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

Supplementary Information

1. FTIR raw data

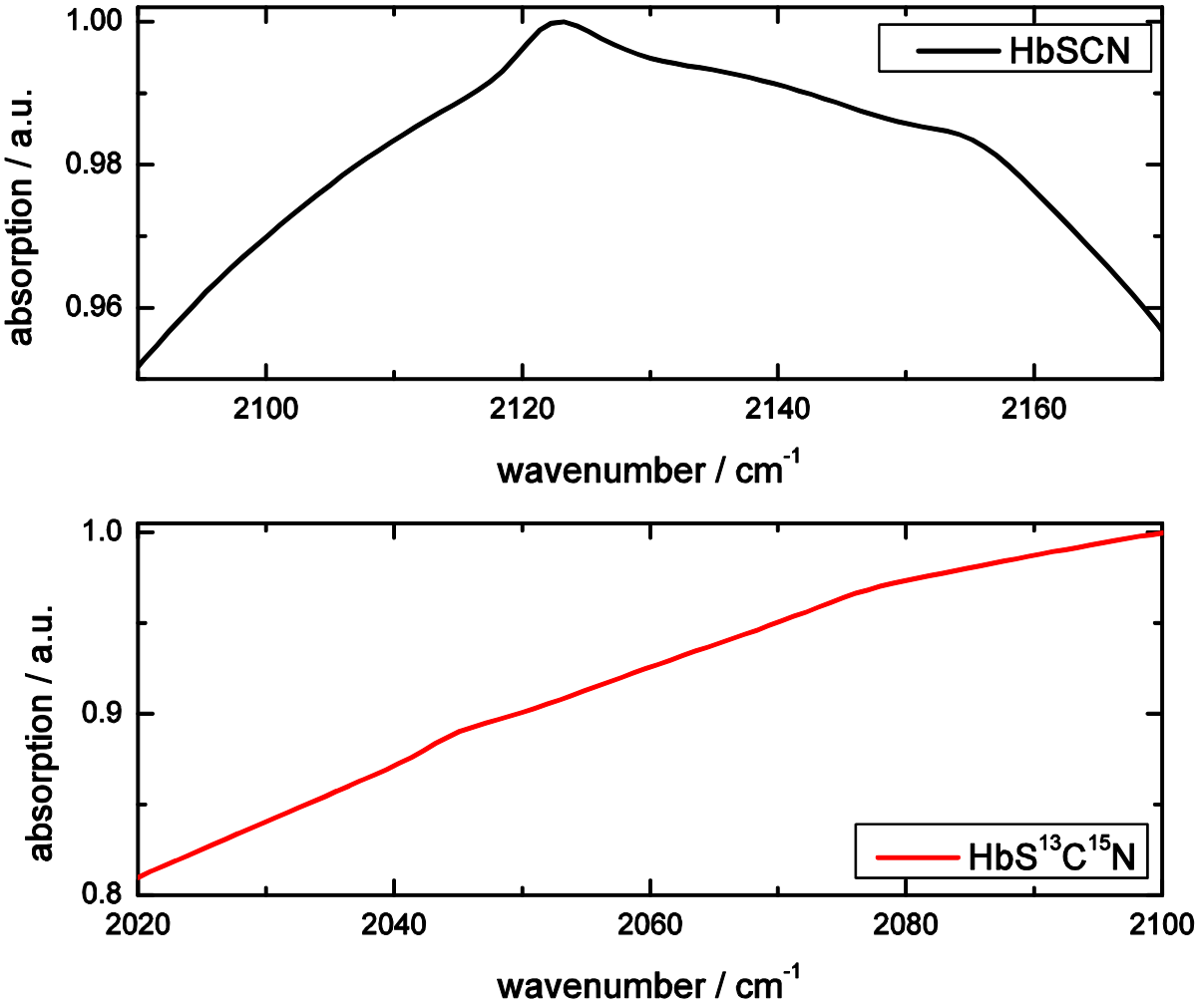


Figure S1. Raw data of the baseline-corrected data shown in Figure 2. Normalized FTIR absorption spectra of the labeled protein in natural abundance (top panel) and of the isotope label (bottom panel). The CN absorption bands (of CN bound to the heme and that of cyanylated cysteine) are superimposed on a water baseline (H_2O in the top panel and D_2O in the lower panel).

