

ARTICLE

Received 2 Mar 2015 | Accepted 23 Jun 2015 | Published 12 Aug 2015

DOI: 10.1038/ncomms8907

OPEN

The photochemical mechanism of a B₁₂-dependent photoreceptor protein

Roger J. Kutta^{1,2,3}, Samantha J.O. Hardman^{3,4}, Linus O. Johannissen^{3,4}, Bruno Bellina^{1,3}, Hanan L. Messiha^{3,4}, Juan Manuel Ortiz-Guerrero⁵, Montserrat Elías-Arnanz⁵, S. Padmanabhan⁶, Perdita Barran^{1,3}, Nigel S. Scrutton^{3,4} & Alex R. Jones^{1,2,3}

The coenzyme B₁₂-dependent photoreceptor protein, CarH, is a bacterial transcriptional regulator that controls the biosynthesis of carotenoids in response to light. On binding of coenzyme B₁₂ the monomeric apoprotein forms tetramers in the dark, which bind operator DNA thus blocking transcription. Under illumination the CarH tetramer dissociates, weakening its affinity for DNA and allowing transcription. The mechanism by which this occurs is unknown. Here we describe the photochemistry in CarH that ultimately triggers tetramer dissociation; it proceeds via a cob(III)alamin intermediate, which then forms a stable adduct with the protein. This pathway is without precedent and our data suggest it is independent of the radical chemistry common to both coenzyme B₁₂ enzymology and its known photochemistry. It provides a mechanistic foundation for the emerging field of B₁₂ photobiology and will serve to inform the development of a new class of optogenetic tool for the control of gene expression.

¹School of Chemistry, The University of Manchester, Oxford Road, Manchester M13 9PL, UK. ²Photon Science Institute, The University of Manchester, Alan Turing Building, Oxford Road, Manchester M13 9PL, UK. ³SYNBIOCHEM, Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester M1 7DN, UK. ⁴Faculty of Life Sciences, The University of Manchester, Carys Bannister Building, Dover Street, Manchester M13 9PL, UK. ⁵Department of Genetics and Microbiology, Area of Genetics (Unidad Asociada al Instituto de Química Física-Consejo Superior de Investigaciones Científicas), Faculty of Biology, Universidad de Murcia, 30100 Murcia, Spain. ⁶Instituto de Química Física 'Rocasolano', Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain. Correspondence and requests for materials should be addressed to R.J.K. (email: roger.kutta@manchester.ac.uk) or to A.R.J. (email: alex.jones-2@manchester.ac.uk).

Coenzyme B₁₂ (5'-deoxyadenosylcobalamin, AdoCbl) is a large, structurally complex organometallic molecule and a biologically active derivative of vitamin B₁₂ (ref. 1). The mechanistic biochemistry of various AdoCbl-dependent enzymes is well described², and recent crystal structures have provided insight into the binding of AdoCbl to riboswitches, which controls the expression of genes involved in its biosynthesis and transport^{3,4}. However, little is known about the photochemical mechanism that facilitates the role of AdoCbl as a photoreceptor chromophore⁵. Although the photosensitivity of AdoCbl has been known for many years⁶, until recently this photochemistry was considered to serve no physiological function. Free in aqueous solution, photoexcitation of AdoCbl leads to Co–C bond homolysis generating the cob(II)alamin/5'-deoxyadenosyl radical pair⁷. This is the same radical pair that triggers catalysis on substrate binding to most AdoCbl-dependent enzymes. The photolysis of AdoCbl has therefore been used to investigate the influence of cofactor structure and environment on the enzymatic reaction dynamics⁸.

Many bacteria respond to light-induced oxidative stress by synthesizing carotenoids⁹. The prospect of B₁₂-dependent photobiology arose when light-dependent activation of the genes that lead to carotenoid biosynthesis in the bacterium *Myxococcus xanthus* was shown to be B₁₂-dependent¹⁰. The transcriptional repressor, CarH, requires B₁₂ for activity *in vivo*¹¹, and specifically AdoCbl for oligomerization, DNA binding and light-triggered tetramer dissociation *in vitro*^{5,12}. Over 200 CarH-like proteins have been identified in various bacterial genomes (InterPro database, EBI)⁵. Furthermore, B₁₂ has been proposed to act as chromophore in the protein AerR, which controls the light-induced synthesis of bacteriochlorophyll in *Rhodobacter capsulatus* and other photosynthetic bacteria¹³. It appears, therefore, that B₁₂ defines a new class of photoreceptor protein.

Here we present the detailed photochemical mechanism over femtoseconds to seconds for CarH from *Thermus thermophilus*. These transient absorption data reveal a chemical pathway quite distinct from anything observed previously for a B₁₂ molecule and one that appears to be independent of the radical intermediates so closely associated with AdoCbl biochemistry and photochemistry. This work therefore introduces a new branch of B₁₂ biochemistry and provides a mechanistic foundation for the emerging field of AdoCbl-dependent photobiology.

Results

Photoexcitation of CarH results in a highly stable adduct. The ultraviolet–visible absorption spectra acquired during stepwise illumination of the CarH tetramer ground state (CarH-GS) show a clean transition directly to a single cobalamin photoproduct (or light state, CarH-LS) without other significant contributions (Fig. 1a). This is supported by singular value decomposition (SVD)-based rank analysis of the data (Fig. 1a, inset), the decomposed mole fraction versus time profiles and pure species spectra from which are shown in Supplementary Fig. 1a,b, respectively. The CarH-LS spectrum has pronounced structure at 478 and 509 nm (ref. 5), and closely resembles the previously published spectrum of a cobalamin with imidazole as upper and lower axial ligands¹⁴. The AdoCbl-binding domains of CarH proteins share sequence similarity with methylcobalamin (MeCbl)-dependent methionine synthase and AdoCbl-dependent methylmalonyl CoA mutase, both of which bind the cofactor base-off with an active-site histidine coordinated to the lower axial position of the Co. Indeed, mutation of a conserved histidine (H177A) in CarH that aligns with those that form the lower axial linkage in these B₁₂-dependent enzymes has been shown to stop

AdoCbl binding and tetramer formation⁵. Moreover, the equivalent H193A mutation in CarH from *M. xanthus* abolishes function *in vivo*¹¹. The putative B₁₂-dependent photoreceptor protein AerR from *Rhodobacter capsulatus* is also thought to bind base-off with a lower axial histidine. Although the details of its photochemical mechanism are not yet known, after illumination, AerR has been shown to bind the cofactor via a strong covalent linkage between the Co and a second upper axial histidine, a bond that is sustained after treatment with 0.04 M HCl (ref. 13).

Our attempts to isolate the cobalamin bound to CarH-LS by heating to 90 °C and by treatment with 6 M urea were unsuccessful (confirmed by size-exclusion chromatography (SEC)), consistent with a similar covalent linkage to that in AerR being formed between cobalamin and the protein during the photochemical reaction in CarH. This is supported by further data from SEC that show that cobalamin bound to CarH-LS cannot be displaced with free AdoCbl in fivefold excess, whereas hydroxocobalamin bound to CarH can be, with subsequent formation of the CarH-GS tetramer (Supplementary Fig. 2). Both AdoCbl and OHcbl bind to CarH with reasonably high affinities ($K_D \sim 250$ and $\sim 5 \mu\text{M}$, respectively; Supplementary Fig. 3). The lower affinity of CarH for OHcbl could be owing to the lack of favourable binding in this case between the protein and the 5'-deoxyadenosyl. OHcbl has been previously shown by SEC to bind to CarH, but unlike AdoCbl binding this does not result in tetramer formation¹² and there is no evidence that it forms the CarH-LS under illumination (Supplementary Fig. 4). The latter point is perhaps unsurprising when one considers the fact that photoexcitation of free OHcbl does not result in dissociation of the upper axial OH (ref. 15). The aqueous buffer we used in our illumination experiments with OHcbl-bound CarH was at pH 7.5, where OHcbl is mostly in the form H₂Ocbl. H₂Ocbl has also been found to not undergo appreciable photolysis¹⁶.

We confirmed the formation of a highly stable adduct between protein and cofactor by native mass spectrometry. On continuous illumination of CarH-GS with white light, strong signals grow in that indicate an accumulation of CarH-LS, the mass of which correspond to an adduct between the CarH monomer and AdoCbl minus its upper axial 5'-deoxyadenosyl ligand (Fig. 1b). CarH-LS was then mass selected using the quadrupole of a Synapt G2 mass spectrometer and fragmented using argon as collision gas in the trap cell region (Supplementary Fig. 5). When no collision energy is applied we only observe the mass distribution of the precursor ion. When the collision energy is increased we observe the formation of two main additional ions corresponding to the loss of cobalamin from the precursor ion. These data support the assignment of a protein/cobalamin adduct in CarH-LS. Owing to its spectral similarity with that of cobalamin with imidazole as upper and lower axial ligands¹⁴, we speculate that this adduct is formed between the cobalamin and a second active-site histidine (Fig. 1c).

