Supporting Information
Exploring the 2D-IR Repertoire of the -SCN Label to Study Site-Resolved Dynamics and Solvation in the Calcium Sensor Protein Calmodulin

Julian M. Schmidt-Engler,[a] Rene Zangl,[b] Patrick Guldan,[a] Nina Morgner,[b] and Jens Bredenbeck[a]

[a] Johann Wolfgang Goethe-University, Institute of Biophysics, Max-von-Laue-Str. 1,

60438, Frankfurt am Main, Germany

[b] Johann Wolfgang Goethe-University, Institute of Physical and Theoretical Chemistry, Max-von-Laue-Str. 7, 60438, Frankfurt am Main, Germany

E-mail: bredenbeck@biophysik.uni-frankfurt.de

Tel: +49 69 798 46428

Fax: +49 69 798 46423

Contents:

1. Mass spectrometry results

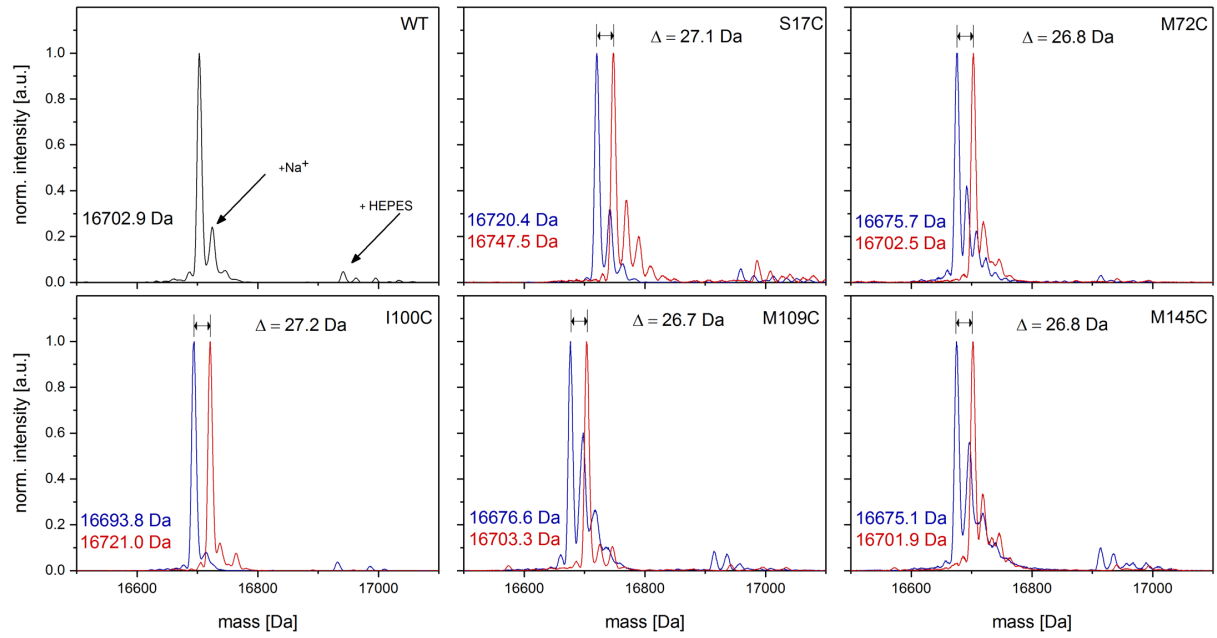

Fig. S1 Deconvoluted mass spectra of CaM wt and both unlabelled and labelled mutants. The $S^{13}C^{15}N$ label at each mutant (red) results in a mass shift which is in excellent agreement with the expected shift of +27 Da, relative to the respective unlabelled mutant (blue). Furthermore, the mass spectra of the labelled mutants contain no signal of unlabelled CaM, which indicates quantitative yields. Sodium and HEPES attachments are exemplarily denoted in the CaM wt spectrum.