2. Raw data of the ultrafast experiments

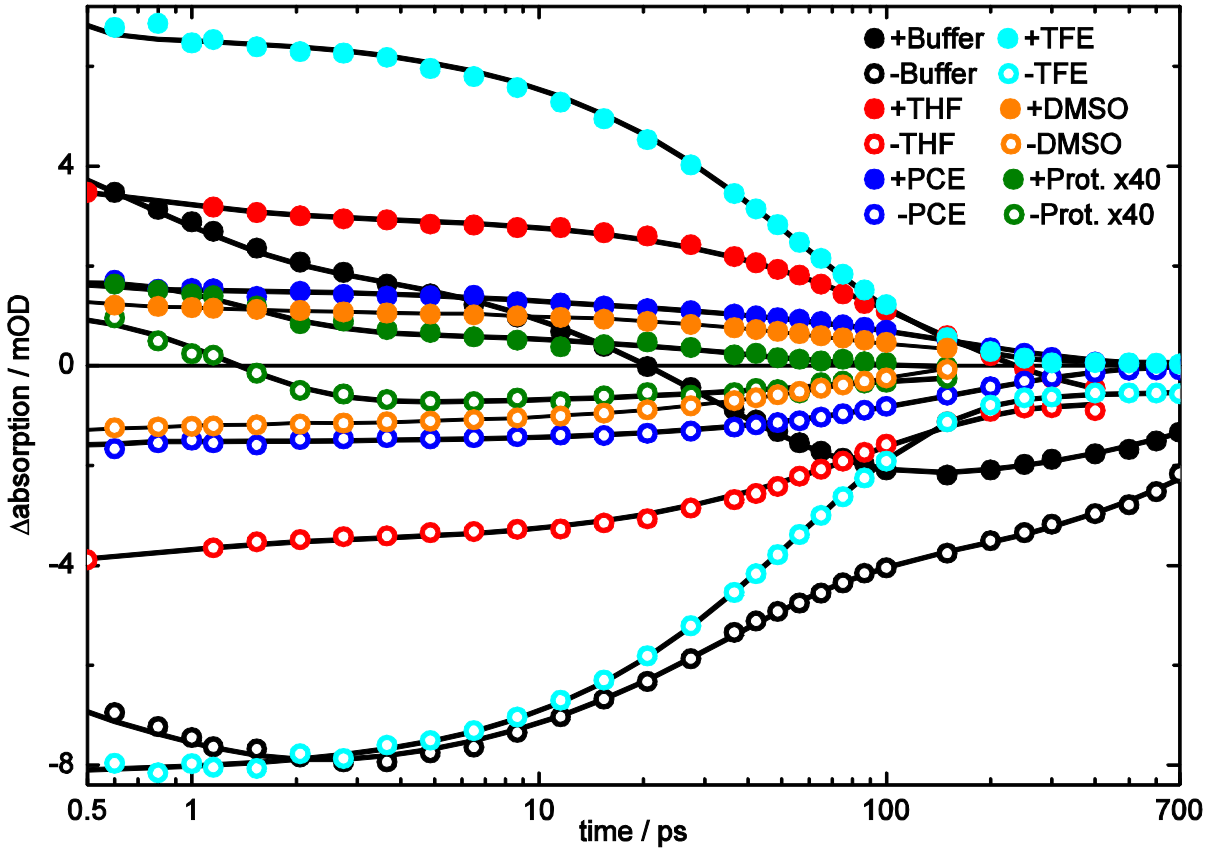


Figure S2. Reproduction of Figure 4A-B in a single plot to show the raw (not normalized) data of the maximum (closed circles, represented by a + in the legend) and minimum (open circles; - in the legend) absorption. The legend denotes which solvent is used for MeSCN. The Hb-S^{13}C^{15}N protein is measured in 100mM phosphate buffer. Note that the protein signal is multiplied by a factor 40. The continuous lines are fits derived from global analysis.