To explore the possible nature of the adduct, we generated a homology model of the CarH-GS structure (Supplementary Figs 6–8). The AdoCbl-binding site (Supplementary Fig. 8) shows that the chromophore is bound 'base-off', with the 5,6-dimethylbenzimidazole displaced by H177. Two histidine residues (H132 and H142) are also adjacent to the upper face of AdoCbl, either of which might plausibly form a covalent linkage with the Co. Molecular dynamics simulations were run on this structural model in the presence and absence of the upper axial 5'-deoxyadenosyl ligand (Supplementary Figs 9,10). Removal of this ligand produces a significant structural change in the protein, which results in a switch in conformation of the corrin ring from the starting structure (Fig. 2a,b) to a position where H132 is directly above the Co (Fig. 2c,d). A glutamate residue (E129) is

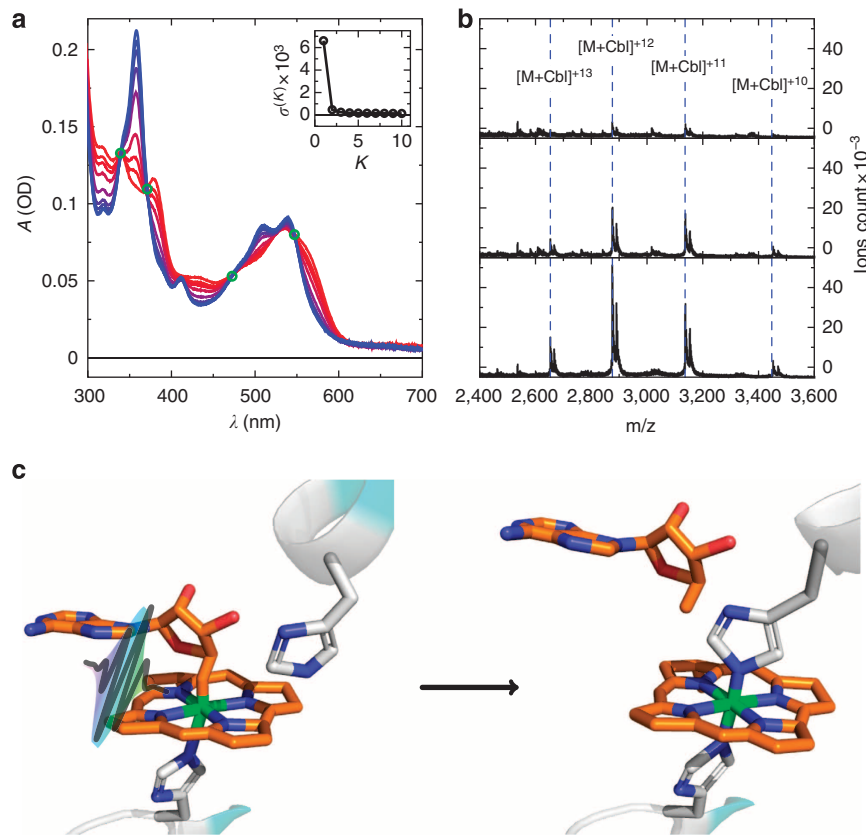


Figure 1 | The photoconversion of CarH. (a) Ultraviolet-visible spectra showing the photoconversion of the CarH tetramer (CarH-GS, red spectra) to the monomeric 'light-state' photoproduct (CarH-LS, blue spectra) under stepwise illumination (pulse duration = 100 ms, λ_{\max} = 530 nm). Isosbestic points (green circles) suggest a transition from a single state to another, supported by SVD-based rank analysis (inset), which shows that only two components are required to represent the data (Supplementary Fig. 1). (b) Nano-electrospray ionization mass spectra of CarH-GS injected in the dark (top), after continuous illumination with white light for ~1 min (middle), and illumination for 2 min (bottom). Signals of the general form $[M + \text{Cbl}]^{n+}$ are annotated, where $10 \leq n \leq 13$, $M = 33,142$ Da and $\text{Cbl} = 1,329$ Da. The total molecular weight is 34,471 Da, which corresponds to the predicted mass of CarH-LS. The secondary peak (higher mass) in each signal is consistent between both CarH-GS and -LS, and is likely a result of an adduct formed during protein purification. (c) Proposed photochemical conversion of base-off AdoCbl in CarH-GS (with incident photon) to a cobalamin adduct with an active-site residue (likely either H132 or H142), which represents the CarH-LS.

well placed to abstract a proton from H132 if necessary for adduct formation (Fig. 2d and Supplementary Fig. 8). The photoconversion is therefore likely to be between base-off AdoCbl, which is bound via a lower axial histidine to CarH-GS, and an adduct similar to that observed in AerR, where the 5'-deoxyadenosyl has been displaced by either H142 or H132 (that is, CarH-LS, Fig. 1c). Although our homology model and molecular dynamics simulations (Supplementary Figs 6–10) favour adduct formation with H132, the precise identity of CarH-LS needs to be established, for example, by crystallography. Adduct formation is followed by dissociation of the CarH tetramer, thus weakening its affinity for DNA and allowing transcription. To identify the various transient intermediates between the CarH-GS and CarH-LS states we have used time-resolved spectroscopic methods measuring ultraviolet-visible absorbance changes between femtoseconds and seconds.

CarH significantly alters the photochemistry of AdoCbl. The ultrafast photochemical dynamics of AdoCbl, in solution and bound to B_{12} -dependent enzymes, are well described^{17–19}. After photoexcitation of AdoCbl in aqueous solution, two excited state intermediates have been fully resolved that decay to a base-off cobalamin intermediate within tens of ps. This dissociates into a base-on cob(II)alamin/5'-deoxyadenosyl radical pair in ~100 ps,

the majority of which (~80%) recombine to the ground state within 10 ns. The photochemistry of AdoCbl bound to B_{12} -dependent enzymes such as glutamate mutase^{20,21} or ethanolamine ammonia lyase²² involves the same radical pair. However, the state that immediately precedes Co–C bond homolysis has been assigned to a cob(III)alamin-like metal-to-ligand charge transfer (MLCT) state, much like that observed during photolysis of MeCbl (ref. 23). Moreover, the quantum yield of longer-lived radicals decreases (to 5–10%) when AdoCbl is confined to an enzyme active site^{20,24}.

The photodynamics of AdoCbl over fs–s are strikingly different when bound to CarH (Fig. 3a–d). False colour representations of the ultraviolet-visible transient absorption difference spectra acquired at time delays up to 3 ns after photoexcitation of CarH-GS are shown in Fig. 3a. Individual examples of raw data difference spectra from the first 15 ps (Fig. 3b), which are not corrected for the pre- t_0 signal, show that signal appears within about 160 fs, comprising two broad, positive absorption bands at wavelengths <490 and >560 nm, and a ground state bleach (that is, a negative signal that represents the loss of a ground-state absorption band) with peaks at 510 and 540 nm. These features decay into a highly structured difference spectrum within 15 ps (Fig. 3b, green spectrum). Apart from the recovery of the bleach at 540 nm over ~400 ps, this spectrum remains constant from 15 ps over the entire acquisition window (Fig. 3a). The linear

dependence of this signal (corrected for the pre- t_0 signal) on laser power confirms that it is the result of single-photon excitation (Supplementary Fig. 11a–c). A significant contribution from multiphoton processes would cause the laser power dependence to deviate from linearity and can therefore be ruled out.

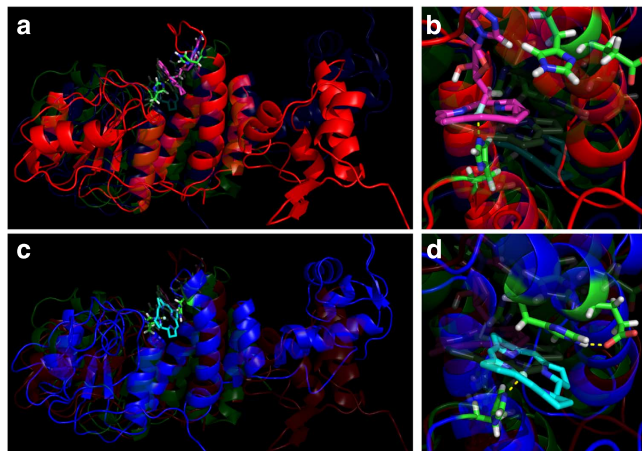


Figure 2 | Molecular dynamics simulations of the CarH photoconversion.

Each structure shown is based on a homology model (Supplementary Figs 6–8); AdoCbl is represented in magenta and active site residues H177 (lower axial coordination), H132 and E129 in green. Molecular dynamics simulations (Supplementary Figs 9,10) were conducted with (a,b, red structure) and without (c,d, blue structure) the 5'-desoxyadenosyl. The snapshots of CarH are representative structures at the minimum of root mean square deviation plots for 10 ns molecular dynamics simulations following 4 ns of relaxation after alignment to the heavy atoms of the cobalamin-binding domain (Supplementary Fig. 9a,c (black lines)).

The effect of CarH on the ultrafast photoresponse of AdoCbl is well illustrated by comparing the decay-associated difference spectra (DADS) from the CarH data to those from an equivalent data set from free AdoCbl (black DADS in Fig. 4a–d,e–h, respectively). DADS were identified by globally fitting the raw data (for CarH, the data in Fig. 3a) to a sum of exponentials. The number of exponentials used was determined by using SVD-based rank analysis²⁵. A good global fit to the data from CarH-GS required one constant component (Fig. 4a) and three dynamic components with lifetimes, τ_{1-3} (Fig. 4b–d). The resulting DADS represent the amplitudes of the exponential components and are used to generate species-associated spectra (SAS) for the various reaction intermediates using a model. The lifetime associated with each DADS is used to calculate the rate constant of each chemical conversion (for SAS and rates see the next section). The DADS with the shortest lifetime from the CarH data set (Fig. 4b $\tau_1 = 1.9$ ps) is quite distinct from that of free AdoCbl (Fig. 4f) and more closely resembles the transient difference spectra observed following photoexcitation of non-alkyl cobalamins such as CNCbl (ref. 16). The homolysis product, cob(II)alamin, is formed (Fig. 4c) but not via a base-off intermediate as observed for free AdoCbl (Fig. 4g). Moreover, these radicals are much shorter lived ($\tau_2 = 794$ ps) than those observed for free AdoCbl photolysis (Fig. 4h), with no evidence of any remaining beyond ~ 1 ns. In fact, the DADS with a lifetime beyond the data acquisition window (Fig. 4d, $\tau_3 \gg 3$ ns) closely resembles the difference spectrum that corresponds to the MLCT state known to precede homolysis during the photolysis of MeCbl (ref. 23) and AdoCbl bound to glutamate mutase²¹. By contrast, there is no evidence within the 3 ns time window of Co–C bond homolysis following the formation of this MLCT state in AdoCbl bound to CarH.