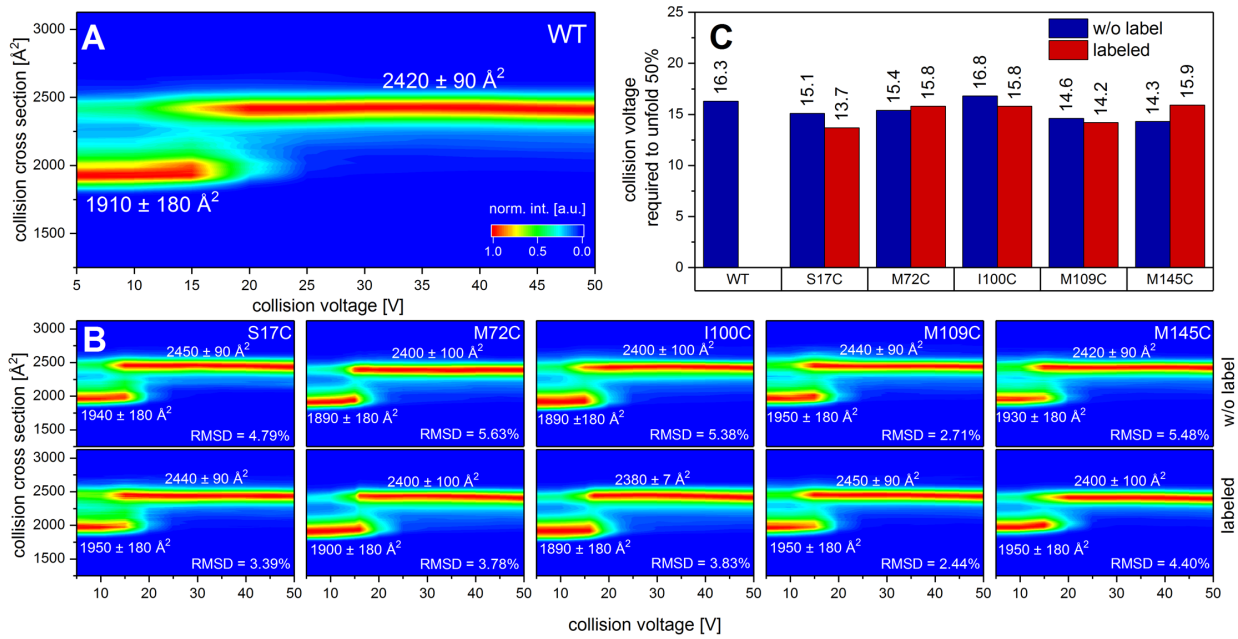

Fig. S2 Collision induced unfolding (CIU)[1] fingerprint contour plots for the 7+ charge state of CaM wt and the respective unlabelled and labelled mutants. By increasing the collision voltage to 16.3 V, the compact structure of 1910 Å² turns into a CIU product of 2420 Å² (A). The collision cross sections of the corresponding unlabelled and labelled mutants correlate with those of CaM wt, which indicates that the mutations and labels do not affect the structure of the protein (B). The calculated RMSD values, estimated by comparing the CIU fingerprints of CaM wt to all variants are comparably low, indicating no significant differences in the respective unfolding behaviours. Additionally, the comparison of the collision voltage required to unfold 50% of the compact structures shows only minor differences, which denotes a similar gas phase stability between all analysed samples (C).

2. Detailed description of FFCF parameter calculation

The procedure to extract the FFCF parameters of an oscillator from a time series of 2D-IR spectra and its linear absorption spectrum has been developed by Kwak *et al.*[2] What follows is a step-by-step protocol applied in this study to extract described characteristics of the FFCF.

In the first step the tilt of the band in the 2D spectrum was quantified using the CLS method that judges the tilt based on the maximum of the band in spectral slices parallel to the detection axis. For each spectral slice one centre point is determined, which in the most intense region of the band is linearly fitted (see fig. 3 grey bar). The CLS is the inverse slope of the fitted line. If frequency correlation between excitation frequency and read out frequency decays the CLS decreases. This decay of the CLS was fitted by a sum of exponentials:

$$CLS(t) = \sum_{i=1}^{n} A_i e^{-t/\tau_i}$$

The decay of the CLS results from the oscillator sampling all possible environments. If these dynamics are slower than the timescale of the experiment i.e. 3-5x the vibrational lifetime of the oscillator its contribution is fitted by a static component i.e. $\tau=\infty$. It is caused by conformational sampling slower than tens of ps, which for example arises from motion of amino acids and collectives of amino acids. Correlation length of processes in proteins on this timescale was found to be about 18 Å.[3] Compared to that is the radius of gyration of CaM Apo 21.6 Å, Holo 23.1 Å and of Holo in presence of the M13 peptide 18.8 Å (calculated using Crysol[4] and PDB: 1CFD, 1CLL and 2BBM). In combination with the FWHM of the band in the linear absorption spectrum and its vibrational lifetime obtained by pump-probe time-resolved IR spectroscopy all parameters of the FFCF can be computed. The normalised FFCF has the form:

$$FFCF(t) = \frac{\delta(t)}{T_2} + \sum_{i=1}^{n} \Delta_i^2 e^{-t/\tau_i}$$

Many CLS decays feature an incomplete relaxation which is modelled by a static contribution Δ_s in the FFCF:

$$FFCF(t) = \frac{\delta(t)}{T_2} + \Delta_s^2 + \sum_{i=1}^{n} \Delta_i^2 e^{-t/\tau_i}$$

Exceedingly fast dynamics around the oscillator lead to an CLS < 1 at T_w= 0. The width of the band is caused by homogeneous broadening in the motionally narrowed regime ($\tau\Delta < 1$). For this component τ and Δ cannot be evaluated separately, instead only its product T_2, the homogeneous dephasing time, can be computed ($T_2 = \tau\Delta^2$).[2] It is the sum of the pure-dephasing time T_2^*, the vibrational lifetime T_1 and the orientational relaxation lifetime T_{OR}, which are all purely homogeneous contributions i.e. do not arise from slow conformational sampling (inhomogeneous broadening).

$$\frac{1}{T_2} = \frac{1}{T_2^*} + \frac{1}{2T_1} + \frac{1}{3T_{OR}}$$

Therefore the deviation of the CLS from 1 at $T_w = 0$ is the fraction of the linewidth caused by homogeneous line broadening and therefore calculated as follows:[2]

$$\frac{1}{\pi T_2} = FWHM_{FTIR} * \left(1 - \sum_{i=1}^{n} A_i\right)$$

From that the pure dephasing homogeneous line width Γ^* can be computed:

$$\Gamma^* = \frac{1}{\pi T_2^*}$$

The other contribution to the observed linewidth arises from inhomogeneous broadening, which is described by the latter terms of the FFCF. For a single exponential decay its amplitude Δ_1 can be obtained using:

$$\Delta_1 = \sqrt{A_1} * \frac{FWHM_{FTIR}}{2\sqrt{2\ln 2}}$$

The overall square root is required because the fit involves the square of the absolute amplitudes. Division by $2\sqrt{2\ln 2}$ is required to change the FWHM into the standard deviation. If the decay is more than single exponential (e.g. is bi-exponential or includes an offset) the fit amplitude needs to be normalised by the total amplitude:

$$\Delta_n = \sqrt{\frac{A_n}{\sum_{i=1}^{n} A_i}} * \frac{FWHM_{FTIR}}{2\sqrt{2\ln 2}}$$

Furthermore we compute the percentage of line broadening that can be ascribed to inhomogeneous broadening we introduce the parameter $Frac^{Inh}$ that is the sum of the inhomogeneous contributions to the CLS (i.e. the CLS at $T_w = 0$):