3. Spectral diffusion for additional pump wavelengths

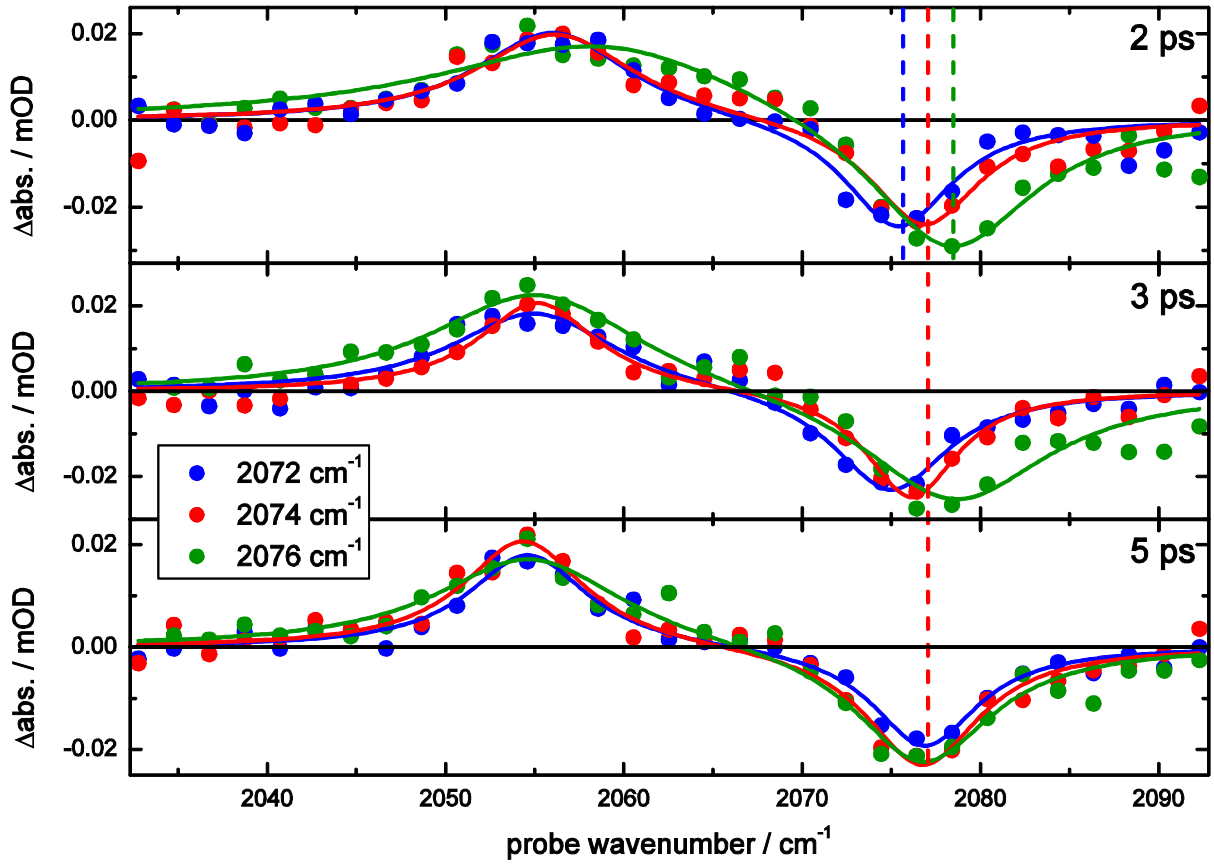


Figure S3. Spectral diffusion of the $S^{13}C^{15}N$ label. Reproduction of Figure 7, but with different pump wavenumbers used (closer to the main bleach).

4. Sample characterization and determination of the extinction coefficient of SCN in bovine hemoglobin

Extensive care has been taken to characterize and analyze the samples. We have subjected the samples to UV-Vis spectroscopy, FTIR spectroscopy, analytical ultracentrifugation, dialysis, and mass spectrometry. UV-Vis spectroscopy (Jasco V-670) allows the determination of the heme concentration via the extinction coefficient for metcyanohemoglobin (Fe^{3+}) at 540 nm (11 mM^{-1} cm^{-1} per heme group[33–36]). Under the same conditions a control experiment with sperm whale myoglobin (Sigma Aldrich) has been done, revealing that all hemoglobin subunits were occupied with a heme group (the absorption at 540 nm was measured to be 3.8 times higher for hemoglobin at equal concentration). Because the heme groups have a high affinity for CN[27] and we used a 100 fold excess concentration of KCN (see Materials and methods), they are considered to be fully saturated with CN. Analytical ultracentrifugation (Beckman Optima-XLA) has revealed that our samples are predominantly (>70%) in tetrameric form, 16% dimers of tetramers and 9% dimers (data not shown). We are confident that our spectroscopy samples were free of free heme groups because our samples have been subjected to two consecutive dialysis steps (having a cut-off of at 14 kDa, with the heme having a mass of roughly 600 Da).