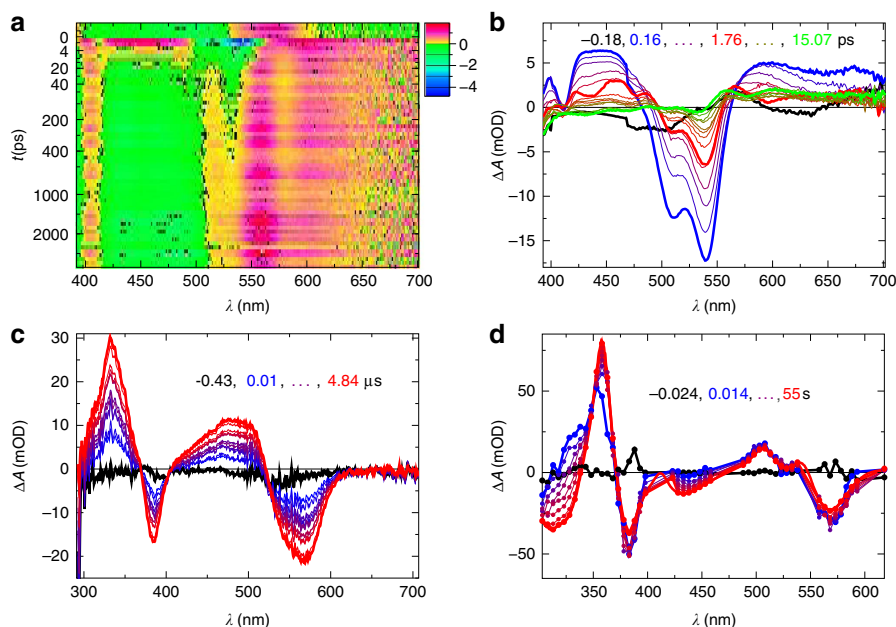


Figure 3 | The transient photoresponse of CarH. (a) False colour plot of the raw data (that is, uncorrected for background data) transient absorption difference spectra that represent the photodynamics of CarH < 3 ns (yellow to red, positive amplitudes; green to blue, negative amplitudes). (b) Examples of raw data transient absorption difference spectra recorded using an ultrafast pump-probe spectrometer up to 15 ps following the photoexcitation of CarH by a fs-pulsed laser. (c) Examples of raw data transient absorption difference spectra recorded using a streak camera up to 5 μ s following the photoexcitation of CarH by a ns-pulsed laser. (d) Examples of raw data transient absorption difference spectra recorded using an ultraviolet-visible spectrometer up to 55 s following the photoexcitation of CarH by a ms-pulsed LED. (b–d) Black lines represent background data, blue lines the early time data and red lines later time data; the green line in (b) represents the contribution of the component that lives from 15 ps to beyond the data acquisition window (that is, $\gg 3$ ns).

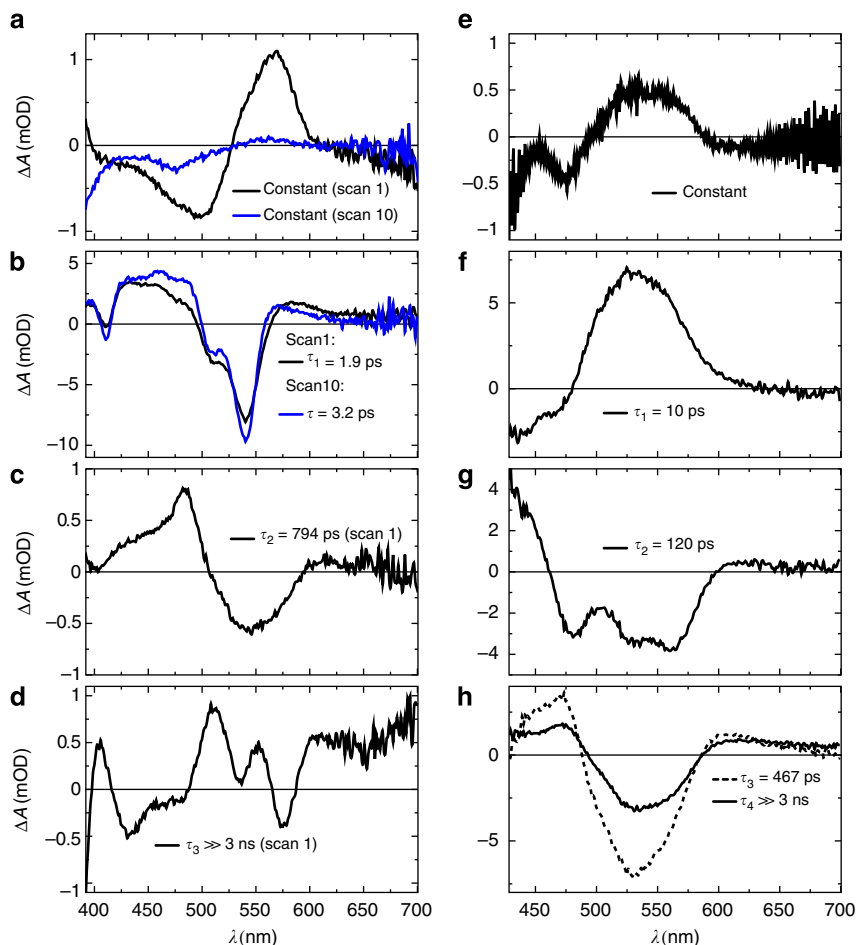


Figure 4 | Decay-associated difference spectra of the ultrafast photoresponse. Comparison of DADS that result from SVD and global analysis of ultrafast transient absorption data acquired between -1.5 ps and ~ 3 ns following photoexcitation of CarH-GS (**a-d**) and free AdoCbl (**e-h**). The constant functions (**a,e**) describe the baseline. The remaining black lines show the dynamic components from fresh sample with lifetimes τ_{1-3} (**b-d**: 1.9 ps, 794 ps and $\gg 3$ ns, respectively), for CarH and τ_{1-4} (**f-h**: 10 ps, 120 ps, 467 ps (dashed) and $\gg 3$ ns, respectively) for AdoCbl. Blue lines are DADS from sample that has mostly converted to the monomeric light state (CarH-LS); they show one constant (**a**) and one dynamic component with lifetime $\tau = 3.2$ ps (**b**).

If a series of repeated, -1.5 ps to 3 ns data acquisition scans are recorded on the same sample, the observed signal changes significantly as the sequence progresses (Supplementary Fig. 12a–c). Example spectral changes from the first and the tenth scans of a sequence are shown in Supplementary Fig. 12d–g, respectively. In the latter case, the photodynamics are complete within 15 ps with no evidence of the longer-lived ($\tau_3 \gg 3$ ns), structured spectrum typical of the MLCT state observed in the earlier scans. This is consistent with the irreversible formation of photoproduct, with the photoresponse of accumulated CarH-LS dominating later scans. The DADS from a later scan (blue DADS, Fig. 4b, and black DADS Supplementary Fig. 13) also bears a close resemblance to the excited state difference spectrum of CNCbl (ref. 16), and similarly decays completely to the ground state within ps.

There is significant signal in the pre-excitation baseline from CarH-GS (that is, in the data acquired before the pump pulse from the laser, Fig. 3a). It is not visible in data from CarH-LS (Supplementary Fig. 12c) and corresponds to the constant component from the analysis of the early scans (black DADS, Fig. 4a). The repetition rate of the laser is 1 kHz; this signal therefore represents an intermediate with a lifetime longer than 1 ms. Our simulations of a hypothetical pump–probe transient absorption experiment (Supplementary Fig. 14) show that species with lifetimes longer than the time difference between two

sequential probe pulses are expected to result in an inverse transient absorption signal before the subsequent laser excitation (Supplementary Fig. 15). On a similar (ms) timescale, a population of radical pairs produced by the photolysis of free and enzyme-bound AdoCbl are known to undergo spin-selective and magnetically sensitive recombination^{17,20,24,26–28}. To further investigate the origin of this pre-excitation signal, transient absorption difference spectra were acquired ns– μ s after the photoexcitation of CarH-GS. They show positive and negative features that grow throughout the acquisition time, with peak shifts at wavelengths < 400 nm (Fig. 3c). SVD and global analysis give two similar but statistically distinct dynamic components (Supplementary Fig. 16), suggesting the same chemical species in a changing environment. The first component has a lifetime of $\tau_1 = 0.6$ μ s and, as predicted by our simulations, the second is an inversion of the constant component from the ultrafast data (black trace, Fig. 4a). This component therefore corresponds to the pre-excitation baseline signal from the ultrafast data (Figs 3a,4a), and has a lifetime of > 1 ms. Unlike intermediates from free AdoCbl photolysis on a similar timescale, however, it does not resemble difference spectrum that represents the conversion from cob(III)alamin to the cob(II)alamin radical (that is, Co–C bond homolysis). Spectral changes continue on ms–s timescale (Fig. 3d), with two further DADS, which are similar to the CarH-LS ground state between 500 and 600 nm

(Supplementary Fig. 17a). However, the absorption at 333 nm does increase initially, followed by a bi-exponential decay, and at 358 nm, increases and continues to grow bi-exponentially into the final, CarH-LS spectrum at 55 s (Supplementary Fig. 17b), after which no further changes are observed. Much like the ns- μ s intermediates, the spectral changes over ms-s are consistent with a single chemical species in a changing environment.

