Comparative analysis of the influence of H-bond strength and solvent on chelate cooperativity in H-bonded supramolecular complexes

Christopher A. Hunter,* Maria Cristina Misuraca and Simon M. Turega

Solvent evaporation rate

5 The evaporation rate for both toluene and 1,1,2,2-tetrachloroethane (TCE) was determined by measuring the change in the mass of the solvent (m/g) as a function of time (t/min). In order to do that, a known volume of solvent was introduced into a small vial with a similar diameter as one of the wells of the 96-well Hellma quartz plate and the solvent was allowed to evaporate at room temperature into the atmosphere. Table S1 and Table S2 summarize the experimental data for the two solvents. The corresponding evaporation rates are listed in Table S3.

10 In the automated titration procedure described in the experimental section, each experiment uses 12 wells and it is completed in 36 min. The maximum volume reached in each of the wells at the end of the experiment is equal to 302 µl. In the case of toluene, 0.22 µl of solvent evaporate per minute, which means that 7.9 µl of solvent evaporate during the time of the experiment. This corresponds to the 2.6% of the total volume of solution in the well. In a similar way, in the case of TCE, 0.03 µl of solvent evaporate per minute, which means that 1.2 µl of solvent evaporate during the time of the experiment, corresponding to the 0.4% of the total volume of the solution in 15 the well. The percentage change determined here for both toluene and TCE is much smaller than the average error reported for the association constants, so that solvent evaporation can be ignored.

Table S 1 Evaporation of toluene as a function of time (t/min) at room temperature.

measurement	t/min	mass of toluene/g
t_0	0	1.4619
t_1	18	1.4591
t_2	32	1.4564
t_3	41	1.4542
t_4	51	1.4522

Table S 2 Evaporation of TCE as a function of time (t/min) at room temperature.

measurement	t/min	mass of toluene/g
t_0	0	2.8042
t_1	15	2.8034
t_2	29	2.8029
t_3	38	2.8022
t_4	48	2.8017

20

Table S 3 Determination of the evaporation rate for toluene and TCE.

	t/min	evaporated fraction		evaporation rate	
		g	ml	g/min	ml/min
toluene	51	0.0097	0.0112	1.9×10^{-4}	2.2×10^{-4}
TCE	48	0.0025	0.0016	5.2×10^{-5}	3.3×10^{-5}

Partially bound states analysis

One H-bond. For complexes that can make a single intramolecular H-bond, the observed association constant is given by the sum 25 of the equilibrium constants for the fully bound and partially bound states (Equation S1).

$$K_{obs} = K_0(1 + \sigma_1 K_1 EM_1) \tag{S1}$$

where K_0 is intermolecular association constant for formation of the zinc-nitrogen interaction, K_1 is the intermolecular association constant for formation of the H-bonding interaction, EM_1 is the effective molarity for the intramolecular interaction, and σ_1 is a statistical factor that describes the degeneracy of the fully bound state. In the complexes discussed here, the porphyrins all have 30 four identical H-bond donor sites, so $\sigma_1 = 4$.

So, the parameter f required for the determination of EM using the double mutant cycle analysis is given by Equation S2.

$$f = 1 + \sigma_1 K_1 EM_1 \tag{S2}$$

The population of the fully bound state, P_b, is given by

$$P_b = \frac{\sigma_1 K_1 EM_1}{1 + \sigma_1 K_1 EM_1} \qquad (S3)$$

and the population of the partially bound state, P_f, where the intramolecular H-bond is not made is given by

$$P_f = \frac{1}{1 + \sigma_1 K_1 EM_1} \qquad (S4)$$

Two H-bonds. For complexes that can make two different intramolecular H-bonds, *i.e.* an ester-phenol and an ether-phenol H-bond, the observed association constant is given by the sum of the equilibrium constants for the fully bound and three partially bound states, so the parameter *f* required for the determination of *EM* using the double mutant cycle analysis is given by Equation S5.

$$f = 1 + \sigma_1 K_1 EM_1 + \sigma_2 K_2 EM_2 + \sigma_{12} K_1 EM_1 K_2 EM_2 \qquad (S5)$$

where K_i is the intermolecular association constant for formation of the ith H-bonding interaction, EM_i is the effective molarity for the ith intramolecular interaction, σ_i is the statistical factor that describes the degeneracy of singly H-bonded complexes, and σ_{12} is the statistical factor that describes the degeneracy of doubly H-bonded complex.

