Supporting Information for
^{15}N SOLID-STATE NMR AS A PROBE OF FLAVIN H-BONDING

Dongtao Cui[1], Ronald L. Koder, Jr.[2,3], P. Leslie Dutton[3] and Anne-Frances Miller[1]

From the Department of Chemistry[1], University of Kentucky, Lexington, KY, U.S.A. 40506-0055 and The Johnson Foundation and Department Biochemistry[2], University of Pennsylvania, Philadelphia, PA, U.S.A.19104-6059, Department of Physics[3], The City College of New York, New York, NY, U.S.A. 10031.

* corresponding author: Phone: (859) 257-9349, fax: (859) 323-1069, email afm@uky.edu

Full citation of Gaussian 03, Revision D.02:

Gaussian 03, Revision D.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.

Details of NMR measurements.

A nominal temperature of -60 °C was chosen on the basis of signal-to-noise obtained, and the requirement that benzene (Bz) be frozen solid. The temperature value was not corrected for frictional heating due to spinning etc. However the choice of temperature was found not to affect the CSPVs of TARF, in previous work. The signal-to-noise of spectra improved as the temperature was lowered from 0 °C to -30 °C to -60 °C. No further improvement was obtained upon cooling to -90 °C.

Table 1S Experimental and calculated ^{15}N Chemical Shift Principal Values[a] (CSPVs) of N5 of solid TPARF, TPARF in Bz, TPARF in Bz with DBAP and TPARF in Bz with H_2O, at different MAS spinning speeds.

	MAS Speed (Hz)	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} (ppm)	span=δ_{11}-δ_{33} (ppm)
TPARF Solid	3000	675.53	384.06	-31.10	342.83	706.63
	4000	669.52	391.60	-32.63	342.83	702.15
	5000	681.655	388.75	-41.91	342.83	723.57
	Average[a]	675 ± 6	388 ± 4	-34 ±6	342.83[c]	710 ±10
TPARF in Bz	3000	673.90	388.32	-35.54	342.23	709.44
	4000	675.63	390.90	-39.85	342.23	715.48
	5000	679.99	391.54	-46.65	341.63	726.64
	Average[a]	676 ± 3	390 ±2	-40 ± 6	342.0 ± 0.3	716 ±9
	Calc. monomer[b]	723	399	-33	363	756
	Calc. dimer	722	398	-33	362	755
TPARF In Bz with DBAP	3000	688.46	391.04	-32.94	348.85	721.40
	4000	681.82	399.97	-33.42	349.46	715.24
	5000	688.53	396.47	-41.15	347.95	729.68
	Average[a]	687 ± 4	395 ± 4	-36 ± 5	348.7 ± 0.8	722 ± 7
	Calc.	734	401	-33	367	767
TPARF In Bz with 5x H_2O	3000	670.364	386.58	-28.45	342.83	698.81
	4000	672.91	392.44	-37.76	342.53	710.67
	5000	667.86	393.20	-34.38	342.23	702.24
	Average[a]	670 ±3	391 ±4	-34 ±5	342.5 ± 0.3	704 ±6
	Calc.(1W Boltzmann)[d]	717	398	-33	361	749

a. Chemical shift relative to liquid ammonia. Averages are weighted according to the chi squared values associated with the Hertzfeld-Berger fits and standard deviations are provided too. Reproducibility was 2, 3, 3, 0.6 ppm in replicate spectra. Standard deviations are provided for the averages. $\delta_{iso}=(\delta_{11}+\delta_{22}+\delta_{33})/3$

b. All calculations employed LF to model the flavin of TPARF. Geometry optimizations and GIAO chemical shifts calculations employed the 6-311++G(2d,2p) basis set. Uncertainties associated with calculated CSPVs are estimated at 5, 1 and 0.7 ppm, respectively for δ_{11}, δ_{22}, and δ_{33}, respectively and that associated with δ_{iso} is estimated at 2 ppm based on calculations performed on 10 frames from a converged MD calculation (see supplemental Fig.s 3S, 4S).

c. Precision limited only by instrumental error, estimated at 0.3 ppm.

d. Boltzmann-weighted average of CSPVs calculated for all 4 optimized one-water complexes with LF taking into account dimerization of LF.

Table 2S: Cartesian coordinates of LF geometry optimized with B3PW91 and 6-311++G(2d,2p)

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms)		
			X	Y	Z
1	6	0	0.000000	0.000000	0.000000
2	6	0	0.000000	0.000000	1.385345
3	6	0	1.196957	0.000000	2.107609
4	6	0	2.419404	0.000000	1.399928
5	6	0	2.402161	0.000000	-0.004472
6	6	0	1.225468	0.000000	-0.717052
7	6	0	3.618931	0.000000	3.337337
8	6	0	2.424939	0.000000	4.161156
9	6	0	3.604188	0.000000	6.147378
10	1	0	-0.947855	0.000000	1.903524
11	1	0	3.362829	0.000000	-0.507087
12	1	0	5.655059	0.000000	5.955290
13	7	0	2.417313	0.000000	5.467071
14	7	0	4.798255	0.000000	5.409662
15	7	0	3.616169	0.000000	2.043392
16	7	0	1.231572	0.000000	3.486037
17	6	0	4.929463	0.000000	4.047387
18	8	0	3.673676	0.000000	7.363037
19	8	0	6.007250	0.000000	3.489235
20	6	0	-1.303024	0.000000	-0.739243
21	1	0	-1.387756	0.876416	-1.386488
22	1	0	-1.387756	-0.876416	-1.386488
23	1	0	-2.149997	0.000000	-0.055244
24	6	0	1.237200	0.000000	-2.217203
25	1	0	0.727073	-0.877772	-2.621472
26	1	0	0.727073	0.877772	-2.621472
27	1	0	2.257370	0.000000	-2.597549
28	6	0	-0.017331	0.000000	4.242048
29	1	0	-0.596100	-0.890588	3.998252
30	1	0	0.232020	0.000000	5.296750
31	1	0	-0.596100	0.890588	3.998252

