Exploring Post-Activation Conformational Changes in Pigeon Cryptochrome 4

Fabian Schuhmann,[†] Daniel R. Kattnig,[‡] and Ilia A. Solov'yov[*,†]

†*Department of Physics, Carl von Ossietzky Universät Oldenburg*

Carl-von-Ossietzky Str. 9-11, 26129 Oldenburg, Germany

‡*University of Exeter, Stocker Rd., Exeter, EX4 4QD, U.K.*

Living Systems Institute and Department of Physics

E-mail: ilia.solovyov@uni–oldenburg.de

Abstract

A wide-spread hypothesis ascribes the ability of migratory birds to navigate over large distances to an inclination compass realized by the protein cryptochrome in the birds' retinae. Cryptochromes are activated by blue light, which induces a radical pair state, the spin dynamics of which may become sensitive to Earth's weak magnetic fields. The magnetic information is encoded and passed on to downstream processes by structural rearrangements of the protein, the details of which remain vague. We utilize extensive all-atom molecular dynamics simulations to probe the conformational changes of pigeon cryptochrome 4 upon light activation. The structural dynamics are analyzed based on a principal component analysis (PCA) and with the help of distance matrices, which reveal significant changes in selected interresidue distances. The results are evaluated and discussed with reference to the protein structure and its putative function as a magnetoreceptor. It is suggested that the phosphate-binding loop could act as a gate controlling the access to the FAD cofactor depending on the redox state of the protein.

Introduction

Sensory proteins often undergo characteristic structural changes upon activation, which is used to signal environmental stimuli to downstream processes, eventually realizing perception. While the mechanisms of many sensing processes have been deciphered in detail, the processes remain largely unknown when it comes to the magnetic compass, putatively realized in the protein cryptochrome.[1,2] The current evidence supports the idea of a magnetic sense that relies on coherent spin dynamics in a radical pair formed in the cryptochrome protein. This capacitates a light-dependent inclination compass that is closely linked to the birds' visual perception pathways,[3] and sensitive to weak radiofrequency magnetic field perturbations.[4,5] Cryptochrome photoactivation has widely been associated with the release of

the C-terminal domain, but how this reorganization is set in motion is unknown.[2,3,6] On the biophysical level, the problem of protein activation translates to one of studying the conformational changes of the photoreceptor protein cryptochrome. An earlier study discussed[7] the activation for the European robin cryptochrome 4 (ErCry4a),[1,8,9] for which strong conformational changes were found in the C-terminal region and in surface-exposed loops (residues 405-415, 276-285, 237-246, 179-186, and 38-48). However, as the crystal structure of the protein had not yet been resolved, this study was based on a homology model, which might not have correctly reproduced some of the protein's pertinent features. Here, the activation of pigeon cryptochrome 4 (ClCry4) by blue light is analyzed. ClCry4 has recently been crystallized,[10] allowing to gain deeper insights into cryptochrome post-activation conformational changes while avoiding the uncertainty of the homology model plaguing the earlier study.[7] Even though pigeons are non-migratory birds, their cryptochrome 4 has many marking of a night-migratory bird present; ClCry4 exhibits high sequence identity to ErCry4a (85.6% identical residues) and was shown to bind flavin adenine dinucleotide (FAD), which is key for the magnetoreception function of the protein.[11] It is therefore expected that the study of ClCry4 will be efficacious in informing further investigations on cryptochromes from migratory species, once those crystal structures become available.

Cryptochrome 4 is activated by photoexcitation of its blue-light sensitive cofactor FAD, which is non-covalently bound in the active site of the protein.[1,2,11–14] When photo-excited, FAD is reduced to a negatively charged radical anion ($FAD^{\bullet-}$), the electron being donated by a nearby tryptophan residue (Trp395, W_A), which is thus converted into a positively charged radical ion, $Trp^{\bullet+}$. In a succession of follow-up electron transfer processes, $W_A^{\bullet+}$ accepts an electron from Trp372 (W_B), which receives an electron from Trp318 (W_C). Lastly, the surface exposed Trp369 (W_D) reduces the tryptophan radical cation of W_C, leaving W_D positively charge. In the course of this charge transfer process, $FAD^{\bullet-}$ and $TRP^{\bullet+}$ form a so-called spin-correlated radical pair,[1,2,7,13,15–18] the recombination of which is magnetosensitive due to

the Radical Pair Mechanism. The concentration of the final radical pair, henceforth denoted Radical Pair D (RPD), is thus expected to be influenced by applied magnetic fields and could therefore allow the bird to sense directions. An integral part of the processes is the structural reorganization of the protein, induced by the charge separation, which is used to signal the direction and intensity of the magnetic field to downstream processes. On a longer timescale, the restructured radical pair is expected to evolve back to its deactivated (resting) state, possibly in processes involving further radicals, such as a long-lived tyrosine radical,[19] and proton transfer.[2,20]