We used the DMC in Figure 8 of the main text to determine the free energy contribution from ether-phenol H-bonding in these complexes. If we consider the partially bound states that are possible in complexes A and B in Figure 9, the value of $\Delta\Delta G°$ measured by this double mutant cycle is given by Equation S6.

$$e^{-\Delta\Delta G°/RT} = \frac{1 + \sigma_1 K_1 EM_1 + \sigma_2 K_2 EM_2 + \sigma_{12} K_1 EM_1 K_2 EM_2}{1 + \sigma_1 K_1 EM_1} \qquad (S6)$$

The value of $K_i EM_i$ for the formation of the intramolecular ester-phenol H-bond can be determined from the double mutant cycle for the relevant singly H-bonded complex. Thus Equation S6 can be rearranged to determine $K_2 EM_2$ for the ether-phenol H-bond (Equation S7).

$$K_2 EM_2 = \frac{(1 + \sigma_1 K_1 EM_1)(e^{-\Delta\Delta G°/RT} - 1)}{\sigma_2 + \sigma_{12} K_1 EM_1} \qquad (S7)$$

We assume that the effective molarity for the formation of the first ester-phenol H-bond is not affected by the formation of the second ether-phenol H-bond: in other words, $K_1 EM_1$ is a constant. Cases where this assumption may not be valid are discussed below. We must also make some assumptions about the values of the statistical factors, σ_1, σ_2 and σ_{12}. For the partially bound, singly H-bonded complexes, there is no ambiguity: $\sigma_1 = \sigma_2 = 4$. For the fully bound, doubly H-bonded complex, there are four degenerate states in which two H-bonds are made with phenol groups on *trans*-related *meso* positions on the porphyrin, and eight degenerate states in which two H-bonds are made with phenol groups on *cis*-related *meso* positions. Models suggest that the geometry of the ester-ether ligands precludes the *trans* arrangement, and so we assume that $\sigma_{12} = 8$. This assumption affects the precise values of *EM* quoted here, but does not qualitatively alter the results.

For complexes that can make two identical intramolecular H-bonds, the observed association constant is given by the sum of the equilibrium constants for the fully bound and two partially bound states (Equation S8).

$$f = 1 + \sigma_1 K_1 EM_1 + \sigma_{11}(K_1 EM_1)^2 \qquad (S8)$$

Again, we assume that the effective molarity for the formation of the first H-bond is not affected by the formation of the second H-bond, so that $K_1 EM_1$ is a constant (Equation S9).

$$K_1 EM_1 = \frac{-\sigma_1 + \sqrt{\sigma_1^2 - 4\sigma_{11}(1 - e^{-\Delta\Delta G°/RT})}}{2\sigma_{11}} \qquad (S9)$$

Again assumptions about the statistical factors are required in order to deduce values of $K_1 EM_1$. For the partially bound, singly H-bonded complex, $\sigma_1 = 8$. For the fully bound, doubly H-bonded complex, there are two degenerate states in which two H-bonds are made with phenol groups on *trans*-related *meso* positions on the porphyrin, and four degenerate states in which two H-bonds are made with phenol groups on *cis*-related *meso* positions. Models suggest that all states are equally accessible, so $\sigma_{11} = 6$, but

this value may be as low as 2, if the *cis* states are inaccessible.

Four H-bonds. For complexes that can make four intramolecular H-bonds, the observed association constant is given by the sum of the equilibrium constants for all possible states (Equation S10).