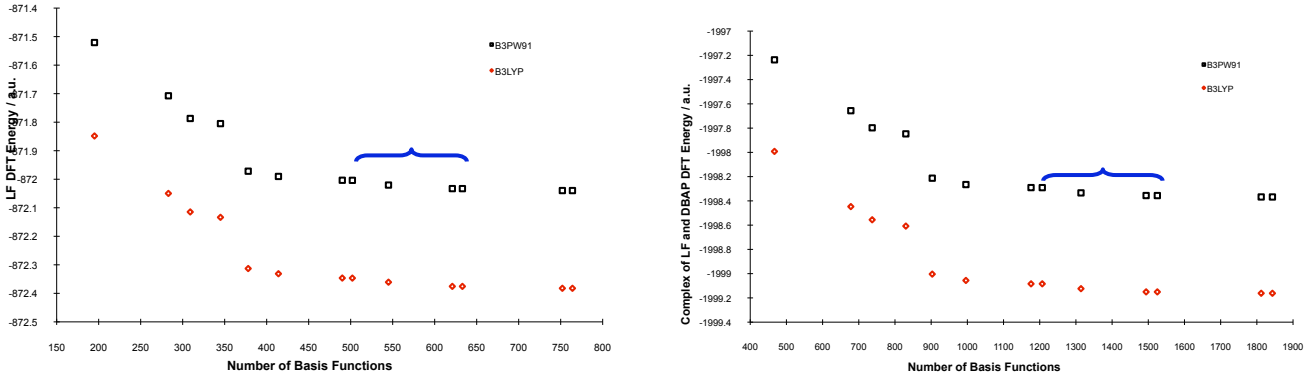
Table 3S: Cartesian coordinates of the complex of LF with DBAP, geometry optimized with B3PW91 and 6-311++G(2d,2p)

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms)		
			X	Y	Z
1	6	0	7.457504	-4.628472	-1.386937
2	6	0	7.63538	-3.409126	-0.745918
3	6	0	6.641772	-2.909209	0.087541
4	6	0	5.459562	-3.618628	0.299133
5	6	0	5.293179	-4.845534	-0.343554
6	6	0	6.282799	-5.342544	-1.182423
7	1	0	8.228706	-5.019965	-2.040557
8	1	0	8.548126	-2.842763	-0.894347
9	1	0	6.790675	-1.956654	0.58599
10	1	0	4.384681	-5.412344	-0.18363
11	1	0	6.134398	-6.296463	-1.67625
12	1	0	4.423431	-1.512773	-0.922775
13	6	0	3.768329	-0.710405	-0.619995
14	6	0	4.062491	0.605518	-0.927669
15	6	0	2.565453	-0.979598	0.033458
16	6	0	3.172818	1.611571	-0.594027
17	1	0	4.974146	0.845467	-1.463544
18	6	0	1.990647	1.252572	0.050677
19	1	0	3.3584	2.637595	-0.872636
20	6	0	3.265325	4.93073	0.512871
21	6	0	2.552895	5.9692	-0.087007
22	6	0	4.645232	4.869048	0.318787
23	6	0	3.206478	6.914722	-0.8675
24	1	0	1.482328	6.035924	0.060238
25	6	0	5.300944	5.818629	-0.455924
26	1	0	5.215777	4.071724	0.784333
27	6	0	4.581906	6.844684	-1.055292
28	1	0	2.635487	7.712962	-1.328644
29	1	0	6.374883	5.753553	-0.59148
30	1	0	5.089321	7.585153	-1.663139
31	6	0	4.417379	-3.064585	1.250253
32	1	0	4.509776	-3.578206	2.211657
33	1	0	4.607263	-2.007241	1.428157
34	6	0	2.976606	-3.314196	0.840719
35	8	0	2.483114	-4.420809	0.976146
36	7	0	2.19585	-2.275476	0.387829
37	1	0	1.192839	-2.493227	0.392886

38	7	0	1.681351	-0.018971	0.334741
39	7	0	1.036442	2.205915	0.412547
40	1	0	0.051558	1.934868	0.336747
41	6	0	1.214748	3.463227	0.936963
42	8	0	0.249452	4.193134	1.09091
43	6	0	2.591083	3.904055	1.401522
44	1	0	3.255987	3.054585	1.550218
45	1	0	2.409188	4.351972	2.382695
46	6	0	-1.563434	-2.129944	0.322067
47	6	0	-2.134607	0.257733	0.162763
48	6	0	-3.550036	-0.180585	0.050064
49	6	0	-3.795349	-1.606749	0.105985
50	6	0	-5.76254	0.30767	-0.176376
51	6	0	-6.110672	-1.061499	-0.125774
52	6	0	-6.774266	1.271798	-0.321992
53	6	0	-7.459174	-1.417347	-0.221616
54	6	0	-8.101027	0.923119	-0.416698
55	1	0	-6.461994	2.309732	-0.355582
56	6	0	-8.444526	-0.454389	-0.364573
57	1	0	-7.752555	-2.456434	-0.185659
58	7	0	-2.866028	-2.518817	0.227801
59	7	0	-1.253711	-0.767757	0.326098
60	7	0	-5.105428	-1.994203	0.01551
61	7	0	-4.472327	0.7183	-0.085429
62	8	0	-0.642665	-2.932482	0.404108
63	8	0	-1.781845	1.424244	0.117369
64	1	0	-0.258741	-0.524194	0.418059
65	6	0	-9.878633	-0.875081	-0.46508
66	1	0	-10.476186	-0.428308	0.333302
67	1	0	-10.319481	-0.544167	-1.408762
68	1	0	-9.982143	-1.956951	-0.402663
69	6	0	-9.160396	1.973735	-0.571059
70	1	0	-9.72879	1.835741	-1.494011
71	1	0	-9.878398	1.941263	0.251999
72	1	0	-8.719133	2.968742	-0.593226
73	1	0	-6.08892	-3.611063	0.916954
74	6	0	-5.436208	-3.415692	0.066649
75	1	0	-5.936562	-3.710739	-0.855276
76	1	0	-4.511518	-3.969985	0.176775