$$Frac^{Inh} = \sum_{i=1}^{n} A_i$$

3. CLS analysis of all studied labels

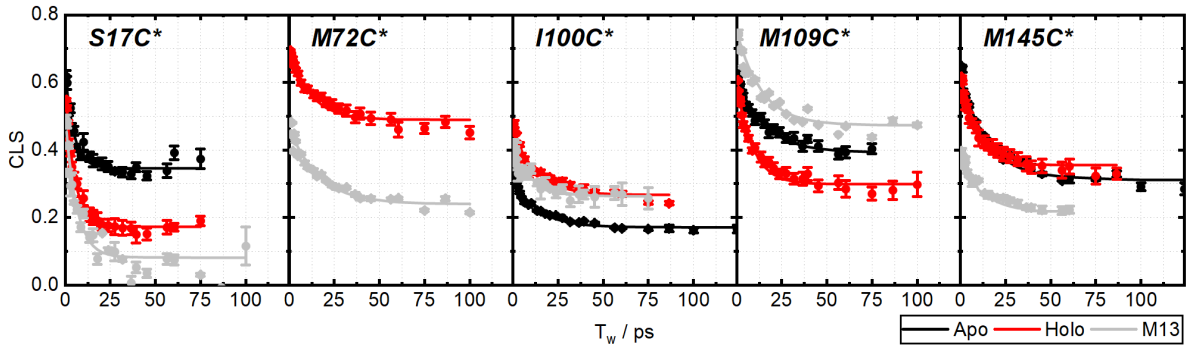

Fig. S3 CLS as function of the waiting time for the studied CaM -SCN labels. The data points are shown as dots, where the error of each point corresponds to the error of the fitted slope. The mono-exponential fits are shown as lines.

Tab. S1 Parameters from mono-exponential CLS analysis performed on the 1-2 transition (ESA) of the SCN labels in CaM. The CLS analysis on MeSCN in D_2O for the 0-1 and 1-2 transition are also included, showing that evaluation of the two transitions yields the same results (within error). The data was recorded with a sample thickness of 150 µm and in deuterated buffer.

Label / State		Γ* / cm^{-1}	Δ_1 / cm^{-1}	τ_1 / ps	Δ_S / cm^{-1}	$Frac^{Inh}$	FWHM / cm^{-1}
S17C*-	*Apo*	5.2±0.2	3.9±0.4	4.7±0.5	4.3±0.2	0.63±0.02	13.7±0.2
	Holo	5.3±0.2	4.4±0.3	5.0±0.3	2.9±0.2	0.58±0.02	12.4±0.4
	M13	6.2±0.5	4.7±0.7	6.2±1.0	2.1±1.1	0.49±0.04	11.2±0.3
M72C*-	*Apo*	-	-	-	-	-	-
	Holo	4.0±0.1	3.0±0.2	11.5±0.9	4.6±0.1	0.70±0.02	12.9±0.8
	M13	5.1±0.2	2.5±0.2	16.3±2.1	2.6±0.2	0.47±0.02	8.6±0.7
I100C*-	*Apo*	4.2±0.1	1.7±0.1	13.3±1.2	1.9±0.1	0.32±0.01	6.0±0.2
	Holo	4.2±0.2	2.1±0.2	12.6±1.9	2.4±0.1	0.46±0.02	7.5±0.5
	M13	5.3±0.2	2.1±0.2	11.4±2.4	3.0±0.1	0.39±0.02	7.1±0.4
M109C*-	*Apo*	4.8±0.1	3.1±0.2	15.4±1.9	4.1±0.1	0.61±0.02	12.0±0.8
	Holo	5.4±0.2	3.9±0.3	8.7±0.4	3.9±0.2	0.60±0.02	13.1±0.5
	M13	4.2±0.1	3.1±0.2	16.6±2.1	4.1±0.2	0.66±0.02	9.5±0.6
M145C*-	*Apo*	4.8±0.2	3.8±0.3	14.5±1.7	3.8±0.3	0.63±0.02	12.7±0.7
	Holo	3.8±0.1	3.3±0.1	9.0±0.7	4.0±0.1	0.69±0.01	13.1±0.5
	M13	4.8±0.2	2.2±0.2	13.9±1.8	2.4±0.1	0.39±0.02	8.2±0.4
MeSCN	*0-1*	6.8±0.2	4.5±0.3	1.6±0.1	1.1±0.2	0.38±0.01	11.0±1.2
	1-2	7.0±0.2	4.6±0.2	1.6±0.1	1.0±0.1	0.36±0.02	11.1±0.9

