Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

Symmetry and ^{1}H NMR chemical shifts of short hydrogen bonds: Impact of electronic and nuclear quantum effects

Shengmin Zhou and Lu Wang*

Department of Chemistry and Chemical Biology, Institute for Quantitative Biomedicine, Rutgers University, Piscataway, NJ 08854

E-mail: lwang@chem.rutgers.edu

1. Calculate the ^{1}H NMR chemical shift of a hydrogen bond network in ketosteroid isomerase

The enzyme ketosteroid isomerase (KSI) and its Asp40Asn mutant KSI^{D40N} contain a network of SHBs in the active site, which is composed of residues Tyr16, Tyr32 and Tyr57.[1] AI-PIMD simulations have been performed on the active-site hydrogen bond network in KSI^{D40N} with a QM/MM setup.[2] We extracted 455 configurations of the hydrogen bond network every 50 fs from the simulations, and calculated their NMR chemical shifts using the same parameters as those for the model molecules. Specifically, the electronic structures were described at the B3LYP level[3] with the D3 dispersion correction[4] and the 6-31+G(d,p) basis set, and the chemical shifts were calculated using the Gauge Independent Atomic Orbital method[5-7] as implemented in the Gaussian 16 software package.[8] In the network, residues Tyr16 and Tyr32 each form a SHB with Tyr57. When we consider the SHB between Tyr16 and Tyr57, we define the proton sharing coordinate as $\nu = d_{O16,H16} - d_{O57,H16}$, where $d_{O16,H16}$ and $d_{O57,H16}$ are the distances from the proton in residue 16 to the oxygen atom in residue 16 and 57, respectively. Similarly, a coordinate of $\nu = d_{O32,H32} - d_{O57,H32}$ is used when treating the SHB between Tyr32 and Tyr57. The resulting ^{1}H NMR chemical shift of each configuration is plotted as a function of ν in Fig. S1.

2. Determine the relation between the ^{1}H chemical shift and the proton position

We assume the probability distribution of ν in a SHB can be written as a linear combination of two Gaussian functions,

$$P(\nu) = \frac{A_1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\nu-\nu_0)^2}{2\sigma^2}} + \frac{A_2}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\nu+\nu_0)^2}{2\sigma^2}} \tag{S1}$$

Here the constants ν_0 and σ define the center and width of the Gaussian functions, which are normalized, $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\nu-\nu_0)^2}{2\sigma^2}} d\nu = 1$. The coefficients A_1 and A_2 represent the symmetry of a SHB and they satisfy $A_1 + A_2 = 1$ to ensure that $P(\nu)$ is normalized, i.e., $\int_{-\infty}^{\infty} P(\nu) d\nu = 1$. As shown in Figs. S2 and S3, the probability distribution of ν from the AIMD and AI-PIMD simulations of the model molecules can be well represented using Eq. S1.

Since $\delta_H(\nu) = 21.9 - 16.1\nu^2$, the average ^{1}H chemical shift can be calculated as

$$\begin{aligned}\langle \delta_H \rangle &= \int_{-\infty}^{\infty} \delta_H(\nu) P(\nu)\, d\nu \\ &= \int_{-\infty}^{\infty} (21.9 - 16.1\nu^2) \left[\frac{A_1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\nu-\nu_0)^2}{2\sigma^2}} + \frac{A_2}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\nu+\nu_0)^2}{2\sigma^2}} \right] d\nu \\ &= 21.9(A_1 + A_2) - 16.1(\nu_0^2 + \sigma^2)(A_1 + A_2) \\ &= 21.9 - 16.1(\nu_0^2 + \sigma^2) \end{aligned} \tag{S2}$$

From the first principles simulations, the proton hopping frequencies between the donor and acceptor groups of a SHB is much faster than the time scale of a typical NMR measurement. Therefore, the experimental ^{1}H NMR chemical shift of a system containing a SHB corresponds to $\langle\delta_H\rangle$ from the simulations. We consider ν_0 as the average proton position in a SHB, which is equivalent to the proton position observed in experiments. When a SHB is highly asymmetric, like in the case of CUA, ν_0 is the average position of the proton from the simulations. When a SHB is symmetric, one cannot distinguish between its donor and acceptor atoms, and so we take the absolute value of ν from the simulations and use its average as ν_0.