$$
\begin{aligned}
f = {} & 1 + \sigma_1 K_1 EM_1 + \sigma_2 K_2 EM_2 + \sigma_3 K_3 EM_3 + \sigma_4 K_4 EM_4 \\
& + \sigma_{12} K_1 EM_1 K_2 EM_2 + \sigma_{13} K_1 EM_1 K_3 EM_3 + \sigma_{14} K_1 EM_1 K_4 EM_4 \\
& + \sigma_{23} K_2 EM_2 K_3 EM_3 + \sigma_{24} K_2 EM_2 K_4 EM_4 + \sigma_{34} K_3 EM_3 K_4 EM_4 \\
& + \sigma_{123} K_1 EM_1 K_2 EM_2 K_3 EM_3 + \sigma_{124} K_1 EM_1 K_2 EM_2 K_4 EM_4 + \sigma_{134} K_1 EM_1 K_3 EM_3 K_4 EM_4 \\
& + \sigma_{234} K_2 EM_2 K_3 EM_3 K_4 EM_4 + \sigma_{1234} K_1 EM_1 K_2 EM_2 K_3 EM_3 K_4 EM_4
\end{aligned}
\tag{S10}
$$

If we assume that the values of $K_i EM_i$ for the ester and ether H-bonds are constants for the symmetric ligands studied here, this equation simplifies to Equation S11.

$$
\begin{aligned}
f = {} & 1 + \sigma_1 K_1 EM_1 + \sigma_2 K_2 EM_2 + \sigma_{11}(K_1 EM_1)^2 + \sigma_{22}(K_2 EM_2)^2 + \sigma_{12} K_1 EM_1 K_2 EM_2 \\
& + \sigma_{112}(K_1 EM_1)^2 K_2 EM_2 + \sigma_{122} K_1 EM_1 (K_2 EM_2)^2 + \sigma_{1122}(K_1 EM_1)^2 (K_2 EM_2)^2
\end{aligned}
\tag{S11}
$$

where the σ subscripts reflect the symmetries of the complexes.

We used the DMC in Figure 8 of the main text to determine the free energy contribution from ether-phenol H-bonding in these complexes. If we consider the partially bound states that are possible in complexes A and B in Figure 8, the value of $\Delta\Delta G°$ measured by the DMC is given by Equation S12.

$$
e^{-\Delta\Delta G/RT} = \frac{\left\{
\begin{array}{l}
1 + \sigma_1 K_1 EM_1 + \sigma_2 K_2 EM_2 + \sigma_{11}(K_1 EM_1)^2 + \sigma_{22}(K_2 EM_2)^2 + \sigma_{12} K_1 EM_1 K_2 EM_2 \\
+ \sigma_{112}(K_1 EM_1)^2 K_2 EM_2 + \sigma_{122} K_1 EM_1 (K_2 EM_2)^2 + \sigma_{1122}(K_1 EM_1)^2 (K_2 EM_2)^2
\end{array}
\right\}}{1 + \sigma_1 K_1 EM_1 + \sigma_{11}(K_1 EM_1)^2}
\tag{S12}
$$

The value of $K_1 EM_1$ for the formation of intramolecular ester-phenol H-bonds can be determined from the DMC for the relevant doubly H-bonded complex, so Equation S12 can be rearranged to determine $K_2 EM_2$ for the ether H-bonds (Equation S13).

$$
K_2 EM_2 = \frac{\left\{
\begin{array}{l}
-\left(\sigma_2 + \sigma_{12} K_1 EM_1 + \sigma_{112}(K_1 EM_1)^2\right) \\[4pt]
+ \sqrt{
\begin{array}{l}
\left(\sigma_2 + \sigma_{12} K_1 EM_1 + \sigma_{112}(K_1 EM_1)^2\right)^2 \\
-4\left(\sigma_{22} + \sigma_{122} K_1 EM_1 + \sigma_{1122}(K_1 EM_1)^2\right)\left(1 + \sigma_1 K_1 EM_1 + \sigma_{11}(K_1 EM_1)^2\right)\left(1 - e^{-\Delta\Delta G/RT}\right)
\end{array}}
\end{array}
\right\}}{2\left(\sigma_{22} + \sigma_{122} K_1 EM_1 + \sigma_{1122}(K_1 EM_1)^2\right)}
\tag{S13}
$$

For the singly H-bonded complexes, $\sigma_1 = \sigma_2 = 8$. For the complexes that form two ester H-bonds or two ether H-bonds, we assume that *trans* and *cis* arrangements are equally likely, so $\sigma_{11} = \sigma_{22} = 6$. However, formation of H-bonds to both the ester and ether groups on the same side chain can only be made in the *cis* arrangement, so $\sigma_{12} = 40$. Using the same assumptions for the triply H-bonded complexes, $\sigma_{112} = \sigma_{122} = 16$, and for the complex where all four H-bonds are made, $\sigma_{1122} = 8$ (see Figure S1 for details of the statistical analysis).