4. Supplementary results from FTIR and 2D-IR

Tab. S2 The band centre was determined by inspection, with the exception of M72C* where the sum of two Gaussian functions was used for fitting the FTIR spectrum. The SASA values shown are determined for the wildtype residues in the according positions before point mutation to cysteine from structures (PDB ID: Apo: 1CFD[6]; Holo: 1CLL[7]; Holo+M13: 2BBM[8]) using the FreeSASA script.[9] To account for the different size of the sidechains the obtained values were normalised to the SASA of alanine of an Gly-X-Gly tripeptide in extended conformation, where X is the according amino acid (A= 67 Å^2, S= 80 Å^2, I= 140 Å^2, M= 160 Å^2).[10] The vibrational lifetimes in either D_2O or H_2O buffer were determined by Global analysis of a spectral slice from the 2D-IR time series parallel to the detection dimension at the centre of the -SCN absorption using the Globe toolbox.[11] The solvent exclusion coefficient SEC measures the change in vibrational lifetime introduced by solvent exchange from H_2O to D_2O relative to the vibrational lifetime of the SCN model compound MeSCN.[12]

Label / State		Band centre / cm^{-1}	SASA / Å^2	τ_{Vib} D_2O / ps	τ_{Vib} H_2O / ps	SEC
S17C-*	*Apo*	2081.9±0.3	50.1	52.6±1.6	31.0±1.4	0.15±0.01
	Holo	2085.0±0.2	57.8	53.5±0.8	31.3±0.9	0.16±0.01
	M13	2083.7	79.6	61.4±5.6	35.6±1.8	0.19±0.01
M72C-*	*Apo α/β*	2074.1 ±0.2 2080.6 ±0.7	2.6	63.0 ±1.2 62.0 ±2.1	55.4 ±5.9 36.0 ±4.1	0.75 ±0.07 0.19 ±0.02
	Holo	2081.6±0.2	19.1	60.8±0.9	51.3±3.3	0.63±0.03
	M13	2076.1	25.6	63.2±1.5	54.0±2.6	0.72±0.03
I100C-*	*Apo*	2079.2±0.1	3.1	60.5±1.1	64.0±7.4	0.94±0.11
	Holo	2081.5±0.2	1.4	65.8±1.4	60.8±6.5	0.95±0.09
	M13	2081.7	0.4	60.6±7.4	69.5±6.2	1.10±0.10
M109C-*	*Apo*	2080.6±0.1	1.4	58.2±1.5	59.0±3.5	0.76±0.04
	Holo	2081.5±0.3	17.2	59.7±0.8	49.9±2.4	0.53±0.02
	M13	2078.7	3.2	62.0±1.7	50.5±2.9	0.60±0.03
M145C-*	*Apo*	2082.4±0.3	9.4	65.5±1.5	53.5±2.0	0.75±0.02
	Holo	2080.1±0.3	36.2	61.7±0.7	50.3±2.0	0.58±0.02
	M13	2075	15.2	61.6±3.0	51.1±7.4	0.60±0.08