Details of methodology used for calculations

Because NMR chemical shifts represent very small energies, the accuracy of their calculation depends on details of the procedures used. We began by evaluating different density functionals and levels of theory, based on their convergence to low energy and abilities to replicate experiment. Supplemental Figure S1 shows the energies obtained upon geometry optimization of lumiflavin (LF) or the complex with DBAP (shown in Scheme 1). The data reveal that basis sets smaller than 6-311G(d,p) produce elevated energies and therefore constitute poor approximations to the expected energy minimum. Hence only basis sets of 6-311G(d,p) (582 primitives) and larger were evaluated for performance in NMR calculations.



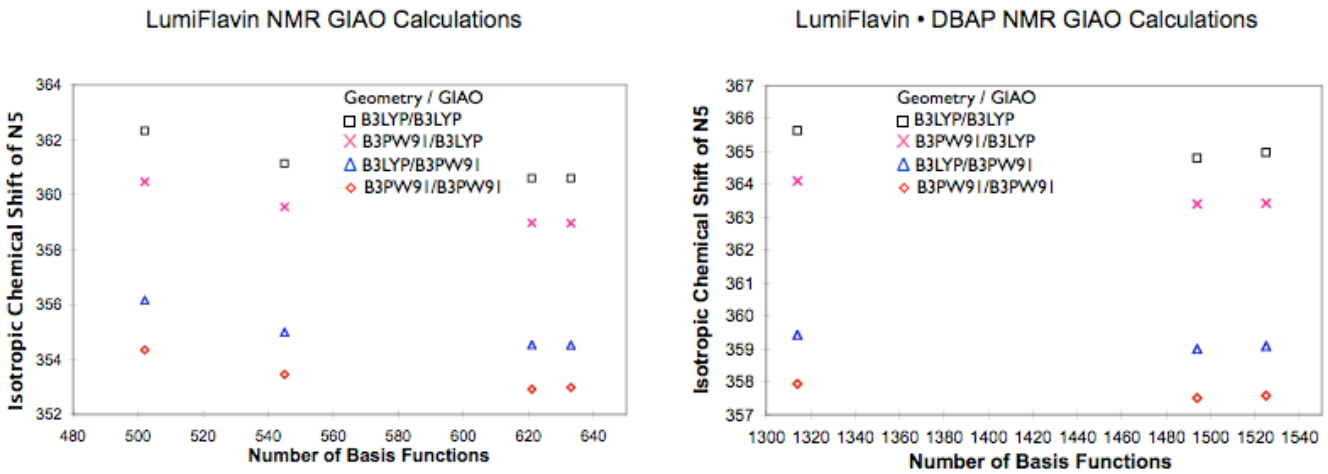
(A) **(B)**

Figure 1S: DFT (B3PW91 and B3LYP) energies versus Basis Set Size. **(A)** LF **(B)** Complex of LF and DBAP.

For calculations of chemical shifts, the GIAO method was chosen based on its success in comparable systems. The basis sets indicated with the blue brackets in Figure 1S were 6-311G(d,p), 6-311G+(d,p), 6-311G++(d,p) and 6-311G(2d,2p). The latter of these and augmented versions 6-311G+(2d,2p) and 6-311G++(2d,2p) were used in GIAO calculations of the chemical shifts (Figure 2S). The 6-311++G(2d,2p) basis produced the best agreement with

the isotropic chemical shift (δ_{iso}) measured for N5 in TPARF powder.

We also compared the performance of two functionals: the commonly used B3LYP and the B3PW91 which was found to best reproduce experiment in previous studies . We performed geometry optimization using each of the two functionals and the 6-311++G(2d,2p) basis set. Each optimized structure was then used for GIAO calculations of NMR shielding, using each of the two functionals. The four methods: PW91:PW91, PW91:B3LYP, B3LYP:PW91 and B3LYP:B3LYP produced significantly different shieldings, of which those produced by PW91:PW91 were in best agreement with experiment (Figure S2). This result suggests that B3LYP should not be the automatic choice for DFT studies.



(A) Lumiflavin (experiment = 341 ppm**)** **(B)** Lumiflavin and DBAP (experiment = 347 ppm)

Figure 2S: Isotropic chemical shifts of N5 versus basis size and choice of functional used for each of GIAO and geometry optimization calculations. Basis sets used were 6-311G(2d,2p), 6-311+G(2d,2p) or 6-311++G(2d,2p). For LF alone, results are also presented for 6.311++G(d,p).

Uncertainty associated with the calculations

To assess the uncertainty associated with our computational results, we took the MM+ geometry-optimized structure of LF and subjected it to 2 ns of molecular dynamics (MD) at 213 K (-60 °C) using the MM+ force-field as implemented in hyperchem®. The calculation equilibrated on a timescale of 100 ps based on the energy (Figure S3). An expansion of the 1 ps including the global energy minimum shows that oscillations nonetheless persist with a period of 0.95 ps and an amplitude of 64 cal/mol even after 2 ns equilibration. Thus we estimate that our calculated energies should be considered to be associated with uncertainties on the order of 130 J/mol. The structure with the global minimum energy was extracted, along with that of the following peak energy, and eight structures from evenly-spaced time points in between. The global minimum structure and the following peak-energy structure are overlaid in Figure S4. The principle difference is the rotational angle of the methyl groups at positions 7 and 8.

