Background Correction of Time-Resolved Infrared Spectroscopy Data for the 'Tricky' Transcriptional Regulator Protein, B_{12}-Dependent CarH

Contact Sam.Hay@Manchester.ac.uk

E. Wall

Department of Chemistry and Manchester Institute of Biotechnology
The University of Manchester, Manchester, UK

I. V. Sazanovich

Central Laser Facility
Rutherford Appleton Laboratory, Didcot, UK

M. Kurttila

Biometrology, Chemical and Biological Sciences
National Physical Laboratory, Teddington, Middlesex, UK

I. S. Camacho

Biometrology, Chemical and Biological Sciences
National Physical Laboratory, Teddington, Middlesex, UK

N. T. Hunt

Department of Chemistry and York Biomedical Research Institute
University of York, York, UK

A. R. Jones

Biometrology, Chemical and Biological Sciences
National Physical Laboratory, Teddington, Middlesex, UK

S. Hay

Department of Chemistry and Manchester Institute of Biotechnology
The University of Manchester, Manchester, UK

Introduction

Time-resolved infrared (TRIR) spectroscopy is a powerful method for studying biological molecules as it can provide information relating to excited state dynamics, conformational changes, intermolecular interactions and solvation.[1] However, TRIR measurements on biological samples can be complicated because they are difficult to prepare in high concentration, as is the case for many proteins. This gives rise to challenges when collecting and analysing TRIR data for such samples, including how to obtain useful results when working close to the detection limit. This article will discuss how we corrected for a delay-dependent step-like background pattern attributed to the detectors from the Central Laser Facilities LIFETIME instrument. The TRIR data were collected for wildtype and variant chromophore binding domains of the photoreceptor protein B_{12}-dependent CarH (i.e., the DNA domain has been truncated), in addition to the full-length protein bound to DNA. This work may be of use to those who study challenging biological samples that may be limited in quantity and/or produce low signals where background patterns become noticeable and hinder a meaningful kinetic analysis. It also builds on our previous study, where we introduced a new method for handling/measuring light sensitive proteins in small volumes.[2]

Briefly, B_{12}-dependent CarH was discovered as a light activated transcriptional regulator, controlling the synthesis of carotenoids to mitigate against photooxidative stress in numerous bacteria.[3,4] Upon the binding of B_{12} (5'-deoxyadenosylcobalamin, AdoCbl), CarH homotetramers form in the dark, which bind to operator DNA and block transcription.[3,5] Absorption of light of wavelengths < 600 nm then causes photodissociation of the upper axial 5'-deoxyadenosyl (Ado) ligand, resulting in conformational changes that drive tetramer disassembly into CarH monomers, leading to DNA release. The oligomerisation state and functionality of this system can be controlled with a high resolution in both space and time using light. Therefore, it has been used in various optogenetic applications and to engineer photo-responsive biomaterials.[6-10]

Interestingly, AdoCbl photochemistry is altered when bound to CarH. Rather than forming cob(II)alamin/Ado radical pairs upon illumination and photocleavage of the Ado group, as occurs for free AdoCbl in solution, CarH appears to stabilise a series of charge transfer (CT) states and steer the reaction away from radical photoproducts.[2,11] This can be rationalised, since it would be counterintuitive for a transcriptional regulator, which helps to protect against photooxidative stress and lies within proximity to DNA, to produce potentially damaging radical species. Instead, the stable product 4',5'-anhydroadenosine is produced in CarH.[12] Previous transient absorption (TA) and TRIR studies show no evidence for the formation of Co(II) species along the productive channel.[2,11] Functionality was rather shown to occur through Co(III) intermediates and a metal-to-ligand CT state similar to that formed when methylcobalamin (MeCbl) is photoexcited.[13]

In our recent TRIR study, we proposed a mechanism describing the initial photochemistry underpinning the activity of B_{12}-dependent CarH.[2] Questions remained, however, as to how exactly the CarH protein tunes the photochemistry of AdoCbl and which protein residues play a key role. A computational study from Goyal and coworkers revealed a series of residues that they believed would encourage the formation of the observed CT states, including the glutamate in the 141 position (E141).[14] It is also known from structural studies that a histidine residue (H132) plays a key role in the conformational changes that lead to monomerisation over longer timescales.[15] Therefore, the CarH mutants E141Q (Q=glutamine to replace the glutamate) and H132A (A=alanine to replace the histidine) were synthesised and purified with the aim of using TRIR to investigate whether the same intermediates are formed and detected spectroscopically in these CarH variants.