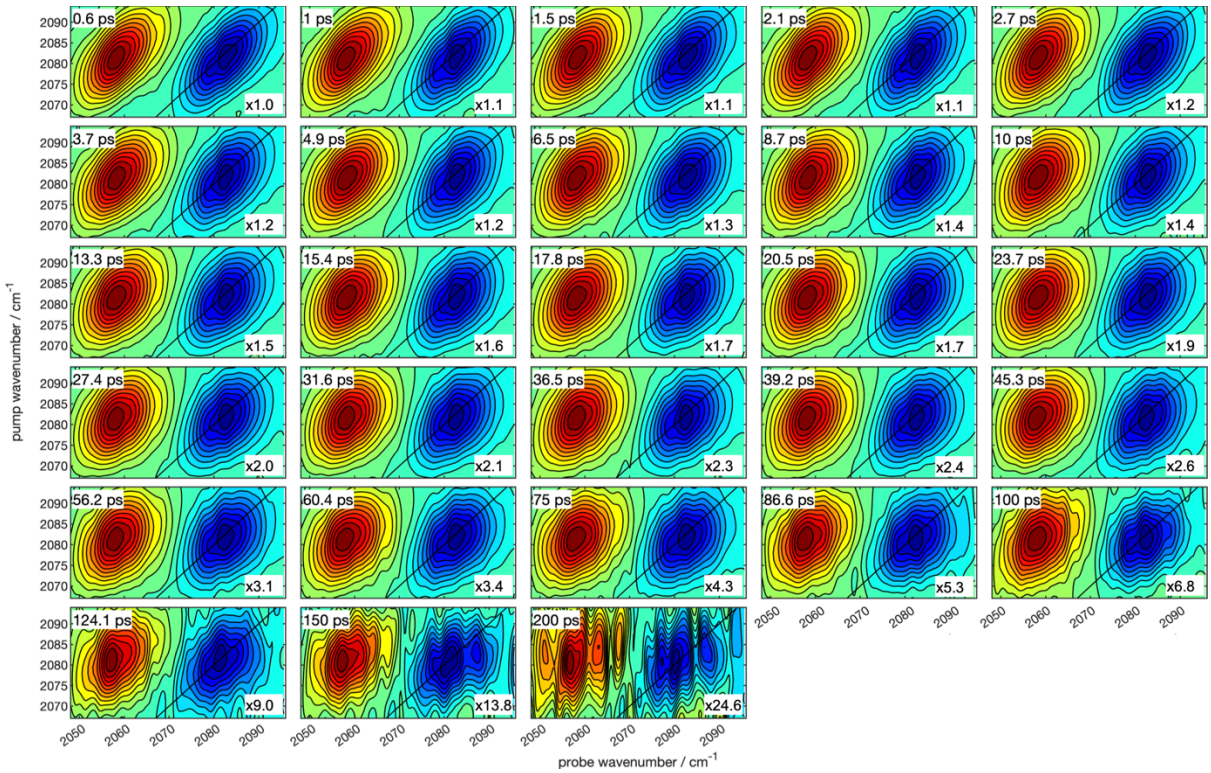

Fig. S4 All 28 2D-IR spectra captured for M145C*-Holo in D_2O. The waiting time T_w for each spectrum is shown in the upper left corner, whereas the scaling parameter of the spectral intensity of each spectrum compared to the time point 0.6 ps is shown in the lower right corner. The diagonal is shown as black line. Each contour level corresponds to 10% of the maximum amplitude ranging from blue (negative) to red (positive).

References

1. J. D. Eschweiler, J. N. Rabuck-Gibbons, Y. Tian and B. T. Ruotolo, *Anal. Chem.*, 2015, **87**, 11516–11522.
2. K. Kwak, S. Park, I. J. Finkelstein and M. D. Fayer, *J. Chem. Phys.*, 2007, **127**, 124503.
3. Z. Bu, D. A. Neumann, S. H. Lee, C. M. Brown, D. M. Engelman and C. C. Han, *J. Mol. Biol.*, 2000, **301**, 525–536.
4. D. Svergun, C. Barberato and M. H. J. Koch, *J. Appl. Crystallogr.*, 1995, **28**, 768–773.
5. S. Bagchi, S. G. Boxer and M. D. Fayer, *J. Phys. Chem. B*, 2012, **116**, 4034–4042.
6. H. Kuboniwa, N. Tjandra, S. Grzesiek, H. Ren, C. B. Klee and A. Bax, *Nat. Struct. Mol. Biol.*, 1995, **2**, 768–776.
7. R. Chattopadhyaya, W. E. Meador, A. R. Means and F. A. Quiocho, *J. Mol. Biol.*, 1992, **228**, 1177–1192.
8. M. Ikura, G. Clore, A. Gronenborn, G. Zhu, C. Klee and A. Bax, *Science*, 1992, **256**, 632–638.
9. S. Mitternacht, *F1000Res.*, 2016, **5**, 189.
10. S. Miller, J. Janin, A. M. Lesk and C. Chothia, *J. Mol. Biol.*, 1987, **196**, 641–656.
11. L. J. G. W. van Wilderen, C. N. Lincoln and J. J. van Thor, *PLoS ONE*, 2011, **6**, e17373.
12. J. M. Schmidt-Engler, L. Blankenburg, B. Błasiak, L. J. G. W. van Wilderen, M. Cho and J. Bredenbeck, *Anal. Chem.*, 2019, doi:10.1021/acs.analchem.9b03997.