The labeling efficiency of our protocol has been tested by maldi-mass spectrometry (Applied Biosystems Voyager-DE PRO). We determined the mass of a fraction of the isotope-labeled protein that has been used for our laser experiments, and repeated this procedure for a freshly prepared pure unlabeled methemoglobin control sample (see Figure S4). We could not detect a significant fraction of unlabeled protein in our labeled samples, nor substantial loss of our label over time.

It has been observed that cyanylation of cysteines involved in disulphide bridges can lead to cleavage of the protein backbone.[30,31] We cannot exclude that a fraction of the protein is cleaved, however, potential fragments

of 10.0 kDa and 6.0 kDa would have been removed by dialysis (cut-off at 14 kDa) and are not visible in the mass spectra (see arrows in the inset of Figure S4). To reduce potential cleavage the buffer conditions could be adjusted to lower pH. Our measurements are carried out with the sample after two fold dialysis and investigation by mass spectrometry and therefore cleavage does not influence the calculation of the extinction coefficient as well as all the other spectroscopic data.

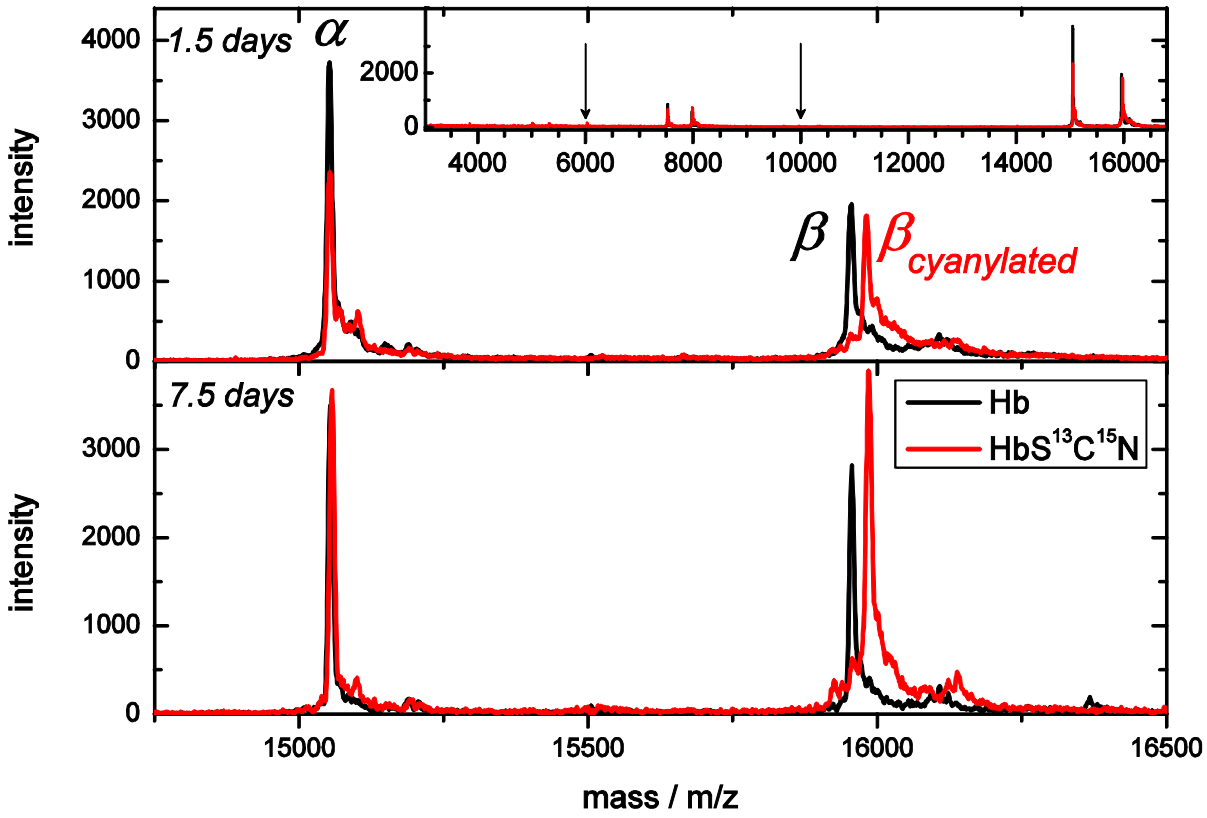
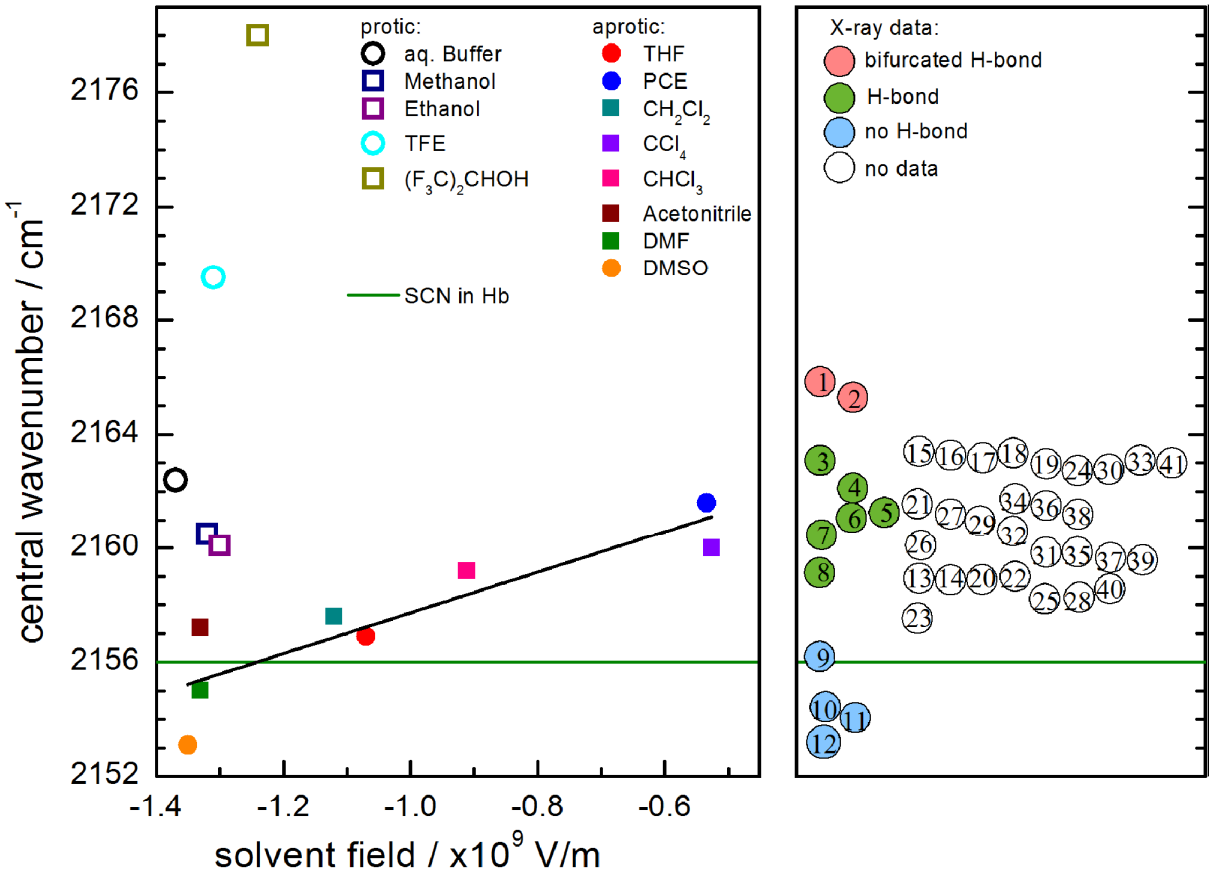


Figure S4. Maldi-mass spectra of Hb and $S^{13}C^{15}N$ –Hb. The samples contain a 100-fold dilution (corresponding to 5-10 μM) of our freshly prepared labeled sample and of freshly prepared unlabeled protein solutions. The time after cyanylation is depicted in each panel. Both panels share the same legend. As matrix a 1:10 ratio of 2,5-dihydroxybenzoic acid and 5-methoxysalicylic acid is used. The labels correspond to hemoglobin's sub-units.. Only the labeled samples have a higher mass for the β-subunits, corresponding to the mass difference due to cyanylation. The theoretical mass difference is 27 mass units for the isotope label, while we measure a difference in mass units of 27 and 29 in the top and bottom panels, respectively. The inset in the upper panel shows a zoom of the same graph in an extended mass range. In the lower mass range only higher-charged protein subunits are seen for both Hb and Hb $S^{13}C^{15}N$, but no significant peaks for cleaved protein due to cyanylation (expected at 10.0 kDa and 6.0 kDa, indicated by the arrows) are observed for the $S^{13}C^{15}N$ samples.

