper bounds for the actual diffusion coefficients.

Assuming fast exchange of species on the NMR timescale, the diffusion coefficient measured for a host species is a population-weighted average of all H_mG_n species present:

$$\overline{D_H} = \frac{1}{[H]_0} \sum_{m=1}^{\infty} \sum_{n=0}^{\infty} m \cdot D(m,n)[\mathrm{H}_m\mathrm{G}_\mathrm{n}] \tag{12}$$

For the case of discrete complexes involving a single host and up to two guests, this can be expanded to

$$\overline{D_H} = \frac{1}{[H]_0}\left([H]D_H + [HG](D_H^{-3} + D_G^{-3})^{-\frac{1}{3}} + [HG_2](D_H^{-3} + 2D_G^{-3})^{-\frac{1}{3}}\right) \tag{13}$$

where $[H]$, $[HG]$, and $[HG_2]$ can be calculated using the total concentrations of host and guest and the association constants K_1 and K_2.[7] If the guest self-associates using isodesmic kinetics (such as $H_2PO_4^-$), then D_G can be replaced with the concentration-dependent diffusion model given in Equation 10.

S6. Experimental determination of the thermodynamics of $H_2PO_4^-$ self-association

Solutions of tetrabutylammonium dihydrogen phosphate ([NBu_4][H_2PO_4]) were prepared in DMSO-d_6 with 0.5% v/v H_2O added (Millipore). 500 µL aliquots were transferred to standard 5 mm NMR tubes (Wilmad) and ^{1}H and ^{31}P diffusion experiments were obtained and processed as previously described (S1 General Experimental). Diffusion coefficients were then obtained by nonlinear regression of $I = I_0 \exp[-bD]$, where b is the effective attenuation strength of the experiment as calculated by the Bruker **diff5** utility ($b = g^2\delta^2\gamma^2\left(\Delta - \frac{\delta}{3}\right)$ for square gradient pulses).

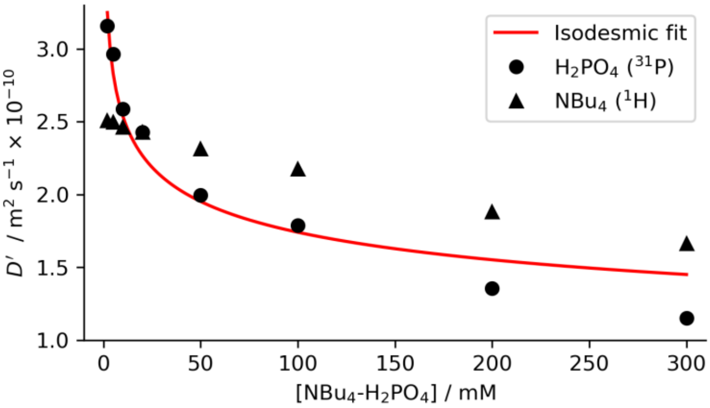


Figure S8. Measured diffusion coefficients of tetrabutylammonium and dihydrogen phosphate uncorrected for changes in viscosity. The model for isodesmic diffusion shown in Eq (10) has been regressed unsuccessfully onto measured D(H_2PO_4), with obtained parameters K_i = (2.5 ± 9) × 10^3 M^{-1} and D_0 = (4.9 ± 2.7) × 10^{-10} m^2 s^{-1}.

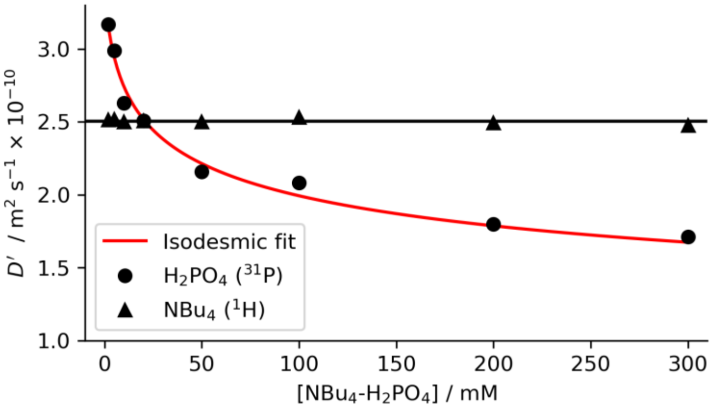


Figure S9. Measured diffusion coefficients of tetrabutylammonium and dihydrogen phosphate corrected for changes in viscosity using the measured viscosities of these solutions (see Table S3; Figure 1 in manuscript). The model for isodesmic diffusion shown in Eq (10) has been regressed onto measured D(H_2PO_4), with obtained parameters K_i = (120 ± 32) M^{-1} and D_0 = (3.39 ± 0.11) × 10^{-10} m^2 s^{-1}. After correction, no change in D(NBu_4) is observed (horizontal black line).

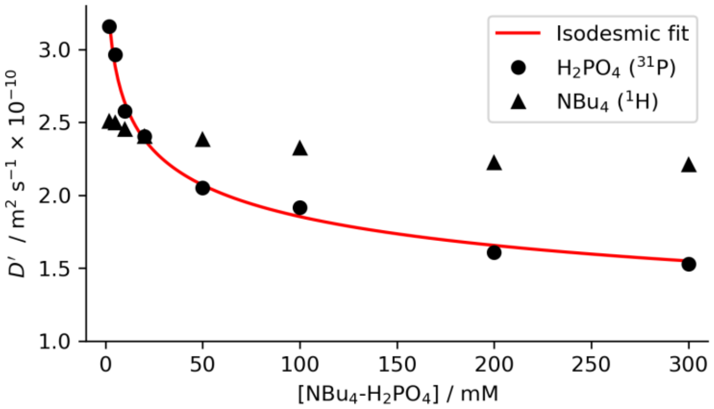


Figure S10. Measured diffusion coefficients of tetrabutylammonium and dihydrogen phosphate corrected for changes in viscosity using the change in the measured diffusion coefficient of the DMSO-d_6 solvent residual. The model for isodesmic diffusion shown in Eq (10) has been regressed onto measured $D(H_2PO_4)$, with obtained parameters K_i = (333 ± 110) M^{-1} and D_0 = (3.71 ± 0.16) × 10^{-10} m^2 s^{-1}.

Table S4. Tabulated data for dihydrogen phosphate isodesmic self-association data.

Correction	K_i / M^{-1}	Relative error (K_i)	D_0 / m^2s^{-1}	Relative error (D_0)
Uncorrected[a]	(2.5 ± 9) × 10^3	360%	(4.9 ± 2.7) × 10^{-10}	55%
Indirect correction[b]	(3.33 ± 1.10) × 10^2	33%	(3.71 ± 0.16) × 10^{-10}	4.3%
Direct correction[c]	(1.20 ± 0.32) × 10^2	27%	(3.39 ± 0.11) × 10^{-10}	3.2%

[a] Isodesmic fit to diffusion data as-measured; [b] Isodesmic fit to diffusion data corrected for viscosity changes by normalising *D* for the DMSO solvent residual; [c] Isodesmic fit to diffusion data corrected for viscosity changes using independently measured viscosity data.

Viscosity changes must be corrected for to allow successful fitting of the isodesmic diffusion model with errors smaller than the fitted value themselves (Table S4). Direct corrections using the change in viscosity measured independently by falling-ball viscometry (SI-S4) gave slightly smaller relative errors than an indirect correction approach using the change in measured *D* for the DMSO-d_5 solvent residual peak. These parameters can be used to calculate the size distribution of oligomeric $H_2PO_4^-$ at arbitrary concentrations, as shown in Figure S11.

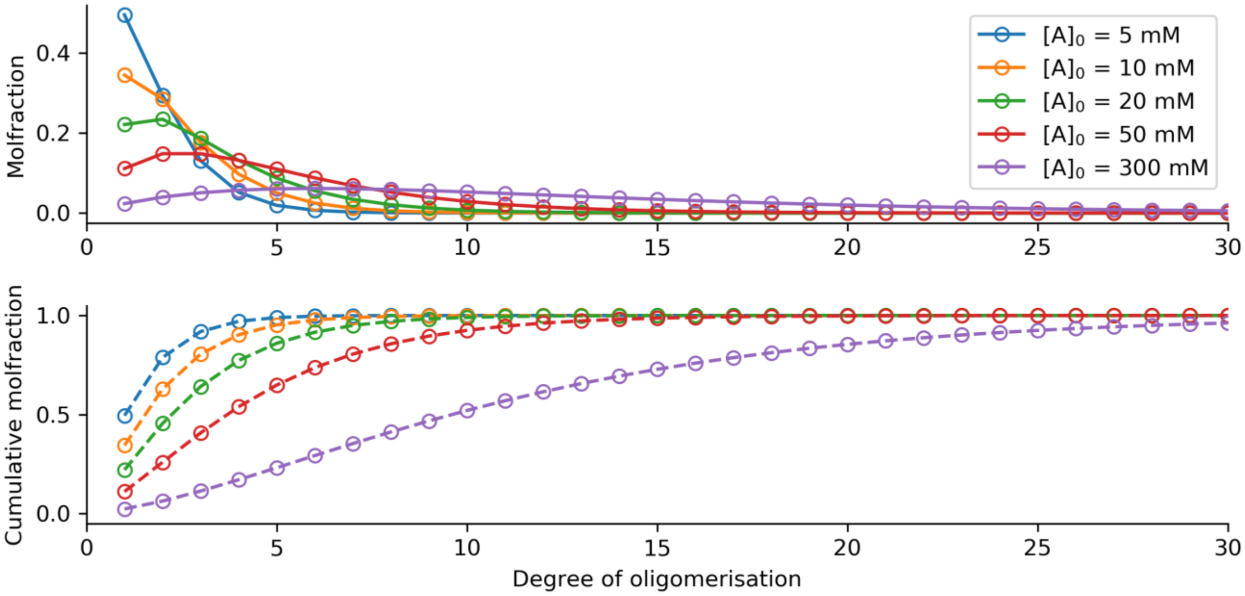


Figure S11. Visualization of the size distribution of phosphate oligomers using experimentally obtained parameters. Top: distribution of oligomers (n = 1 to 30) for solutions of $[A]_0$ = 5, 10, 20, 50, and 300 mM with isodesmic K_i = 120 M^{-1} (as experimentally obtained for $[NBu_4][H_2PO_4]$ in DMSO-d_6). Bottom: cumulative distribution functions under the same conditions. For the 50 mM solutions of $[NBu_4][H_2PO_4]$ used repeatedly in this work (red) the median phosphate oligomer contains n = 4 units, the 90th percentile contains n = 10 units, and the 99th percentile oligomer contains 16 units.

S6.1 Control experiments with tetrabutylammonium acetate

Due to the large quantities of DMSO-d_6 required for falling-ball viscometry, no direct viscosities were measured for solutions of tetrabutylammonium acetate ([NBu$_4$][OAc]). Instead, an indirect viscosity calibration was performed from the measured diffusion coefficient of the DMSO-d$_6$ solvent residual, under the assumption that changes in measured *D* for the solvent residual would be primarily a result of changing viscosity.

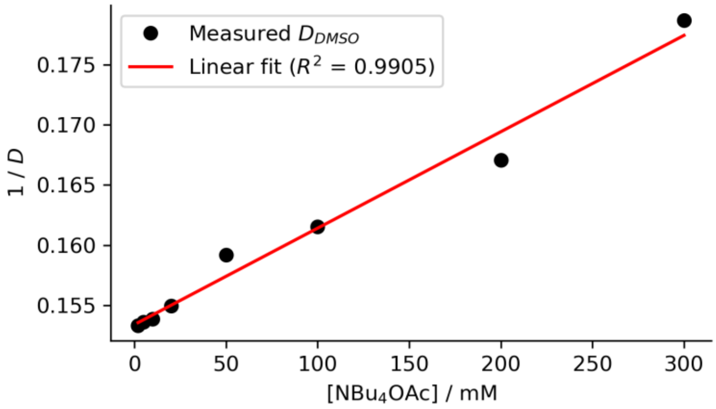


Figure S12. Indirect viscosity calibration for tetrabutylammonium acetate, using the diffusion coefficient of the DMSO solvent residual peak. Linear fit for implied relative change in dynamic viscosity: $\frac{\eta}{\eta_0} = 1 + (0.521 \times 10^{-3})[\mathrm{NBu_4OAc}]$.

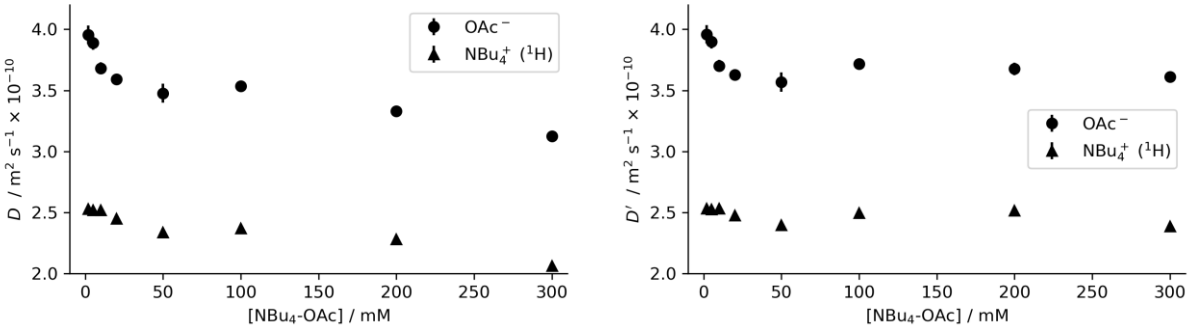


Figure S13. Measured diffusion coefficients of tetrabutylammonium acetate over the concentration range 0 – 300 mM without (left) and with (right) correction for changes in viscosity (see Figure S12). While some acetate self-association appears to occur, the <10% decrease in the measured diffusion coefficient is much smaller than that seen for dihydrogen phosphate.

Table S5. Measured diffusion coefficients for DMSO, tetrabutylammonium, and acetate over a range of tetrabutylammonium acetate concentrations. No correction for viscosity has been applied to these data. ^1H, 500 MHz, DMSO-d_6 with 0.5% added H_2O.

$[NBu_4][OAc]$ / mM	**D_{DMSO}[a] / $m^2 s^{-1} \times 10^{-10}$**	**$D(OAc)$[a] / $m^2 s^{-1} \times 10^{-10}$**	**$D(NBu_4)$ [a] / $m^2 s^{-1} \times 10^{-10}$**
2	6.52 ± 0.01	3.96 ± 0.08	2.53 ± 0.01
5	6.51 ± 0.01	3.89 ± 0.05	2.52 ± 0.01
10	6.50 ± 0.02	3.69 ± 0.05	2.52 ± 0.01
20	6.45 ± 0.02	3.59 ± 0.04	2.45 ± 0.01
50	6.28 ± 0.05	3.48 ± 0.08	2.33 ± 0.02
100	6.19 ± 0.02	3.54 ± 0.04	2.38 ± 0.01
200	5.98 ± 0.03	3.33 ± 0.05	2.27 ± 0.01
300	5.60 ± 0.01	3.12 ± 0.03	2.07 ± 0.01

[a] 500 MHz ^1H PGSTE, δ = 4 ms, Δ = 50 ms, $g = 0 – 53.45$ G cm^{-1}

Table S6. Changes in diffusion coefficients for solutions of 2–300 mM tetrabutylammonium dihydrogen phosphate ($[NBu_4][H_2PO_4]$; $D(H_2PO_4)$) and tetrabutylammonium acetate (OAc; D_{OAc}), corrected for change in viscosity.

$[NBu_4]$[Anion] / mM	**$D(H_2PO_4)$ [b] / $m^2 s^{-1} \times 10^{-10}$**	**$D(NBu_4)$[c] (H_2PO_4) / $m^2 s^{-1} \times 10^{-10}$**	**$D(OAc)$[b] / $m^2 s^{-1} \times 10^{-10}$**	**$D(NBu_4)$[b] (OAc) / $m^2 s^{-1} \times 10^{-10}$**
2	3.17 ± 0.28	2.52 ± 0.01	3.96 ± 0.1	2.53 ± 0.01
5	2.99 ± 0.17	2.52 ± 0.01	3.90 ± 0.08	2.53 ± 0.01
10	2.63 ± 0.08	2.50 ± 0.01	3.70 ± 0.07	2.53 ± 0.01
20	2.51 ± 0.10	2.51 ± 0.02	3.63 ± 0.06	2.48 ± 0.01
50	2.16 ± 0.03	2.50 ± 0.02	3.57 ± 0.1	2.40 ± 0.03
100	2.08 ± 0.02	2.53 ± 0.02	3.72 ± 0.05	2.49 ± 0.01
200	1.80 ± 0.01	2.49 ± 0.01	3.68 ± 0.07	2.52 ± 0.01
300	1.71 ± 0.02	2.48 ± 0.03	3.61 ± 0.05	2.39 ± 0.01

[a] At concentrations below 20 mM the $H_2PO_4^-$ anion (MW = 97) diffuses more rapidly than the $[NBu_4]^+$cation (MW = 242), but at concentrations higher than 20 mM the diffusion coefficient of $H_2PO_4^-$is lower than that of $[NBu_4]^+$. In contrast, the diffusion coefficient of the acetate anion drops slightly until 50 mM concentration, after which it does not lower any further. Measurements compensated for changes in viscosity using the measured viscosity ($[NBu_4][H_2PO_4]$), or the measured diffusion of residual DMSO ($[NBu_4][OAc]$) as an internal standard. DMSO-d_6 with 0.5% added H_2O, 500 MHz ^1H. [b] ^{31}P PGSTE, δ = 7 ms, Δ = 100 ms, $g = 0 – 53.45$ G cm^{-1}. [c] ^1H PGSTE, δ = 4 ms, Δ = 50 ms, $g = 0 – 53.45$ G cm^{-1}.

S7. Host-guest association data

S7.1 Method

Dilution-corrected NMR titrations were performed by adding aliquots of a solution of guest containing 5 mM host to 5 mM solutions of host (in DMSO-d_6/0.5%H_2O. The spectra obtained were processed and analysed using MestReNova software to obtain chemical shift data, which were used to obtain association constants using the processing routines given in Script S1 (see section S11). In all cases, the data was fitted to 1:2 H:G binding models[7] giving K_1 and K_2, which are defined as:

$$K_1 = \frac{[\mathrm{HG}]}{[\mathrm{H}][\mathrm{G}]}$$

and

$$K_2 = \frac{[\mathrm{HG_2}]}{[\mathrm{HG}][\mathrm{G}]}$$

with cooperativity α defined as

$$\alpha = \frac{K_1}{4K_2}$$

For the *E*-host, the urea binding sites were treated as equal, i.e. a non-cooperative (α = 1) model was used. For the *Z*-host, a cooperative model with a floating α was used as two urea groups may be involved in binding one anion.

S7.2 Addition of tetrabutylammonium dihydrogen phosphate to *E*-**1**

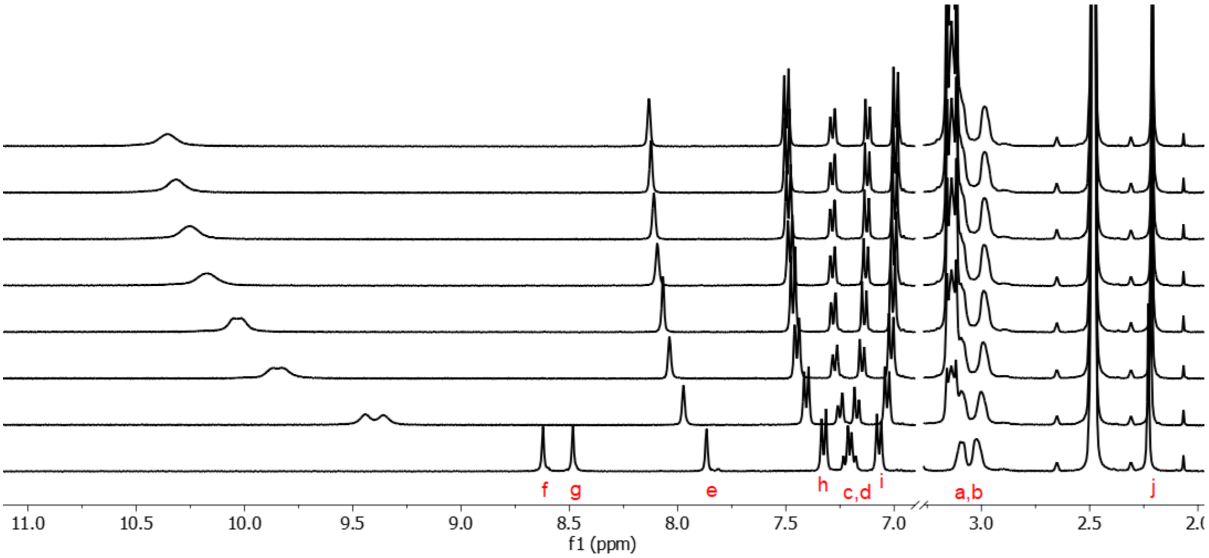


Figure S14. Titration of 50 mM [NBu_4][H_2PO_4] into 5 mM of *E*-**1**. 400 MHz, DMSO-d_6 + 0.5% H_2O.

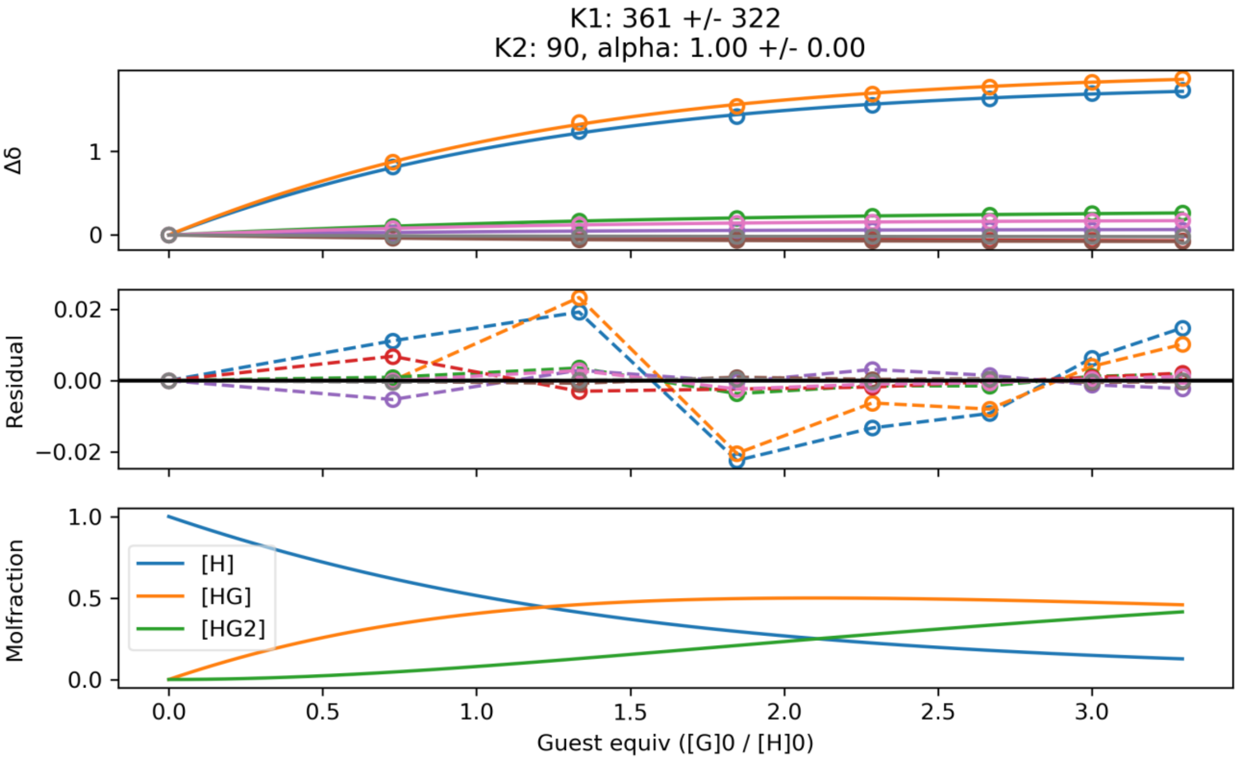


Figure S15. Data from Figure S14 fitted to a **non-cooperative** 1:2 host:guest association model. K_1 = 3.6 × 10^2 M^{-1}, K_2 = 9.0 × 10^1 M^{-1}.