Due to difficulties with purifying some of the full-length (FL) CarH variants, the DNA binding domain (DBD) was removed and truncated (T) versions comprising just the cobalamin-binding domain (CBD) were synthesised (*via* recombinant protein expression in *E. coli* and subsequent protein purification). It has been shown from X-ray crystallography studies that the structure of the wild-type (WT) CBD, which is light-sensitive and responsible for tetramer formation, is unaffected by removal of the DBD when comparing against the FL structure bound to DNA.[16] Thus, understanding the light-dependent changes using the truncated protein serves as a good

model for FL CarH. The T-CarH variants showed greater stability, are soluble at higher concentrations, and are less prone to sample precipitation than FL-CarH.

However, despite sample concentrations being increased more than two-fold on average for the T-CarH variants compared to our earlier work, and the quantum yields of the FL- *vs.* T- WT CarH and variants remaining similar, our most recent TRIR data for the T- WT, E141Q and H132A CarH variants suffered from a delay-dependent step-like background pattern attributed to the detector. In addition, the FL protein bound to DNA was also studied. It was hoped that complimentary TRMPS could be used to detect protein and DNA structural changes *i.e.* the protein releasing the DNA at longer times (ms) after photoexcitation. It was also expected that the photochemistry of B_{12}-dependent CarH when bound to DNA would be the same as DNA-free samples, and this work hoped to test this hypothesis. The lower concentrations required for the FL protein to avoid precipitation and scattering resulted in this data set being most affected by background pattern.

Experimental Methods

The FL-WT, T-WT, T-E141Q and T-H132A CarH variants were expressed and purified using protocols as described previously,[2] but with the AdoCbl cofactor added prior to size exclusion chromatography to aid further with stability. Before TRIR measurements, protein samples were exchanged into deuterated buffer (50 mM sodium phosphate buffer, 150 mM NaCl, adjusted to pH 7.91 (pD 7.5) with NaOD) since H_2O absorbs strongly in the IR region of interest (1700 – 1300 cm^{-1}). This buffer exchange was done using pre-equilibrated Bio-Rad (BioGel® P-60) size-exclusion columns. The CarH samples were then brought up to higher working concentrations (~500 µM for the truncated samples and ~100 µM for the full-length) using centrifugal filter devices (Vivaspin 10 kDa MWCO). For DNA binding, FL WT-CarH tetramers were incubated with their complimentary 26 base pair DNA ligand at room temperature for 1 hour.[16] All sample preparation following addition of AdoCbl was carried out under red light (strictly > 600 nm) to avoid irreversible photoconversion to the light-adapted state prior to measurement.

As we knew from our earlier study, CarH is a difficult system for study using TRIR.[2] To obtain reasonable signal/noise with the short 100 µM path length needed for such measurements, CarH concentrations of >100 µM are required. The issue is that upon photoexcitation of CarH tetramers, the light-adapted monomers produced can precipitate at the concentrations used.[2] This would result in scattering and the signal being obscured. It was therefore necessary to use a one-way flow system from a syringe pump, where the 'waste' sample could be collected, centrifuged to remove any precipitate and the remaining sample made available for subsequent measurement. The flow rate and experimental set-up was optimised as described previously.[2]

The TRIR data were collected using the LIFETIME instrument at the Central Laser Facility, STFC, Rutherford Appleton Laboratory, UK. The following experimental parameters were typically used: an excitation pump laser at 525 nm with pulse energy 0.44 µJ and a 1 kHz repetition rate, 3 scans per measurement, and an averaging time of 7.5 s. Due to limitations in the amount of CarH available, the experiment setup was first tested using free AdoCbl (4 mM). An estimate of the absorbance in the green at the concentrations used for FL- and T- CarH samples, and free AdoCbl are 0.01, 0.04 and 0.3, respectively, for a 100 µm pathlength. The absorption is therefore linear and avoids inner-filter effects. Estimations were calculated using the published molar extinction coefficient of 8.0 x 10^3 M^{-1} cm^{-1} for AdoCbl at 522 nm.[17]