CarH does not employ a radical mechanism. The model in Fig. 5a was applied to the DADS resulting from the global analysis over fs-s. From this modelling we generated a SAS for each intermediate observed (Fig. 3b-e) from which we propose the photochemical mechanism in Fig. 6. Photoexcitation of CarH-GS gives an initial excited state A, the signal from which decays with the rate k_1 . From our proposed model, A either relaxes to the ground state with rate $\alpha k_1 = k_D = 0.471 \times 10^{12} \text{ s}^{-1}$ or decays into one of two intermediates B or C with rates $\beta k_1 = k_{\text{HomoC}} = 0.012 \times 10^{12} \text{ s}^{-1}$ and $\gamma k_1 = k_{\text{CT}} = 0.044 \times 10^{12} \text{ s}^{-1}$, respectively (Fig. 5a). α , β and γ are branching ratios, where $\alpha + \beta + \gamma = 1$. The green SAS of intermediate B in Fig. 5b is distinctive of the Co-C bond homolysis product, the cob(II)alamin radical (Fig. 6, green box)¹⁸. This is a non-productive

channel, with a recombination rate $k_2 = k_{\text{RPR}} = 1.26 \times 10^9 \text{ s}^{-1}$, after which all of the radical pairs recombine to the ground state. The brown SAS of C in Fig. 5b closely resembles the estimated spectrum of the MLCT state reported previously^{21,23}, where the Co(III) effectively receives dative coordination from an anionic ligand (Fig. 6, brown box). This intermediate has a lifetime of $\gg 3 \text{ ns}$ and leads to a productive channel ending in the photoproduct. The branching ratios in our model are $\alpha = 0.894$, $\beta = 0.022$ and $\gamma = 0.084$. They suggest that, following a single-fs laser pulse, the majority of excited states decay to the ground state, with only 8.4% leading to the photoproduct and tetramer dissociation. A relatively low yield of productive channel would be consistent with the need for carotenoid biosynthesis only under high light intensity conditions, when the bacteria that express CarH are likely to experience significant photo-oxidative stress^{5,11}.

The dark green SAS of D* in Fig. 5c represents the next observed intermediate formed from C between 3 and 15 ns with an approximate rate $k_{\text{HeteroC}} \sim 0.1 \times 10^9 \text{ s}^{-1}$. It resembles neither the spectrum of cob(II)alamin nor cob(I)alamin, the Co-C bond dissociation products observed in reactions catalysed by AdoCbl and MeCbl-dependent enzymes, respectively. The spectrum has lost some of the structure evident at $\sim 550 \text{ nm}$

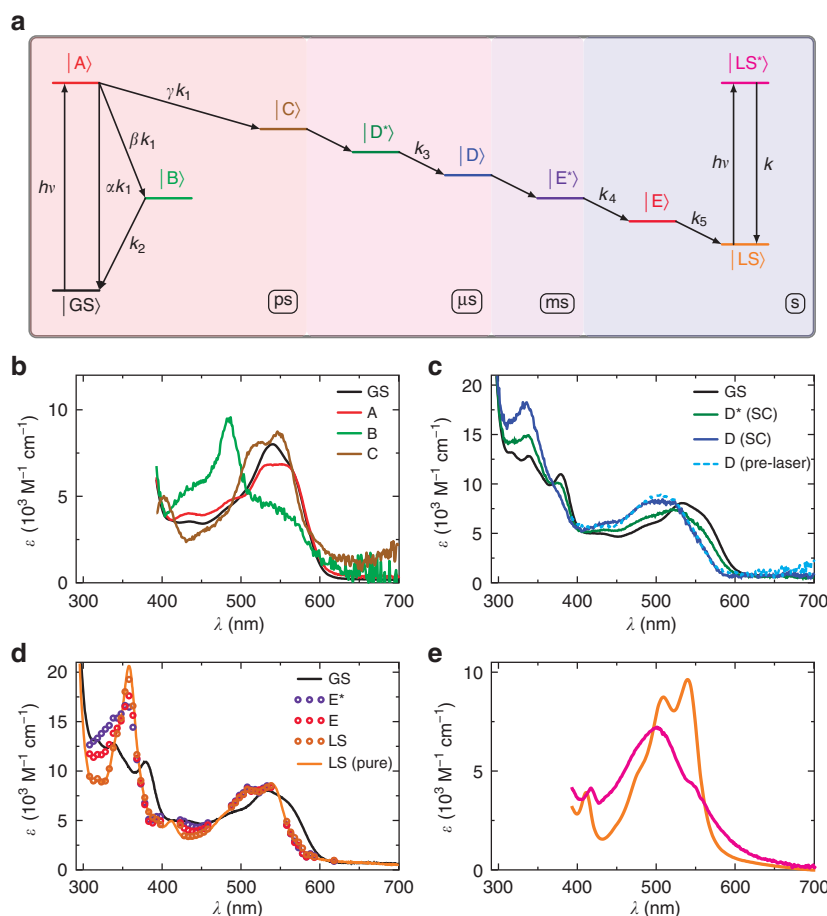


Figure 5 | Model and species-associated spectra representing the CarH reaction intermediates. (a) Schematic representation of the model used to generate the SAS from the ground-state (GS) spectrum and DADS (Supplementary Figs 18–21), which correspond to the excited states, reaction intermediates (a–e) and photoproduct (or ‘light state’, LS). Each intermediate decays with rate constant, k_{1-5} and α , β and γ represent branching ratios. ps, μ s, ms and s time domains are indicated. (b–e) Coloured SAS correspond to the colour code of each intermediate in (a), and the black line in each is the GS spectrum. (b) ps intermediates observed by ultrafast transient absorption. (c) ns- μ s intermediates observed in the streak camera data; the pre-excitation baseline signal from the ultrafast pump-probe experiments (pale blue, dashed line) closely resembles intermediate, D (dark blue line). (d) ms-s intermediates; the final state, LS (brown dots), corresponds to the spectrum known⁵ for the CarH-LS (pure, orange line). (e) SAS of the CarH (photoproduct) excited state (magenta line) built from the ground-state spectrum (orange line) and corresponding DADS.