All ten extracted structures were used as input for GIAO chemical shift calculations (without geometry re-optimization). The chemical shifts obtained had standard deviations of 4.9, 1.4, 0.68 and 1.8 ppm for $\delta_{11,}$ δ_{22} and δ_{33} , and δ_{iso}, , respectively. These values therefore provide estimates of the uncertainty associated with our chemical shift calculations.

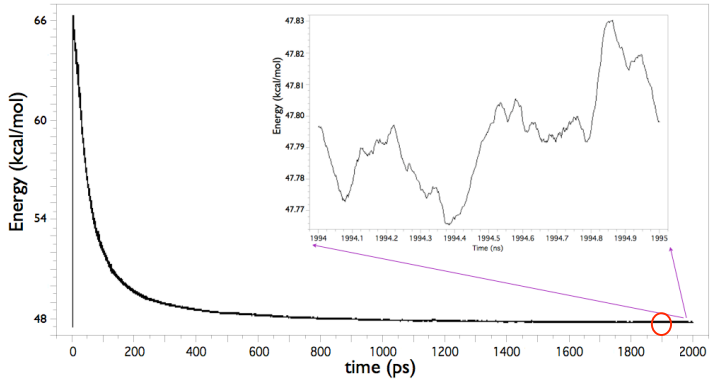


Figure 3S: Energy vs. time trajectory of MD calculation beginning with the MM+ geometry-optimized structure. Entire 2 ns trajectory, insert: 1 ps of the trajectory surrounding the global energy minimum.

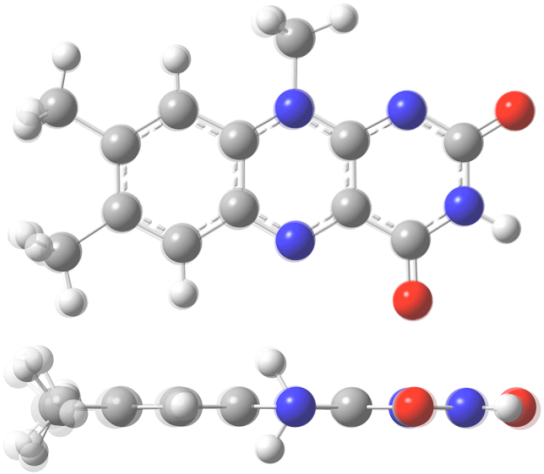


Figure 4S: Comparison of the global energy-minimum structure and the structure representing the peak of the energy oscillations that persisted after equilibration (red circles in figure 3S).

Association between flavins at high concentrations

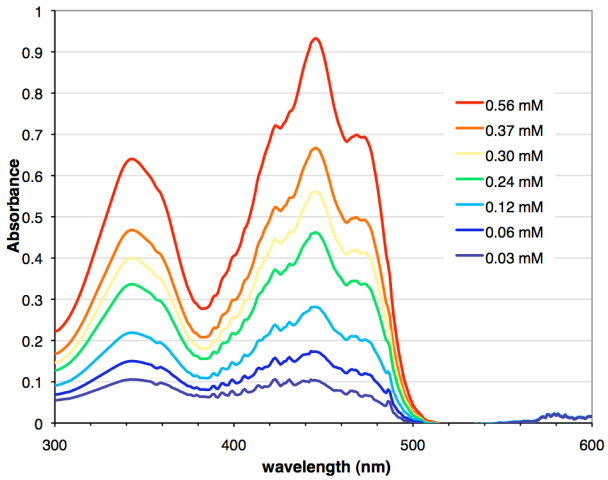


Figure 5S. Concentration dependence of the visible spectrum of TPARF in dry benzene, in a 1 mm path-length cuvette.

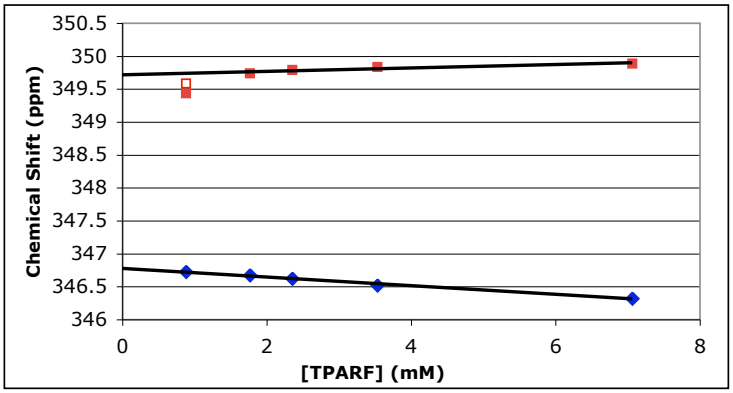


Figure 6S: Chemical shift changes upon serial dilution of a solution of TPARF in benzene (◆) and TPARF+DBAP in benzene (■). At the lowest concentration of TPARF + DBAP, dissociation of the complexes was evident in a sharper decrease in chemical shift, which could be attenuated by addition of more DBAP (□).