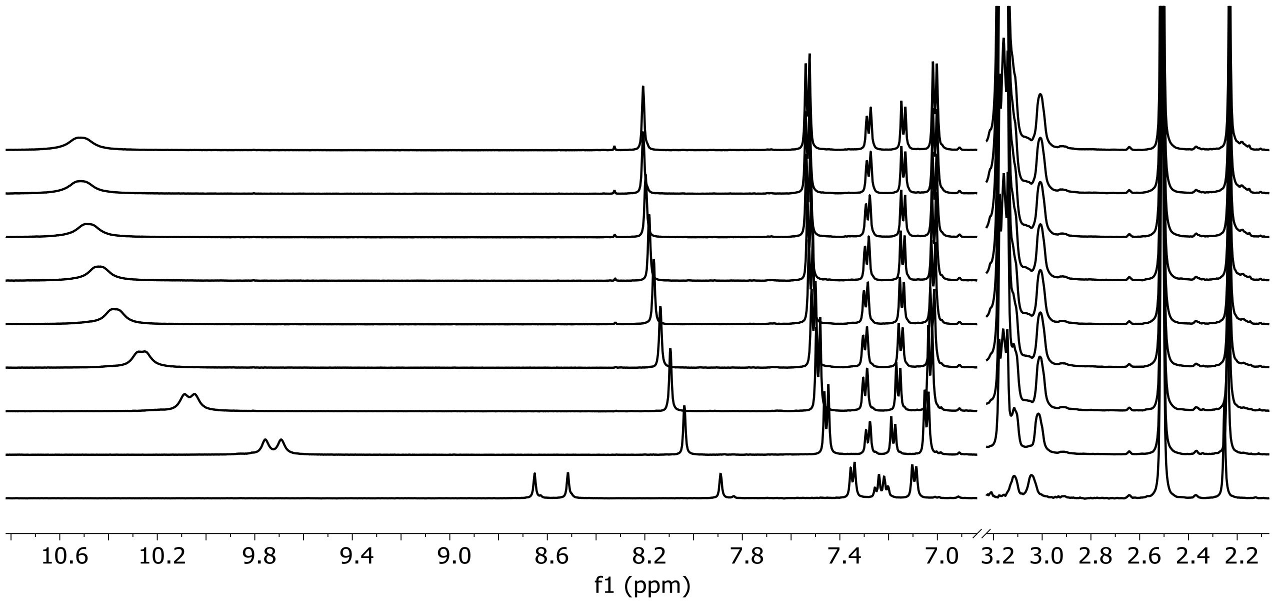


Figure S16. Titration of 50 mM $[NBu_4][H_2PO_4]$ into 5 mM of *E*-**1**. 500 MHz, DMSO-d_6 + 0.5% H_2O.

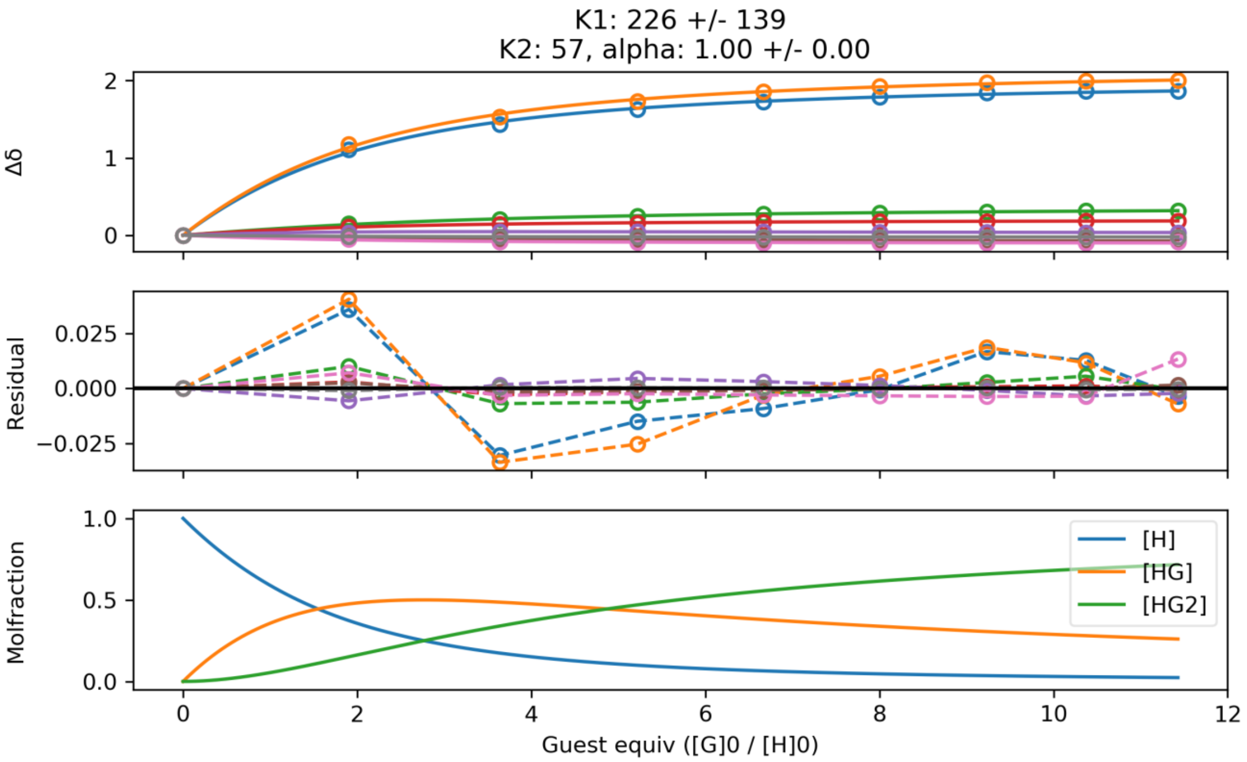


Figure S17. Data from Figure S16 fitted to a **non-cooperative** 1:2 host:guest association model. $K_1 = 2.3 \times 10^2$, M^{-1}, $K_2 = 6.0 \times 10^1$ M^{-1}.

S7.3 Addition of tetrabutylammonium dihydrogen phosphate to *Z*-1 (up to 3 equiv.)

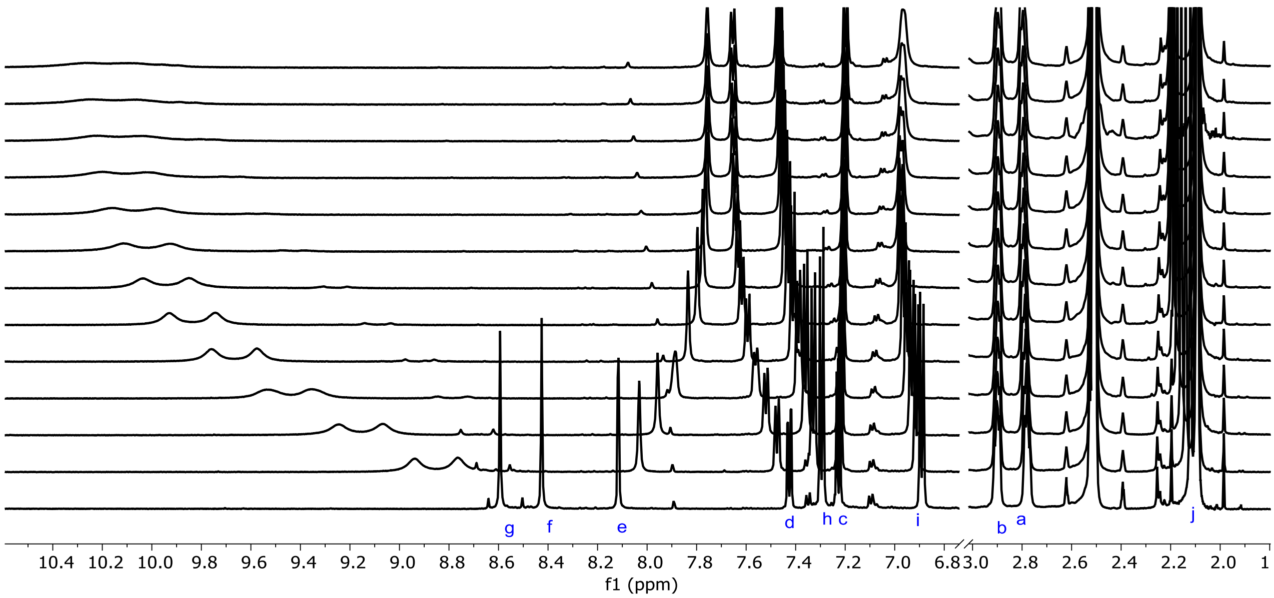


Figure S18. Titration of 50 mM $[NBu_4][H_2PO_4]$ into 5 mM of *Z*-**1**, to 9.7 mM guest concentration. 600 MHz, DMSO-d_6 + 0.5% H_2O.

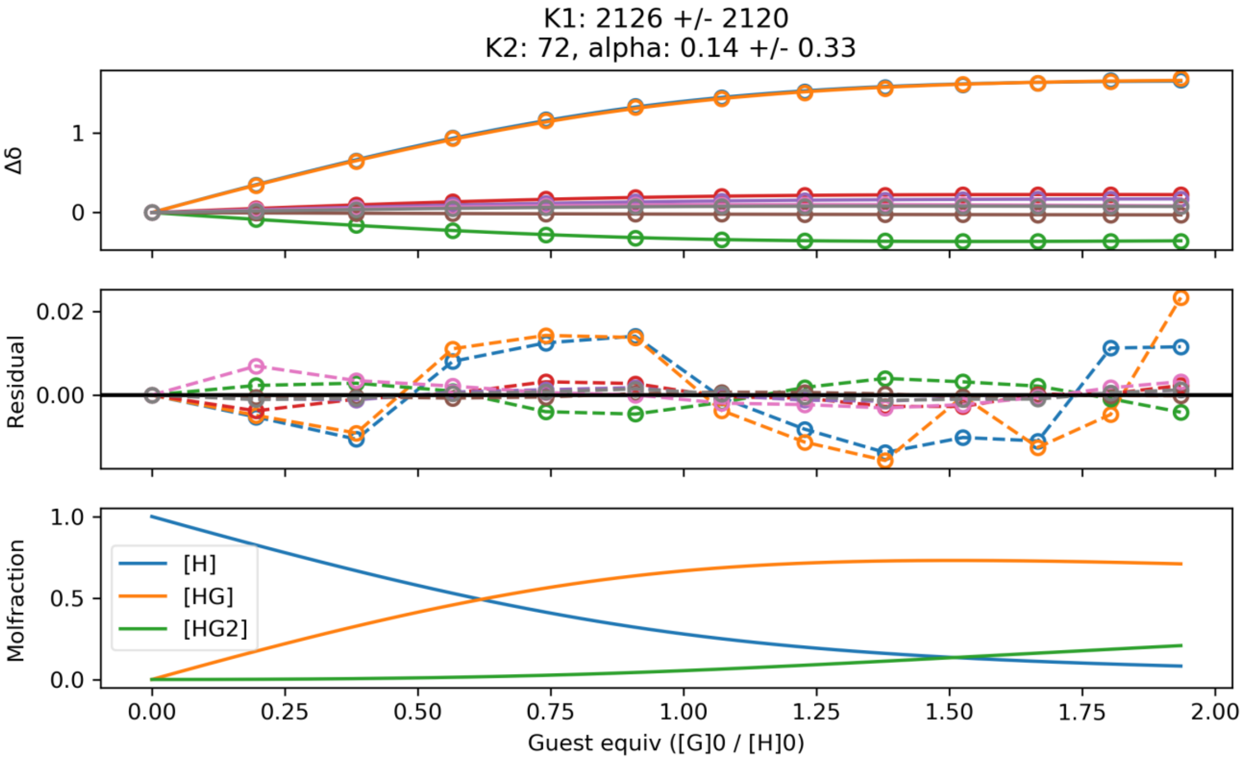


Figure S19.Data from Figure S18 fitted to a 1:2 host:guest association model. $K_1 = 2.1 \times 10^3\ M^{-1}$, $K_2 = 7.0 \times 10^1\ M^{-1}$.

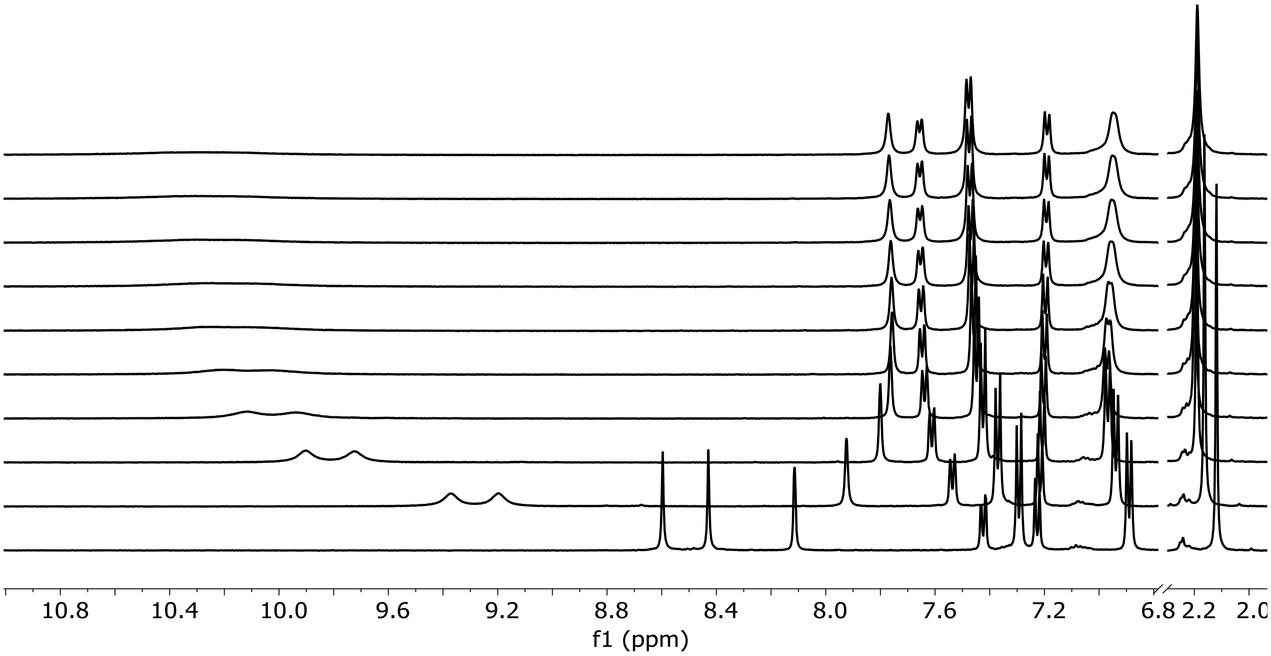


Figure S20. Titration of 50 mM $[NBu_4][H_2PO_4]$ into 5 mM of *Z*-**1** to 15.5 mM guest concentration. 500 MHz, DMSO-d_6 + 0.5% H_2O.

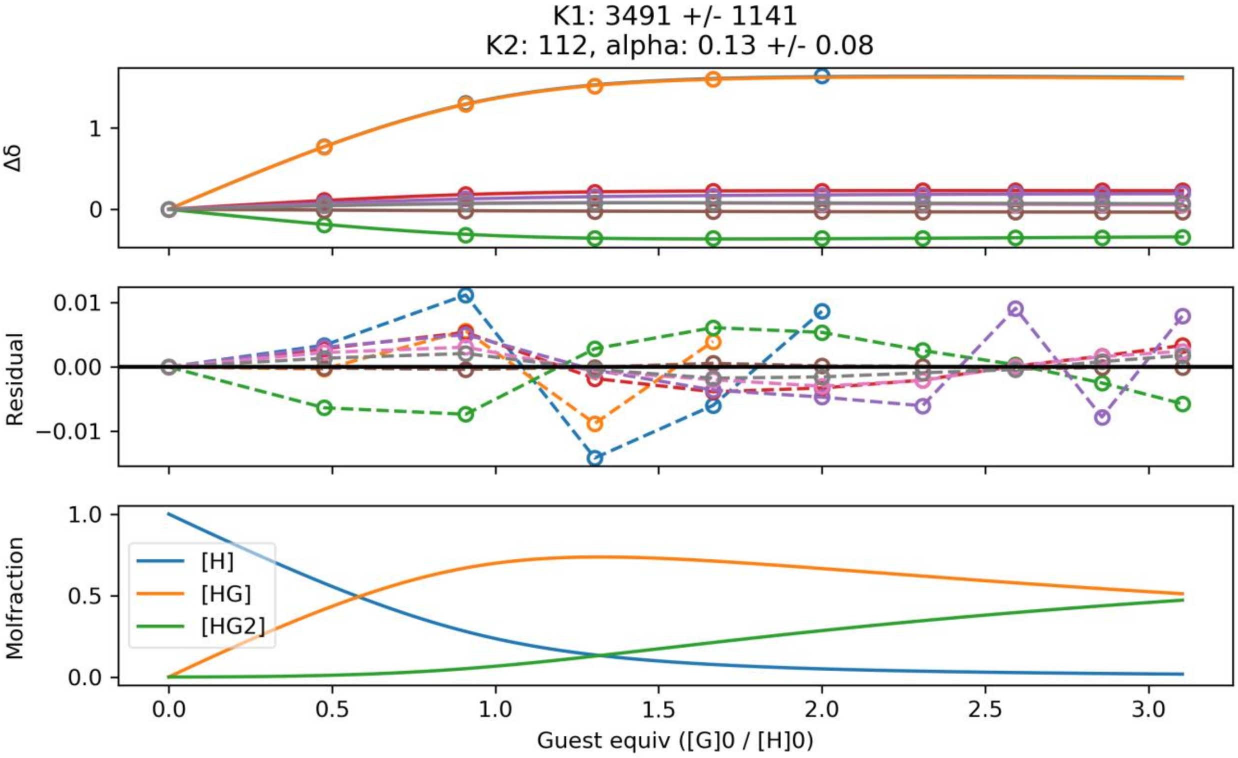


Figure S21. Data from Figure S20 fitted to a 1:2 host:guest association model. $K_1 = 3.5 \times 10^3$ M^{-1}, $K_2 = 1.1 \times 10^2$ M^{-1}.

S7.4 *Z*-1 dihydrogen phosphate binding fitted to a 1:1 model

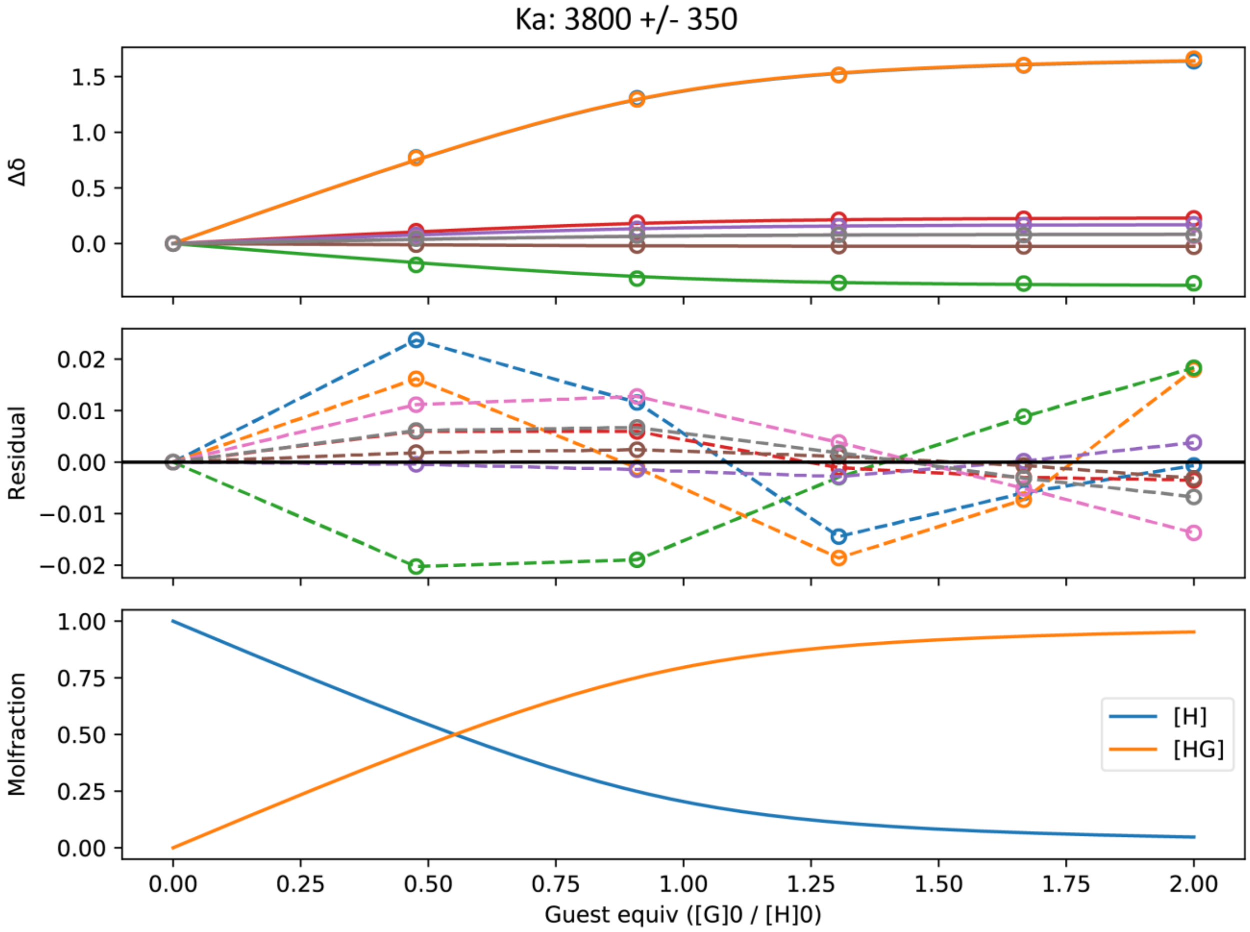


Figure S22. Data from Figure S20 over a low concentration range (5 mM *Z*-**1**, 0 – 10 mM $NBu_4H_2PO_4$) fitted to a **1:1** host:guest association model. K_1 = 3.8 ± 0.3 × 10^3 M^{-1}, within error of the value found for a 2:1 binding model in Figure S21.

S7.5 Addition of tetrabutylammonium dihydrogen phosphate to *Z*-**1** (beyond 3 equiv.)

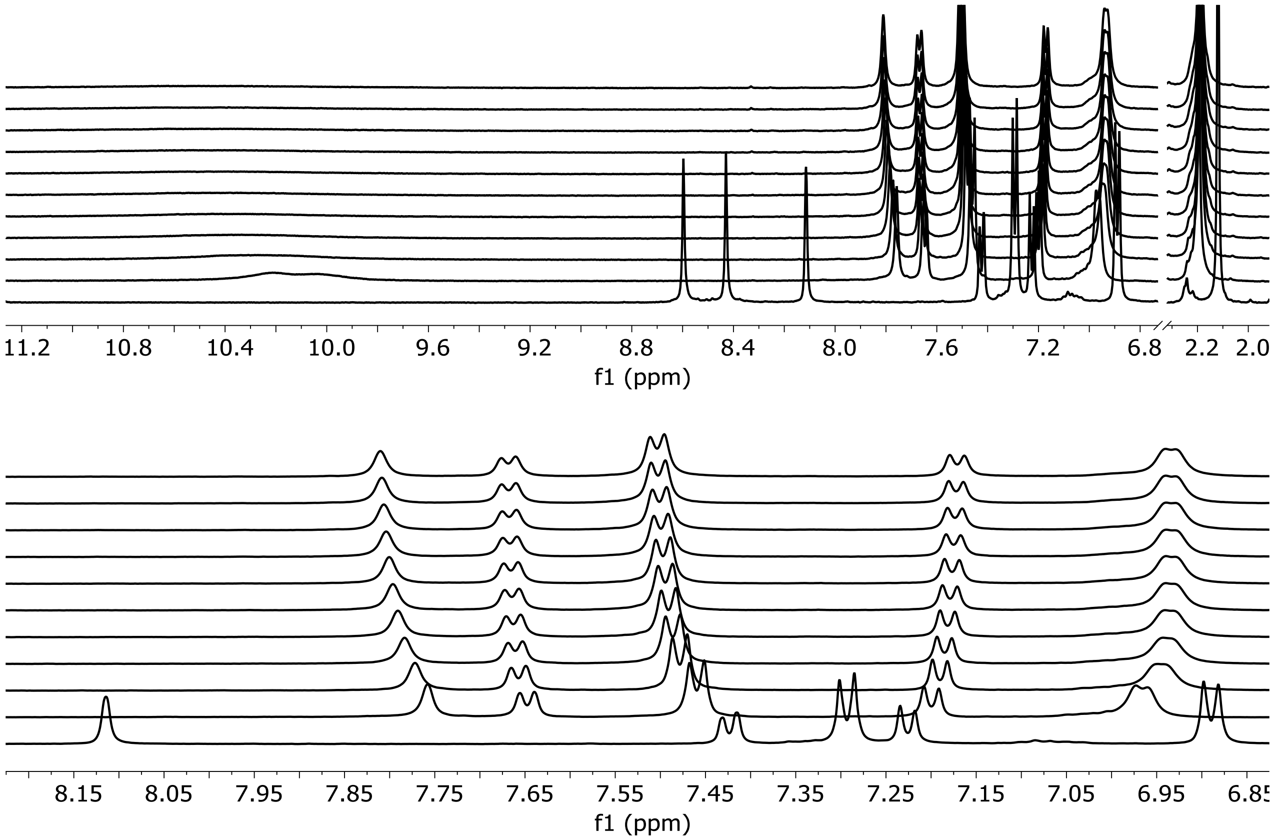


Figure S23. Titration of 200 mM [NBu_4][H_2PO_4] into 5 mM of *Z*-**1** to 66.7 mM guest concentration. 500 MHz, DMSO-d_6 + 0.5% H_2O. Bottom: the same spectrum, zoomed to focus on the aromatic region and exclude the coalesced urea NH peaks.