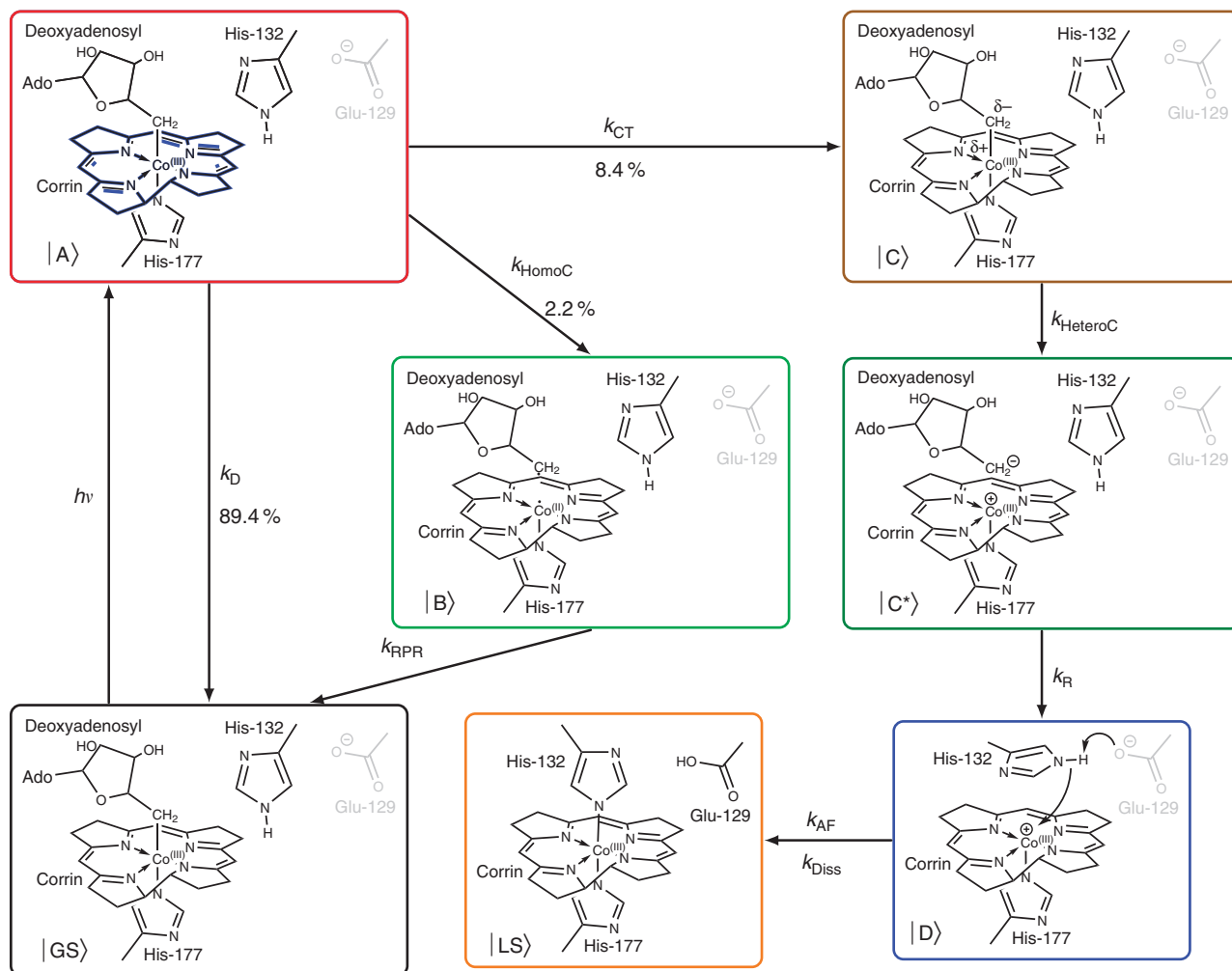


Figure 6 | Proposed photochemical mechanism of CarH. The coloured boxes and letters correspond to the colour-coded ground state (GS), intermediates (A–D) and product (or ‘light state’, LS) in Fig. 5. Intermediates E* and E in Fig. 5 have not been shown here because they represent the adduct between protein and the cobalamin before complete dissociation of the tetramer and therefore resemble LS. AdoCbl is represented with a simplified corrin ring for clarity, with the Co coordinated by a lower axial H177 from the protein. The involvement of residues H132 and E129 are those favoured by our molecular dynamic simulations based on our homology model of the CarH structure (Fig. 2). After photoexcitation ($h\nu$) of the GS, each step proceeds with a rate constant, k_i , where i : CT, charge transfer; HomoC, homolytic cleavage; HeteroC, heterolytic cleavage; R, rearrangement; AF, adduct formation; Diss, tetramer dissociation. See main text for discussion of the mechanism.

in the preceding MLCT state (C in Fig. 5b), but still retains cob(III)alamin character. D* decays into the dark blue SAS of intermediate D (Fig. 5c) with a rate $k_3 = k_R = 1.79 \times 10^6 \text{ s}^{-1}$. This decay involves a blue shift in the broad feature around 500–550 nm and a slight growth of the peak at ~ 350 nm. Such relatively subtle spectral shifts are consistent with a change in the chromophore environment, with a blue shift suggesting it has become more hydrophobic (as observed for the chromophore in active-site variants of AdoCbl-dependent ethanolamine ammonia lyase)²⁹. Consistent with the analysis of the ns– μ s data is Co–C bond heterolysis to give five-coordinate cob(III)alamin and the 5'-deoxyadenosyl anion (Fig. 6, D*, dark green box), the latter of which is then displaced by the residues involved in adduct formation (Fig. 6, D, dark blue box). Generation of the MLCT state of the preceding intermediate (C) has shifted electron density onto the ligand, which might favour generation of the 5'-deoxyadenosyl anion. Displacement of this polar moiety after dissociation would also be expected to result in the spectral blue shift observed on transition from D* to D. The molecular dynamic simulations (Fig. 2 and Supplementary Figs 9,10)

also support this; they reveal a switch in conformation of the corrin ring to a position that displaces the recently dissociated 5'-deoxyadenosyl and favours adduct formation. The 5'-deoxyadenosyl anion can then undergo β -hydride elimination to generate 4',5'-anhydroadenosine, which has been recently identified as the only counterpart photoproduct to the CarH-LS (ref. 30). Although electron transfer from the 5'-deoxyadenosyl anion to cob(III)alamin to generate the radical pair might be considered possible, there are no spectral intermediates consistent with this happening. A potential explanation is that it is precluded by the kinetics of the structural rearrangement that follows 5'-deoxyadenosyl dissociation.

An alternative assignment of spectral intermediates D* and D is hydridocobalamin³¹, a poorly characterized six-coordinate, cob(III)alamin species with an upper axial H. We were prompted to consider this possibility by the identification of 4',5'-anhydroadenosine as the sole photoproduct following aerobic photolysis of CarH (ref. 30). This product has been previously observed after the thermal homolysis of AdoCbl, where β -hydrogen transfer from the 5'-deoxyadenosyl radical to

the cob(II)alamin radical represents a possible reactive fate to give 4',5'-anhydroadenosine and hydridocobalamin^{32,33}. We have no evidence of radicals subsequent to the MLCT intermediate C in the CarH photoresponse. However, it cannot yet be ruled out that Co–C cleavage and H transfer in CarH somehow occurs in a concerted fashion or that a hydride is transferred from the 5'-deoxyadenosyl anion to the 5-coordinate cob(III)alamin. Either way, the 4',5'-anhydroadenosine could then be displaced by the structural changes described above, followed by substitution of the upper axial H to form the CarH-LS adduct. Because of the resemblance of D from the ns– μ s data to the SAS from the pre-excitation baseline signal from the ultrafast data (Fig. 5c, pale blue dashed) it is likely that this intermediate lives for > 1 ms.

The SAS with lifetimes of ms–s (Fig. 5d, E*, purple dots and E, dark red dots) are almost identical to that of the CarH-LS between 500 and 600 nm, suggesting that after intermediate D, cobalamin does indeed form the stable adduct with an active-site residue (Fig. 6, orange box, $k_{AF} \sim 0.1 \times 10^3 \text{ s}^{-1}$). Owing to the close resemblance of the CarH-LS spectrum (Fig. 3a) to that of a cobalamin with imidazole as upper and lower axial ligands¹⁴, we speculate that this is a histidine, and our homology model of the CarH structure (Supplementary Figs 6–8) places both H132 or H142 close by. Molecular dynamics simulations (Supplementary Figs 9,10) favour H132 and show that an active-site glutamate E129 is well positioned in case proton abstraction from H132 is necessary for nucleophilic attack (Figs 2b,d,6, dark blue box). As stated above, the identity of the adduct needs to be confirmed experimentally. Despite the lack of significant changes between 500 and 600 nm, there are spectral changes in the blue and ultraviolet between the SAS representing intermediates E* \rightarrow E \rightarrow LS (Fig. 5d). These are likely to be a result of changes to the cofactor environment caused by structural changes in the protein that are triggered by adduct formation and which ultimately result in dissociation of the CarH tetramer into monomers. On the basis of the slow rates of interconversion between each of these states ($k_4 = k_{Diss} = 1.71 \text{ s}^{-1}$ and $k_5 = k_{Diss'} = 0.06 \text{ s}^{-1}$) from the global fit of the ms–s data, we propose two alternative pathways: (i) structural changes internal to each monomer (k_{Diss}) that trigger complete tetramer dissociation (k_{Diss}); or (ii) dissociation into dimers (k_{Diss}) followed by dissociation into monomers ($k_{Diss'}$). Further study is required to distinguish between these dissociative pathways. Finally, the SAS of the excited state of CarH-LS (Fig. 5e) bears a close resemblance to the estimated excited state spectrum of CNCbl, which has been previously attributed to a ligand to metal charge transfer state¹⁶. For a justification of the model chosen in Fig. 5a from the various different models explored, see Supplementary Figs 18–21 and the Supplementary Methods section.

Discussion

We have described the detailed photochemical mechanism over the fs–s that follow photoexcitation of a new class of photoreceptor protein that is dependent on AdoCbl. In doing so we have revealed that AdoCbl is capable of diverse biochemistry to enable its increasingly diverse known biology. In summary, after photoexcitation of AdoCbl, electron density is moved away from the Co towards the upper axial 5'-deoxyadenosyl ligand forming a MLCT state, which then leads to the dissociation of the Co–C bond. The dissociation products are 4',5'-anhydroadenosine³⁰ and a cob(III)alamin species that could potentially be the five-coordinate product of heterolysis or six-coordinate hydridocobalamin³¹. Nucleophilic attack by an active-site residue, likely either H132 or H142, then forms a stable

adduct with the cob(III)alamin, the irreversible formation of which is followed by dissociation of the CarH tetramer. Therefore, our data suggest that the productive pathway of CarH does not proceed via the cob(II)alamin/5'-deoxyadenosyl radical pair intermediates so closely associated with both AdoCbl enzymology and photochemistry⁸. The fact that formation of the stable cobalamin adduct appears to be irreversible is perhaps surprising considering the high energetic cost of AdoCbl biosynthesis³⁴. However, although CarH proteins only bind AdoCbl to form a tetramer *in vitro*⁵, exogenously supplying other cobalamins such as CNCbl, MeCbl or OHCbl *in vivo* is sufficient for CarH repressor activity¹¹. This implies the existence of intracellular machinery that can replace the upper axial 5'-deoxyadenosyl after photoconversion and thus recycle the cofactor, possibly after proteolysis and release of cobalamin. The adenosyltransferase enzyme capable of activating cobalamins such as CNCbl (ref. 34) is a likely candidate⁵.

We have thus shown that CarH tunes AdoCbl photochemistry to facilitate the response of bacteria to photo-oxidative stress. It also appears that other B₁₂-dependent bacterial proteins (for example, AerR) play an intimate role in the light-dependent control of photosynthetic genes¹³. A compelling new area of molecular and mechanistic photobiology is therefore beginning to emerge based on B₁₂ as photoactive chromophore. Our findings represent the first photochemical mechanism for a B₁₂-dependent photoreceptor protein and should serve to guide the development of this class of photoreceptors as optogenetic tools for the control of gene expression in cells and organisms.

Note added in proof: crystal structures of the CarH-GS tetramer and CarH-LS monomer states are about to be published⁴⁰. These structures reveal the bis-His linkage involving H132 in CarH-LS and the large conformational change associated with its formation from the CarH-GS tetramer.