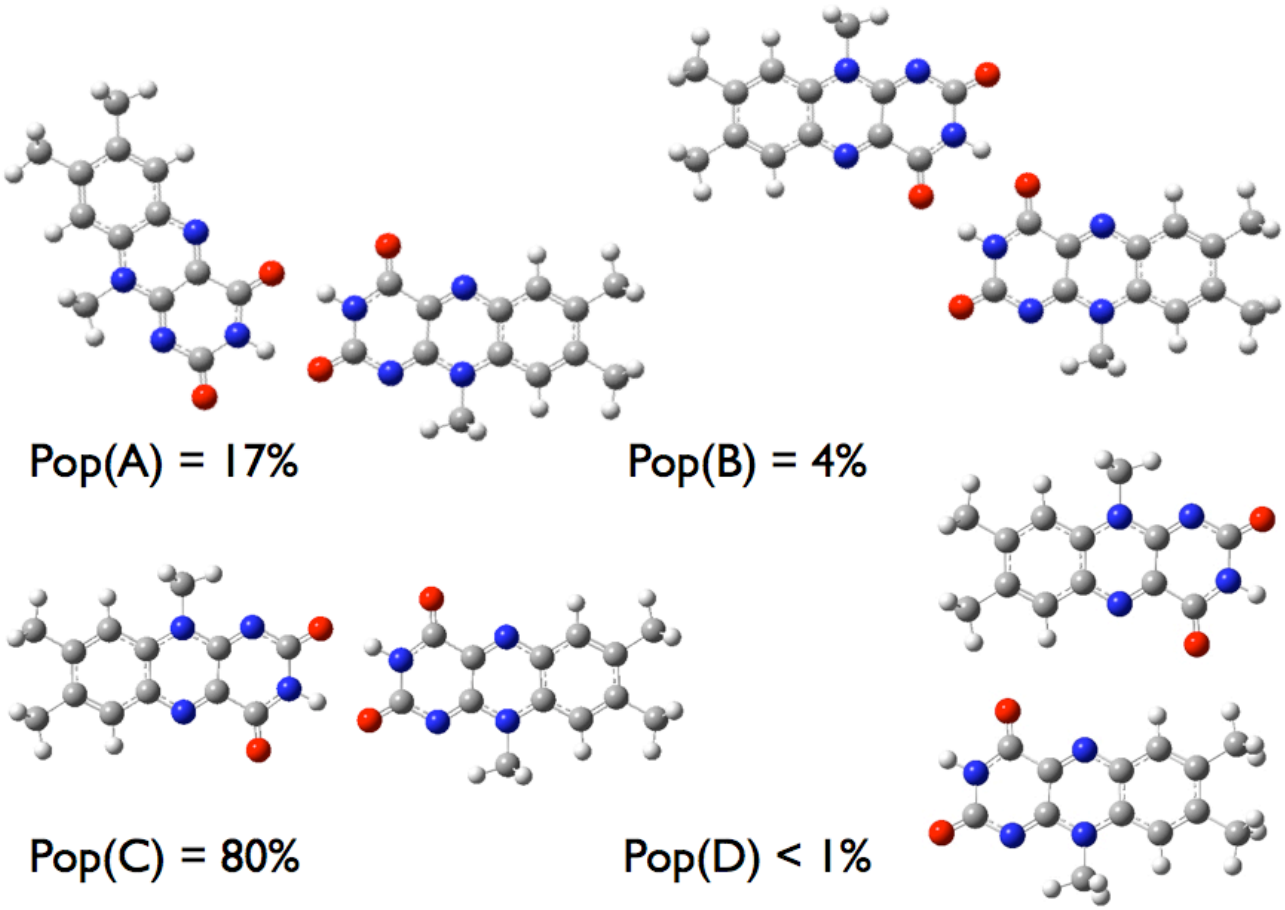


Figure 7S: Four optimized coplanar flavin dimer structures with Boltzmann populations calculated based on the dimer energies. DFT geometry optimizations were performed using B3PW91 and the 6-311++G(2d,2p) basis set.

Table 4S Calculated chemical shifts for energetically viable LF dimers[a,b]

Chemical shift	δ_{iso}	δ_{11}	δ_{22}	δ_{33}	Rel. energy (kJ/mol)	Frac^{n} Pop
Configuration A[c]	362.8	722	398	-32	2.77	.166
Configuration B[c]	362.9	724	399	-33	5.38	.038
Configuration C[c]	362.1	722	398	-33	0	.796
Boltzmann Pop. Avg.	362.2	722	398	-33		
Monomer	362.8	723	399	-33		
Effect of dimerization	-0.6	-1	-1	0		

[a] The fourth configuration had a relative energy of 12 kJ/mol and therefore made a negligible contribution to the population (<0.1 %) and the average CSPVs.

[b] Uncertainties associated with calculated CSPVs are estimated at 5, 1 and 0.7 ppm for δ_{11}, δ_{22}, and δ_{33}, respectively, and that associated with δ_{iso} is estimated at 2 ppm.

[c] The configurations are shown in Figure 7S.

Note that the chemical shifts are very similar for the different complexes, indicating that effects cannot be very large or specific. The effects of binding a LF are much smaller than the effects of binding DBAP, possibly because DBAP accepts some electron density from LF.

All calculations were performed using B3PW91 and the 6-311++G(2d,2p) basis set.

CSPV dependence on temperature and physical state

Table 5S: Comparison of chemical shifts[a] of TPARF powder and frozen TPARF in benzene, at room temperature and -60 °C

Sample	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)	δ_{iso} [b] (ppm)
TPARF solid, RT	670 ± 4	387 ± 2	-31 ± 3	342.1 ± 0.2
TPARF solid, -60°C	675 ± 6	388 ± 4	-34 ± 6	342.8
TPARF in benzene	676 ± 3	390 ±2	-40 ± 6	342.0 ± 0.4

[a] Chemical shift relative to liquid ammonia.

[b] $\delta_{iso}=(\delta_{11}+\delta_{22}+\delta_{33})/3$

[c] Weighted averages weighted each set of CSPVs according to the chi-squared value associated with its Hertzfeld-Berger fit. Standard deviations are also provided for each.