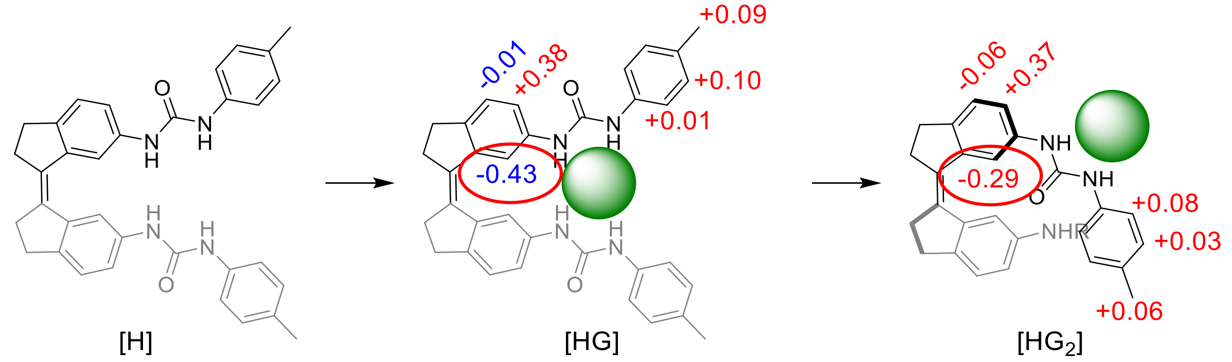


Figure S24. Changes in ^{1}H NMR chemical shifts of of the [HG] and [HG_2] complexes *of Z*-**1** relative to the free host (ppm, relative to those of free [H]) as obtained from the 1:2 nonlinear fit of the ^{1}H NMR titration data. The fitted chemical shift changes support the two-step binding fit proposed for the *Z* isomer in Figure S25. Most notably, the inner aryl proton (circled) is first shielded and then deshielded by the two steps of binding. This is consistent with the expected chemical shift induced by the carbonyl group magnetic anisotropy.

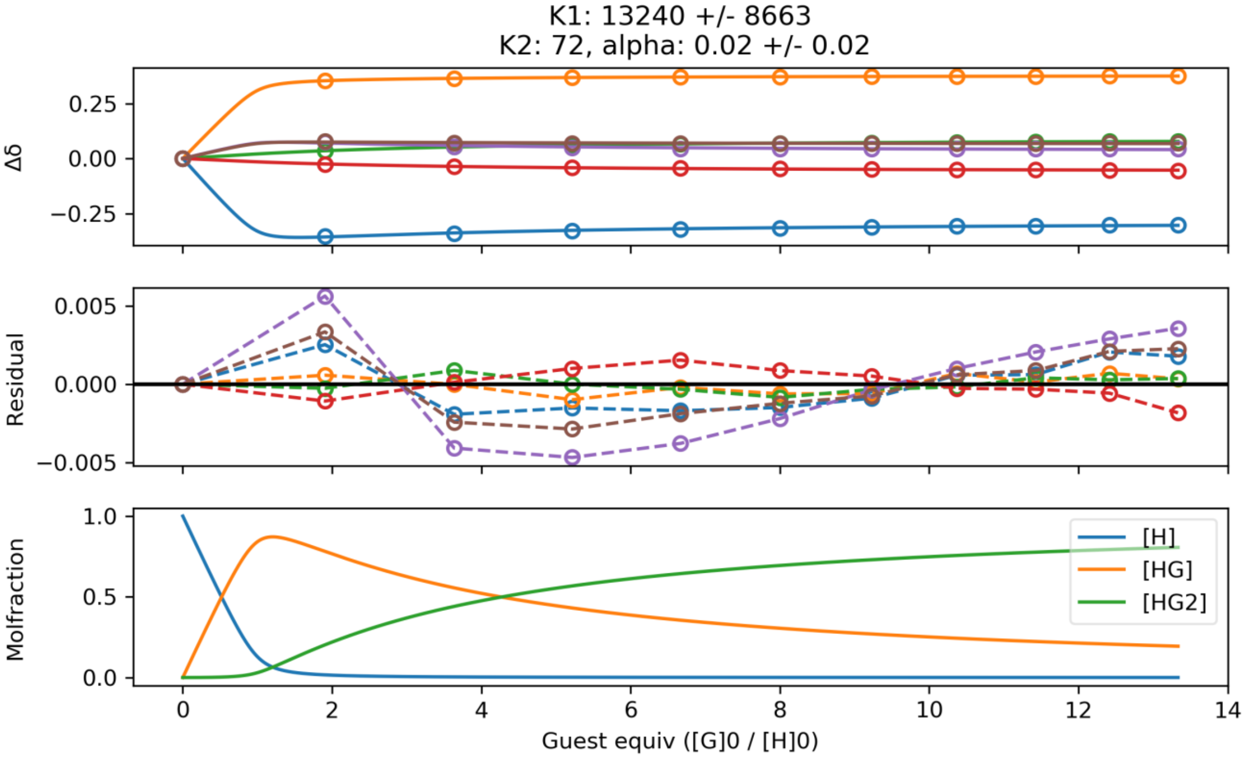


Figure S25. Data shown in Figure S23 fitted to a 1:2 host:guest association model. K_1 = 13 × 10^3 M^{-1}, K_2 = 7 × 10^1 M^{-1}.

S7.6 Addition of tetrabutylammonium acetate to *E*-1

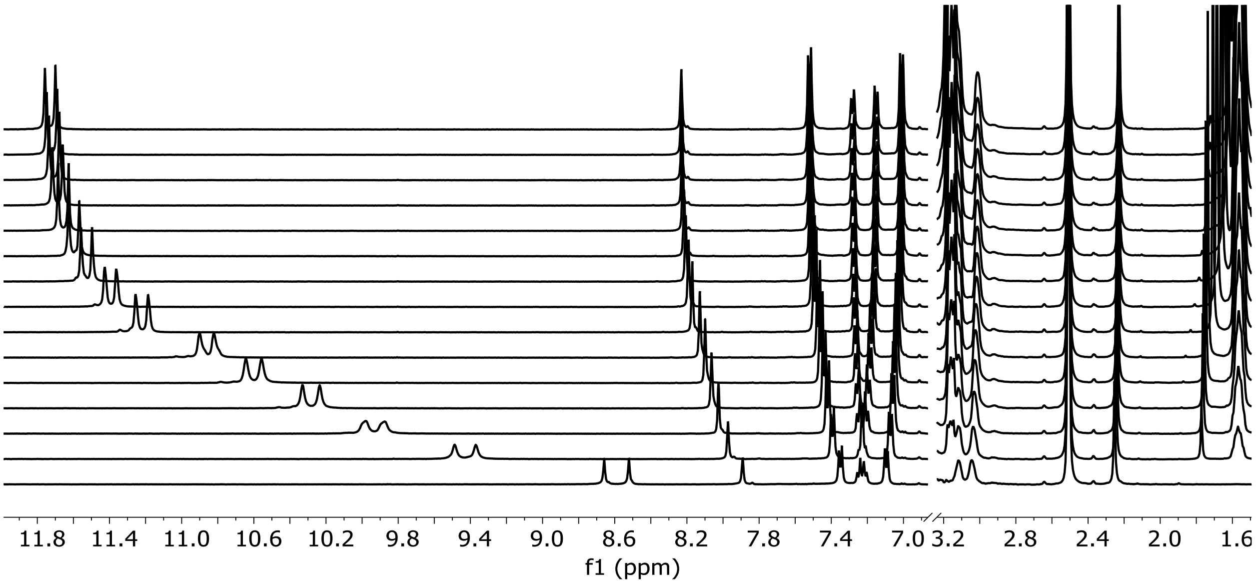


Figure S26. Titration of 200 mM [NBu_4][OAc] into 5 mM of *E*-**1**. 500 MHz, DMSO-d_6 + 0.5% H_2O.

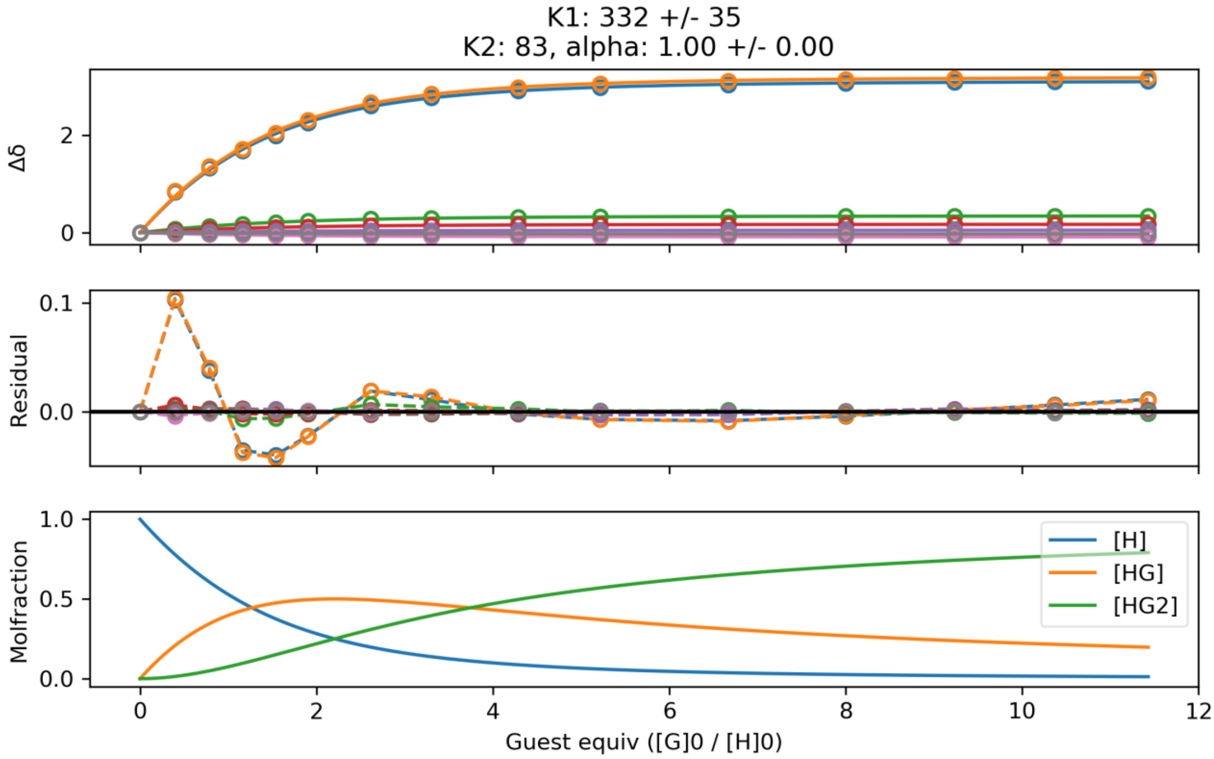


Figure S27. Data from Figure S26 fitted to a non-cooperative 1:2 host:guest association model. $K_1 = 3.3 \times 10^2\ M^{-1}$, $K_2 = 8.3 \times 10^1\ M^{-1}$.

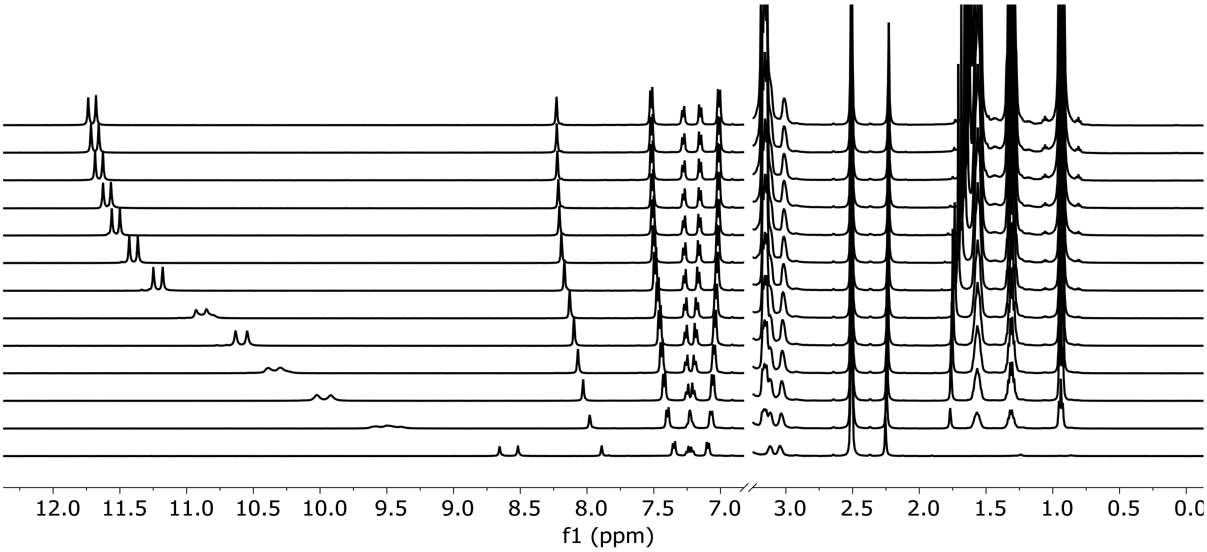


Figure S28. Titration of 200 mM $[NBu_4][OAc]$ into 5 mM of *E*-**1**. 500 MHz, DMSO-d_6 + 0.5% H_2O.

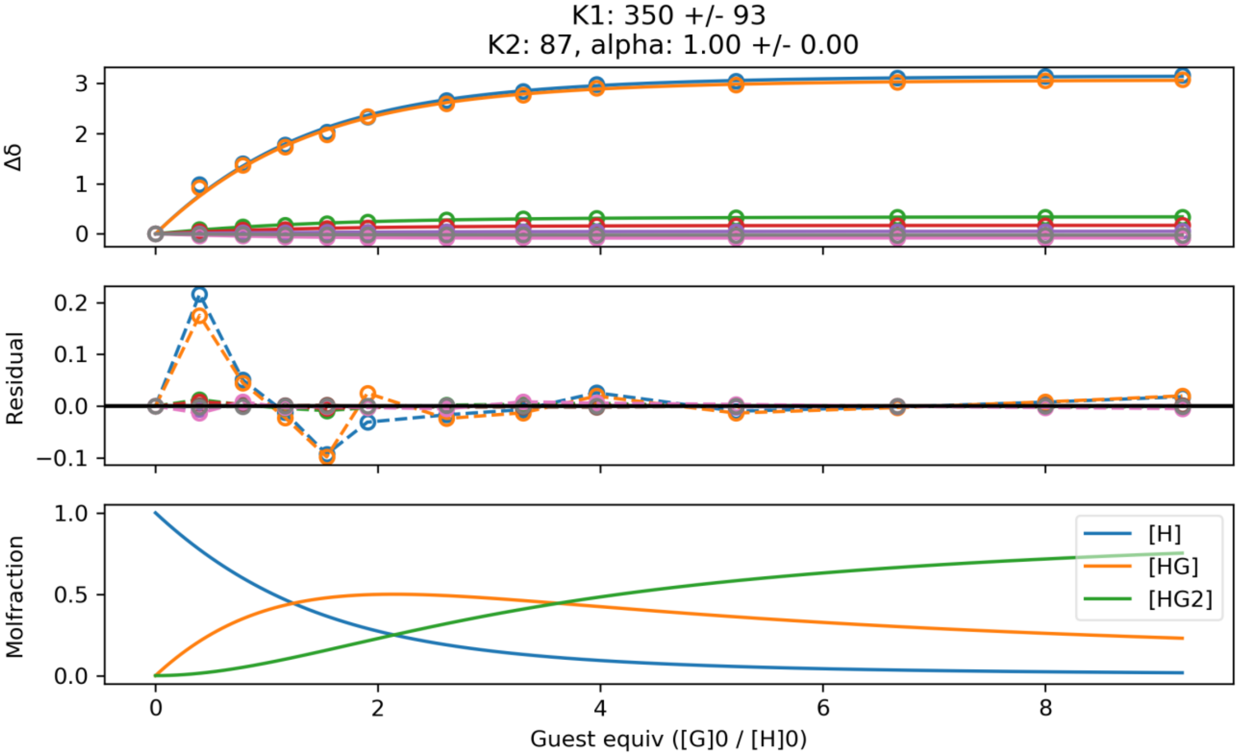


Figure S29. Data from Figure S28 fitted to a non-cooperative 1:2 host:guest association model. $K_1 = 3.5 \times 10^2$ M^{-1}, $K_2 = 8.7 \times 10^1$ M^{-1}.

S7.7 Addition of tetrabutylammonium acetate to *Z*-1

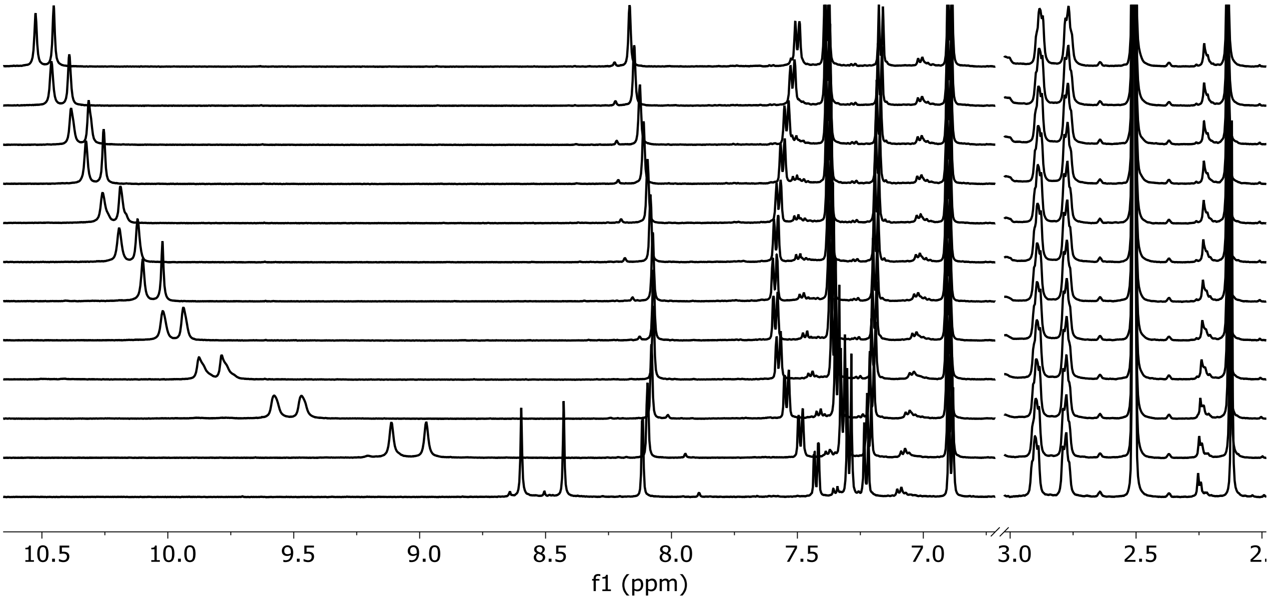


Figure S30. Titration of 200 mM $[NBu_4][OAc]$ into 5 mM of *Z*-**1**. 500 MHz, DMSO-d_6 + 0.5% H_2O.

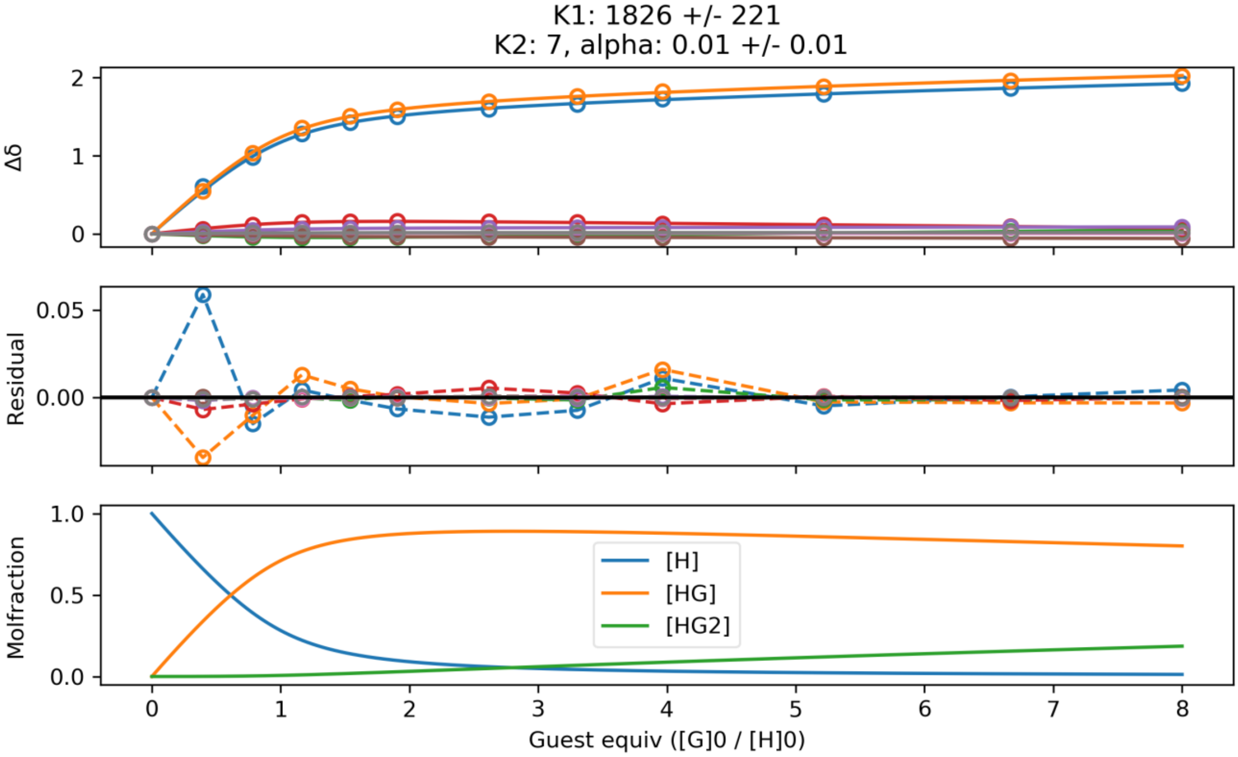


Figure S31. Data from Figure S30 fitted to a 1:2 host:guest association model. $K_1 = 1.8 \times 10^3$, M^{-1}, $K_2 = 7$ M^{-1}.

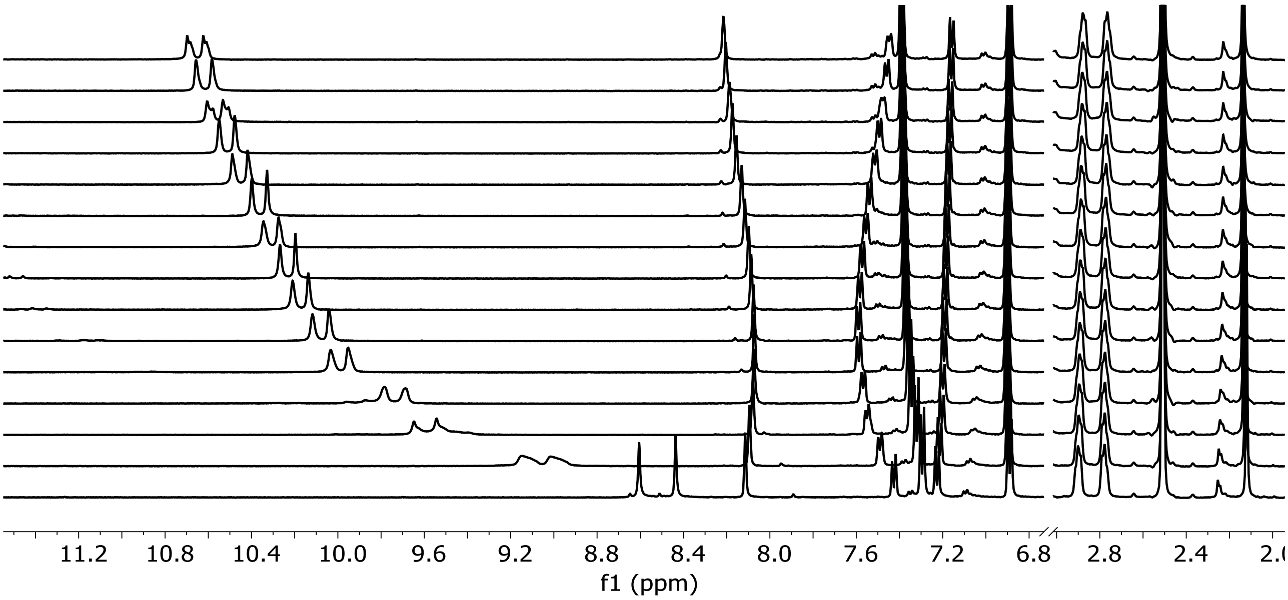


Figure S32. Titration of 200 mM [NBu$_4$][OAc] into 5 mM of *Z*-**1**. 500 MHz, DMSO-*d*$_6$ + 0.5% H_2O.