From Eq. S2, $\langle\delta_H\rangle = 21.9 - 16.1\nu_0^2 - 16.1\sigma^2$. Compared to AIMD simulations, AI-PIMD simulations include both electronic and nuclear quantum effects and mimic the experimental conditions more closely. Therefore, we will use the ν_0 values from the AI-PIMD simulations of the model molecules, as listed in Table S1. As the BLYP density functional is known to overstructure hydrogen bonded systems,[9, 10] it overestimates the ^{1}H chemical shift in the AI-PIMD simulations. To alleviate the problem, we determine σ by globally minimizing the differences between the predicted chemical shift, $\langle\delta_H\rangle_{\mathrm{predict}}$, and the experimental chemical shift of all the systems. The resulting σ is 0.3 Å and the root-mean-square deviation is 0.8 ppm. The $\langle\delta_H\rangle_{\mathrm{predict}}$ values are also listed in Table S1.

Table S1. ν_0 from AI-PIMD simulations and the experimental ^{1}H chemical shifts of the model molecules. The chemical shifts predicted from Eq. S3 are also included.

Model molecule	DMANH		HM		DHND	CUA
solvent	acetonitrile	water	acetone	water	water	DMSO
$\nu_{0,\mathrm{AI-PIMD}}$ (Å)	0.32	0.32	0.23	0.28	0.42	-0.34
$\delta_{H,exp}$ (ppm)	18.7	18.5	20.7	20.2	17.7	17.4
$\langle\delta_H\rangle_{\mathrm{predict}}$ (ppm)	18.9	18.9	19.6	19.2	17.7	18.6

Therefore, Eq. S2 becomes

$$\langle\delta_H\rangle = 20.5 - 16.1\nu_0^2. \qquad \text{(S3)}$$

3. SI figures

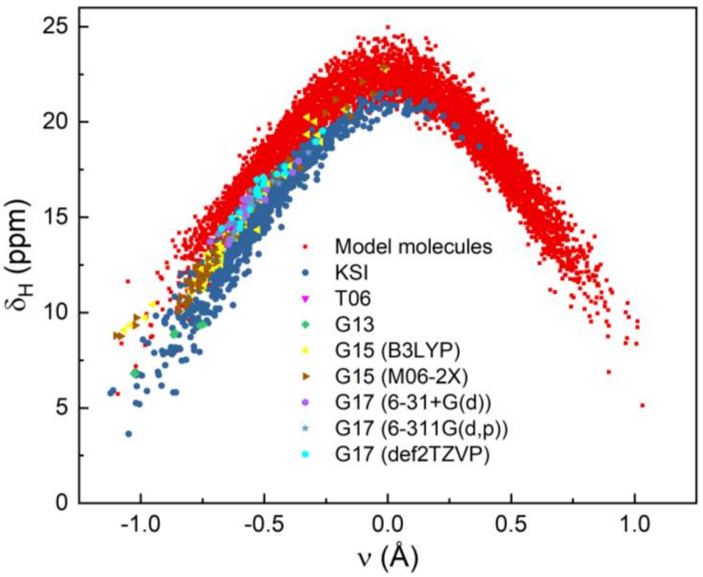

Fig. S1. Correlation of δ_{H} and ν from the model molecules and KSI. Results from previous computational studies are also included. In the data set T06, δ_{H} is for a proton in the N–H···O hydrogen bond in a pyrrole derivative.[11] In G13, NMR calculations are performed for the hydrogen bond between phenol and different solvents.[12] In G15, NMR chemical shifts are calculated for a set of organic compounds with intermolecular or intramolecular hydrogen bonds using the B3LYP[3] and the M06-2X[13] density functionals and the 6-31+G(d) basis set.[14] In G17, the NMR chemical shifts of the enol-enol tautomers of β-dicarbonyl compounds are computed using the B3LYP density functional[3] and the 6-31+G(d), 6-311G(d,p) and def2TZVP basis set.[15]