Methods

Materials. Coenzyme B₁₂ (>98% Sigma) was used as purchased without further purification. The plasmid, pET15b-CarH, encoding CarH from *Thermus thermophilus* was transformed into *Escherichia coli* and the apoprotein monomer overexpressed. The His₆-tagged protein was purified under native conditions using Talon metal affinity resin (GE Healthcare) and imidazole elution^{5,12}. All spectroscopic experiments were conducted in 50 mM phosphate buffer, 150 mM NaCl, pH 7.5.

Preparation of cobalamin-bound CarH samples. All steps were conducted under dim red light. CarH tends to precipitate if its concentration is brought above $\sim 30 \mu\text{M}$ before AdoCbl binding and tetramer formation. A dilute solution of the apoprotein monomer ($< 1 \mu\text{M}$) was therefore incubated with a large excess of AdoCbl or OHCbl ($\sim 250 \mu\text{M}$) for 10–15 min. Unbound cobalamin was then removed from the resulting CarH tetramer dark state using a CentriPure P100 gel filtration column (emp BIOTECH). Each sample of the cobalamin-bound CarH was then concentrated as required using centrifugal filtration devices (10 kDa molecular weight cutoff).

Stationary ultraviolet-visible absorption spectroscopy. Stationary ultraviolet-visible absorption spectra were recorded using an Agilent Cary 50 UV-Vis Spectrophotometer. Samples of the CarH-GS tetramer or OHCbl-bound CarH were illuminated using a pulsed, high-power light-emitting diode (LED; M530L3, Thorlabs) with $\lambda_{\text{max}} = 530 \text{ nm}$. The 100 ms rectangular excitation pulse (2 mJ) was collimated using an anti-reflection-coated aspheric lens (Thorlabs), and delivered along the 2 mm pathlength of the cuvette orthogonal to the detection beam. After each excitation an absorption spectrum was recorded resulting in a sequence of spectra over illumination time. The sample volume was 150 μl with an optical density over 1 cm of ~ 0.1 at the excitation wavelength. The entire sample volume was therefore homogeneously illuminated, avoiding possible effects of diffusion during the recording of the spectra.

Analysis of the photoconversion of CarH-GS to CarH-LS. The data matrix A shown in Fig. 1a represents the overall photoconversion of CarH-GS to CarH-LS, and was decomposed into its time profiles and corresponding spectra using a method employed previously³⁵. First, the number of principle components in A

was determined using SVD-based rank analysis. Second, a model matrix \mathbf{X} was generated based on known features such as isosbestic points, known spectra and assumptions about the characteristics of the unknown spectra (for example, wavelengths where the extinction coefficient is assumed to be zero). Finally, the model matrix \mathbf{X} allows the reconstruction of the species spectra \mathbf{S} and their corresponding concentration time profiles \mathbf{C} via:

$$\mathbf{S} = \mathbf{U}\mathbf{X} \quad (1)$$

$$\mathbf{C}^t = \mathbf{X}^{-1}\mathbf{W}^{(K)}\mathbf{V}^t \quad (2)$$

where the matrices \mathbf{U} , \mathbf{V} and $\mathbf{W}^{(K)}$ are the result of the SVD, with K principle components used to reconstruct the data matrix \mathbf{A} :

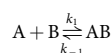
$$\mathbf{A} = \mathbf{U}\mathbf{W}\mathbf{V}^t \cong \Delta\mathbf{A}' = \mathbf{U}\mathbf{W}^{(K)}\mathbf{V}^t \quad (3)$$

SVD analysis of the CarH photoconversion shows that only two principle components are necessary to represent the data (inset in Fig. 1a). The species spectra and time profiles were therefore calculated using these components and the following assumptions: (i) the CarH-GS species spectrum is the first spectrum of the sequence; (ii) the final spectrum of the sequence is the CarH-LS species spectrum contaminated by the CarH-GS spectrum. Accounting for ~2% of the ground state in the final spectrum allowed the sum of both concentration time profiles to be 1 over the entire temporal range. The mole fraction versus time profiles and pure species spectra are shown in Supplementary Fig. 1a,b, respectively.

Size-exclusion chromatography. To assess the stability of the protein/cobalamin adduct in CarH-LS, we attempted to displace the cobalamin with an excess of free AdoCbl. The following samples were separately incubated before eluting each by SEC: hydroxocobalamin (OHCbl)-bound CarH with fivefold excess of AdoCbl (black), CarH-LS with fivefold excess of AdoCbl (red).

All procedures were conducted under safe light. The elution profiles recorded at 280 and 522 nm for each pre-incubation condition are shown in Supplementary Fig. 2a,b, respectively.

Cobalamin-binding titrations. The K_D of AdoCbl and OHCbl binding to CarH were estimated using a spectral method. Samples were prepared containing different concentrations of CarH apoprotein (up to 27 μM) and a constant concentration (5 μM) of each cobalamin. An ultraviolet-visible absorption spectrum was acquired for each concentration of apoprotein used (Supplementary Fig. 3a,c) and the spectral evolution was analysed using a similar method to that described above for the photoconversion of the CarH-GS. Briefly, the data matrix was decomposed into its principle species spectra and corresponding mole fraction profiles (Supplementary Fig. 3b,d). Here the mole fraction profiles show a typical binding curve of the following chemical equilibrium:



where,

$$K_D = \frac{k_{-1}}{k_1} = \frac{[\text{A}][\text{B}]}{[\text{AB}]} \quad (4)$$

The amount of complex AB, x_{AB} , is dependent on [A] and [B], the concentration of the two binding partners, and is given by:

$$x_{\text{AB}}([\text{A}]) = \frac{K_D + [\text{B}] + [\text{A}] - \sqrt{(K_D + [\text{B}] + [\text{A}])^2 - 4[\text{B}][\text{A}]}}{2[\text{B}]} \quad (5)$$

In each case [B] was assigned to that of the cobalamin and therefore held at 5 μM and the K_D for each cobalamin was determined by least squared fitting using the Levenberg-Marquardt algorithm.

Mass spectrometry. Mass spectra were acquired using a Waters Synapt G2 mass spectrometer operating in time-of-flight mode. The tuning parameters were set to produce a mass resolving power of 20,000. The ions were produced using the Z-spray source and nano-electrospray ionization. The spray voltage ranged from 1.3 to 1.8 kV and the source temperature was set to 80 °C. Collision-induced dissociation experiments were performed in the trap cell region of the Synapt using argon as a collision gas.

CarH homology model. The programme MODELLER 9.11 was used to generate a homology model structure of full-length CarH (ref. 36). It is based on the available structures of methionine synthase from *E. coli* (PDB code: 1BMT)³⁷, corrinoid (factor IIIm)-binding protein from *Moorella thermoacetica* (PDB code: 1Y80) and a transcriptional regulator of the MerR family from *Bacillus cereus* (PDB code: 3HH0). The structures 1BMT and 1Y80 include the CarH C-terminal B₁₂-binding site, and 1BMT includes structural information about the linkage between the CarH photoreceptor and DNA-binding domains. Structural information about the CarH N-terminal DNA-binding domain is given by 3HH0. All three structures

together cover the entire sequence of full-length CarH. An initial sequence alignment (Supplementary Fig. 6a,b) was made using the programme ClustalX 1.83 (Gap Opening: 10; Gap Extension: 0.2; Delay Divergent Sequence: 30%; Use Negative Matrix: yes; Protein Weight Matrix: BLO-SUM series)^{38,39}. It shows CarH sequence identity of 16.5% to 1BMT, 22.8% to 1Y80 and 22.9% to 3HH0 (Supplementary Fig. 6a). The overall similarity in terms of the physicochemical properties of the amino acids is high for the corresponding domains (Supplementary Fig. 6b). Therefore, these templates are an acceptable basis for the model.

Molecular dynamics simulations. Molecular dynamics simulations were performed on the CarH homology model structure using the Gromacs Package with the gromos53a6 force field, a solvation box of minimum 1 nm around the protein and periodic boundary conditions. After energy minimization the system was initially thermalized to 300 K for 100 ps using NVT dynamics, and the pressure was then equilibrated for 100 ps using NPT dynamics. The protein, cobalamin and 5'-deoxyadenosyl were constrained during these steps. All constraints and pressure couplings were then switched off and the system relaxed using NPT at 250, 280, 290 and 300 K for 1 ns each. Finally, 10 ns of NPT dynamics were run at 300 K. Representative structures were chosen as those with the smallest root mean squared deviations for the cobalamin and H132 heavy atoms relative to the average, following alignment to the heavy atoms of the cobalamin-binding domain (Supplementary Fig. 9).

Ultrafast pump-probe transient absorption. All measurements were performed at room temperature using a Helios (Ultrafast Systems LLC) broadband pump-probe transient absorption spectrometer with an instrument response function of ~170 fs. The Ti:sapphire amplifier (a hybrid Coherent Legend Elite-F-HE) is pumped by a Q-switched Nd:YLF laser (Positive Light Evolution-30) and seeded by a Ti:sapphire oscillator (Spectra-Physics Mai Tai). The amplifier output (1 kHz, 800 nm, with a typical pulse duration of ~120 fs) is split to provide pump and probe beams. The pump beam is tuned using a non-collinear optical parametric amplifier (Light Conversion TOPAS-White) with a typical pulse duration of <70 fs. The concentration of CarH samples was adjusted to an optical density between 0.2 and 0.8 at the excitation wavelength (370 nm; 0.5 μJ per pulse; 2 mm pathlength). Data were recorded in a stirred quartz cuvette at time delays between -1.5 ps and 3 ns after excitation, the order of which was randomized. At each delay time, 1,000 single shots of pump and probe cycles were averaged.