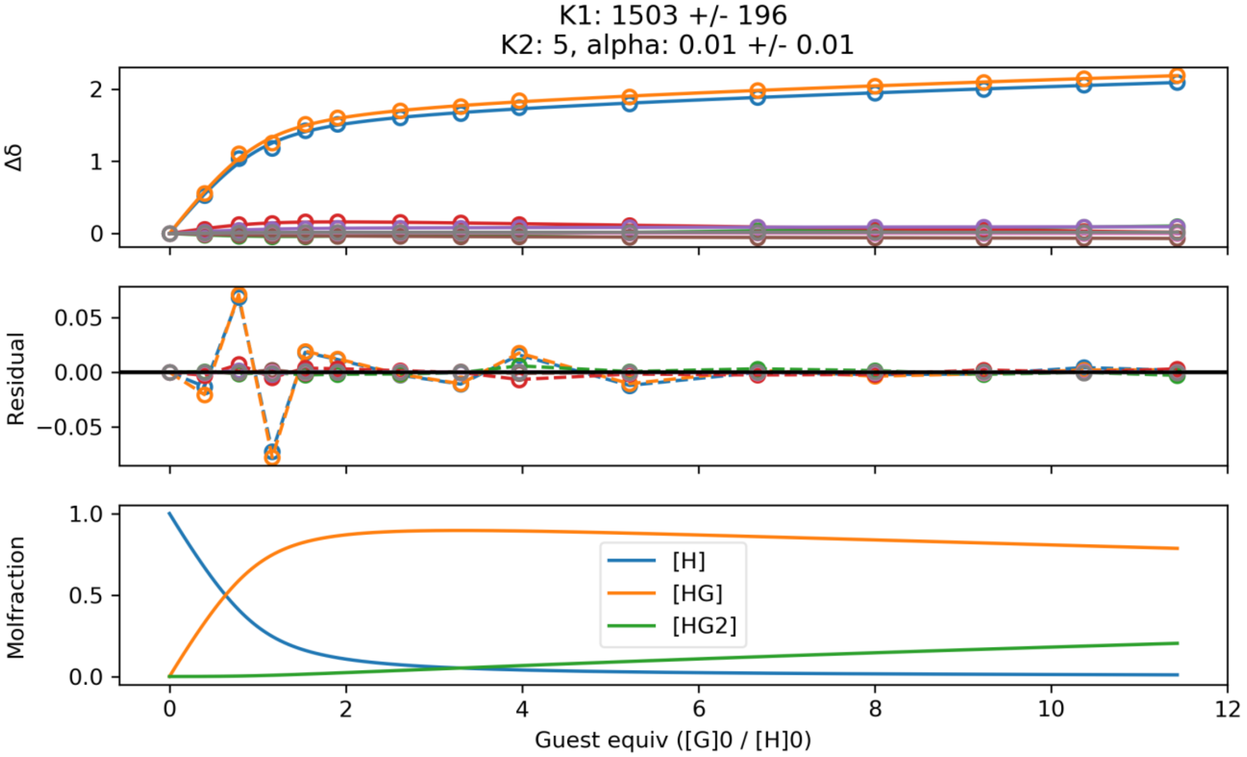
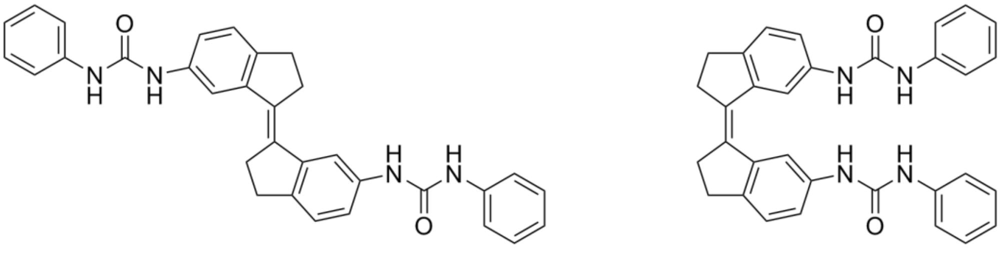


Figure S33. Data from Figure S32 fitted to a 1:2 host:guest association model. K_1 = 1.5 × 10^3, M^{-1}, K_2 = 5 M^{-1}.

S7.8 Summary of binding data

Calculated association constants for binding of acetate and dihydrogen phosphate to **1** are summarized here and compared to literature values for the previously reported phenyl-substituted stiff stilbene host **S2** (Figure S34).



E-**S2** *Z*-**S2**

Figure S34. Anion-binding stiff stilbene hosts characterized in the literature.[3] These diphenyl bisurea hosts only differ from *p*-tolylurea host **1** in the absence of a methyl group on the terminal aryl rings, and are thus expected to possess similar guest-binding properties.

S7.9 Summary of dihydrogen phosphate binding data

As discussed in detail in S7.11, the standard models of host-guest association are inapplicable here due to the self-associating nature of the $H_2PO_4^-$ anion. One consequence of this is that the estimated errors obtained through nonlinear least-squares regression are nonsensical, as the experimental data is not randomly distributed about the model used. The results obtained for **1** binding $H_2PO_4^-$ are presented here for completeness and for comparison to previous studies, but we again caution that they are obtained using an approach fundamentally inadequate for this associative process.

×

Table S7. Summary of data for host **1** binding $H_2PO_4^-$ (as $[NBu_4][H_2PO_4]$; DMSO-d_6 + 0.5% H_2O).

Species	***E*-1**[a]	***E*-1**[a]	***E*-S2**[b]	***Z*-1**[a]	***Z*-1**[a]	***Z*-1**[a]	***Z*-S2**[b]
$H_2PO_4^-$ endpoint / mM	18	57		9.8	15.5	67	
K_1 / M^{-1}	$3.6 \pm 3.2 \times 10^2$	$2.3 \pm 3.2 \times 10^2$	1.5×10^2	$2.1 \pm 2.1 \times 10^3$	$3.5 \pm 1.1 \times 10^3$	$1.3 \pm 0.8 \times 10^4$	2.0×10^3
K_2 / M^{-1}	90	60	39	70	110	72	-
α	Forced as 1			0.14 ± 0.33	0.13 ± 0.08	0.02 0.02	-

[a] Measured association constants for compound **1**; [b] Literature association constants for analogous compound **S2**.[3]

S7.10 Summary of acetate binding data

Table S8. Summary of data for host **1** binding OAc^- (as $[NBu_4][OAc]$; DMSO-d_6 + 0.5% H_2O).

	***E*-1**[a]	***E*-1**[a]	***E*-S2**[b]	***Z*-1**[a]	***Z*-1**[a]	***Z*-S2**[b]
OAc^- Endpoint / mM	57	8.6		40	57	
K_1 / M^{-1}	$3.3 \pm 0.35 \times 10^2$	$3.5 \pm 0.9 \times 10^2$	2.1×10^2	$1.8 \pm 0.2 \times 10^3$	$1.5 \pm 0.2 \times 10^3$	1.4×10^3
K_2 / M^{-1}	83	87	52	7	5	-
α	Forced as 1			0.01 ± 0.01	0.01 ± 0.01	

[a] Measured association constants for compound **1**; [b] Literature association constants for analogous compound **S2**.[3]

S7.11 Host-guest binding models with an aggregating guest

The self-association of dihydrogen phosphate described here can make determination of binding constants challenging. Considering a minimal case of host-guest binding with a 1:1 stoichiometry and no self-association for either host or guest, the concentration of free guest [G] can be found by solving the quadratic:[7]

$$0 = [\mathrm{G}]^2 - [\mathrm{G}]\left([\mathrm{G}]_0 - [\mathrm{H}]_0 - \frac{1}{K_a}\right) - \frac{[\mathrm{G}]_0}{K_a} \tag{17}$$

where $[\mathrm{G}]_0$ and $[\mathrm{H}]_0$ are respectively the total concentration of guest and host, and K_a is the association constant. This quadratic derives from the definition of an association constant and the mass balances for each of host and guest, as shown below:

$$K_a = \frac{[\mathrm{HG}]}{[\mathrm{H}][\mathrm{G}]}$$

$$[\mathrm{G}]_0 = [\mathrm{G}] + [\mathrm{HG}]$$

$$[\mathrm{H}]_0 = [\mathrm{H}] + [\mathrm{HG}]$$

If the guest is capable of indefinite self-association, constructing such a model for host-guest binding becomes more challenging. With infinite oligo-guest species present in solution, if the host binding site is capable of binding not only monomeric G but also G_2 dimers, G_3 trimers, G_4 tetramers *etc* there is no longer a single guest but instead many different guest oligomer species that cannot be distinguished. Therefore, there is not a single host-guest association constant, K_a, capable of reflecting the infinite possible host-guest interactions in solution. The mass balance equations for such a system are also complex:

$$[\mathrm{G}]_0 = \sum_{n=1}^{\infty} n[\mathrm{G}_n] + \sum_{n=1}^{\infty} n[\mathrm{HG}_n]$$

$$[\mathrm{H}]_0 = [\mathrm{H}] + \sum_{n=1}^{\infty} [\mathrm{HG}_n]$$

Modelling the titration data to such complicated competitive systems is beyond the scope of this work. Instead, we use the standard equations of host-guest binding (S7.1[7]) that assume that the guest can only exist in free or host-bound forms and that all host-guest complexes are identical.

It may be possible to measure K_a if the speciation due to one process (i.e. host-guest binding) can be varied while the other process (i.e. the isodesmic self-association) does not change. In practice this requires the two association constants to be sufficiently different that one process can be ignored over a limited concentration range where changes due to the other process can be observed (e.g. measurement of strong association at low concentrations).

S8. Diffusion of static host-guest systems

S8.1 Diffusion measurements with a single host

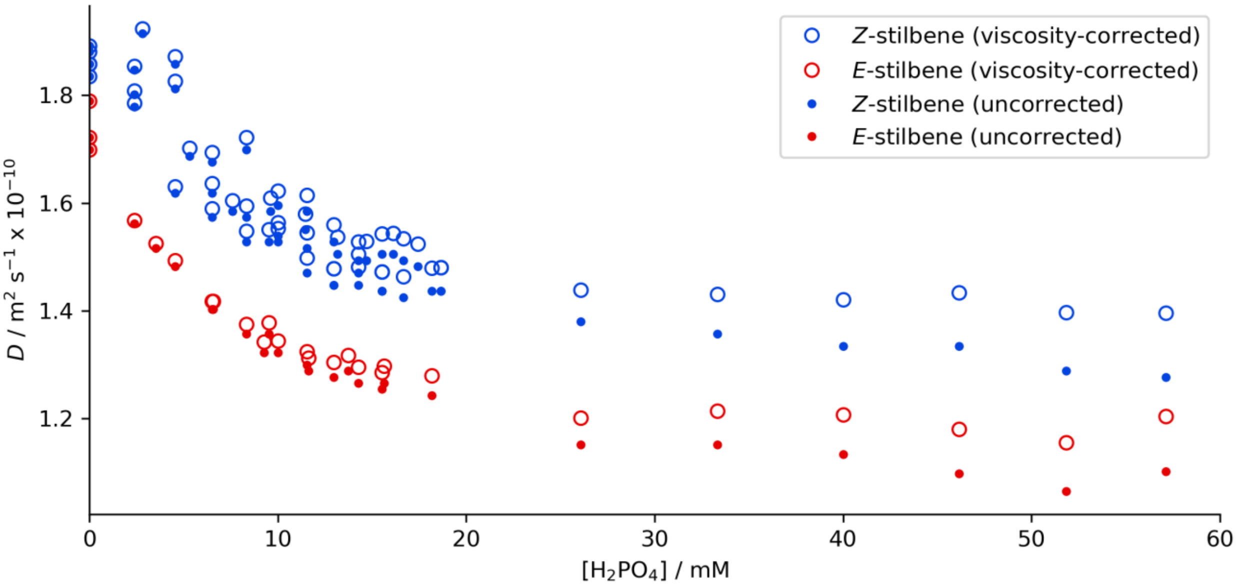


Figure S35. Measured diffusion coefficients for separate 5 mM solutions of pure *E* or *Z* hosts titrated with [H_2PO_4], presented on a single axis for clarity. Diffusion coefficients as-acquired are shown as solid points; diffusion coefficients corrected for changes in viscosity from to increasing [H_2PO_4] are shown as empty circles (linear correction from directly measured viscosities of [NBu_4][H_2PO_4] solutions; see Figure S6). Each point represents one diffusion experiment. 500 MHz ^{1}H, Δ = 50 ms, δ = 4 ms, linear ramp of 12 gradients from g = 0 – 53.45 g cm^{-1}.

S8.2 Control experiments with tetrabutylammonium acetate

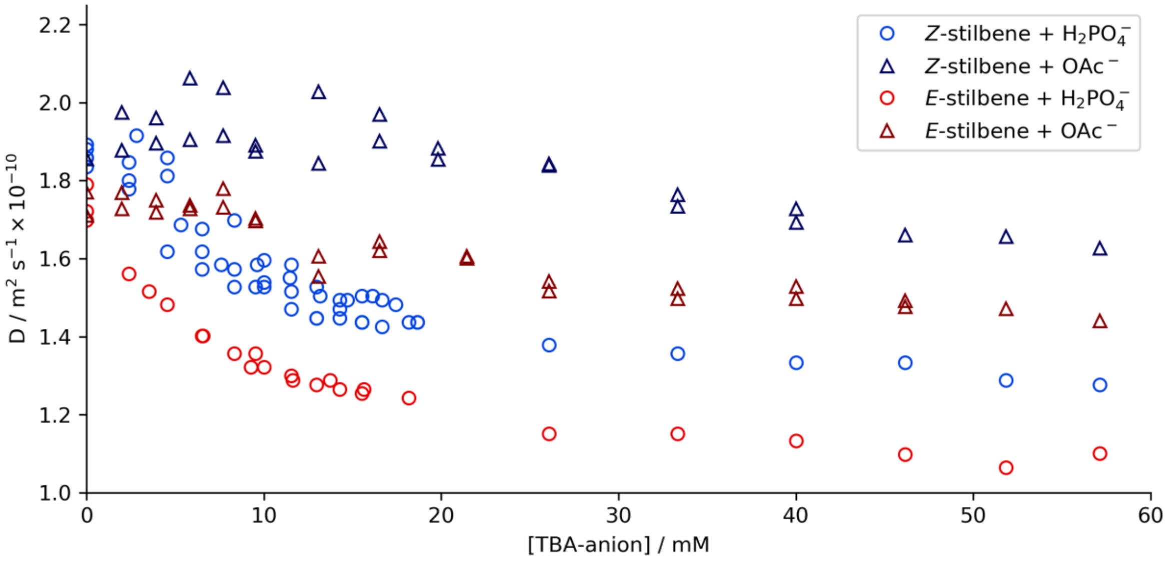


Figure S36. Comparison of guest-dependent diffusion for separate 5 mM solutions of pure *E* or *Z* hosts in the presence of [NBu_4][H_2PO_4] or [NBu_4][OAc], presented on a single axis for clarity. No corrections have been made for changes in viscosity due to the lack of viscometry data for solutions of tetrabutylammonium acetate. Each point represents one diffusion experiment. 500 MHz ^{1}H, Δ = 50 ms, δ = 4 ms, linear ramp of 12 gradients from g = 0 – 53.45 g cm^{-1}.

S8.3 Diffusion measurements for 1:1 *E*:*Z* mixed systems

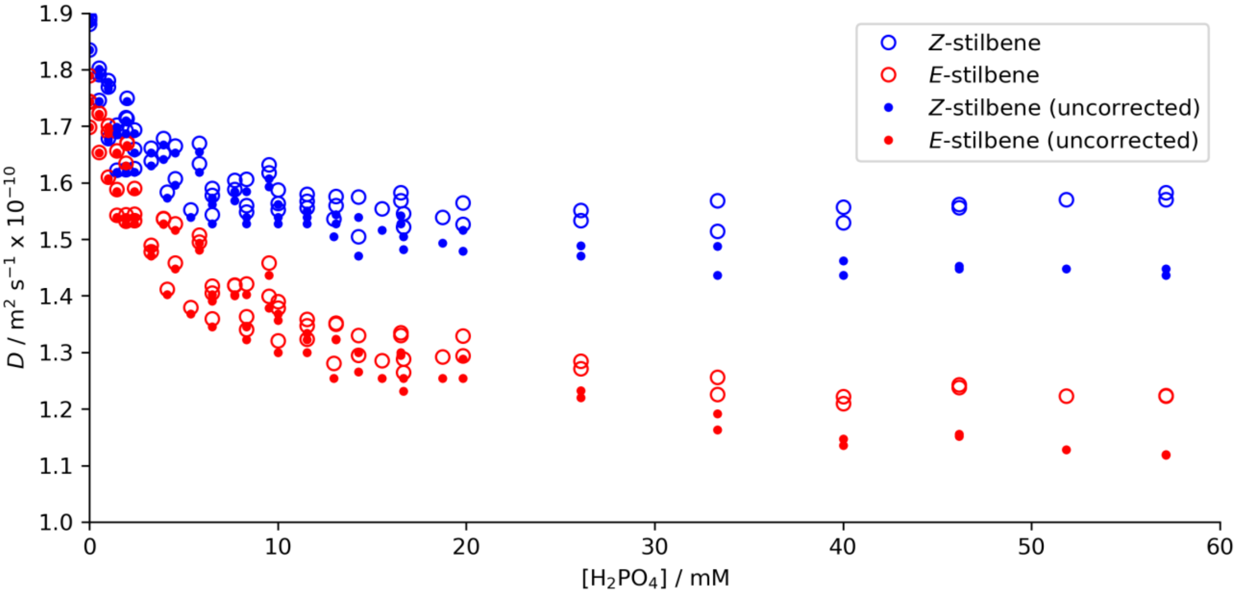


Figure S37. Diffusion of hosts **1** in 1 mM mixed solution (1:1 *E*:*Z*), in the presence of 0 – 70 mM [NBu$_4$][H$_2$PO$_4$]. Diffusion coefficients as-acquired are shown as solid points; diffusion coefficients corrected for changes in viscosity from to increasing [NBu$_4$][H$_2$PO$_4$]] are shown as empty circles (linear correction from directly measured viscosities of [NBu$_4$][H$_2$PO$_4$] solutions; see Figure S6). Each point represents one diffusion experiment. 500 MHz ^{1}H, Δ = 50 ms, δ = 4 ms, linear ramp of 12 gradients from *g* = 0 – 53.45 g cm^{-1}.

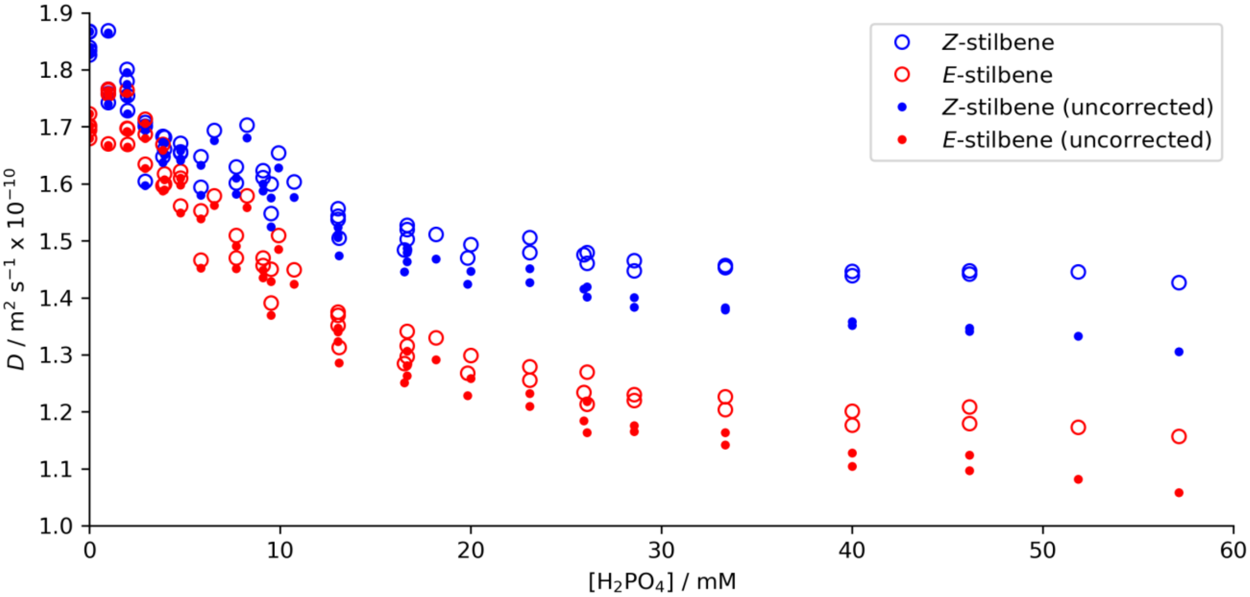


Figure S38. Diffusion of hosts **1** in 5 mM mixed solution (1:1 *E*:*Z*), in the presence of 0 – 70 mM [NBu$_4$][H$_2$PO$_4$]. Diffusion coefficients as-acquired are shown as solid points; diffusion coefficients corrected for changes in viscosity from to increasing [NBu$_4$][H$_2$PO$_4$]] are shown as empty circles (linear correction from directly measured viscosities of [NBu$_4$][H$_2$PO$_4$] solutions; see Figure S6). Each point represents one diffusion experiment. 500 MHz ^{1}H, Δ = 50 ms, δ = 4 ms, linear ramp of 12 gradients from *g* = 0 – 53.45 g cm^{-1}.

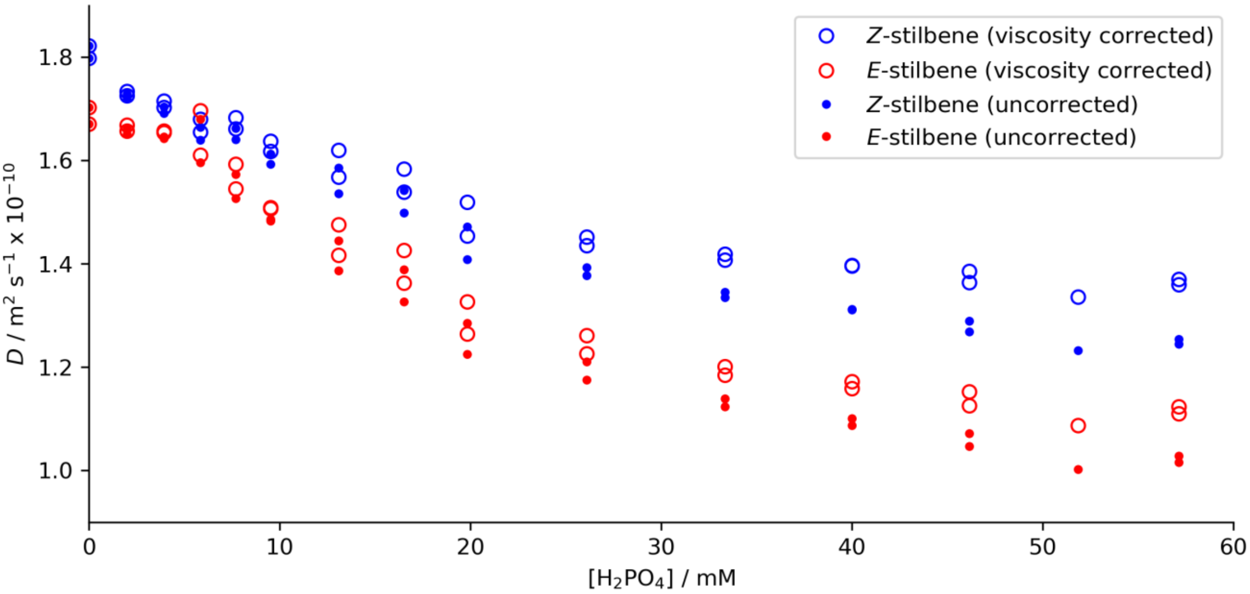


Figure S39. Diffusion of hosts **1** in 10 mM mixed solution (1:1 E:Z), in the presence of 0 – 70 mM [NBu$_4$][H$_2$PO$_4$]. Diffusion coefficients as-acquired are shown as solid points; diffusion coefficients corrected for changes in viscosity from to increasing [H$_2$PO$_4$] are shown as empty circles (linear correction from directly measured viscosities of [NBu$_4$][H$_2$PO$_4$] solutions; see Figure S6). Each point represents one diffusion experiment. 500 MHz ^{1}H, Δ = 50 ms, δ = 4 ms, linear ramp of 12 gradients from g = 0 – 53.45 g cm^{-1}.