Results and Discussion

Whilst it was possible to average across data from at least two repeated acquisitions for each CarH sample, the data were to some degree obscured by the fixed-pattern and delay-dependent background. Figure 1 shows the averaged raw data for T-WT CarH, whilst the data for the other samples can be found in the Appendix (Figures S1-S3). Block-like features always appear at the same wavenumbers (detector pixels) and vary in magnitude with delay time.

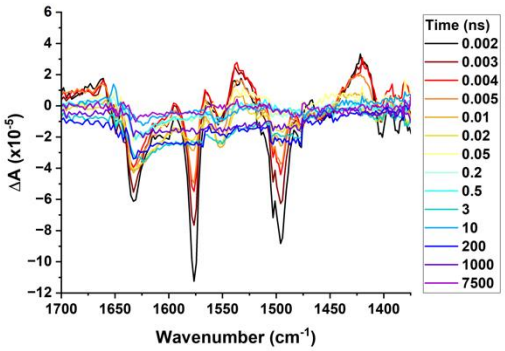


Figure 1*: Raw TRIR difference spectra for T-WT CarH relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over five repeats. A block-like background pattern is evident, which originates from the detector and varies with delay time.*

The delay-dependent background pattern is a result of the structure of the detectors and the experimental set up and has a similar amplitude to the weak signals from our samples, which approach the detection limit. The LIFETIME instrument uses two mercury cadmium telluride (MCT) detectors, which contain 128 pixels each, and the detector chips are assembled in blocks of 32 pixels. The blocks are joined together in an alternating pattern, with neighboring blocks flipped upside down relative to one another. Since the IR sensitivity in the vertical direction is not quite uniform, the alternating pattern of orientations causes slight differences in signal intensity between blocks.

There are several factors contributing to the time delay dependence of the background pattern. These include: electromagnetic interference (EMI) pickup from the pump laser electronics, which may appear at longer delays (µs to ms) due to the gated reading of the detectors; thermal lensing in the sample induced by the pump pulse causing lateral displacement of the probe beam across the detector chip; and the refractive index change in the sample induced by the pump beam leading again to the lateral displacement of the probe beam across the detector. These effects are time-dependent and delay-dependent in different ways, which makes correction for the fixed-pattern background a challenging task.

When signals are weak, the effects described above become noticeable as evident in our results for CarH, particularly at later delay times. Whereas for the B_{12} test, the background pattern is largely insignificant (Figure 2). The data in Figure 2 were acquired using B_{12} concentrations ~8x higher than the protein bound samples, which gave stronger signals as seen previously for free AdoCbl.[18] For the truncated CarH variants and FL-WT CarH bound to DNA, interpretation of the raw differential spectra and any sort of kinetic analysis is obscured by the block-like features. Therefore, to obtain meaningful information about the intermediate states formed, such as their spectral profiles and lifetimes, it was necessary to correct the data for this background pattern.

Prior to background correction, pre-excitation spectra were subtracted. To remove the step-like features, which always occur at the same pixel numbers in line with the grouping in the detector, the absorbance change for neighboring pixels from adjacent blocks (*i.e.*, the last pixel in block A *vs.* first in block B) were made equivalent, with the rest of the pixels in one

block being offset by the initial difference between neighbors. This was carried out systematically for each block and for every delay time for a given sample, using an automated Excel spreadsheet (available from the authors). This approach was taken rather than averaging across blocks, because offsetting adjacent blocks by the average difference in absorption change between them would have resulted in a discontinuity at the boundary. This is because the sample signal varies with wavenumber and the discontinuity would reflect the average difference in signal between and across the two blocks.