Table S4 Statistical factors for complexes formed with ligands **L1d-L6d**.

	Ligand					
	L1d	**L2d**	**L3d**	**L4d**	**L5d**	**L6d**
$\sigma_1 = \sigma_2$	4	4	8	4	4	8
$\sigma_{11} = \sigma_{22}$			12			12
σ_{12}					8	40
$\sigma_{112} = \sigma_{122}$						32
σ_{1122}						16

Table S5 Statistical factors for complexes formed with ligands **L1b-L6b**.

	Ligand					
	L1b	**L2b**	**L3b**	**L4b**	**L5b**	**L6b**
$\sigma_1 = \sigma_2$					4	8
$\sigma_{11} = \sigma_{22}$						12

Free energy contribution from ester-phenol intramolecular H-bonds

Table S6 Free energy contributions from ester-phenol H-bonds ($\Delta\Delta G°$/kJ mol^{-1}) in the complexes formed between **P1a-P4a** and **L1b-L6b** determined in 1,1,2,2,-tetrachloroethane at 298 K using the appropriate DMC. Complexes that make no H-bonds are shaded.[a]

Porphyrin		Ligand					
		L1b	**L2b**	**L3b**	**L4b**	**L5b**	**L6b**
	P1a	0	0	-1	0	-2	-5
	P2a	0	0	0	0	-2	-4
	P3a	0	0	-1	-1	-3	-6
	P4a	-1	0	0	0	0	0

[a] Average error over the data set ±1 kJ mol^{-1}.

Effective molarity for ester-phenol intramolecular H-bonds

Table S6 Effective molarities (EM/mM) for intramolecular ester-phenol H-bonds measured in 1,1,2,2,-tetrachloroethane at 298 K.[a]

Porphyrin		Ligand					
		L1b	**L2b**	**L3b**	**L4b**	**L5b**	**L6b**
	P1a					180	240
	P2a					130	140
	P3a					290	310
	P4a						

[a] Average error over the data set ±50%.

Statistically-corrected equilibrium constants for formation of intramolecular H-bonds

Table S6 $K EM$ values for formation of an intramolecular ether-phenol H-bond.[a]

Porphyrin		Ligand					
		L1d	**L2d**	**L3d**	**L4d**	**L5d**	**L6d**
	P1a		1.2	1.1			0.4
	P2a						0.2
	P3a		0.3	0.2			0.3
	P4a						0.4

[a] Errors are ±50%.

Table S7 $K EM$ values for formation of an intramolecular ester-phenol H-bond.[a]

Porphyrin		Ligand					
		L1b	**L2b**	**L3b**	**L4b**	**L5b**	**L6b**
	P1a					0.4	0.5
	P2a					0.3	0.3
	P3a					0.6	0.6
	P4a						

[a] Errors are ±50%.

H-bond occupancy

Table S8 Occupancy of the H-bonded state for ether-phenol H-bonds (%).

Porphyrin		Ligand					
		L1d	L2d	L3d	L4d	L5d	L6d
	P1a		84	76			47
	P2a						33
	P3a		49	43			40
	P4a						55