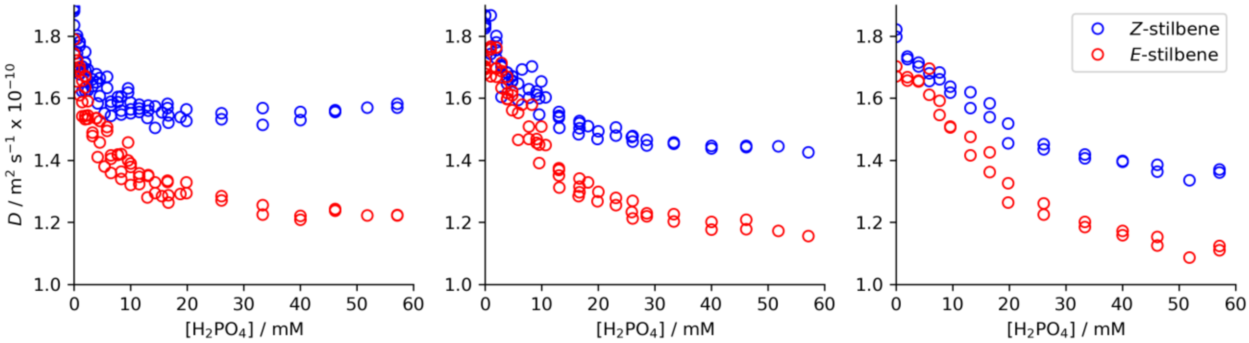


Figure S40. Comparison of measured diffusion coefficients of hosts **1** in 1:1 E:Z mixtures of 1mM, 5 mM, and 10 mM total concentration (as above). All data corrected for changes in viscosity due to the addition of [NBu$_4$][H$_2$PO$_4$] (SI S4). Each point represents one diffusion experiment. 500 MHz ^{1}H, Δ = 50 ms, δ = 4 ms, linear ramp of 12 gradients from g = 0 – 53.45 g cm^{-1}.

S8.4 Viscosity measurements for host-guest systems

Viscosities and densities for 5 mM solutions of host 1 in combination with 50 mM of $[NBu_4][H_2PO_4]$ were measured at 25 °C using a Lovis 2000M rolling ball viscometer (Ø1.59 mm capillary; 1 mm stainless steel ball), with results shown in Table S9. The presence of 50 mM $[NBu_4][H_2PO_4]$ caused a 9.6% increase in dynamic viscosity (η) relative to DMSO-*d*$_6$ with 0.5% water. Subsequent addition of 5 mM of host **1** to the 50 mM $[NBu_4][H_2PO_4]$ solution as either pure *E*-**1**, pure *Z*-**1**, or 1:1 *E*-**1**:*Z*-**1** caused minor increases in η of 1.44%, 0.51%, or 0.81% respectively with reference to 50 mM $[NBu_4][H_2PO_4]$, indicating that the addition of host **1** has only a minor effect on viscosity.

Table S9. Viscosity measurements of solutions of host and $[NBu_4][H_2PO_4]$ in DMSO-*d*$_6$ with 0.5% v/v added water.

$[NBu_4][H_2PO_4]$ / mM	[*E*-**1**] /mM	[*Z*-**1**] /mM	Density / g/cm³	Temperature / °C	Dyn. Viscosity (η) / mPa·s	Error / %	η/η_0
-	-	-	1.1833	25.04	2.149	0.02	1.000
50	-	-	1.1819	25.06	2.356	0.04	1.096
50	5	-	1.1820	25.05	2.390	0.08	1.112
50	-	5	1.1818	25.06	2.368	0.03	1.102
50	2.5	2.5	1.1820	25.06	2.375	0.02	1.105

While viscosities were not measured directly for 1:1 *E*:*Z* host mixtures at concentrations other than 5 mM, a linear interpolation between η = 2.356 in the absence of host and η = 2.375 for the 5 mM mixture was used to estimate η = 2.360 for 50 mM $[H_2PO_4]$ plus 1 mM 1:1 host and η = 2.394 for 50 mM $[H_2PO_4]$ plus 10 mM 1:1 host. This interpolated and extrapolated viscosity data was only used for the viscosity-corrected data shown in Table S12.

S8.5 Tabulated diffusion data at 50 mM [NBu$_4$][H$_2$PO$_4$]

Table S10. Changes in diffusion coefficients of pure or 1:1 mixed solutions of *E*-**1** and *Z*-**1** in the presence of 50 mM [NBu$_4$][H$_2$PO$_4$]. DMSO-*d$_6$* with 0.5% added water, 500 MHz. No correction applied for viscosity changes.

[*E*-1] /mM	**[*Z*-1] /mM**	**D(H$_2$PO$_4$)[a] / 10^{-10} m^{2}s^{-1}**	**D_E[b] / 10^{-10} m^{2}s^{-1}**	**D_Z[b] / 10^{-10} m^{2}s^{-1}**	**D(NBu$_4$)[b] / 10^{-10} m^{2}s^{-1}**	**D_{DMSO}[b] / 10^{-10} m^{2}s^{-1}**
-	-	1.99 ± 0.02	-	-	2.32 ± 0.01	6.32 ± 0.04
0.5	0.5	1.90 ± 0.02	1.15 ± 0.02	1.45 ± 0.01	2.33 ± 0.01	6.35 ± 0.02
2.5	2.5	1.82 ± 0.06	1.10 ± 0.01	1.34 ± 0.01	2.26 ± 0.01	6.28 ± 0.03
5	5	1.69 ± 0.08	1.04 ± 0.01	1.26 ± 0.01	2.14 ± 0.01	6.11 ± 0.03
5	-	1.79 ± 0.04	1.08 ± 0.02	-	2.21 ± 0.01	6.06 ± 0.02
-	5	1.86 ± 0.03	-	1.29 ± 0.01	2.19 ± 0.01	6.02 ± 0.02

[a] ^{31}P PGSTE experiment, δ = 7 ms, Δ = 100 ms, *g* = 0 – 53.45 G cm^{-1}. [b] ^{1}H PGSTE experiment, δ = 4 ms, Δ = 50 ms, *g* = 0 – 53.45 G cm^{-1}.

Table S11. Diffusion data from Table S10, corrected for changes in viscosity using *η* for 50 mM [NBu$_4$][H$_2$PO$_4$] addition only, ignoring changes in *η* caused by host.

[*E*-1] /mM	**[*Z*-1] /mM**	**η/η_0**	**D'(H$_2$PO$_4$)[a] / 10^{-10} m^{2}s^{-1}**	**D'_E[b] / 10^{-10} m^{2}s^{-1}**	**D'_Z[b] / 10^{-10} m^{2}s^{-1}**	**D'(H$_2$PO$_4$)[b] / 10^{-10} m^{2}s^{-1}**	**D'_{DMSO}[b] / 10^{-10} m^{2}s^{-1}**
-	-	1.096	2.18 ± 0.02	-	-	2.54 ± 0.01	6.93 ± 0.04
0.5	0.5	1.096	2.08 ± 0.02	1.26 ± 0.02	1.59 ± 0.01	2.55 ± 0.01	6.96 ± 0.02
2.5	2.5	1.096	2.00 ± 0.07	1.21 ± 0.01	1.47 ± 0.01	2.48 ± 0.01	6.88 ± 0.03
5	5	1.096	1.85 ± 0.09	1.14 ± 0.01	1.38 ± 0.01	2.35 ± 0.01	6.70 ± 0.03
5	-	1.096	1.96 ± 0.04	1.18 ± 0.02	-	2.42 ± 0.01	6.64 ± 0.02
-	5	1.096	2.04 ± 0.03	-	1.41 ± 0.01	2.4 ± 0.01	6.60 ± 0.02

[a] ^{31}P PGSTE experiment, δ = 7 ms, Δ = 100 ms, *g* = 0 – 53.45 G cm^{-1}. [b] ^{1}H PGSTE experiment, δ = 4 ms, Δ = 50 ms, *g* = 0 – 53.45 G cm^{-1}.

Table S12. Diffusion data from **Table S10**, corrected for changes in viscosity caused by host and guest

[*E*-1] /mM	**[*Z*-1] /mM**	**η/η_0**	**D'(H$_2$PO$_4$)[a] / 10^{-10} m^{2}s^{-1}**	**D'_E[b] / 10^{-10} m^{2}s^{-1}**	**D'_Z[b] / 10^{-10} m^{2}s^{-1}**	**D'(H$_2$PO$_4$)[b] / 10^{-10} m^{2}s^{-1}**	**D'_{DMSO}[b] / 10^{-10} m^{2}s^{-1}**
-	-	1.096	2.18 ± 0.02	-	-	2.54 ± 0.01	6.93 ± 0.04
0.5	0.5	1.098[c]	2.09 ± 0.02	1.26 ± 0.02	1.59 ± 0.01	2.56 ± 0.01	6.97 ± 0.02
2.5	2.5	1.105	2.01 ± 0.07	1.22 ± 0.01	1.48 ± 0.01	2.50 ± 0.01	6.94 ± 0.03
5	5	1.114[c]	1.88 ± 0.09	1.16 ± 0.01	1.40 ± 0.01	2.38 ± 0.01	6.81 ± 0.03
5	-	1.112	1.99 ± 0.04	1.20 ± 0.02	-	2.46 ± 0.01	6.74 ± 0.02
-	5	1.102	2.05 ± 0.03	-	1.42 ± 0.01	2.41 ± 0.01	6.63 ± 0.02

[a] ^{31}P PGSTE experiment, δ = 7 ms, Δ = 100 ms, *g* = 0 – 53.45 G cm^{-1}. [b] ^{1}H PGSTE experiment, δ = 4 ms, Δ = 50 ms, *g* = 0 – 53.45 G cm^{-1}. [c] Not measured directly but estimated by linear interpolation from other points: see S8.4.

S9. Comparison of modelled and measured guest-dependent diffusion coefficients for pure *E* and *Z* systems

Table S13. Parameters used for modelling concentration-dependent diffusion coefficients of *E*-**1** and *Z*-**1**.

	D_0 / m^{2}s^{-1} × 10^{-10}	K_i / M^{-1}	K_1 / M^{-1}	K_2 / M^{-1}
Monomeric $H_2PO_4^-$ (see S6)$^-$	3.39	120	-	-
5 mM *E*-1	1.74	-	360	90
5 mM *Z*-1	1.87	-	2100	70

Host-guest speciation and diffusion coefficients for pure *E* or *Z* hosts binding up to two $H_2PO_4^-$ oligomers using the parameters given in Table S13 were modelled using Equation 13 (S5.3), and plotted in Figure S41. The modelled dependence of D_E and D_Z on $H_2PO_4^-$ concentration significantly overestimates *D* at all concentrations studied, particularly for the *E*-**1** isomer, indicating that this model does not accurately reflect the structures formed in solution.

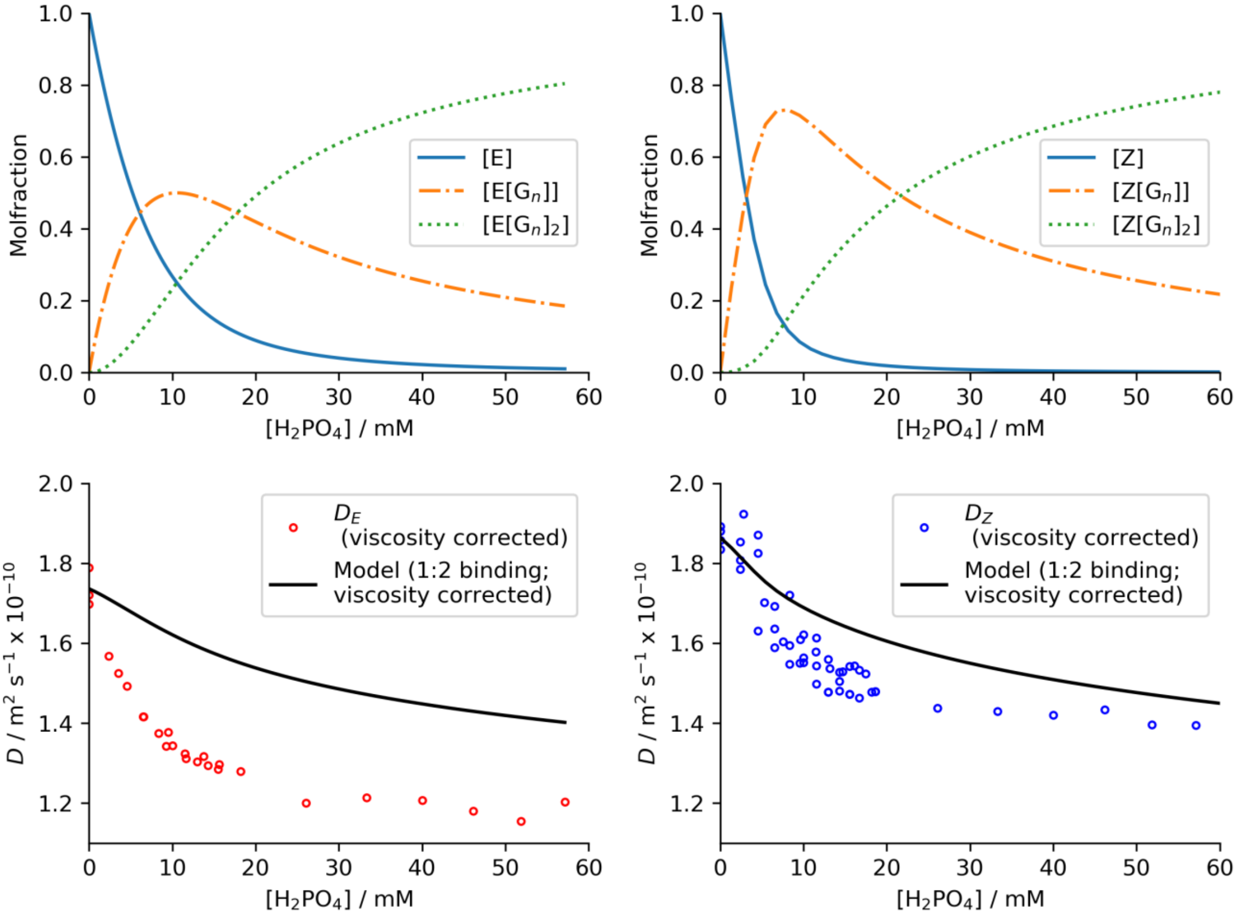


Figure S41. Comparison of experimental and modelled [H_2PO_4]-dependent diffusion coefficients of hosts, using Equation 13 and the parameters shown in Table S13. Top: speciation of *E* (left) and *Z* (right) hosts using a 1:2 host:guest binding model and the association parameters shown in Table S13. Bottom: modelled average diffusion coefficients of hosts binding up to 2 $H_2PO_4^-$ oligomers (see S5.3 for details). This model is clearly insufficient to explain the large decrease in *D* observed on addition of $H_2PO_4^-$, particularly for the *E*-**1** isomer.

S9.1 Estimated diffusion coefficients of potential supramolecular host-guest complexes

To aid in comparing the sizes of proposed complexes, approximate molecular weights are shown in Table S14. As $H_2PO_4^-$ oligomers have been modelled as forming with a continuous distribution of sizes, no single definitive molecular weight can be given for complexes involving oligomeric $H_2PO_4^-$. Instead, we provide molecular weights for the calculated median and mean oligomer sizes.

Table S14. Estimated diffusion coefficients for representative supramolecular structures of the form $[H_mG_n]$. Estimated diffusion coefficients for complexes (calculated using Eq (11)) are shown in normal text. While a wide distribution of host-guest systems are expected to form in solution, the measured host diffusion coefficients at 5 mM host and 50 mM $[NBu_4][H_2PO_4]$ suggest complexes of approximately 1.5 – 2.5 kDa in size under these conditions.

Complex	**Median MW / kDa**	**Mean MW / kDa**	$\boldsymbol{D(H_2PO_4)}$[a] **/ 10^{-10} m^{2}s^{-1}**	$\boldsymbol{D_E}$[b] **/ 10^{-10} m^{2}s^{-1}**	$\boldsymbol{D_Z}$[b] **/ 10^{-10} m^{2}s^{-1}**
Free host $[H_1G_0]$	0.5		-	1.74[a]	1.87[a]
Oligomeric $H_2PO_4^-$ $[G_n]$ (50 mM $[NBu_4][H_2PO_4]$, K_i = 120)	0.4 (4 $H_2PO_4^-$ units)	0.5 (5 $H_2PO_4^-$ units)	2.16[b]		
Monomeric G ($H_2PO_4^-$, infinite dilution)	0.1		3.39		
Host with monomeric phosphate ($\boldsymbol{D(H_2PO_4)}$ = 3.39)					
$[H_1G_1]$	0.6			1.67	1.78
$[H_1G_2]$	0.7			1.61	1.70
Host with phosphate oligomers (50 mM, $\boldsymbol{D(H_2PO_4)}$ = 2.16)					
$[H_1G'_1]$	0.9	1.0		1.51	1.58
$[H_1G'_2]$	1.3	1.5		1.37	1.42
$[H_2G'_1]$	1.4	1.5		1.22	1.35
$[H_2G'_2]$	1.8	2.0		1.20	1.26
$[H_2G'_3]$	2.2	2.5		1.14	1.18
$[H_3G'_2]$	2.4	2.6		1.09	1.15
$[H_3G'_3]$	2.8	3.0		1.05	1.10
$[H_3G'_4]$	3.1	3.5		1.01	1.05
5 mM host, 50 mM guest			1.93[b] (*E*), 2.01[b] (*Z*)	1.17[a]	1.39[a]

Experimentally obtained values: [a] ^{31}P PGSTE, δ = 7 ms, Δ = 100 ms, g = 0 – 53.45 G cm^{-1}; [b] ^{1}H PGSTE experiment, δ = 4 ms, Δ = 50 ms, g = 0 – 53.45 G cm^{-1}.

S10. Photoswitching studies

S10.1 Time-dependent diffusion with *in situ* irradiation

Photo-diffusion studies were conducted by combining *in situ* light irradiation with time-dependent diffusion experiments previously reported by our group.[8] It is well known that NMR diffusion measurements are often disrupted by convection, generally caused by the presence of temperature gradients within the sample.[9] Convection can be challenging at the best of times, but here the use of *in situ* irradiation (and subsequent heating) deliver by fiber to the center of the sample can be almost guaranteed to generate temperature gradients during irradiation. This can be seen in the time-dependent diffusion data presented in Figure S42, where measured *D* increases sharply to an unrealistic value during periods of irradiation.

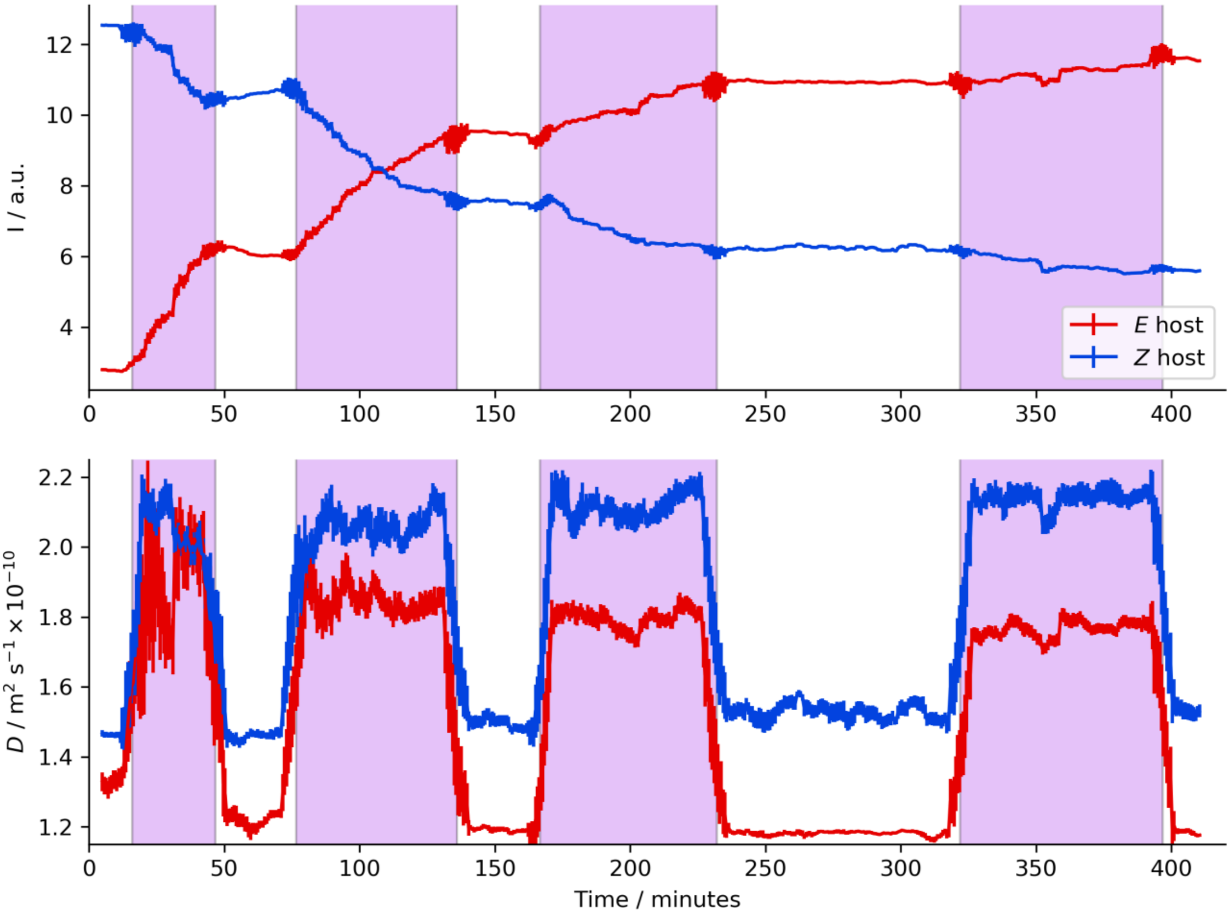


Figure S42. Time-resolved diffusion study while photoswitching under *in situ* irradiation (405 nm LED source, shaded areas). 5 mM host (initially ~90% *Z*-**1**) in solution with 50 mM [NBu_4][H_2PO_4] was irradiated with 405 nm light during the shaded periods of time, while diffusion information was continuously acquired using a long list of random gradients as previously reported.[8] Measured diffusion coefficients can be seen to increase sharply under irradiation, but in reality this is convection resulting from temperature gradients caused by uneven sample heating under light irradiation.

S10.2 Monitoring temperatures and suppressing convection

There are many approaches to suppressing convection in diffusion NMR experiments. We chose the crude but effective method of introducing physical obstructions to the sample to disrupt connective flow. We use glass TLC spotting pipettes broken up into 30 – 40 mm irregular lengths to pack the sample and restrict flow. These open-ended capillaries were added to the sample around the 1.5 mm quartz optic fibre and flame-sealed methanol capillary used for thermometry (Figure S43).

To our pleasant surprise, NMR lineshapes were retained and few difficulties were encountered in shimming after packing our samples with glass and quartz (signal intensity was reduced somewhat by the displacement of liquid from the active volume). Representative time-resolved diffusion data are shown in Figure S44. The 'noise' observed for the time-dependent diffusion coefficient of **1**-*E* in the early stages of the isomerization is an artifact resulting from the low signal-to-noise and rapidly changing concentration of this species, as previously observed by us for other systems under similar circumstances.[8]

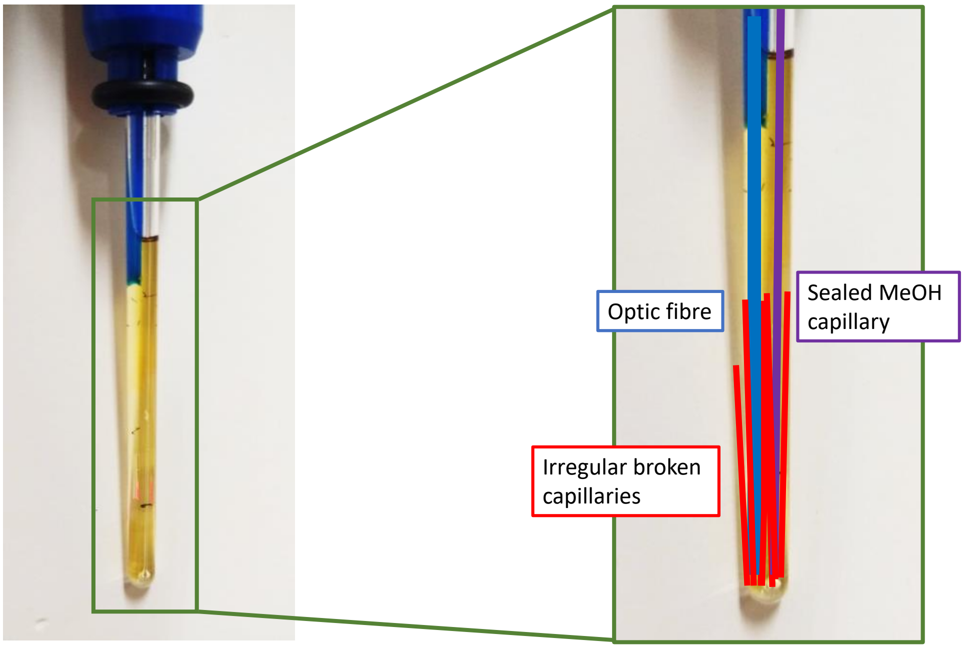


Figure S43. NMR sample used for photo-diffusion studies with *in situ* irradiation and temperature measurements. In the expansion the optic fiber (quartz, 1.5 mm diameter), flame-sealed methanol capillary (Sigma-Aldrich, 5 µL capacity glass micropipette), and broken lengths of glass TLC capillaries are highlighted by colored lines.