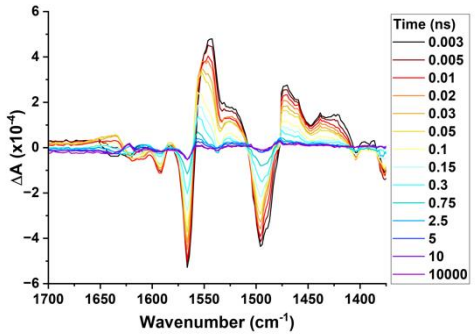


***Figure 2**: Raw TRIR difference spectra for free AdoCbl relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm. There is no evidence of the block-like background pattern at these signal amplitudes.*

For this method to work, there needed to be a zero-reference point at which there is zero signal throughout the time course of the measurement. After the relative offsets have been made between blocks of pixels for a given time delay, the whole differential spectrum can then be shifted to cause the absorbance change at the reference point to equal zero. This point can be chosen by visibly inspecting the 3D map of wavenumber, delay time and absorbance change. It can also help to refer to literature data on similar systems. In this case, points around 1478 cm^{-1} were averaged and used as the reference point. An example 3D map for T-WT CarH, illustrating the justification for this, can be found in the Appendix (Figure S4). Ideally, a region where no signal is expected would be measured and the reference would be taken from there. In our case, this was not possible and so a suitable reference point was chosen near the middle of the spectrum.

The results of the background correction for T-WT CarH are shown in Figure 3, whilst the corrected data for the other CarH variants and the FL protein bound to DNA can be found in the Appendix (Figures S5-S7). Features in the T-WT corrected spectra are in good agreement with those observed for FL-WT CarH in our earlier study.[2] This gives confidence that the background correction does not obscure the data and produce significant artifacts. Furthermore, it enables more meaningful kinetic analyses to be performed, which will be the subject of future work.

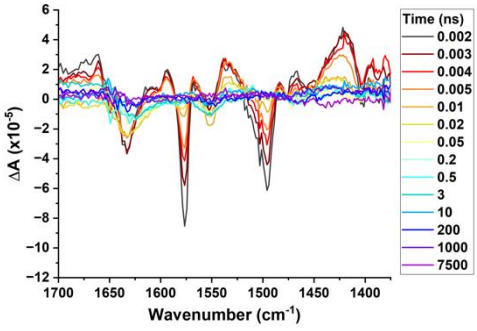


***Figure 3**: Corrected TRIR difference spectra for T-WT CarH relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over five repeats. The block-like background pattern evident in the raw data has been removed.*

Hence, this methodology may be useful to those who would like to study difficult biological samples using the powerful technique of TRIR but are limited by sample quantities and weak signals, where predictable background patterns are noticeable and obscure the signals of interest. To summarise, performing a background correction in such scenarios has the following benefits:

- Kinetic analysis will not be obscured by large block-like features.
- Differential spectra become easier to interpret as they will not be dominated by background.
- Samples that produce weak signal (due to concentration, yield, *etc.*) can be studied using TRIR.

However, the following must also be considered:

- Overprocessing can lead to issues. One must take care when interpreting features at the boundaries between blocks of pixels that have been offset with respect to one another. Comparisons between raw and corrected data when the signal is greatest in magnitude (often at early times) is essential.
- If a sample always produces a signal across the spectral range, the zero-reference point is difficult to define.
- The process presented here was automated to a degree but was still time consuming.

Conclusions

TRIR spectroscopy can reveal a wealth of information, but not all samples are easily amenable to investigation using this method, particularly if they are low concentration and result in a weak signal that approaches detection limits. This can be the case for biological samples, which are often limited in quantity due to challenging preparation and/or sample solubility. Background effects can become noticeable and can obscure the data when measuring these samples. For example, pixel grouping in the MCT detector as part of the LIFETIME instrument at the UK's CLF can result in predictable block-like intensity variations. These intensity variations were evident in our data for truncated B_{12}-dependent CarH variants and the full-length protein when bound to DNA. However, we have illustrated that such background patterns can be corrected without introducing significant artifacts. This will aid towards easier interpretation and more meaningful data analysis in future work.