H-bonding distances

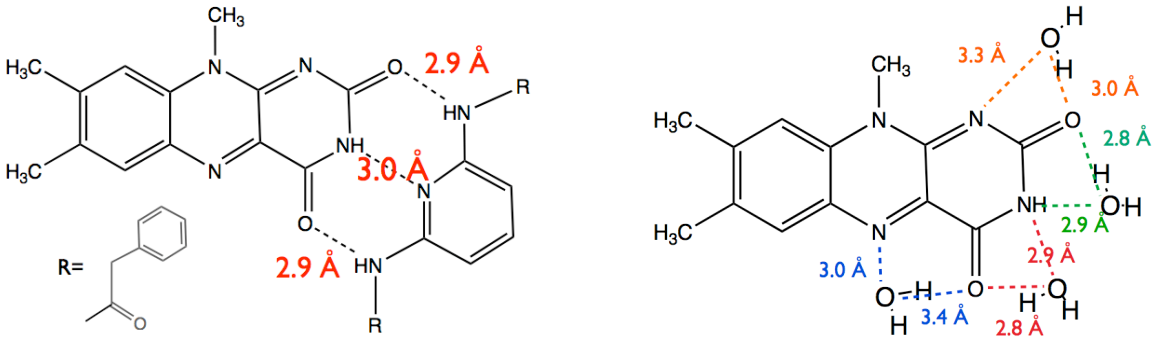


Figure 8S: Lengths of H-bonds measured from the LF•DBAP and the LF•5H_2O structures optimized using B3PW91 with 6-311++G(2d,2p).

NBO orbitals of LF alone and LF complexed with DBAP

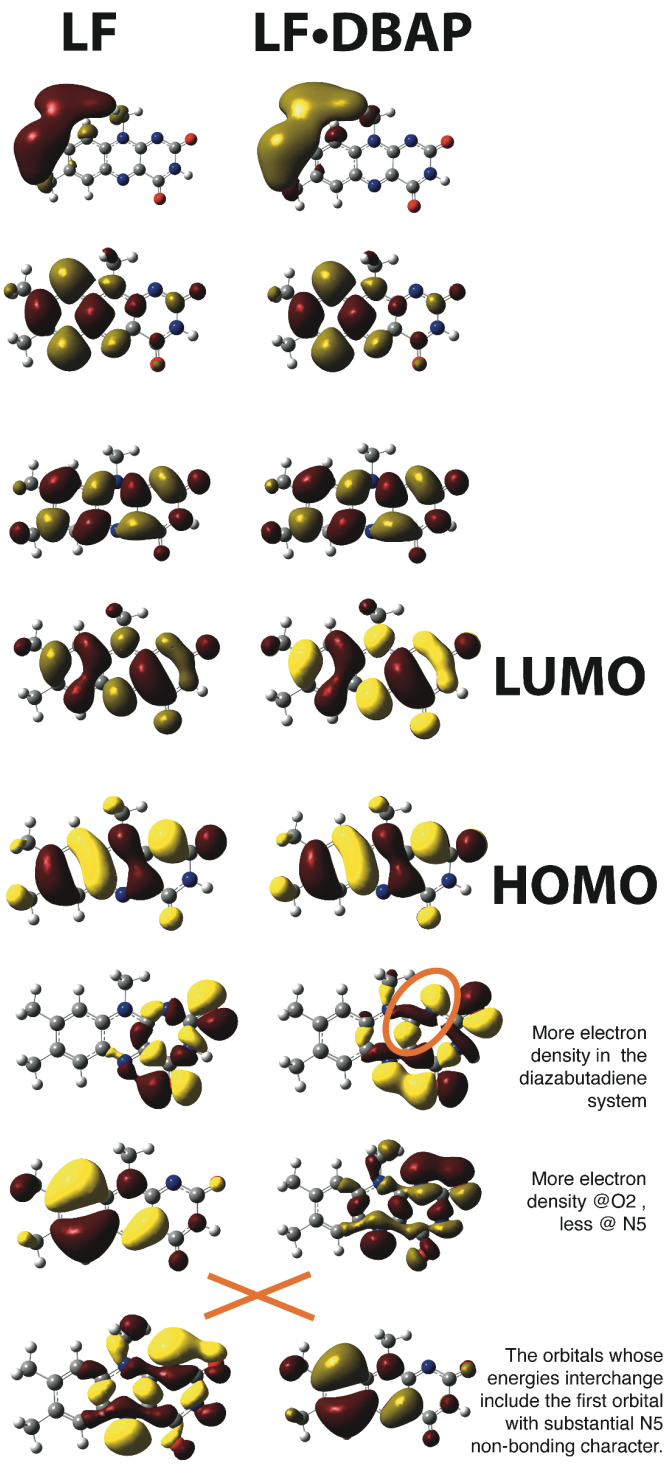


Figure 9S Four highest-occupied and four lowest unoccupied molecular orbitals of LF alone (left-hand column) and when complexed with DBAP (right-hand column). Figures were generated by Gaussian based on NBO analyses of computations performed using B3PW91 and the 6-311++G(2d,2p) basis set.