In order to calculate the cyanylated cysteine concentration in our Hb samples for the determination of its IR extinction coefficient, an accurate determination of the pathlength in the IR was performed using the spectral fringes produced by an empty measurement cell. This pathlength and the heme concentration (calculated from the absorption at 540 nm) where used to calculate the IR extinction coefficient for one HbSCN and for two $HbS^{13}C^{15}N$ samples (The sample characterization reported above allows the calculation of the concentration of SCN and $S^{13}C^{15}N$ from the heme concentration). The extinction coefficient for CN and $^{13}C^{15}N$ bound to the heme (30 $M^{-1}cm^{-1}$) could be determined from the same samples and from a concentration series for metcyanohemoglobin. The extinction coefficient of heme bound CN was then used to determine the cysteine concentration for two additional samples for which IR spectra had been measured. Assuming equal SCN and $S^{13}C^{15}N$ extinction coeffcients, the 5 samples resulted in an averaged value for the SCN extinction coefficient of $41 \pm 5 M^{-1}cm^{-1}$ (see Table 1).

5. Extended Onsager plot with overview of protein values reported in literature



1 KSI D40N F86C
2 KSI D40 F86C
3 Ralβ R20C
4 Ralβ N54C
5 KSI D40 M116C
6 Ralβ Y31C
7 Ralβ K52C
8 Ralβ K32C
9 RNase S pH 8.0
10 RNase S pH 4.5
11 KSI D40N M105C
12 KSI D40 M105C
13 RC
14 hALR2
15 NTAIL S407C
16 NTAIL L498C
17 NTAIL S491C
18 NTAIL V517C
19 MBP F87C (buffer)
20 MBP F87C (DPC)
21 CM 15 L4C (buffer)
22 CM 15 L4C (DPC)
23 CM 15 L4C (E.coli lipids)
24 CM 15 I8C (buffer)
25 CM15 I8C (DPC)
26 CM 15 I8C (E.coli lipids)
27 CM 15 A10C (buffer)
28 CM 15 A10C (DPC)
29 CM 15 A10C (E.coli lipids)
30 CM 15 V14C (buffer)
31 CM 15 V14C (DPC)
32 CM 15 V14C (E.coli lipids)
33 Ralβ I18C
34 Ralβ N27C
35 Ralβ G28C
36 Ralβ N29C
37 Ralβ M30C
38 Ralβ S33C
39 KSI D40 M116C
40 KSI D40N L61C
41 DHFR

Figure S5. Central wavenumber of the nitrile group of MeSCN in different solvents (left) and cyanylated cysteine in different proteins (right). Left: Central wavenumber of MeSCN's thiocyanate group in protic (open symbols) and aprotic (solid symbols) as function of the solvent field $\vec{F}_{Onsager}$. A linear fit (black line; R^2 = 0.76) of the central wavenumber as function of the solvent field ($\tilde{\nu} = \tilde{\nu}_0 + |\Delta\vec{\mu}| \times F_{Onsager}$) is shown, resulting in $\tilde{\nu}_0 = 2165(2)\mathrm{cm}^{-1}$ and a Stark tuning rate of $|\Delta\vec{\mu}| = 7(2) \cdot 10^{-9}\ \mathrm{cm}^{-1}/(\mathrm{V/m})$ for the aprotic solvents reported here together with those reported in previous publications. The measured central wavenumber of SCN in Hb is depicted by a green line. The data points symbolized by circles are measured in this work, the data in squares is taken from ref. [38] with the exception of DMF, which has been taken from [22]. Right: Central wavenumber of thiocyanate incorporated into proteins (data points 1, 2, 5, 11, 12, 39 and 40 are taken from [48], 3, 4, 6, 7, 8 and 33 to 38 from [21], 9 and 10 from [49], 13 and 14 from [15], 15 to 18 from [2], 19 and 20 from [39], 21 to 32 from [57], and 41 from [22]. Note that our value of 2156 cm^{-1} is among the lowest reported so far.