Acknowledgements

This work was supported by: STFC-funded access to the Ultra Facility; the National Measurement System of the UK Government Department for Science, Innovation, and Technology; a BBSRC CASE studentship to EW (ref. 2777745).

References

1. N. T. Hunt, Biomolecular infrared spectroscopy: making time for dynamics, Chem. Sci., 2024, 15, 414-430. DOI: 10.1039/D3SC05223K
2. I. S. Camacho, E. Wall, I. V. Sazanovich, E. Gozzard, M. Towrie, N. T. Hunt, S. Hay and A. R. Jones, Tuning of B12 photochemistry in the CarH photoreceptor to avoid radical photoproducts, Chem. Commun., 2023, 87, 13014-13017. DOI: 10.1039/D3CC03900E

3. J. M. Ortiz-Guerrero, M. C. Polanco, F. J. Murillo, S. Padmanabhan and M. Elías-Arnanz, Light-dependent gene regulation by a coenzyme B12-based photoreceptor, Proc. Natl Acad. Sci, 2011, 108, 7565–7570. DOI: 10.1073/pnas.1018972108

4. J. Fernández-Zapata, R. Pérez-Castaño, J. Aranda, F. Colizzi, M. C. Polanco, M. Orozco, S Padmanabhan, M. Elías-Arnanz, Plasticity in oligomerization, operator architecture, and DNA binding in the mode of action of a bacterial B12-based photoreceptor, J. Biol.Chem, 2018, 293(46), 17888–17905. DOI: 10.1074/jbc.RA118.004838

5. I. S. Camacho, R. Black, D. J. Heyes, L. O. Johannissen, L. A. I. Ramakers, B. Bellina, P. E. Barran, S. Hay and Alex R. Jones, Interplay between chromophore binding and domain assembly by the B12-dependent photoreceptor protein, CarH., Chem. Sci., 2021, 12, 8333–8341. DOI: 10.1039/D1SC00522G

6. C. Chatelle, R. Ochoa-Fernandez, R. Engesser, N. Schneider, H. M. Beyer, A. R. Jones, J. Timmer, M. D. Zurbriggen, W. Weber, A Green-Light-Responsive System for the Control of Transgene Expression in Mammalian and Plant Cells, ACS Synthetic Biology, 2018, 7(5), 1349-1358. DOI: 10.1021/acssynbio.7b00450

7. M. Mansouri, M. Hussherr, T. Strittmatter, P. Buchmann, S. Xue, G. Camenisch and M. Fussenegger, Smart-watch-programmed green-light-operated percutaneous control of therapeutic transgenes, Nat. Commun, 2021, 12, 3388, DOI: 10.1038/s41467-021-23572-4

8. B. Jiang, X. Liu, C. Yang, Z. Yang, J. Luo, S. Kou, K. Liu and F. Sun, Injectable, photoresponsive hydrogels for delivering neuroprotective proteins enabled by metal-directed protein assembly, SCIENCE ADVANCES, 2020, 6(41), DOI: 10.1126/sciadv.abc4824

9. R. Wang, Z. Yang, J. Luo, I-M. Hsing and F. Sun, B12-dependent photoresponsive protein hydrogels for controlled stem cell/protein release, PNAS, 2017, 114(23), 5912-5917. DOI: 10.1073/pnas.1621350114

10. O. Narayan, X. Mu, O. Hasturk and D. L. Kaplan, Dynamically tunable light responsive silk-elastin-like proteins, Acta Biomaterialia, 2021, 121, 214-223. DOI: 10.1016/j.actbio.2020.12.018

11. R. J. Kutta, S. J. O. Hardman, L. O. Johannissen, B. Bellina, H. L. Messiha, J. M. Ortiz-Guerrero, M. Elías-Arnanz, S. Padmanabhan, P. Barran, N. S. Scrutton and A. R. Jones, The photochemical mechanism of a B12-dependent photoreceptor protein, Nat. Commun., 2015, 6, 7907. DOI: 10.1038/ncomms8907