Simulations of pre- t_0 signals from pump-probe experiments. To investigate how a long-lived signal might manifest in pump-probe transient absorption data acquired before the subsequent laser pulse, we performed a simulation of a hypothetical transient absorption experiment. In the simulation the frequency for collecting a single pair of signal and reference data sets was set to 0.5 kHz, and the simulation was based on four subsequent pump-probe cycles (8 ms). For simplicity, it is assumed that: (i) the sample volume does not change between each cycle; (ii) the hypothetical sample features only a ground |A> and excited state |B>; (iii) the sample consists of wavelengths at which the time evolution of the transmission arises only due to either pure absorbance of a new species (Supplementary Fig. 14a,b) or pure bleach of the ground-state absorbance (Supplementary Fig. 14c,d). The extinction coefficients of the ground and excited state were both set to 1 M⁻¹ (cm)⁻¹, and the optical pathlength was set to 1 cm. Photoexcitation of the sample was simulated by Gaussian-shaped laser pulses. The following differential equations were used to calculate the signals:

$$|\dot{\mathbf{A}}(t)\rangle = -\sum_i^4 I_{\text{RF},i}(t)|\mathbf{A}(t)\rangle + k|\mathbf{B}(t)\rangle \quad (6)$$

$$|\dot{\mathbf{B}}(t)\rangle = I_{\text{RF},i}(t)|\mathbf{A}(t)\rangle - k|\mathbf{B}(t)\rangle \quad (7)$$

where $I_{\text{RF},i}(t) = \frac{f_0}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{t-t_{0,i}}{\sigma})^2)$; f_0 = fraction of the ground state; σ = pulse width; and $t_{0,i}$ = time zero of pulse i .

Three simulations for three different rate constants, $k_1 = 10$ (ms)⁻¹, $k_2 = 3$ (ms)⁻¹ and $k_3 = 1$ (ms)⁻¹, of the excited state decay are shown in Supplementary Fig. 14.

Assuming that the total intensity applied to the sample is 1 at both wavelengths for pure absorption and pure bleach, the transmission values from the simulations are the measured intensities at these wavelengths. The transient absorption signals for a single pump-probe cycle can therefore be calculated using:

$$\Delta A = -\log\left(\frac{T_{\text{Sig}}}{T_{\text{Ref}}}\right) = -\log\left(\frac{I_{\text{Sig}}}{I_{\text{Ref}}}\right) \quad (8)$$

The results are shown in Supplementary Fig. 15. They demonstrate that species with lifetimes longer than the time difference between two probe pulses result in an inverse transient absorption signal before the subsequent laser excitation. They also show that such long-lived signals overlay with the transient absorption signals that evolve after the subsequent excitation pulse. The time window covered by the pump-probe experiment is typically two orders of magnitude smaller than the lifetime of the long-lived species. These contributions therefore manifest as a constant component in each data acquisition window, and these simulations

therefore nicely illustrate the need to use a constant function over all acquisition wavelengths in the global fit analysis of our data.

Nanosecond to microsecond transient absorption. The development and performance of the transient absorption spectrometer using streak camera detection employed here has been described in detail previously²⁵. Briefly, the sample is flowed through a fused silica cuvette and the peristaltic pump was synchronized with the timing of the measuring process so that the sample was exchanged stepwise between each individual laser excitation. The sample is excited through a cylindrical lens ($f = 150$ mm) along a 2-mm optical path by a single outcoupled laser flash (530 nm, 8 ns, 10 mJ) from an optical parametric oscillator pumped by a Nd:YAG laser (Surelite II, Continuum) at a repetition rate of 10 Hz. White light from a 150-W pulsed Xe flash lamp of ~ 1 ms duration (MSP-05, Müller Elektronik-Optik) is passed through the sample along a 10 mm optical path and subsequently analysed by an imaging spectrograph (Bruker 200is, grating 100 grooves per mm) coupled to a streak camera (C7700, Hamamatsu Photonics). The data were registered by a charge-coupled device camera (ORCA-CR, Hamamatsu Photonics) and stored to a computer. Each transient absorption data set was calculated from four images taken with a frequency of 0.5 Hz in the following sequence: D_{FL} —both flash lamp and laser; D_0 —without any incoming light; D_F —only with the flash lamp; D_0 repeated. This sequence is repeated 100 times, averaged and the transient absorption calculated as:

$$A(t, \lambda) = \log\left(\frac{D_F - D_0}{D_{FL} - D_0}\right) \quad (9)$$

Millisecond to second transient absorption. Samples were excited using the collimated, 10 ms rectangular pulse (2 m) from a high-power LED (M530L3, Thorlabs) with $\lambda_{max} = 530$ nm. Sample exposure was conducted as described above for the stationary spectra to avoid effects of diffusion. Absorbance changes were recorded at single wavelengths ranging from 300 to 700 nm using an Agilent Cary 50 UV-Vis Spectrophotometer in 5 nm steps with a time resolution of 12.5 ms and time window of 55 s. The LED flash was manually triggered for each time trace and all data sets corrected to the same $t = 0$ during data processing. The raw data, which comprised 4,800 data points, were reduced to 500 points by averaging on a logarithmic time axis before analysis.

Transient absorption data analysis and modelling. SVD-based rank analysis and global fitting were performed using an in-house written programme described previously²². Briefly, time-resolved data sets from transient absorption measurements are a series of difference spectra recorded at a number of delay times, $\Delta A(t, \lambda)$. For a set of discrete N_T time points t_i and N_L wavelengths λ_j , these data form a rectangular matrix ΔA of dimension $N_T \times N_L$ with matrix elements $\Delta A_{ij} = \Delta A(t_i, \lambda_j)$. Each column of the matrix is a time trace for a fixed wavelength, each row is a spectrum at a given delay time. The data matrix can be decomposed into a sum of products of one-dimensional functions by a global fit:

$$\Delta A_{ij} = \sum_{k=1}^N c_k(t_i) S_k(\lambda_j) = \sum_{k=1}^N c_{ik} S_{kj} \quad (10)$$

The global fit assumes that the data can be modelled by a factorization into a sum of products of spectra, $S_k(\lambda)$, and concentration time profiles, $c_k(t)$, as defined in equation (10), where the time profiles can be written as a linear combination of known analytic functions $f_j(t)$:

$$D_{ij} = \sum_{k=1}^{N_C} c_k(t_i) S_k(\lambda_j) \quad (11)$$

$$c_k(t) = \sum_{l=1}^{N_F} f_l(t) X_{lk} \quad (12)$$

Here N_C is the number of distinct spectral species and N_F is the number of analytic time functions. Defining \mathbf{C} as a $N_T \times N_C$ matrix with elements, $C_{ik} = c_k(t_i)$, and \mathbf{F} as a $N_T \times N_C$ matrix with elements, $F_{il} = f_l(t_i)$, the model data matrix \mathbf{D} can be written as:

$$\mathbf{D} = \mathbf{C}\mathbf{S} = \mathbf{F}\mathbf{X}\mathbf{S} = \mathbf{F}\mathbf{B} \quad (13)$$

The k -th row of the matrix \mathbf{B} , with elements $B_{kj} = b_k(\lambda_j)$, corresponds to the spectral changes associated with the time function, $f_k(t)$. When these time functions are exponential decays (convoluted with the instrument response), the corresponding spectra are called DADS. Optimization includes solving the least squares optimization problem:

$$\chi^2 = \|\Delta A - \mathbf{F}\mathbf{B}\|^2 = \text{Min} \quad (14)$$

Efficient algorithms exist that solve equation (14) for fixed matrices ΔA and \mathbf{F} . The value of χ^2 optimized in this way is further minimized by a nonlinear least squares algorithm by variation of the rate constants in \mathbf{F} . The DADS and the corresponding rate constants are the unique result of this global fit. This treatment does not require any model for the kinetics involved in the transient processes. The number of exponentials to use in the global fit is determined by SVD-based rank analysis, which is described in detail in the Supplementary Information of Lanzl *et al.*³⁵. The details of a model will be entirely defined in the matrix \mathbf{X} that relates the actual species

kinetics to the elementary function, $f_k(t)$. The appropriate matrix, \mathbf{X} , was chosen based on the model in Fig. 5a, and the SAS in matrix \mathbf{S} was calculated by:

$$\mathbf{S} = \mathbf{X}^{-1}\mathbf{B} \quad (15)$$

We tested the validity of each mechanistic model in two steps (see Supplementary Figs 18–21 and Supplementary Methods for full details): (i) the unique DADS are determined by the global lifetime analysis described above; (ii) various kinetic schemes are projected onto these DADS according to equation (15). Since the second step does not change the χ^2 value found in the first step, this procedure has the advantage that all interpretation is performed with the same quality of fit.