SSNMR spectra before and after of water to TPARF in benzene

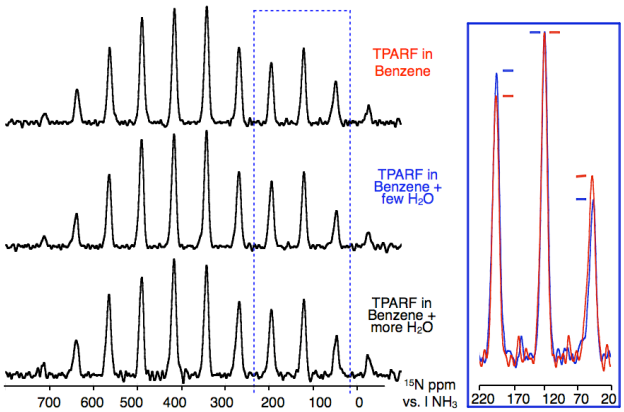


Figure 10S. Comparison of the ^{15}N CP-MAS spectra of [^{15}N-N5] TPARF in benzene, and [^{15}N-N5] TPARF in benzene with 5 and then 15 stoichiometric equivalents of water, at 3000 Hz MAS speed. (Side bands in the blue box are enlarged for comparison of their intensities: red = before and blue = after addition of water.)

Treatment of water molecules H-bonding with LF

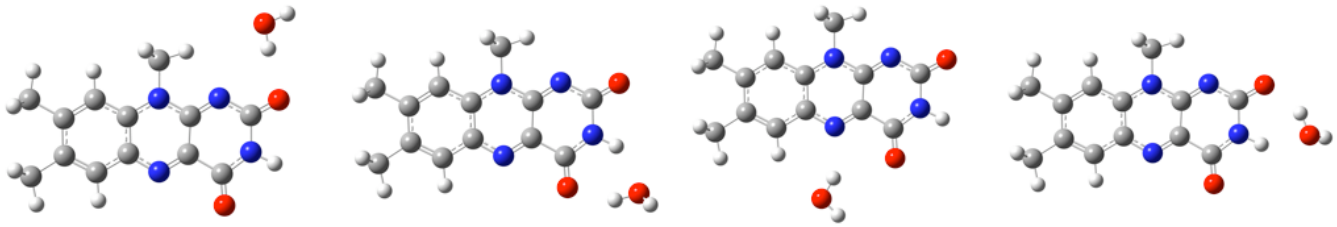
A water molecule can both donate and accept H-bonds, and so can the flavin. Thus, there are numerous possible H-bonding configurations. For H-bonding between a flavin and one H_2O, we employed starting structures placing water molecules at many plausible positions, such as close to N1, or N5. However upon geometry optimization, only the four configurations shown in Figure 11S were obtained. Numerous additional calculations modelled H-bonding with two, three, four, five, six, seven or eight H_2O molecules. Upon optimization of these systems, the positions occupied by the H_2O water molecules reproduced the positions observed in single-H_2O calculations 50 % of the time. The configurations produced upon optimization with two, five and eight H_2Os are shown in figure 12S.

None of the single-H_2O, two-H_2O, five-H_2O or even the eight-H_2O configuration reproduced the changes in chemical shift observed upon addition of 5 equivalents of H_2O to TPARF (Figure 13S). Moreover we found that as the number of water molecules increased the water molecules displayed increased tendencies to not be associated with the flavin, especially near N5, but to preferentially associate with one-another instead (Figure 12S). Waters also began

migrating to positions out of the flavin plane. However the experimental CSPV values fall within those obtained computationally. Therefore, based on the relatively similar energies calculated for the different configurations with the same number of waters bound, we propose that our NMR samples contain an ensemble representing multiple different configurations and that all the water-binding locations are not all occupied all the time, even when excess water is present (otherwise the 8 H_2O: 1 LF calculation would have come closest to experiment).

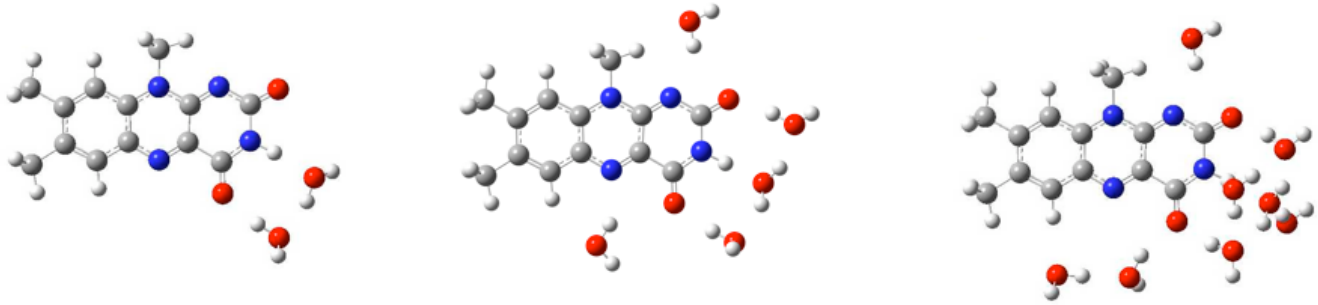
Beginning with the simplest systems (one water per LF), the relative energies of the four single-water configurations were used to calculate the relative probabilities with which one water would be expected to bind in each of the four configurations. The Boltzmann probabilities were used along with a total probability of 1 to calculate the contributions of all four positions to the ensemble with one water bound to LF. The chemical shifts calculated for each water configuration were weighted by the applicable population before summing to obtain Boltzmann-weighted average CSPVs. The result displays improved agreement, within error of our experimental results (Figure 13S). Thus, while numerous more sophisticated and physically detailed models can be imagined, the precision of our experiment would not distinguish a better model from the simple one we begin with here, which suffices to explain the changes in the three CSPVs observed upon addition of water. The predictions of our statistical model come extremely close for δ_{iso}, and moreover much closer than individual configurations other than the one with a water between N3 and O4, which does better for δ_{iso} but makes larger compensating errors in δ_{11} and δ_{33}.

We anticipate that the current quality of spectroscopic results would permit adjudication between more models, if data for additional sites on the flavin ring were measured (especially N1, N3, C2 and C4).