12. M. Jost, J. H. Simpson and C. L. Drennan, The Transcription Factor CarH Safeguards Use of Adenosylcobalamin as a Light Sensor by Altering the Photolysis Products, Biochemistry, 2015, 54,3231–3234. DOI: 10.1021/acs.biochem.5b00416

13. L. A. Walker, J. T. Jarrett, N. A. Anderson, S. H. Pullen, R. G. Matthews and R. J. Sension, Time-Resolved Spectroscopic Studies of B12 Coenzymes: The Identification of a Metastable Cob(III)alamin Photoproduct in the Photolysis of Methylcobalamin, J. Am. Chem. Soc., 1998, 120, 3597–3603. DOI: 10.1021/ja974024q

14. C. L. Cooper, N. Panitz, T. A. Edwards and P. Goyal, Role of the CarH photoreceptor protein environment in the modulation of cobalamin photochemistry, Biophys. J., 2021, 120, 3688–3696. DOI: 10.1016/j.bpj.2021.07.020

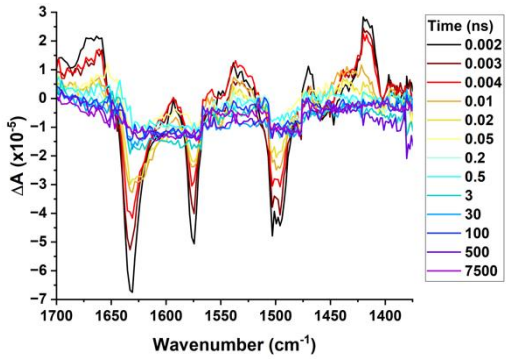
15. H. Poddar, R. Rios-Santacruz, D. J. Heyes, M. Shanmugam, A. Brookfield, L. O. Johannissen, C. W. Levy, L. N. Jeffreys, S. Zhang, M. Sakuma, J-P. Colletier, S. Hay, G. Schirò, M. Weik, N. S. Scrutton and D. Leys, Redox driven B12-ligand switch drives CarH photoresponse, Nat. Commun., 2023, 14, 5082. DOI: 10.1038/s41467-023-40817-6

16. M. Jost, J. Fernández-Zapata, M. C. Polanco, J. M. Ortiz-Guerrero, P. Y-T. Chen, G. Kang, S. Padmanabhan, M. Elías-Arnanz and C. L. Drennan, Structural basis for gene regulation by a B12-dependent photoreceptor, Nature, 2015, 526, 536-541. DOI: 10.1038/nature14950

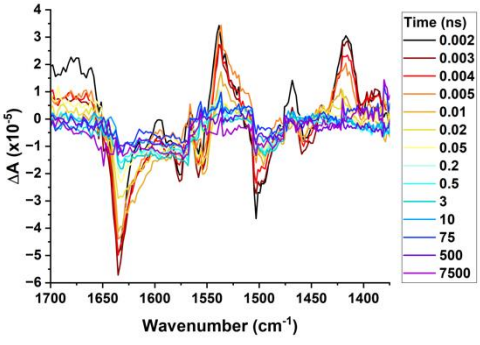
17. H. P. Hogenkamp and S. Holmes, Polarography of cobalamins and cobinamides, Biochemistry, 1970, 9(9), 1886–1892. DOI: 10.1021/bi00811a004

18. A. R. Jones, H. J. Russell, G. M. Greetham, M. Towrie, S. Hay and N. S. Scrutton, Ultrafast Infrared Spectral Fingerprints of Vitamin B12 and Related Cobalamins, J. Phys. Chem. A, 2012, 116, 5586−5594. DOI: 10.1021/jp304594d