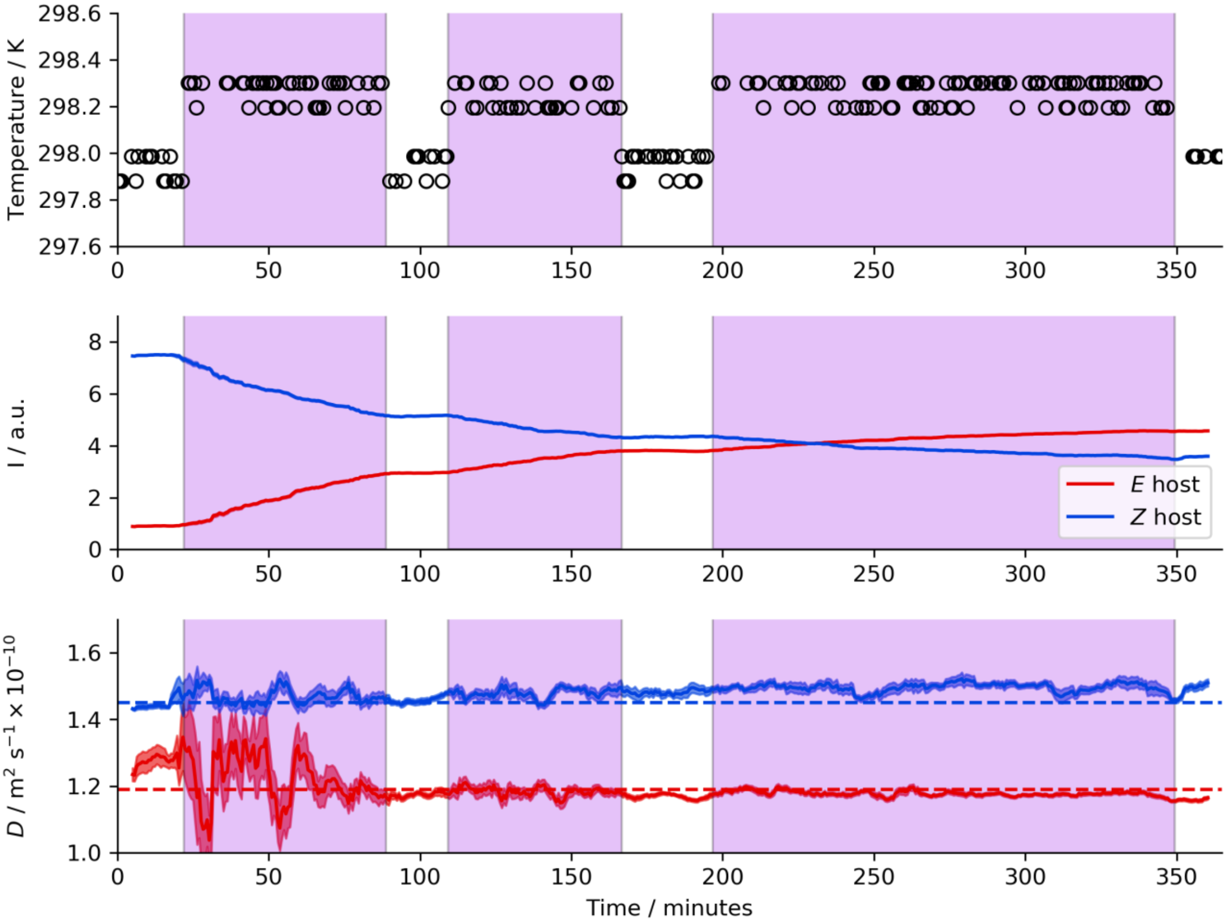


Figure S44. Time-resolved diffusion study[8] while photoswitching under *in situ* irradiation (400 nm LED source), with flame-sealed methanol capillary added for *in situ* thermometry and short (30-40 mm) sections of broken glass capillaries added to suppress convection. 5 mM host (initially ~90% *Z*-**1**) in solution with 50 mM [NBu_4][H_2PO_4] was irradiated with 405 nm light during the shaded periods of time, while diffusion information was continuously acquired using a long list of random gradients. **Top**: temperature measurements taken from the separation of the methanol CH_3 and OH peaks (top) show an increase in sample temperature of approximately 0.35 K while under irradiation.[10] The apparent quantization of temperature measurements results from the ~0.5 Hz resolution of the NMR spectra. **Middle**: signal intensity changes of the tolyl peaks of *E*-**1** and *Z*-**1** during irradiation. **Bottom**: time-dependent diffusion measurements of the same peaks over time. Measured *D* for each species taken from measurements of the 1:1 mixture of photoisomers (SI-S8.5) are shown as dashed horizontal lines for comparison.

S10.3 Demonstration of reversible photoswitching by UV-vis

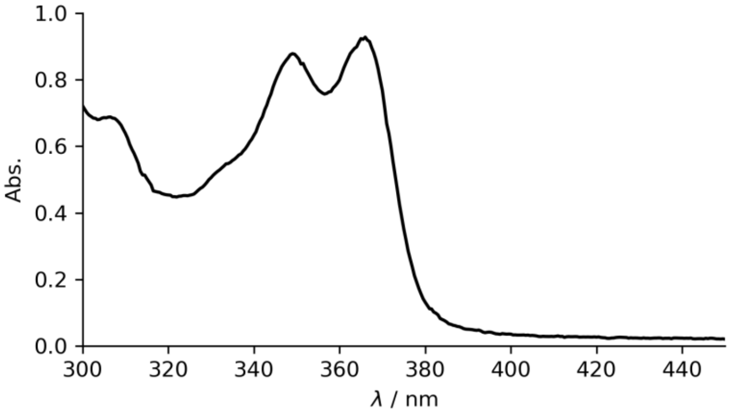


Figure S45. UV-vis absorption spectrum of *E*-**1** (approx. 25 μM) in DMSO + 0.5% water. The absorption maximum λ_{max} occurs at 365 nm, identical to that previously reported for analogous host *E*-**S1.**

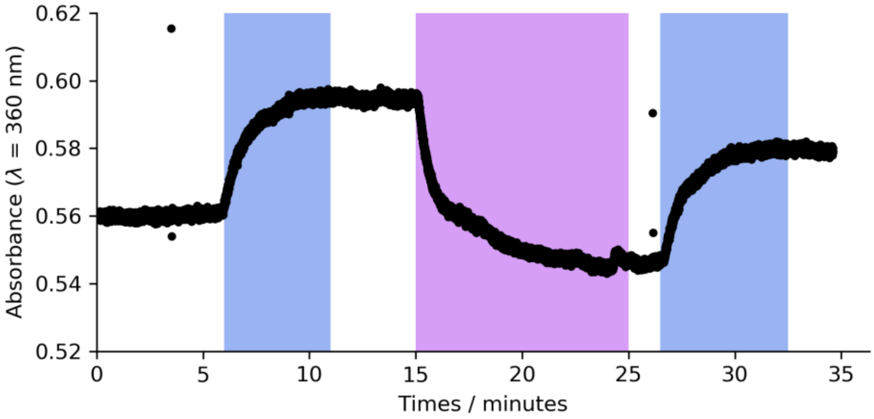


Figure S46. Demonstration of reversible photoswitching of compound **1** by UV-vis spectroscopy under irradiation from 385 nm (blue) and 375 nm (violet) LEDs. Solution of pure **1** (approx. 25 μM) in DMSO + 0.5% water.

The reversibility of photoswitching in DMSO was studied by UV-vis spectroscopy with *in situ* irradiation delivered by 385 nm and 375 nm LEDs. These wavelengths are not ideal for selectively photoswitching this class of compound, but are sufficient to demonstrate the reversibility of photoswitching as shown in Figure S46.

S11. Geometry optimizations by DFT

Input geometries were generated using Avogadro[11] (UFF optimization). The Gaussian 03 program[12] was used for further geometry optimization: First, energy minimization was performed at the semi-empirical PM3 level of theory, and subsequently at the DFT B3LYP/6-31G(d,p) level of theory using tight convergence criteria. The DFT optimized geometries were found to have zero imaginary frequencies. The $H_2PO_4^-$ complexes were calculated for the phenyl derivative **S2** (Figure S34), as for the tolyl derivative **1** convergence was hampered by free rotation over the C(aryl)-CH_3 bond. The optimized structure of *Z*-**S2**@$H_2PO_4^-$ has been reported elsewhere.[3] For the *E*-isomer, two possible geometries, with urea groups pointing in opposite directions, were optimized.

S11.1 Calculated structure of $H_2PO_4^-$

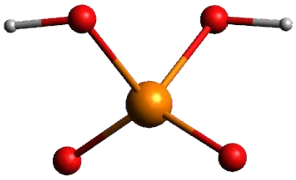


Table S15. Cartesian coordinates of $H_2PO_4^-$ (approximate radius = 2.5 Å).

atom	X	Y	Z
P	0.0000000000	0.0000180000	-0.1653310000
O	-1.0101600000	-0.8033300000	0.9083060000
H	-1.8339000000	-0.3034530000	0.8348330000
O	1.0101720000	0.8031570000	0.9084420000
O	-0.8685690000	1.0143440000	-0.8576320000
H	1.8339010000	0.3032610000	0.8349340000
O	0.8685580000	-1.0141790000	-0.8578410000

Sum of electronic and zero-point Energies = -643.562195

S11.2 Calculated structure of *Z*-1

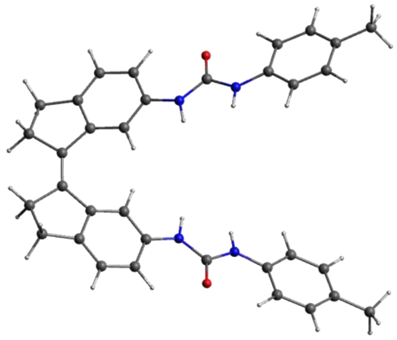


Table S16. Cartesian coordinates of *Z*-**1** (approximate width × height × depth = 15.5 × 10.1 × 13.2 Å).

atom	X	Y	Z
C	-4.4187180000	-2.6292070000	1.3266160000
C	-3.9527390000	-1.5791220000	0.5111400000
C	-5.8729570000	-2.4232700000	1.6806280000
C	-6.3342430000	-1.3578480000	0.6529360000
H	-5.9755270000	-2.0360570000	2.7036870000
H	-6.4586460000	-3.3468050000	1.6303780000
C	-5.0527180000	-0.6315450000	0.2447640000
H	-7.0910580000	-0.6809600000	1.0584210000
H	-6.7811660000	-1.8572210000	-0.2186010000
C	-3.5683980000	-3.6730620000	1.6727960000
C	-2.2574410000	-3.7131230000	1.1939130000
C	-1.8023360000	-2.6928510000	0.3403620000
C	-2.6518610000	-1.6288650000	0.0019210000
H	-3.9213350000	-4.4811900000	2.3083480000
H	-1.5913490000	-4.5217820000	1.4561660000
H	-2.3033540000	-0.8637980000	-0.6848360000
C	-3.5684300000	3.6730450000	-1.6727910000
C	-2.2574770000	3.7131210000	-1.1938980000
H	-1.5913930000	4.5217880000	-1.4561450000
C	-1.8023660000	2.6928510000	-0.3403460000
C	-2.6518800000	1.6288550000	-0.0019140000
N	-0.5165810000	-2.6862610000	-0.2443320000
C	0.5623550000	-3.5111720000	0.0458160000
N	1.6615730000	-3.2297750000	-0.7522750000
C	2.9335330000	-3.8420110000	-0.7334530000
O	0.5451870000	-4.3803600000	0.9072120000
C	3.8773100000	-3.3837580000	-1.6653740000
C	3.3017570000	-4.8591470000	0.1588180000
C	4.5892410000	-5.3906590000	0.0976570000
C	5.1558110000	-3.9296050000	-1.7080860000
C	5.5408340000	-4.9487480000	-0.8284620000
C	6.9185680000	-5.5637080000	-0.8958930000

H	6.9497470000	-6.4025140000	-1.6029560000
H	7.2313860000	-5.9506340000	0.0785830000
H	7.6670300000	-4.8361060000	-1.2250050000
H	2.5852350000	-5.2186980000	0.8826370000
H	4.8566950000	-6.1770440000	0.7994330000
H	3.6100390000	-2.5891770000	-2.3594030000
H	5.8663780000	-3.5520500000	-2.4391240000
H	-0.3719810000	-1.9445230000	-0.9136010000
H	1.5454680000	-2.5424810000	-1.4824010000
C	0.5623220000	3.5111820000	-0.0458070000
N	1.6615350000	3.2298050000	0.7522970000
C	2.9334970000	3.8420380000	0.7334640000
O	0.5451590000	4.3803460000	-0.9072270000
C	3.3017300000	4.8591420000	-0.1588410000
C	3.8772670000	3.3838140000	1.6654070000
C	5.1557690000	3.9296600000	1.7081080000
C	4.5892150000	5.3906520000	-0.0976900000
C	5.5408010000	4.9487710000	0.8284510000
C	6.9185360000	5.5637280000	0.8958700000
H	7.6670010000	4.8361240000	1.2249700000
H	6.9497240000	6.4025310000	1.6029380000
H	7.2313440000	5.9506590000	-0.0786070000
H	2.5852140000	5.2186700000	-0.8826770000
H	3.6099890000	2.5892580000	2.3594610000
H	5.8663300000	3.5521280000	2.4391640000
H	4.8566760000	6.1770120000	-0.7994920000
H	1.5454180000	2.5425490000	1.4824570000
N	-0.5166160000	2.6862790000	0.2443590000
H	-0.3720100000	1.9445390000	0.9136240000
C	-3.9527540000	1.5790980000	-0.5111430000
H	-2.3033680000	0.8637900000	0.6848440000
C	-4.4187400000	2.6291790000	-1.3266200000
H	-3.9213720000	4.4811700000	-2.3083440000
C	-5.8729740000	2.4232240000	-1.6806430000
C	-5.0527240000	0.6315070000	-0.2447770000
C	-6.3342550000	1.3577940000	-0.6529580000
H	-6.7811920000	1.8571600000	0.2185760000
H	-7.0910570000	0.6808970000	-1.0584510000
H	-5.9755300000	2.0360130000	-2.7037050000
H	-6.4586740000	3.3467520000	-1.6303960000

Sum of electronic and zero-point Energies = -1683.883953

S11.3 Calculated structure of *E*-1 (conformer A)

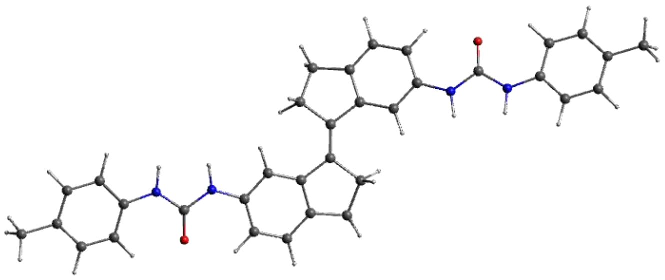


Table S17. Cartesian coordinates of *E*-**1** (approximate width × height × depth = 26.9 × 10.2 × 2.8 Å).

atom	X	Y	Z
C	1.7307360000	2.4455110000	-0.8581430000
C	1.6501510000	1.0686670000	-0.5668060000
C	0.3615640000	3.0232440000	-1.1207740000
C	-0.6039440000	1.9032300000	-0.6565820000
H	0.2273660000	3.2315030000	-2.1906950000
H	0.1930620000	3.9688890000	-0.5944620000
C	0.2418030000	0.6338950000	-0.5797530000
H	-1.4559230000	1.7864390000	-1.3321670000
H	-1.0104210000	2.1500260000	0.3341140000
C	2.9642930000	3.0830780000	-0.8889100000
C	4.1393680000	2.3780140000	-0.6207050000
C	4.0668470000	1.0096400000	-0.3109790000
C	2.8229460000	0.3590410000	-0.2835020000
H	3.0281880000	4.1438660000	-1.1176240000
H	5.1016020000	2.8678410000	-0.6376280000
H	2.7850880000	-0.6897100000	-0.0104510000
C	-2.9642860000	-3.0830350000	-0.8890290000
C	-4.1393650000	-2.3779770000	-0.6208240000
H	-5.1015980000	-2.8678060000	-0.6377670000
C	-4.0668490000	-1.0096120000	-0.3110630000
C	-2.8229490000	-0.3590120000	-0.2835520000
N	5.1973210000	0.2210930000	-0.0055450000
C	6.5296460000	0.6070880000	0.0516980000
N	7.3613870000	-0.4460710000	0.4046280000
C	8.7647880000	-0.4394010000	0.5522650000
O	6.9250110000	1.7412930000	-0.1826970000
C	9.3723340000	-1.6428400000	0.9415430000
C	9.5719080000	0.6871470000	0.3363880000
C	10.9512800000	0.5845210000	0.5096270000
C	10.7510330000	-1.7214050000	1.1083700000
C	11.5736270000	-0.6093010000	0.8925950000
C	13.0737000000	-0.7017300000	1.0410440000
H	13.5028480000	0.2517150000	1.3638680000

H	13.5553750000	-0.9684740000	0.0916330000
H	13.3580790000	-1.4636510000	1.7733150000
H	9.1156970000	1.6215230000	0.0441740000
H	11.5597710000	1.4698500000	0.3409770000
H	8.7595420000	-2.5245240000	1.1197800000
H	11.1925570000	-2.6667620000	1.4137680000
H	4.9830960000	-0.7459300000	0.1878470000
H	6.9311790000	-1.3379360000	0.5996190000
C	-6.5296460000	-0.6070850000	0.0516520000
N	-7.3614020000	0.4460800000	0.4045310000
C	-8.7647970000	0.4393820000	0.5522210000
O	-6.9249950000	-1.7413120000	-0.1826660000
C	-9.5718880000	-0.6872200000	0.3365310000
C	-9.3723670000	1.6428540000	0.9413660000
C	-10.7510600000	1.7213950000	1.1082450000
C	-10.9512570000	-0.5846180000	0.5098200000
C	-11.5736260000	0.6092320000	0.8926570000
C	-13.0736950000	0.7016400000	1.0411630000
H	-13.3580500000	1.4633410000	1.7736740000
H	-13.5028490000	-0.2518970000	1.3637020000
H	-13.5553850000	0.9686890000	0.0918450000
H	-9.1156580000	-1.6216180000	0.0444170000
H	-8.7595970000	2.5245820000	1.1194580000
H	-11.1926020000	2.6667780000	1.4135370000
H	-11.5597260000	-1.4699900000	0.3413160000
H	-6.9312170000	1.3379880000	0.5993750000
N	-5.1973290000	-0.2210700000	-0.0056360000
H	-4.9831070000	0.7459480000	0.1877850000
C	-1.6501490000	-1.0686310000	-0.5668570000
H	-2.7850950000	0.6897320000	-0.0104750000
C	-1.7307300000	-2.4454680000	-0.8582290000
H	-3.0281770000	-4.1438170000	-1.1177700000
C	-0.3615540000	-3.0231940000	-1.1208550000
C	-0.2418010000	-0.6338580000	-0.5797730000
C	0.6039470000	-1.9031910000	-0.6566220000
H	1.0104100000	-2.1500110000	0.3340740000
H	1.4559360000	-1.7863830000	-1.3321920000
H	-0.2273410000	-3.2314270000	-2.1907790000
H	-0.1930590000	-3.9688520000	-0.5945640000

Sum of electronic and zero-point Energies = -1683.887605

S11.4 Calculated structure of *E*-1, (conformer B)

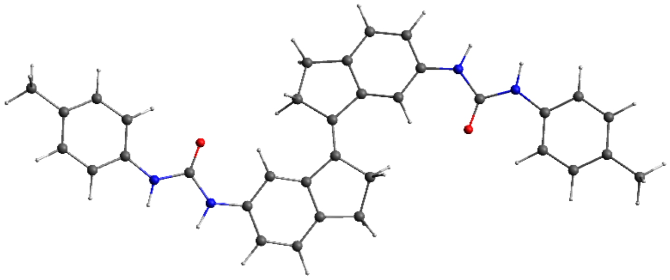


Table S18. Cartesian coordinates of *E*-**1** (approximate width × height × depth = 24.7 × 10.3 × 2.8 Å).

atom	X	Y	Z
C	-0.6788410000	-2.9166680000	-0.7954180000
C	-1.1267910000	-1.6087630000	-0.5203710000
C	0.8103000000	-2.9423890000	-1.0363010000
C	1.2801190000	-1.5350630000	-0.5893440000
C	0.0174970000	-0.6784450000	-0.5270240000
C	-1.5825220000	-3.9717170000	-0.8331920000
C	-2.9341830000	-3.7314300000	-0.5926510000
C	-3.3881740000	-2.4337490000	-0.3016880000
C	-2.4827830000	-1.3649790000	-0.2556370000
C	1.5825380000	3.9717620000	-0.8331180000
C	2.9341930000	3.7314710000	-0.5925460000
C	3.3881720000	2.4337870000	-0.3015790000
C	2.4827780000	1.3650210000	-0.2555530000
N	-4.7682980000	-2.2877990000	-0.0332130000
C	-5.5009660000	-1.1139240000	0.0688470000
N	-6.8442440000	-1.3595730000	0.3201040000
C	-7.8874010000	-0.4235440000	0.4804490000
O	-5.0212820000	0.0067240000	-0.0436320000
C	-9.1582900000	-0.9286310000	0.7953970000
C	-7.7287390000	0.9618180000	0.3299690000
C	-8.8296160000	1.7999530000	0.5000860000
C	-10.2424450000	-0.0729670000	0.9598780000
C	-10.1023650000	1.3130020000	0.8190670000
C	-11.2717200000	2.2452690000	1.0299940000
C	5.5009680000	1.1139370000	0.0687810000
N	6.8442210000	1.3595560000	0.3202080000
C	7.8873820000	0.4235040000	0.4803990000
O	5.0213170000	-0.0066920000	-0.0440190000
C	7.7288040000	-0.9618010000	0.3293590000
C	9.1581980000	0.9285240000	0.7957730000
C	10.2423530000	0.0728450000	0.9601220000
C	8.8296880000	-1.7999590000	0.4993550000
C	10.1023540000	-1.3130820000	0.8187520000

C	11.2717280000	-2.2453560000	1.0295380000
N	4.7682830000	2.2878240000	-0.0330380000
C	1.1267950000	1.6088060000	-0.5203260000
C	0.6788550000	2.9167140000	-0.7953760000
C	-0.8102800000	2.9424390000	-1.0363010000
C	-0.0174930000	0.6784880000	-0.5270200000
C	-1.2801140000	1.5351070000	-0.5893750000
H	1.0266380000	-3.1046930000	-2.1008260000
H	1.3139390000	-3.7467790000	-0.4894570000
H	2.0332730000	-1.1185850000	-1.2626550000
H	1.7478440000	-1.5954080000	0.4021720000
H	-1.2461920000	-4.9828430000	-1.0471920000
H	-3.6430960000	-4.5563500000	-0.6227160000
H	-2.8528610000	-0.3857800000	-0.0044270000
H	3.6431100000	4.5563880000	-0.6225910000
H	-11.1666740000	3.1583050000	0.4361760000
H	-11.3580040000	2.5503740000	2.0807410000
H	-12.2182240000	1.7708280000	0.7526280000
H	-6.7563510000	1.3629950000	0.0847960000
H	-8.6866660000	2.8709200000	0.3771600000
H	-9.3001540000	-2.0018910000	0.9077810000
H	-11.2153170000	-0.4942280000	1.2008360000
H	-5.2820350000	-3.1564440000	-0.0373630000
H	-7.1216510000	-2.3174680000	0.4755130000
H	12.2181250000	-1.7712020000	0.7513020000
H	11.1662060000	-3.1587400000	0.4363460000
H	11.3586290000	-2.5498110000	2.0804230000
H	6.7564750000	-1.3629250000	0.0838670000
H	9.2999950000	2.0017470000	0.9085970000
H	11.2151650000	0.4940530000	1.2014200000
H	8.6868040000	-2.8708850000	0.3759950000
H	7.1215520000	2.3174000000	0.4760610000
H	5.2820250000	3.1564680000	-0.0371280000
H	2.8528470000	0.3858170000	-0.0043460000
H	1.2462150000	4.9828900000	-1.0471200000
H	-1.7478760000	1.5954400000	0.4021240000
H	-2.0332410000	1.1186350000	-1.2627190000
H	-1.0265880000	3.1047560000	-2.1008290000
H	-1.3139330000	3.7468220000	-0.4894600000