(A) **(B)** **(C)** **(D)**

Figure 11S: Four optimized structures of lumiflavin with one H_2O molecule. Geometry optimizations were performed using B3PW91 with 6-311++G(2d,2p)



2 water molecules (dominant) 5 water molecules 8 water molecules

Figure 12S: Optimized structures for LF:2 H_2O, LF: 5 H_2O and LF: 8 H_2O. The configuration shown for 2 waters is the one that dominates the ensemble, at a calculated 99 % of the population. Note that water molecules begin moving to above and below the flavin plane, and form a cluster at the H-bonding edge of the flavin as the number of water molecules increases. Geometry optimizations were performed using B3PW91 with 6-311++G(2d,2p)

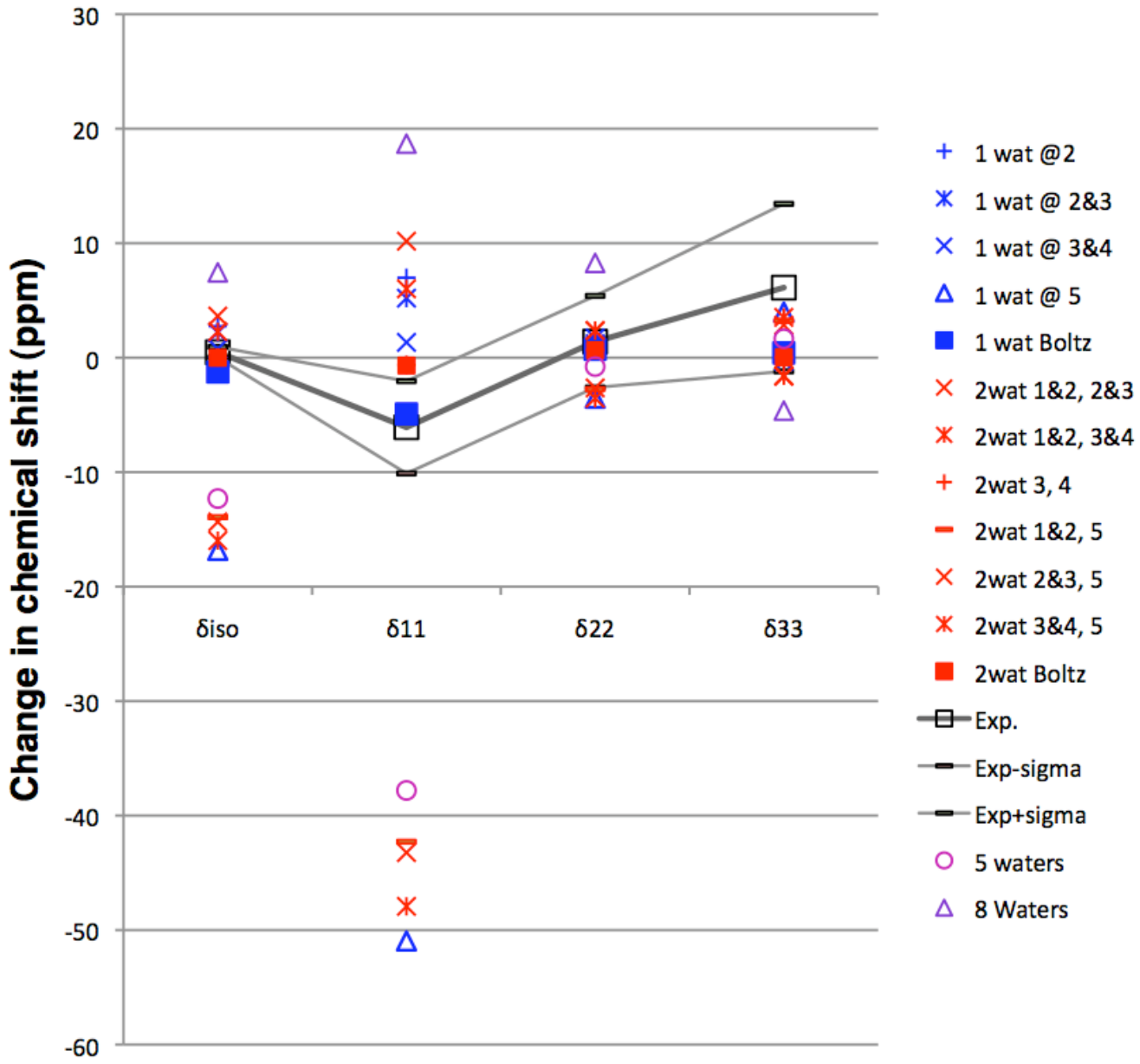


Figure 13S: Changes in GIAO NMR CSPVs and δ_{iso} calculated from optimized LF•nH_2O complexes, n = 1 (blue), n=2 (red), n=5 (magenta) or n=8 (mauve) . The configurations yielding large decreases in δ_{11} have a water H-bonding to N5 (5; 3&4,5; 2&3,5; 1&2,5) whereas the configurations yielding increases in δ_{11} have no water H-bonding to N5 (3,4; 1&2,3&4; 1&2,2&3; 3&4; 2&3; 2). In the notation used for water positions, 1&2,5 indicates that one water refined to a place between N1 and O2 and the other refined to a position near N5. 3 and 4 refer to positions near N3 and O4. Boltzmann-weighted averages are in filled symbols, ■ for 1-water and ■ for 2-water population averages. The black □ denotes the experimental change in chemical shift upon addition of water and the associated experimental errors are used to generate error margins shown as black dashes connected by lines between the columns. Calculations used B3PW91 for GIAO and geometry optimization calculations with the 6-311G++(2d,2p) basis set.