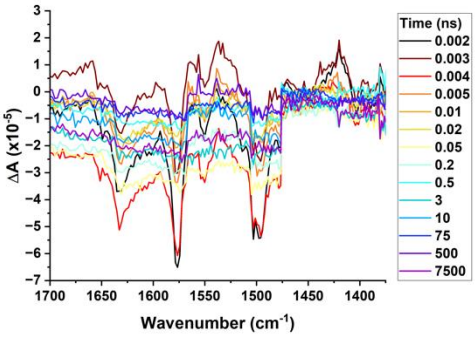
Appendix



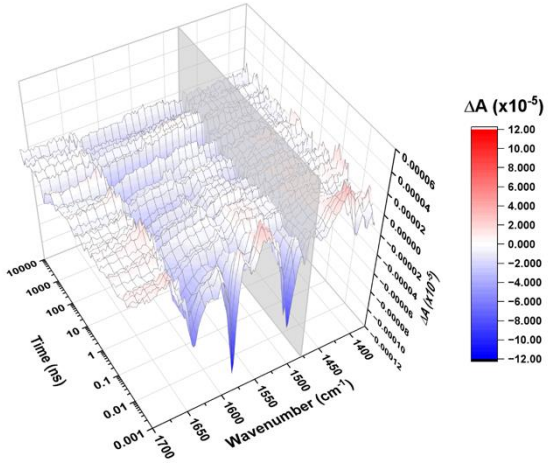
***S1**: Raw TRIR difference spectra for T-E141Q CarH relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over four repeats. A block-like background pattern is evident, which originates from the detector and varies with delay time.*



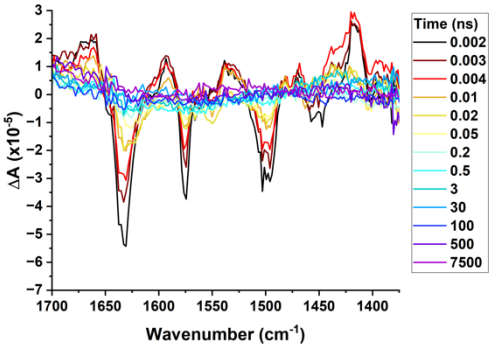
***S2**: Raw TRIR difference spectra for T-H132A CarH relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over six repeats. A block-like background pattern is evident, which originates from the detector and varies with delay time.*



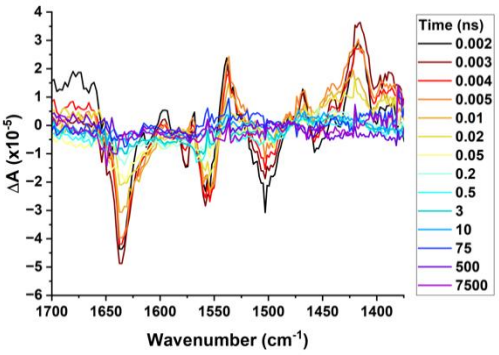
***S3**: Raw TRIR difference spectra for FL-WT CarH bound to DNA relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over two repeats. A block-like background pattern is evident, which originates from the detector and varies with delay time.*



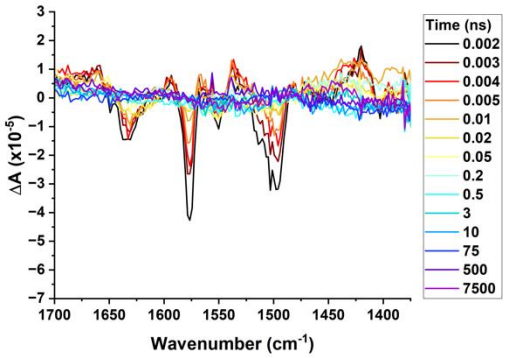
***S4**: Raw TRIR data matrix between 2 ps–7.5 µs for T-WT CarH following photoexcitation at 525 nm and averaged over five repeats. The plane at 1478 cm^{-1} illustrates the zero-reference points by which the differential spectra at each time delay were shifted to cause these points to equal zero after background correction was performed to remove block-like features.*



***S5**: Corrected TRIR difference spectra for T-E141Q CarH relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over four repeats. The block-like background pattern evident in the raw data has been removed.*



***S6**: Corrected TRIR difference spectra for T-H132A CarH relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over six repeats. The block-like background pattern evident in the raw data has been removed.*



***S7**: Corrected TRIR difference spectra for FL-WT CarH bound to DNA relative to the ground state between 2 ps–7.5 µs following photoexcitation at 525 nm and averaged over two repeats. The block-like background pattern evident in the raw data has been removed.*