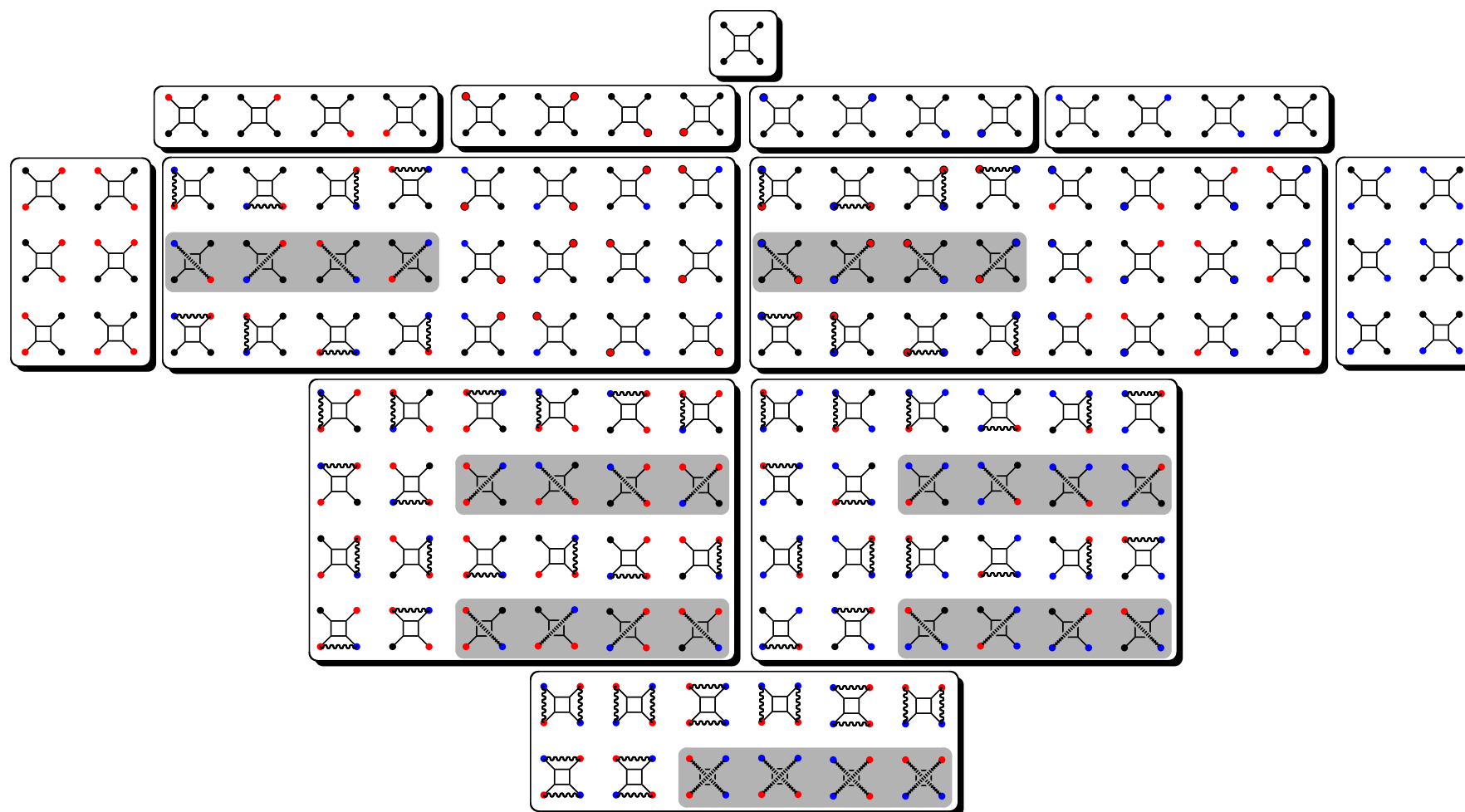**Table S7** Occupancy of the H-bonded state for ester-phenol H-bonds in the control ligand complexes (%).

Porphyrin		Ligand					
		L1b	L2b	L3b	L4b	L5b	L6b
	P1a					58	61
	P2a					52	49
	P3a					70	67
	P4a						

5

Table S8 Occupancy of the H-bonded state for ester-phenol H-bonds in the ether ligand complexes (%).

Porphyrin		Ligand					
		L1a	L2a	L3a	L4a	L5a	L6a
	P1a					54	50
	P2a					48	41
	P3a					66	57
	P4a						

Fig. S1 Degenerate states for the complex where four intramolecular H-bonds are possible. Unbound phenol OH groups on the porphyrin are represented by black dots, phenol groups H-bonded to phosphonate diesters are represented by red dots and phenol groups H-bonded to esters are represented by blue dots. The connectivity between phosphonate diester and ester groups on the same side chain are represented by wiggly lines. In complexes where the two side chains are non-equivalent, i.e. one is bound and one is free, the side chains are distinguished by large and small dots. The shaded complexes are structures that models suggest are geometrically impossible. The panels are orders zero, one, two, three and four H-bonds from top to bottom and grouped by symmetry.