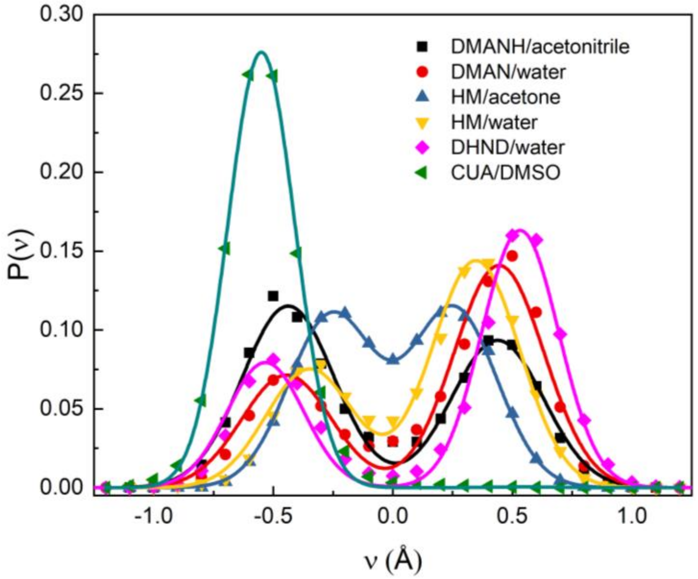

Fig. S2. The probability distribution of ν from the AIMD simulations of the model molecules in different solvents (symbols) and the least squares fitting of them using Eq. S1 (lines). The R^2 values for the fitting range between 0.971 and 0.999.

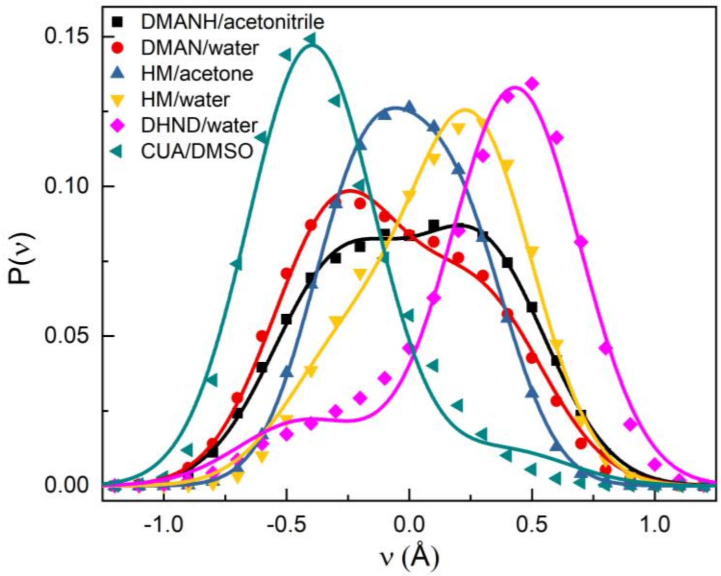

Fig. S3. The probability distribution of ν from the AI-PIMD simulations of the model molecules in different solvents (symbols) and the least squares fitting of them using Eq. S1 (lines). The R^2 values for the fitting range between 0.986 and 0.998.

4. References

1. N.-C. Ha, M.-S. Kim, W. Lee, K. Y. Choi and B.-H. Oh, *J. Biol. Chem.*, 2000, **275**, 41100-41106.
2. L. Wang, S. D. Fried, S. G. Boxer and T. E. Markland, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 18454-18459.
3. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
4. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
5. F. London, *J. Phys. Radium*, 1937, **8**, 397-409.
6. R. Ditchfield, *Mol. Phys.*, 1974, **27**, 789-807.
7. K. Wolinski, J. F. Hinton and P. Pulay, *J. Am. Chem. Soc.*, 1990, **112**, 8251-8260.
8. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Journal*, 2016.
9. J. A. Morrone and R. Car, *Phys. Rev. Lett.*, 2008, **101**, 017801.
10. M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales and T. E. Markland, *Chem. Rev.*, 2016, **116**, 7529-7550.
11. A. V. Afonin, I. A. Ushakov, L. N. Sobenina, Z. V. Stepanova, O. g. V. Petrova and B. A. Trofimov, *Magn. Reson. Chem.*, 2006, **44**, 59-65.

12. M. G. Siskos, V. G. Kontogianni, C. G. Tsiafoulis, A. G. Tzakos and I. P. Gerothanassis, *Org. Biomol. Chem.*, 2013, **11**, 7400-7411.
13. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215-241.
14. M. G. Siskos, A. G. Tzakos and I. P. Gerothanassis, *Org. Biomol. Chem.*, 2015, **13**, 8852-8868.
15. M. G. Siskos, M. I. Choudhary and I. P. Gerothanassis, *Org. Biomol. Chem.*, 2017, **15**, 4655-4666.