6. Kamlet-Taft's solvatochromic relationship for MeSCN in different solvents

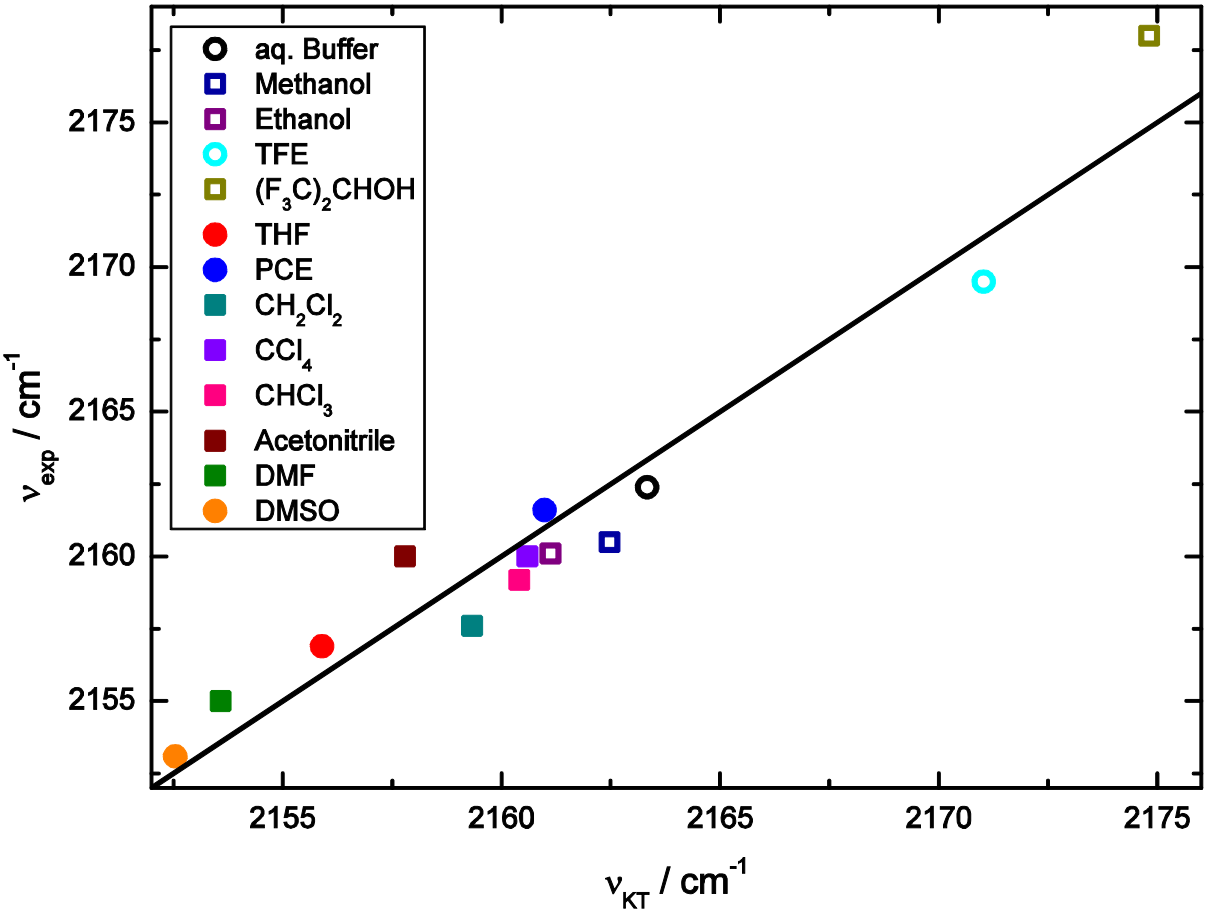


Figure S6. The measured wavenumber shifts for MeSCN in different solvents are plotted against the shifts predicted by the Kamlet-Taft solvatochromic relationship equation for MeSCN (error in brackets) $\nu_{KT} = 2163(2) + 7.6(0.8) \cdot \alpha - 8(2) \cdot \beta - 4(3) \cdot \pi^*$ ($R^2 = 0.94$), where α is the hydrogen bond donor aciditiy, β the hydrogen bond acceptor basicitiy and π^* the solvent dipolarity/polarizability.[46,47] The black line represents the diagonal.

7. References

The reference numbers refer to those in the main paper