References

- Brown, K. L. Chemistry and enzymology of vitamin B₁₂. *Chem. Rev.* **105**, 2075–2149 (2005).
- Reed, G. H. Radical mechanisms in adenosylcobalamine-dependent enzymes. *Curr. Opin. Chem. Biol.* **8**, 477–483 (2004).
- Johnson, Jr J. E., Reyes, F. E., Polaski, J. T. & Batey, R. T. B₁₂ cofactors directly stabilize an mRNA regulatory switch. *Nature* **492**, 133–137 (2012).
- Peselis, A. & Serganov, A. Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. *Nat. Struct. Mol. Biol.* **19**, 1182–1184 (2012).
- Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S. & Elias-Arnanz, M. Light-dependent gene regulation by a coenzyme B₁₂-based photoreceptor. *Proc. Natl Acad. Sci. USA* **108**, 7565–7570 (2011).
- Weissbach, H., Ladd, J. N., Volcani, B. E., Smyth, R. D. & Barker, H. A. Structure of the adenylobamide coenzyme: degradation by cyanide, acid, and light. *J. Biol. Chem.* **235**, 1462–1473 (1960).
- Pratt, J. M. The chemistry of vitamin B₁₂. Part II. Photochemical reactions. *J. Chem. Soc.* 5154–5160 (1964).
- Jones, A. R., Levy, C., Hay, S. & Scrutton, N. S. Relating localized protein motions to the reaction coordinate in coenzyme B₁₂-dependent enzymes. *FEBS J.* **280**, 2997–3008 (2013).
- Armstrong, G. A. Genetics of eubacterial carotenoid biosynthesis: a colorful tale. *Annu. Rev. Microbiol.* **51**, 629–659 (1997).
- Cervantes, M. & Murillo, F. J. Role for vitamin B₁₂ in light induction of gene expression in the bacterium *Myxococcus xanthus*. *J. Bacteriol.* **184**, 2215–2224 (2002).
- Pérez-Marín, M. C., Padmanabhan, S., Polanco, M. C., Murillo, F. J. & Elías-Arnanz, M. Vitamin B₁₂ partners the CarH repressor to downregulate a photoinducible promoter in *Myxococcus xanthus*. *Mol. Microbiol.* **67**, 804–819 (2008).
- Diez, A. *et al.* Analytical ultracentrifugation studies of oligomerization and DNA-binding of TtCarH, a *Thermus thermophilus* coenzyme B₁₂-based photosensory regulator. *Eur. Biophys. J.* **42**, 463–476 (2013).
- Cheng, Z., Li, K., Hammad, L. A., Karty, J. A. & Bauer, C. E. Vitamin B₁₂ regulates photosystem gene expression via the CrtJ antirepressor AerR in *Rhodospirillum rubrum*. *Mol. Microbiol.* **91**, 649–664 (2014).
- Marques, H. M., Marsh, J. H., Mellor, J. R. & Munro, O. Q. The coordination of imidazole and its derivatives by aquocobalamin. *Inorg. Chim. Acta* **170**, 259–269 (1990).
- Rury, A. S., Wiley, T. E. & Sension, R. J. Energy cascades, excited state dynamics, and photochemistry in cob(III)alamins and ferric porphyrins. *Acc. Chem. Res.* **48**, 860–867 (2015).
- Shiang, J. J. *et al.* Ultrafast excited-state dynamics in vitamin B₁₂ and related cob(III)alamins. *J. Am. Chem. Soc.* **128**, 801–808 (2006).
- Walker, L. A., Shiang, J. J., Anderson, N. A., Pullen, S. H. & Sension, R. J. Time-resolved spectroscopic studies of B₁₂ coenzymes: the photolysis and geminate recombination of adenosylcobalamin. *J. Am. Chem. Soc.* **120**, 7286–7292 (1998).
- Yoder, L. M., Cole, A. G., Walker, L. A. & Sension, R. J. Time-resolved spectroscopic studies of B₁₂ coenzymes: influence of solvent on the photolysis of adenosylcobalamin. *J. Phys. Chem. B* **105**, 12180–12188 (2001).
- Jones, A. R. *et al.* Ultrafast infrared spectral fingerprints of vitamin B₁₂ and related cobalamins. *J. Phys. Chem. A* **116**, 5586–5594 (2012).
- Sension, R. J. *et al.* Photolysis and recombination of adenosylcobalamin bound to glutamate mutase. *J. Am. Chem. Soc.* **126**, 1598–1599 (2004).
- Sension, R. J. *et al.* Time-resolved measurements of the photolysis and recombination of adenosylcobalamin bound to glutamate mutase. *J. Phys. Chem. B* **109**, 18146–18152 (2005).
- Jones, A. R., Hardman, S. J. O., Hay, S. & Scrutton, N. S. Is there a dynamic protein contribution to the substrate trigger in coenzyme B₁₂-dependent ethanolamine ammonia lyase? *Angew. Chem. Int. Ed.* **50**, 10843–10846 (2011).
- Walker, L. A. *et al.* 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.* **120**, 3597–3603 (1998).

24. Robertson, W. D. & Warncke, K. Photolysis of adenosylcobalamin and radical pair recombination in ethanolamine ammonia-lyase probed on the micro- to millisecond time scale by using time-resolved optical absorption spectroscopy. *Biochemistry* **48**, 140–147 (2009).
25. Kutta, R.-J., Langenbacher, T., Kensy, U. & Dick, B. Setup and performance of a streak camera apparatus for transient absorption measurements in the ns to ms range. *Appl. Phys. B* **111**, 203–216 (2013).
26. Chen, E. & Chance, M. R. Nanosecond transient absorption spectroscopy of coenzyme B₁₂. Quantum yields and spectral dynamics. *J. Biol. Chem.* **265**, 12987–12994 (1990).
27. Chagovetz, A. M. & Grissom, C. B. Magnetic Field Effects in adenosylcob(III)alamin photolysis: relevance to B₁₂ enzymes. *J. Am. Chem. Soc.* **115**, 12152–12157 (1993).
28. Jones, A. R., Woodward, J. R. & Scrutton, N. S. Continuous wave photolysis magnetic field effect investigations with free and protein-bound alkylcobalamins. *J. Am. Chem. Soc.* **131**, 17246–17253 (2009).
29. Chen, Z.-G. *et al.* Dynamic, electrostatic model for the generation and control of high-energy radical intermediates by a coenzyme B₁₂-dependent enzyme. *ChemBioChem* **14**, 1529–1533 (2013).
30. Jost, M., Simpson, J. H. & Drennan, C. L. The transcription factor carH safeguards use of adenosylcobalamin as a light sensor by altering the photolysis products. *Biochemistry* **54**, 3231–3234 (2015).
31. Schrauzer, G. N. & Holland, R. J. Hydridocobalamin and a new synthesis of organocobalt derivatives of vitamin B₁₂. *J. Am. Chem. Soc.* **93**, 4060–4062 (1971).
32. Garr, C. D. & Finke, R. G. Radical cage effects in adocobinamide (axial-base-off coenzyme B₁₂): a simple method for trapping [Ado••Co^{II}] radical pairs. A new β-H elimination product from the radical pair and measurement of an unprecedentedly large cage-recombination efficiency factor, $F_c \geq 0.94$. *J. Am. Chem. Soc.* **114**, 10440–10445 (1992).
33. Garr, C. D. & Finke, R. G. Adocobalamin (AdoCbl or coenzyme B₁₂) cobalt-carbon bond homolysis radical-cage effects: product, kinetic, mechanistic, and cage efficiency factor (F_c) studies, plus the possibility that coenzyme B₁₂-dependent enzymes function as “ultimate radical cages” and “ultimate radical traps”. *Inorg. Chem.* **32**, 4414–4421 (1993).
34. Warren, M. J., Raux, E., Schubert, H. L. & Escalante-Semerena, J. C. The biosynthesis of adenosylcobalamin (vitamin B₁₂). *Nat. Prod. Rep.* **19**, 390–412 (2002).
35. Lanzl, K., Sanden-Flohe, M. v., Kutta, R.-J. & Dick, B. Photoreaction of mutated LOV photoreceptor domains from *Chlamydomonas reinhardtii* with aliphatic mercaptans: implications for the mechanism of wild type LOV. *Phys. Chem. Chem. Phys.* **12**, 6594–6604 (2010).
36. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
37. Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G. & Ludwig, M. L. How a protein binds B₁₂: A 3.0 Å X-ray structure of B₁₂-binding domains of methionine synthase. *Science* **266**, 1669–1674 (1994).
38. Higgins, D. G. & Sharp, P. M. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237–244 (1988).
39. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882 (1997).
40. Jost, M. *et al.* Structural basis for gene regulation by a B₁₂-dependent photoreceptor. *Nature*. doi: 10.1038/nature14950 (2015).

Acknowledgements

We are grateful to Prof. Bernhard Dick (Institut für Physikalische und Theoretische Chemie, Universität Regensburg) for use of the streak camera transient absorption apparatus, and Marco Jost and Prof. Catherine Drennan (MIT, USA) for helpful discussions regarding the CarH structure. A.R.J. and N.S.S. thank The University of Manchester, the Electromagnetic Fields Biological Research Trust and the Colt Foundation for funding. B.B. and P.B. thank the BBSRC for funding (grant BB/L002655/1 BBSRC). R.J.K. is a Deutsche Forschungsgemeinschaft Research Fellow. N.S.S. is a Royal Society Wolfson Merit Award awardee and EPSRC Established Career Fellow. Funding was also provided by the Ministerio de Economía y Competitividad (Spain) grants BFU2012-40184-C02-01 (co-financed by the European Union) to M.E.-A. and BFU2012-40184-C02-02 to S.P.

Author contributions

A.R.J. and R.J.K. wrote the paper with input from all authors; all experimental work was conducted by R.J.K. with S.J.O.H., B.B., H.L.M., J.M.O.-G., M.E.-A., S.P. and A.R.J.; computational work conducted by L.O.J. and R.J.K.; experimental data were analysed by R.J.K. with B.B. and A.R.J.; and results discussed by all the authors.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Kutta, R. J. *et al.* The photochemical mechanism of a B₁₂-dependent photoreceptor protein. *Nat. Commun.* **6**:7907 doi: 10.1038/ncomms8907 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>