Sum of electronic and zero-point Energies = -1683.888335

S11.5 Calculated structure of *E*-S2@2$H_2PO_4^-$ (conformer A)

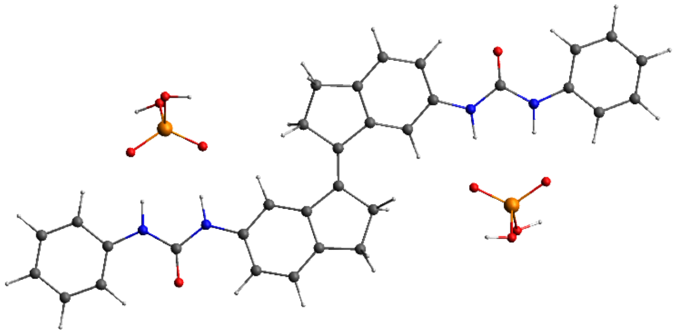


Table S19. Cartesian coordinates of *E*-**S2**@2$H_2PO_4^-$ (conformer 1)

atom	X	Y	Z
C	1.3851010000	-2.6672590000	-0.6297560000
C	1.4996610000	-1.2774600000	-0.4113940000
C	-0.0552660000	-3.0753900000	-0.8214480000
C	-0.8588490000	-1.8005080000	-0.4592270000
C	0.1573110000	-0.6605500000	-0.4142580000
C	2.5234400000	-3.4615890000	-0.6648860000
C	3.7920570000	-2.9052350000	-0.4834570000
C	3.9201630000	-1.5214440000	-0.2560660000
C	2.7666080000	-0.7123930000	-0.2142370000
C	-2.5234400000	3.4615880000	-0.6648890000
C	-3.7920570000	2.9052340000	-0.4834590000
C	-3.9201630000	1.5214440000	-0.2560660000
C	-2.7666080000	0.7123920000	-0.2142370000
N	5.1509210000	-0.8636620000	-0.0673850000
C	6.4028750000	-1.4389150000	-0.0169610000
N	7.4164710000	-0.4986560000	0.1267330000
C	8.7846820000	-0.7546170000	0.2539480000
O	6.6127160000	-2.6539330000	-0.0918210000
C	9.6384540000	0.3692640000	0.3207680000
C	9.3581320000	-2.0404140000	0.3304230000
C	10.7393580000	-2.1782900000	0.4703170000
C	11.0121720000	0.2063620000	0.4601300000
C	11.5818580000	-1.0681950000	0.5371300000
C	-6.4028750000	1.4389150000	-0.0169610000
N	-7.4164710000	0.4986560000	0.1267330000
C	-8.7846820000	0.7546170000	0.2539490000
O	-6.6127160000	2.6539320000	-0.0918210000
C	-9.3581320000	2.0404140000	0.3304270000
C	-9.6384540000	-0.3692640000	0.3207680000
C	-11.0121720000	-0.2063620000	0.4601310000
C	-10.7393580000	2.1782900000	0.4703230000
C	-11.5818580000	1.0681950000	0.5371340000

N	-5.1509210000	0.8636620000	-0.0673840000
C	-1.4996610000	1.2774600000	-0.4113950000
C	-1.3851010000	2.6672580000	-0.6297580000
C	0.0552660000	3.0753890000	-0.8214510000
C	-0.1573110000	0.6605500000	-0.4142590000
C	0.8588490000	1.8005070000	-0.4592290000
P	5.7502910000	2.8698040000	0.2719310000
O	5.7139340000	3.8042470000	1.6330600000
O	7.1615720000	2.3386460000	0.1510490000
O	4.5799780000	1.9141540000	0.2549260000
O	5.4647640000	3.9994340000	-0.8949070000
P	-5.7502910000	-2.8698030000	0.2719320000
O	-5.4647550000	-3.9994330000	-0.8949040000
O	-4.5799790000	-1.9141530000	0.2549350000
O	-7.1615710000	-2.3386450000	0.1510400000
O	-5.7139440000	-3.8042460000	1.6330610000
H	-0.3379180000	-3.9343100000	-0.2008720000
H	-0.2424660000	-3.3733680000	-1.8634170000
H	-1.3499110000	-1.9233620000	0.5138940000
H	-1.6632880000	-1.6053380000	-1.1735700000
H	2.4365390000	-4.5339660000	-0.8325720000
H	4.6808860000	-3.5180050000	-0.5024210000
H	2.9020200000	0.3407810000	-0.0080800000
H	-4.6808860000	3.5180050000	-0.5024240000
H	8.7110760000	-2.9032440000	0.2744530000
H	11.1592780000	-3.1808190000	0.5275700000
H	9.1935470000	1.3576970000	0.2615720000
H	11.6450460000	1.0901740000	0.5082870000
H	5.0649890000	0.1557060000	0.0655630000
H	7.1780590000	0.5054450000	0.1080570000
H	-8.7110760000	2.9032440000	0.2744590000
H	-9.1935480000	-1.3576970000	0.2615700000
H	-11.6450460000	-1.0901740000	0.5082880000
H	-11.1592770000	3.1808190000	0.5275780000
H	-7.1780590000	-0.5054450000	0.1080540000
H	-5.0649890000	-0.1557060000	0.0655670000
H	-2.9020200000	-0.3407810000	-0.0080790000
H	-2.4365390000	4.5339660000	-0.8325760000
H	1.6632880000	1.6053370000	-1.1735720000
H	1.3499110000	1.9233620000	0.5138920000
H	0.3379180000	3.9343100000	-0.2008760000
H	0.2424660000	3.3733670000	-1.8634200000
H	6.5902940000	3.7033220000	2.0275870000
H	4.6056650000	3.7636760000	-1.2705120000
H	-4.6056540000	-3.7636740000	-1.2705040000
H	-6.5903070000	-3.7033220000	2.0275810000
H	-12.6566820000	1.1914100000	0.6454170000

H	12.6566820000	-1.1914110000	0.6454110000

Sum of electronic and zero-point Energies = -2892.517617

S11.6 Calculated structure of *E*-S2@2$H_2PO_4^-$ (conformer B)

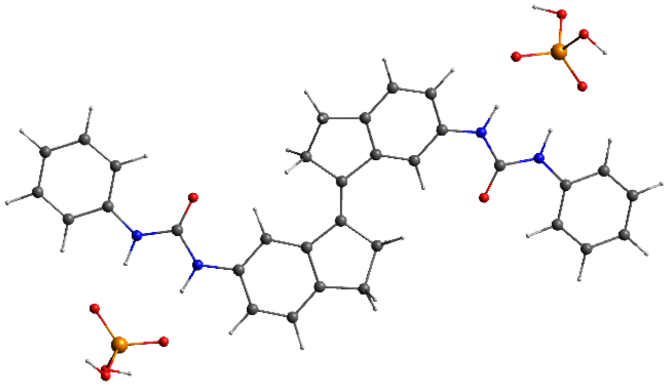


Table S20. Cartesian coordinates of *E*-**S2**@2$H_2PO_4^-$ (conformer 2)

atom	X	Y	Z
C	1.6079490000	2.5315980000	0.7449160000
C	1.6047820000	1.1464790000	0.4864990000
C	0.2092750000	3.0490810000	0.9759150000
C	-0.6971700000	1.8687520000	0.5456550000
C	0.2144800000	0.6441480000	0.4918160000
C	2.8101210000	3.2308300000	0.7806040000
C	4.0138710000	2.5674500000	0.5580780000
C	4.0260420000	1.1818680000	0.2882990000
C	2.8132700000	0.4709070000	0.2442360000
C	-2.8101230000	-3.2308450000	0.7805450000
C	-4.0138730000	-2.5674610000	0.5580310000
C	-4.0260430000	-1.1818750000	0.2882780000
C	-2.8132720000	-0.4709130000	0.2442280000
N	5.2914570000	0.6083820000	0.0616480000
C	5.5795150000	-0.7271630000	-0.1304790000
N	6.9391590000	-0.9600870000	-0.2989490000
C	7.5576670000	-2.1944930000	-0.5245660000
O	4.7295210000	-1.6216710000	-0.1530560000
C	8.9617610000	-2.1824850000	-0.6786540000
C	6.8813200000	-3.4284180000	-0.6064920000
C	7.6016940000	-4.6010260000	-0.8357810000
C	9.6578000000	-3.3644830000	-0.9063090000
C	8.9884680000	-4.5891100000	-0.9879890000
C	-5.5795140000	0.7271640000	-0.1304830000
N	-6.9391600000	0.9600940000	-0.2989310000
C	-7.5576630000	2.1944990000	-0.5245630000

O	-4.7295160000	1.6216680000	-0.1530720000
C	-6.8813120000	3.4284190000	-0.6065130000
C	-8.9617590000	2.1824930000	-0.6786440000
C	-9.6577940000	3.3644900000	-0.9063130000
C	-7.6016820000	4.6010260000	-0.8358180000
C	-8.9884570000	4.5891140000	-0.9880170000
N	-5.2914590000	-0.6083830000	0.0616390000
C	-1.6047830000	-1.1464890000	0.4864790000
C	-1.6079510000	-2.5316130000	0.7448700000
C	-0.2092770000	-3.0491010000	0.9758590000
C	-0.2144820000	-0.6441590000	0.4918040000
C	0.6971680000	-1.8687640000	0.5456200000
H	-0.0038380000	3.9667520000	0.4142130000
H	0.0517130000	3.2928410000	2.0367510000
H	-1.1279370000	2.0673460000	-0.4446610000
H	-1.5434840000	1.7267380000	1.2224560000
H	2.8190840000	4.3009250000	0.9799150000
H	4.9592890000	3.1004400000	0.5773800000
H	2.8475850000	-0.5780810000	0.0049290000
H	-4.9592910000	-3.1004510000	0.5773230000
H	5.8078240000	-3.4443370000	-0.4898570000
H	7.0582630000	-5.5419080000	-0.8955540000
H	9.4794970000	-1.2304900000	-0.6149310000
H	10.7390160000	-3.3250240000	-1.0219870000
H	6.0668420000	1.2869540000	0.0812290000
H	7.5950430000	-0.1663550000	-0.2552880000
H	-5.8078150000	3.4443360000	-0.4898830000
H	-9.4794980000	1.2305020000	-0.6149030000
H	-10.7390110000	3.3250340000	-1.0219840000
H	-7.0582480000	5.5419060000	-0.8956080000
H	-7.5950490000	0.1663690000	-0.2552240000
H	-6.0668420000	-1.2869560000	0.0811980000
H	-2.8475860000	0.5780790000	0.0049390000
H	-2.8190860000	-4.3009440000	0.9798360000
H	1.5434830000	-1.7267630000	1.2224220000
H	1.1279330000	-2.0673400000	-0.4447010000
H	0.0038360000	-3.9667620000	0.4141410000
H	-0.0517150000	-3.2928790000	2.0366910000
P	8.6610820000	2.3956780000	-0.0075640000
O	9.5389310000	2.9811690000	1.2632740000
O	9.0762910000	0.9532660000	-0.1888610000
O	7.1984520000	2.7441630000	0.1419050000
O	9.2479590000	3.3329380000	-1.2337070000
H	10.0069650000	2.2113350000	1.6131250000
H	8.4751880000	3.8015520000	-1.5767240000
P	-8.6610860000	-2.3956680000	-0.0075430000
O	-9.5388920000	-2.9812490000	1.2632830000

O	-9.0762860000	-0.9532370000	-0.1887070000
O	-7.1984530000	-2.7441810000	0.1418400000
O	-9.2480210000	-3.3328240000	-1.2337380000
H	-10.0068950000	-2.2114360000	1.6132180000
H	-8.4752700000	-3.8014260000	-1.5768160000
H	-9.5359220000	5.5113230000	-1.1664490000
H	9.5359360000	-5.5113200000	-1.1664100000

Sum of electronic and zero-point Energies = -2892.519619

S12. Diffusion data for tetrabutylammonium dihydrogen phosphate in $CDCl_3$

Table S21. Measured diffusion coefficients of $[Bu_4N][H_2PO_4]$ in $CDCl_3$.

$[Bu_4N][H_2PO_4]$	^{1}H δ	^{1}H Δ	^{31}P δ	^{31}P Δ	*D*($CHCl_3$)[a]	*D*(Bu_4N)[b]	(+/-)	*D*(H_2PO_4)[c]	(+/-)
		/ ms					/ 10^{10} m^2s^{-1}		
0	2	14			24				
2	2	65	4	65	26	7.2	2.2	5.9	0.6
10	2	65	4	65	23	5.7	0.2	5.7	0.3
50	2.5	65	6	65	-[d]	4.3	0.1	3.9	0.2
100	3	65	6.5	65	-[d]	3.7	0.2	3.3	0.2
300	5	65	8	65	19	2.4	0.1	2.2	0.1

Experimentally obtained values: [a] ^{1}H PGSTE experiment, δ = 4 ms, Δ = 14 ms, $g = 0 – 50$ G cm^{-1}, in all cases fitted using peak intensities for consistency. [b] ^{1}H PGSTE experiment using the δ and Δ values given in the table and $g = 0 – 50$ G cm^{-1}. [c] ^{31}P PGSTE using the δ and Δ values given in the table and $g = 0 – 50$ G cm^{-1}. [d] Due to peak overlap between $CHCl_3$ and OH, these values could not be accurately determined. All errors are given as the largest of either the standard deviation of two replicates or 3% of the diffusion coefficient.

S12.1 Viscosity calibration for $[Bu_4N][H_2PO_4]$ in $CDCl_3$

Viscosity changes were corrected for indirectly by assuming a linear relationship between concentration and viscosity, eg $\eta/\eta_0 = 1 + b\times[Bu4N.H2PO4]$. See section S6.1 for details.

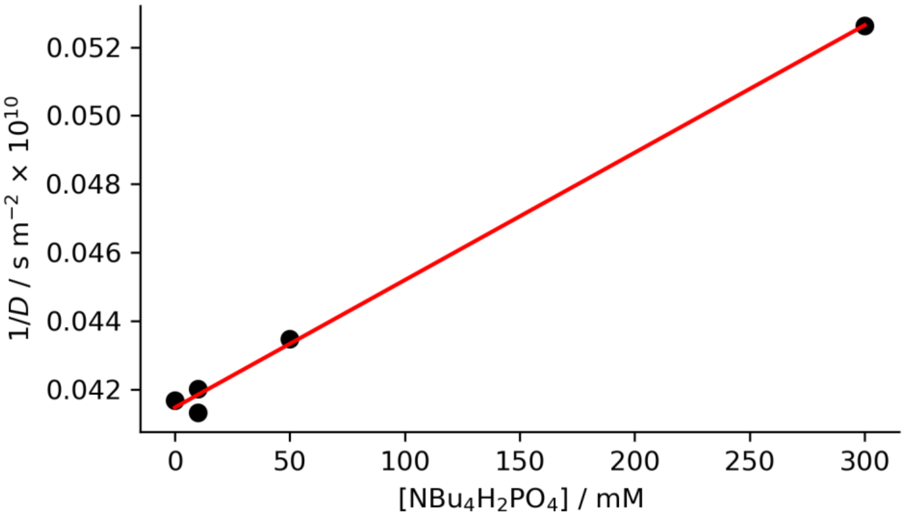


Figure S47. Indirect viscosity calibration for tetrabutylammonium dihydrogen phosphate in $CDCl_3$, using the diffusion coefficient of the $CHCl_3$ solvent residual peak. Linear fit for implied relative change in dynamic viscosity: $\frac{\eta}{\eta_0} = 1 + (0.899 \times 10^{-3})([\mathrm{NBu_4OAc}]$ / mM)

S12.2 Fitted isodesmic self-association of $[Bu_4N][H_2PO_4]$ in $CDCl_3$

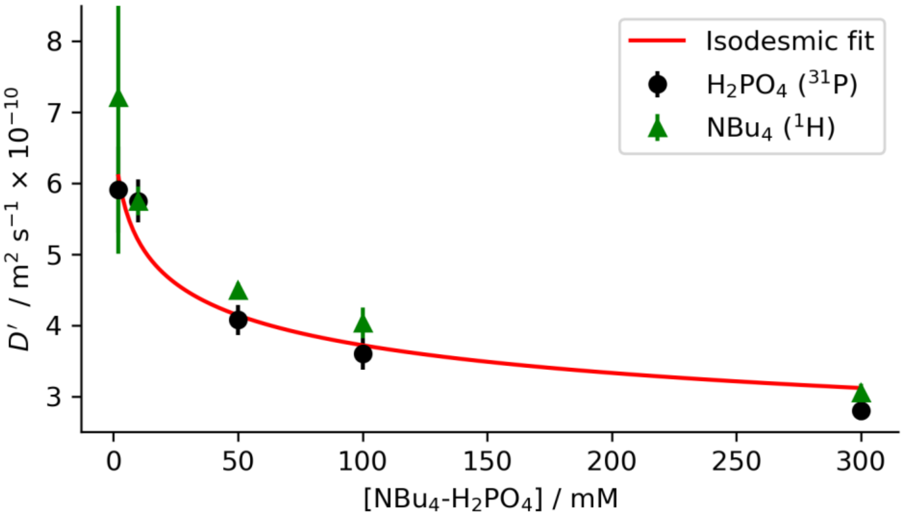


Figure S48. Measured diffusion coefficients of tetrabutylammonium and dihydrogen phosphate corrected for changes in viscosity using the relationship shown in **Figure S47**. The model for isodesmic diffusion shown in Eq (10) has been regressed onto measured $D'(H_2PO_4)$, with obtained parameters K_i = (180 ± 190) M^{-1} and D_0 = (6.7 ± 0.9) × 10^{-10} m^2 s^{-1}. Unlike the concentration-independent results obtained for $D'(NBu_4)$ in DMSO (**Figure S9**), in $CDCl_3$ $D'(NBu_4)$ displays a similar concentration dependence to $D'(H_2PO_4)$. We ascribe this to the stronger ion pairing in $CDCl_3$ than in DMSO.

The isodesmic self-association model described previously (SI S5) was fitted to the viscosity-corrected diffusion coefficients *D'* of $[NBu_4]$ and $[H_2PO_4]$ in $CDCl_3$, with results shown in **Figure S48**. The isodesmic association constant was obtained as $K_i(CDCl_3)$ = 180 ± 190 M^{-1} (taking the diagonal-squared σ_{xx}^2 term of the covariance matrix of the nonlinear fit as the error), a value well within error of the constant K_i(DMSO) = (120 ± 32) M^{-1} obtained for self-association in DMSO.

Despite the large difference in size between the free ionic species, $D'(NBu_4)$ is close to $D'(H_2PO_4)$ at all concentrations measured and shows a similar concentration-dependent decrease. This contrasts with the concentration-independent $D'(NBu_4)$ observed in DMSO (**Figure S9**) and is suggestive of tight association between the NBu_4^+ cation and the $H_2PO_4^-$ anion in $CDCl_3$. This ion pairing may also be responsible for the surprisingly low diffusion coefficient of the 'H_2PO_4' monomer obtained in **Figure S48** of $D_0(CDCl_3)$ = (6.7 ± 0.9) × 10^{-10} m^2 s^{-1}, which is barely higher than the value found in DMSO of D_0(DMSO) = (3.39 ± 0.11) × 10^{-10} m^2 s^{-1} despite the four-fold different in viscosity between these two solvents. The fitted D_0 obtained in DMSO may represent a true D_0 of the H_2PO_4 monomer, while the 'D_0' obtained for $CDCl_3$ may represent a tightly ion-paired electrostatically neutral species.

S13. Scripts used for data processing

S13.1 Analysis of association data from chemical shifts

Script S1. Python code used for fitting association constants to chemical shift data.

```
import pandas as pd
import numpy as np
import scipy.optimize
import matplotlib.pyplot as plt
from lmfit import minimize, Parameters, report_fit
import itertools
import matplotlib.cm as cm
import seaborn as sns



#Global fit for 2:1 H:G complex data



def Global21(data,figsize=(8,6)):
    '''Fitting function for 2:1 Host:Guest stoichiometry.
    This function calculates association constants from NMR titration data as a Pandas dataframe with the following columns
    [Host concentration / Molar] [Guest concentration / Molar] [Peak 1 / ppm] [ Peak 2 / ppm]...
    ...for an arbitry number of peaks. If desired, this can be called from the clipboard using pd.read_clipboard(), eg:
    params = Global21(pd.read_clipboard())

    Binding parameters are calculated using a global nonlinear least-squares fit over all provided peaks.
    '''
    def H21(G0,H0,K1,K2):
        '''H21 calculate the concentration of free host, [H], given G0 (total gues concentration), H0 (total host concentration), K1 (H + G -> [HG] association constant), and K2 (H + [HG] -> [H2G] association constant).
        '''
        G0,H0,K1,K2 = np.asarray(G0), np.asarray(H0), np.asarray(K1), np.asarray(K2)
        try:
            length = len(G0)
            coeffs = np.zeros((length,4))
            for i in range(length):
                coeffs[i] = [K1*K2, K1*(2*K2*G0[i] - K2*H0[i] + 1), K1*(G0[i]-H0[i])+1, -H0[i]]
            roots = np.zeros((len(G0),3))
            R = np.zeros(len(G0))
            for i in range(len(G0)):
                roots[i] = np.roots(coeffs[i])
                R[i] = min(roots[i][roots[i] >= 0])
```

```
            return R
        except TypeError:
            coeffs = [K1*K2, K1*(2*K2*G0 - K2*H0 + 1), K1*(G0-H0)+1, -H0]
            R =  np.roots(coeffs)
        return min(R[R >= 0])

    def nmr21(X,K1,K2,delHG,delH2G):
        '''nmr21: calculates chemical shift difference delDel for a given peak from the free host chemical shift as a function of X = H0, G0, K1, K2, delHG = chemical shift difference on first association, delH2G = chemical shift difference on second association.'''
        H0, G0 = X
        H = H21(G0,H0,K1,K2)
        delDel = (delHG*K1*G0*H + 2*delH2G*G0*K1*K2*H**2)/(H0*(1+K1*H+K1*K2*H**2))
        return delDel

    def nmr21_dataset(X,params,i):
        '''Given X = H0, G0 and params = lmfit parameters element for multiple peaks, calculates chemical shifts for each peak, for each [G0, H0] pairing'''
        K1 = params['K1_%i' % (i+1)].value
        K2 = params['K2_%i' % (i+1)].value
        delHG = params['delDHG1_%i' % (i+1)].value
        delHG2 = params['delDHG2_%i' % (i+1)].value
        return nmr21(X,K1,K2,delHG,delHG2)

    def objective(params, X, data):
        """Calculate total residual for fits to several data sets"""
        dataT = np.array(data.T[2:])
        ndata, nx = dataT.shape
        resid = 0.0*dataT[:]
        # make residual per data set
        for i in range(ndata):
            dataT[i] = dataT[i] - dataT[i][0]
            resid[i, :] = dataT[i, :] - nmr21_dataset(X,params,i)
        # now flatten this to a 1D array, as minimize() needs
        return resid.flatten()

    X = data.iloc[:,0],data.iloc[:,1]

    dataT = np.array(data.T[2:])
    fit_params = Parameters()
    for iy, y in enumerate(dataT):
        fit_params.add('K1_%i' % (iy+1), value=1000, min=0.0,  max=1e6)
        fit_params.add('K2_%i' % (iy+1), value=100, min=0.0,  max=1e4)
        fit_params.add('delDHG1_%i' % (iy+1), value=1.0, min=-4.0,  max=4.0)
        fit_params.add('delDHG2_%i' % (iy+1), value=-1.0, min=-4.0,  max=4.0)
```

```
    for iy in range(2,len(dataT)+1):
        fit_params['K1_%i' % iy].expr='K1_1'
        fit_params['K2_%i' % iy].expr='K2_1'


    result = minimize(objective, fit_params, args=(X, data),nan_policy='omit')
    #report_fit(result.params)

    palette = itertools.cycle(sns.color_palette())
    fig, (ax1, ax2, ax3) = plt.subplots(3,1,sharex=True,figsize=figsize)
    plt.rcParams["lines.markeredgewidth"] = 1.5
    interp = np.linspace(min(X[1]),max(X[1]),1000)
    Xinterp = (interp*0.0+X[0][0]), interp

    K1, K2 = result.params['K1_1'].value, result.params['K2_1'].value
    Hinterp = H21(Xinterp[1],Xinterp[0],K1,K2)
    #Ginterp = (-K1+np.sqrt(K1**2 - 8*K1*K2*(1-Xinterp[0]/Hinterp)))/(4*K1*K2)
    Ginterp = Xinterp[1]/(1+K1*Hinterp + K1*K2*Hinterp**2)
    HGinterp = K1*Hinterp*Ginterp
    H2Ginterp = K2*Hinterp*HGinterp


    for i in range(len(dataT)):
        c = next(palette)
        y_fit = nmr21_dataset(X, result.params, i)
        Yinterp = nmr21_dataset(Xinterp, result.params, i)
        ax1.plot(X[1]/X[0], dataT[i, :]-dataT[i,0],'o',color=c,markerfacecolor='None')
        ax1.plot(Xinterp[1]/Xinterp[0], Yinterp, '-', color=c)
        ax2.plot(X[1]/X[0],dataT[i,:]-dataT[i,0]-y_fit, 'o--',color=c,markerfacecolor='None')
        ax1.set_ylabel('Δδ')
        ax2.set_ylabel('Residual')

        ax2.axhline(y=0,color='k')
    ax3.plot(Xinterp[1]/Xinterp[0], Hinterp/Xinterp[0],label='[H]')
    ax3.plot(Xinterp[1]/Xinterp[0], HGinterp/Xinterp[0],label='[HG]')
    ax3.plot(Xinterp[1]/Xinterp[0], H2Ginterp/Xinterp[0],label='[H2G]')
    ax3.legend(loc='best',frameon=True,framealpha=0.5)
    ax3.set_xlabel('Guest equiv ([G]0 / [H]0)')
    ax3.set_ylabel('Molfraction')

    safeVal = lambda x: "ERROR" if result.params[x].value==None else result.params[x].value
    safeErr = lambda x: "ERROR" if result.params[x].stderr==None else result.params[x].stderr

    ax1.set_title("K1: {:.0f} +/- {:.0f}\n K2: {:.0f} +/- {:.0f}".format(safeVal("K1_1"),
                                                                         safeErr('K1_1'),
                                                                         safeVal("K2_1"),
```

```
                                                                         safeErr("K2_1")))

    plt.tight_layout()
    fig.align_labels()
    plt.show()
    return result, fig


#Global fit for 1:2 H:G complex data
def Global12(data,figsize=(8,6),alpha=None,fmt='o--'):
'''data:    A pandas dataframe (with header) containining host concentrations in the first column, guest
concentrations in the second column, and chemical shifts in subsequent columns.
figsize:    Optional. a tuple passed to matplotlib for plotting of association data.
alpha:      Optional. Cooperativity used for fitting 1:2 model. Allowed to float if set to None (default).
For a non-cooperative 1:2 fit, set alpha=1.
fmt:        Formatting string used for residuals of plot.'''
    def G12(G0,H0,K1,K2):
        G0,H0,K1,K2 = np.asarray(G0), np.asarray(H0), np.asarray(K1), np.asarray(K2)
        try:
            length = len(G0)
            coeffs = np.zeros((length,4))
            for i in range(length):
                coeffs[i] = [K1*K2, K1*(2*K2*H0[i] - K2*G0[i] + 1), K1*(H0[i]-G0[i])+1, -G0[i]]
            roots = np.zeros((len(G0),3))
            R = np.zeros(len(G0))
            for i in range(len(G0)):
                roots[i] = np.roots(coeffs[i])
                R[i] = min(roots[i][roots[i] >= 0])
            return R
        except TypeError:
            coeffs = [K1*K2, K1*(2*K2*H0 - K2*G0 + 1), K1*(H0-G0)+1, -G0]
            R =  np.roots(coeffs)
        return min(R[R >= 0])

    def nmr12(X,K1,K2,delHG,delHG2):
        H0, G0 = X
        G = G12(G0,H0,K1,K2)
        delDel = (delHG*K1*G + delHG2*K1*K2*G**2)/(1+K1*G+K1*K2*G**2)
        return delDel

    def nmr12_dataset(X,params,i):
        K1 = params['K1_%i' % (i+1)].value

        K2 = K1*params['alpha_%i' % (i+1)].value/4

        delHG = params['delDHG1_%i' % (i+1)].value
```

```
    delHG2 = params['delDHG2_%i' % (i+1)].value
    return nmr12(X,K1,K2,delHG,delHG2)

def objective(params, X, data):
    """ calculate total residual for fits to several data sets"""
    dataT = np.array(data.T[2:])
    ndata, nx = dataT.shape
    resid = 0.0*dataT[:]
    # make residual per data set
    for i in range(ndata):
        dataT[i] = dataT[i] - dataT[i][0]
        resid[i,:] = dataT[i, :] - nmr12_dataset(X,params,i)
    # now flatten this to a 1D array, as minimize() needs
    return resid.flatten()

X = data.iloc[:,0],data.iloc[:,1]

dataT = np.array(data.T[2:])
fit_params = Parameters()
for iy, y in enumerate(dataT):
    fit_params.add('K1_%i' % (iy+1), value=1000, min=0.0,  max=1e5)
    if alpha == None:
        fit_params.add('alpha_%i' % (iy+1), value=1, min=0.0,  max=10)
    else:
        fit_params.add('alpha_%i' % (iy+1), value=alpha,vary=False)

    fit_params.add('delDHG1_%i' % (iy+1), value=1.0, min=-4.0,  max=4.0)
    fit_params.add('delDHG2_%i' % (iy+1), value=-1.0, min=-4.0,  max=4.0)

for iy in range(2,len(dataT)+1):
    fit_params['K1_%i' % iy].expr='K1_1'
    fit_params['alpha_%i' % iy].expr='alpha_1'

result = minimize(objective, fit_params, args=(X, data),nan_policy='omit')
#report_fit(result.params)

palette = itertools.cycle(sns.color_palette())
fig, (ax1, ax2, ax3) = plt.subplots(3,1,sharex=True,figsize=figsize)
plt.rcParams["lines.markeredgewidth"] = 1.5
interp = np.linspace(min(X[1]),max(X[1]),1000)
Xinterp = (interp*0.0+X[0][0]), interp

K1 = result.params['K1_1'].value
K2 = K1*result.params['alpha_1'].value/4

Ginterp = G12(Xinterp[1],Xinterp[0],K1,K2)
```

```
    Hinterp = Xinterp[0]/(1+K1*Ginterp + K1*K2 * Ginterp * Ginterp)
    HGinterp = K1*Ginterp * Hinterp
    HG2interp = K2 * HGinterp * Ginterp

    for i in range(len(dataT)):
        c = next(palette)
        y_fit = nmr12_dataset(X, result.params, i)
        Yinterp = nmr12_dataset(Xinterp, result.params, i)
        ax1.plot(X[1]/X[0], dataT[i, :]-dataT[i,0],'o',color=c,markerfacecolor='None')
        ax1.plot(Xinterp[1]/Xinterp[0], Yinterp, '-', color=c)
        ax2.plot(X[1]/X[0],dataT[i,:]-dataT[i,0]-y_fit, fmt,color=c,markerfacecolor='None')
        ax1.set_ylabel('Δδ')
        ax2.set_ylabel('Residual')

        ax2.axhline(y=0,color='k')
    ax3.plot(Xinterp[1]/Xinterp[0], Hinterp/Xinterp[0],label='[H]')
    ax3.plot(Xinterp[1]/Xinterp[0], HGinterp/Xinterp[0],label='[HG]')
    ax3.plot(Xinterp[1]/Xinterp[0], HG2interp/Xinterp[0],label='[HG2]')
    ax3.legend(loc='best',frameon=True,framealpha=0.5)
    ax3.set_xlabel('Guest equiv ([G]0 / [H]0)')
    ax3.set_ylabel('Molfraction')
    safeVal = lambda x: "ERROR" if result.params[x].value==None else result.params[x].value
    safeErr = lambda x: "ERROR" if result.params[x].stderr==None else result.params[x].stderr
    ax1.set_title("K1: {:.0f} +/- {:.0f}\n K2: {:.0f}, alpha: {:.2f} +/-
{:.2f}".format(safeVal("K1_1"),safeErr('K1_1'),

safeVal("K1_1")*safeVal("alpha_1")/4,

safeVal("alpha_1"),safeErr("alpha_1")))
    plt.tight_layout()
    fig.align_labels()
    plt.show()
    return result, fig

#GLOBAL ANALYSIS 1:1 complex
def Global11(data,figsize=(8,6),plotaxis='eq',fmt='o--'):
    import seaborn as sns
    def HG11(G0,H0,K):
            return 0.5*(G0 + H0 + 1/K - np.sqrt((G0-H0-1/K)**2 + 4*G0/K))
    def nmr11(X,K,delDHG):
        H0, G0 = X
        HG = HG11(G0,H0,K)
        out = delDHG*HG/H0
        return out
    def nmr11_dataset(X,params,i):
        K = params['K_%i' % (i+1)].value
```

```
    delDHG = params['delDHG_%i' % (i+1)].value
    return nmr11(X,K,delDHG)

def objective(params, X, data):
    """Calculate total residual for fits to several data sets"""
    dataT = np.array(data.T[2:])
    ndata, nx = dataT.shape
    resid = 0.0*dataT[:]
    # make residual per data set
    for i in range(ndata):
        dataT[i] = dataT[i] - dataT[i][0]
        resid[i, :] = dataT[i, :] - nmr11_dataset(X,params,i)
    # now flatten this to a 1D array, as minimize() needs
    return resid.flatten()

X = data.iloc[:,0],data.iloc[:,1]

dataT = np.array(data.T[2:])
fit_params = Parameters()
for iy, y in enumerate(dataT):
    fit_params.add('K_%i' % (iy+1), value=200, min=0.0,  max=1e6)
    fit_params.add('delDHG_%i' % (iy+1), value=0.4, min=-4.0,  max=4.0)
for iy in range(2,len(dataT)+1):
    fit_params['K_%i' % iy].expr='K_1'



result = minimize(objective, fit_params, args=(X, data),nan_policy='omit')



import matplotlib.cm as cm
import seaborn as sns
palette = itertools.cycle(sns.color_palette())
plt.rcParams["lines.markeredgewidth"] = 1.5
fig, (ax1, ax2, ax3) = plt.subplots(3,1,sharex=True,figsize=figsize)

interp = np.linspace(min(X[1]),max(X[1]),1000)
Xinterp = (interp*0.0+X[0][0]), interp

HGinterp = HG11(Xinterp[1],Xinterp[0],result.params['K_1'].value)
Hinterp = Xinterp[0] - HGinterp

for i in range(len(dataT)):
    c = next(palette)
    y_fit = nmr11_dataset(X, result.params, i)
    Yinterp = nmr11_dataset(Xinterp, result.params, i)
    if plotaxis =='eq':
```

```
            ax1.plot(X[1]/X[0], dataT[i, :]-dataT[i,0],'o', color=c,markerfacecolor='None')
            ax1.plot(Xinterp[1]/Xinterp[0], Yinterp, '-', color=c)
            ax2.plot(X[1]/X[0],dataT[i,:]-dataT[i,0]-y_fit, fmt,color=c,markerfacecolor='None')
        elif plotaxis =='abs':
            ax1.plot(X[1], dataT[i, :]-dataT[i,0],'o', color=c,markerfacecolor='None')
            ax1.plot(Xinterp[1], Yinterp, '-', color=c)
            ax2.plot(X[1],dataT[i,:]-dataT[i,0]-y_fit, fmt,color=c,markerfacecolor='None')
    if plotaxis == 'eq':
        ax3.plot(Xinterp[1]/Xinterp[0],Hinterp/Xinterp[0],label = '[H]')
        ax3.plot(Xinterp[1]/Xinterp[0],HGinterp/Xinterp[0],label = '[HG]')
        ax3.set_xlabel('Guest equiv ([G]0 / [H]0)')
    elif plotaxis == 'abs':
        ax3.plot(Xinterp[1],Hinterp/Xinterp[0],label = '[H]')
        ax3.plot(Xinterp[1],HGinterp/Xinterp[0],label = '[HG]')
        ax3.set_xlabel('Guest concentration / M')
    ax3.legend(loc='best',frameon=True,framealpha=0.5)
    ax1.set_ylabel('Δδ')
    ax2.set_ylabel('Residual')

    ax2.axhline(y=0,color='k')
    ax3.set_ylabel('Molfraction')
    safeVal = lambda x: "ERROR" if result.params[x].value==None else result.params[x].value
    safeErr = lambda x: "ERROR" if result.params[x].stderr==None else result.params[x].stderr
    ax1.set_title("Ka: {} +/- {}".format(safeVal('K_1'),safeErr('K_1')))
    plt.tight_layout()
    fig.align_labels()
    plt.show()
    return result, fig
```

S13.2 Processing and analysis of diffusion data

Script S 2. Functions used for analysis and processing of diffusion NMR data. MestReNova was used to obtain integrals of peaks of interest, and

```
'''Module for global fitting of (time-dependent) diffusion NMR data.
This module provides the following functions:
GlobalDiff(data)
    Global nonlinear fitting of the S-T equation for an arbitrary number of gradients and chemical shift environments.
    Returns an lmfit parameters object containing fitted I0 intensities for all peaks, as well as a single globally fitted D value.
MovingDiff(data,slicelength=10)
    Global moving average fit for diffusion, using GlobalDiff to obtain a D value for each time point.
S
    Takes as input a pandas dataframe with the first column containing B values, and subsequent columns containing integrals for the peaks of interest.
```

```
    Slicelength sets the number of experiments used for each D(time) point. Slicetime sets the time (in minutes) taken to acquire each gradient slice.
    Returns (Dpoints, I0points,Derr,Ierr): four pandas dataframes, each with indices corresponding to time. I0points and Ierr contain a column for each fitted NMR peak, while Dpoint and Der contain a single column with globally fitted diffusion data.
    For systems involving multiple chemical species, use SeparateMovingDiffusion to obtain individual (non-globally fitted) diffusion coefficients.
    Dpoints, I0points, Derr, and Ierr contain diffusion points, extrapolated intensities, and respective errors for each from fitting.
MovingDiff_csv(fname,slicelength=10,slicetime=2/3)
    A wrapper for MogvingDiff to act on a similarly formatted .csv file.
SeparateMovingDiffusion(data,slicelength=10,slicetime=2/3)
    Returns (Dpoints, I0points,Derr,Ierr): four pandas dataframes, each with indexes corresponding to time and a column for each peak.
    Dpoints, I0points, Derr, and Ierr contain diffusion points, extrapolated intensities, and respective errors for each from fitting.
SeparateMovingDiffusion_csv(fname,slicelength=10,slicetime=2/3)
    Generates a pair of pandas dataframes [D,I] containing calculation time-dependent diffusion coefficients and unattenuated integrals.
    Acts on a .csv file with the first column containing B-values, and each subsequent column containing the corresponding integrals for a particular chemical shift.
MeOHTemp(dDelta)
    Calculates temperature from methanol OH-CH3 chemical shift separation (in ppm)
MeOHDiff(dDelta)
    Calculates expected diffusion coefficient from methanol OH-CH3 chemical shift separation (in ppm)
'''


def GlobalDiff(data):
    '''Function to globally fit a single diffusion coefficient to data from a list of peaks
    Input: a pandas dataframe consisting of:

    B_0 I0_0 I1_0 ... In_0
    B_1 I0_1 I1_1 ... In_1
    ...
    B_m I0_m I1_m ... In_m
    where B is the list of B-parameters for all experiments, and each column In_ contains the integrals measured for a single peak.
    The function returns a single lmfit Parameters object.
    '''
    import numpy as np
    import pandas as pd
    from lmfit import minimize, Parameters, report_fit

    def STExp(B,I0,D):
        I0,B,D = np.asarray(I0), np.asarray(B), np.asarray(D)
        return I0*np.exp(-B*D)
```

```
    def STExp_dataset(B,params,i):
        I0 = params['I0_%i' % (i+1)].value
        D = params['D_%i' % (i+1)].value
        return STExp(B,I0,D)
    def objective(params,B,data):
        dataT = np.array(data.T[1:])
        ndata, nx = dataT.shape
        resid = 0.0*dataT[:]
        #Residual per data set:
        for i in range(ndata):
            resid[i,:] = dataT[i,:] - STExp_dataset(B,params,i)
        #Flatten to a 1d array:
        return resid.flatten()
    B = data.iloc[:,0]
    dataT = np.array(data.T[1:])
    fit_params = Parameters()
    I0guesses = data.max()[1:]
    for iy, y in enumerate(dataT):
        fit_params.add('D_%i' % (iy+1), value = 1e-9, min = 1e-12, max = 1e-8)
        fit_params.add('I0_%i' % (iy+1), value = I0guesses[iy], min = 1, max = 100*I0guesses[iy]) #Give
each I0 parameter a unique guessed I0
    for iy in range(2,len(dataT)+1):
        fit_params['D_%i' % iy].expr='D_1'
    return minimize(objective,fit_params,args=(B,data))

def MovingDiff(data,slicelength=10,slicetime=2/3):
    '''Fitting for time-dependent diffusion + concentration data.
    Input: pandas dataframe formatted as
    B_0 I0_0 I1_0 ... In_0
    B_1 I0_1 I1_1 ... In_1
    ...
    B_m I0_m I1_m ... In_m

    '''
    import numpy as np
    import pandas as pd
    from tqdm import tqdm_notebook as tqdm
    from lmfit import minimize, Parameters, report_fit

    npoints = data.shape[0]-slicelength
    ipoints = np.arange(0,npoints)
    tpoints = ipoints*slicetime+slicetime*slicelength/2

    cols = data.columns[1:]
    D = pd.DataFrame(index = tpoints,columns = [cols[0]])
    Derr = pd.DataFrame(index = tpoints,columns = [cols[0]])
```

```
    I0 = pd.DataFrame(index = tpoints,columns = cols)
    I0err = pd.DataFrame(index = tpoints,columns = cols)

    for i in tqdm(range(npoints),desc='Progress:',position=1,leave=False):
        params = GlobalDiff(data.iloc[i:i+slicelength])
        I0slice,I0errslice = [],[]
        for Ival in range(0,len(cols)):
            ParamName = 'I0_{}'.format(Ival+1)
            #I0.insert(params.params[ParamName].value,index=i,col)
            I0slice.append(params.params[ParamName].value)
            I0errslice.append(params.params[ParamName].stderr)
        D.loc[tpoints[i],cols[0]] = params.params['D_1'].value
        Derr.loc[tpoints[i],cols[0]] = params.params['D_1'].stderr
        I0.loc[tpoints[i],cols] = I0slice
        I0err.loc[tpoints[i],cols] = I0errslice
    return D,I0,Derr,I0err



def MovingDiff_csv(fname,slicelength=10,slicetime=2/3):
    '''A simple wrapper of  MovingDiff() to act on .csv files'''
    import pandas as pd
    return MovingDiff(pd.read_csv(fname),slicelength,slicetime)

def SeparateMovingDiffusion(data,slicelength=10,slicetime=2/3):
    '''Moving average diffusion processing for multiple separate chemical species.
    Acts on a pandas dataframe containing a list of B-values in teh first column, and corresponding peak
integrals in subsequent columns.
    Returns a pair of pandas dataframes [D,I] containing the calculated diffusion coefficients
    and concentrations for each peak present in the input array. '''
    import pandas as pd
    import numpy as np
    from tqdm import tqdm_notebook as tqdm
    npoints = data.shape[0]-slicelength
    ipoints = np.arange(0,npoints)
    tpoints = ipoints*slicetime+slicetime*slicelength/2
    D = pd.DataFrame(index = tpoints,columns=data.columns[1:])
    I = pd.DataFrame(index = tpoints,columns=data.columns[1:])
    Derr = pd.DataFrame(index = tpoints,columns=data.columns[1:])
    Ierr = pd.DataFrame(index = tpoints,columns=data.columns[1:])

    for peak in tqdm(data.columns[1:],desc='Peak-by-peak progress',position=2):
        td,ti,tderr,tierr =
MovingDiff(pd.concat([data.iloc[:,0],data[peak]],axis=1),slicelength=slicelength,slicetime=slicetime
)
        D[peak] = td
        I[peak] = ti
```

```
        Derr[peak] = tderr
        Ierr[peak] = tierr
    return D, I, Derr, Ierr
def SeparateMovingDiffusion_csv(fname,slicelength=10,slicetime=2/3):
    '''A wrapper of SeparateMovingDiffusion to act on .csv files'''
    import pandas as pd
    return SeparateMovingDiffusion(pd.read_csv(fname),slicelength=slicelength,slicetime=slicetime)

def MeOHTemp(dDelta):
    '''Converts a methanol CH3-OH chemical shift separation (in ppm) to a temperature (in K).
    See J. Magn. Reson. 1982, 46, 319-321'''
    return 409-36.54*dDelta-21.85*dDelta**2
def MeOHDiff(dDelta):
    '''Calculates the expected self-diffusion coefficient of methanol for a given OH-CH3 peak chemical shift separation.
    See MacDonald et al, ChemPhysChem 2019, 20, 926-930'''
    import numpy as np
    return 5.124e-7 * np.exp((-1601)/(MeOHTemp(dDelta)))

def MeanGlobularDiffusionD0(A0,K,D0,rho=1,dps=25):
    '''
    Function for modelling the concentration-dependent diffusion coefficient of a supramolecular oligomer.
    A0 = species concentration in M (scalar or array)
    K = association constant. If alpha = 1 (isodesmic), K is the association constant for each step of oligomerisation.
    D0 = diffusion coefficient of monomer
    rho = cooperativity of system.
        rho = 1: isodesmic model, with each oligomer associating with the same constant.
        rho > 1: dimer (n=2 oligomer) forms more easily (eg electrostatic repulsion), with K(dimer) = rho*K
        '''
    #from mpmath import mp, polylog,sqrt,nsum
    from mpmath import fp
    import numpy as np
    #mp.dps = dps #Set mpmath precision. dps = 25 seems to work well.
    import scipy.constants

    #a0k = np.array(A0)*np.array(K)

    try: #First, try to calculate an array of [A] free-monomer concentrations. This should work if A0 was given as an array
        length = len(A0)
        coeffs = np.zeros((length,4))
        for i in range(length):
            coeffs[i] = [1-rho,
```

```
                     2*rho-2-K*A0[i],
                     2*K*A0[i]+1,
                     -K*A0[i]]
        roots = np.zeros((length,3))
        R = np.zeros(length)
        for i in range(length):
            roots = np.roots(coeffs[i])
            R[i] = min(roots[roots >= 0])
        A = R/K

    except TypeError: #If A0 was given as a single float, not as an array, calculate a single [A]
        coeffs = [1-rho,
                  2*rho-2-K*A0,
                  2*K*A0+1,
                  -K*A0]
        R = np.roots(coeffs)
        A = min(R[R >= 0])/K

    try:
        if A == 0:
            sumval = D0 #If there is no monomer present (either A0 = 0 or K = 0, D = D0
        else:
            sumval = D0/(K*A0)*float(A*K + rho*(fp.polylog(-2/3,(K*A))-A*K))
    except:
        j = 0
        sumval = []
        while j<len(A):
            if A[j] == 0:
                sumval.append(D0)
            else:
                sumval.append(float(D0/(A0[j]*K)*(A[j]*K + rho*(fp.polylog(-2/3,K*A[j])-A[j]*K))))
            j+=1
    return sumval
```

S14. Additional References

[1] D. Barišić, V. Tomišić, N. Bregović, *Anal. Chim. Acta* **2019**, *1046*, 77-92.

[2] K. Izutsu, T. Adachi, T. Fujinaga, *Electrochim. Acta* **1970**, *15*, 135-145.

[3] S. J. Wezenberg, B. L. Feringa, *Org. Lett.* **2017**, *19*, 324-327.

[4] R. B. Martin, *Chem. Rev.* **1996**, *96*, 3043-3064.

[5] W. Sutherland, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1905**, *9*, 781-785.

[6] W. S. Price, F. Tsuchiya, Y. Arata, *J. Am. Chem. Soc.* **1999**, *121*, 11503-11512.

[7] P. Thordarson, *Chem. Soc. Rev.* **2011**, *40*, 1305-1323.

[8] T. S. C. MacDonald, W. S. Price, J. E. Beves, *ChemPhysChem* **2019**, *20*, 926-930.

[9] a) T. M. Barbosa, R. Rittner, C. F. Tormena, G. A. Morris, M. Nilsson, *RSC Adv.* **2016**, *6*, 95173-95176; b) I. Swan, M. Reid, P. W. A. Howe, M. A. Connell, M. Nilsson, M. A. Moore, G. A. Morris, *J. Magn. Reson.* **2015**, *252*, 120-129.

[10] C. Ammann, P. Meier, A. Merbach, *J. Magn. Reson.* **1982**, *46*, 319-321.

[11] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminformatics* **2012**, *4*, 17. Avogadro: an open-source molecular builder and visualization tool. Version 11.12.10, http://avogadro.cc, .

[12] G. W. T. M. J. Frisch, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Revision D.01, Wallingford CT, 2004