Employing extensive molecular dynamics (MD) simulation of ClCry4 in its inactive dark state (DS) and in the RPD state, the conformational changes throughout the protein were sampled, characterized, and visualized. The DS had FAD bound in its fully oxidized resting state and all protein residues in their usual form; for the RPD state, the FAD and W_D were radicalized. Details of the structures, their configuration for MD and simulation protocols are provided in the Supplementary Material (SM). In total, three pairs of simulations, each comprising a DS and RPD simulation, were produced. Each simulation set included ClCry4 in the DS and RPD state and were carried out for a duration of 1 μs, the expected lifetime of a radical pair.[21,22] The stability of the DS and RPD structures was verified by analysis of the root mean square displacement (RMSD). The modeled structures ascribed to the two states were reasonably stable during the studied time frame. To quantify the early rearrangements in ClCry4 arising upon photoexcitation, we have analyzed distance matrices of the average any-to-any distance for each residue for the DS and the RPD states for all replica simulations, which, when combined, reveal the relative average distance by which the RPD state changes with respect to the DS. Furthermore, a principal component analysis (PCA) was performed. The PCA allows a reduction in dimension, permitting to focus on the most relevant components of structural fluctuations with regard to the covariance relative to the average of the DS and RPD state of ClCry4. The performed investigation revealed

that the PCA can not only be used to identify the essential conformational changes within a protein but also allows a discussion of the structure evolution over time. More specifically, the results illuminate that activated ClCry4 exhibits a noteworthy conformational change in residues 220-245. It is suggested that this change, involving the so-called phosphate-binding loop, might act as a gate controlling the access to the FAD cofactor.[11,23]

Methods

The crystal structure of ClCry4 was taken from the protein data bank (6PU0).[10] The CHARMM36 force field with CMAP corrections was used to describe the interatomic interactions.[24–31] Parameters for FAD and FAD$^{\bullet-}$ were adopted from earlier studies.[1,20,32] MD simulations were performed using NAMD.[33,34] The simulations were conducted in an NPT (constant number of particles, pressure, and temperature) statistical ensemble at a temperature of 310 K and atmospheric pressure of 1 bar. Simulations were set up using the online platform VIKING.[35] The crystal structure of ClCry4[10] is missing residues 228-244. In order to simulate a connected protein structure, the gap was reconstructed with the aid of data obtained from the authors of the original crystal structure, the poorer resolution of which did not meet the publication requirement[1]. The protonation states of amino acid residues in the ClCry4 structure, specifically histidines, were checked with the use of multiple pK_a predictors (DelPhiPka, pdb2pqr); further checks of the protonation states of histidines were done through visual inspection of the hydrogen bonding network at the sites of potential protonation which turned out to be consistent with the automated predictors. His64, His353, His405, and His471 residues were protonated at their ε-position, and the His3 and His7 residues were doubly protonated. The remaining histidines were assumed δ-protonated. The missing hydrogen atoms were added using the program VMD.[36] Subsequently, the structure of the

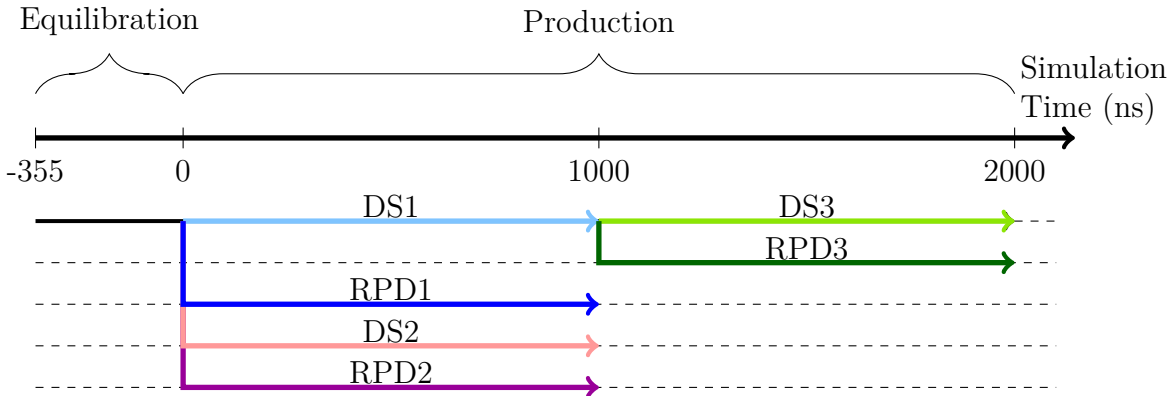_____

[1]Brian Zoltowski, private communication

Figure 1: A timeline for the conducted simulations of ClCry4 in the DS and RPD state. Three sets of replica simulations were performed. The equilibration simulation of only a DS structure was conducted until the time instance 0 ns. At this point in time, the first two replica production simulations were initialized (DS1 and DS2), and two RPD structures were forked off the original DS structure (RPD1 and RPD2). A third set of simulations (DS3 and RPD3) was initiated at a time instance of 1 μs, i.e., at the end of the DS1 simulation.

protein was solvated in a water box with a padding distance of 20 Å and NaCl was added to a final concentration of 0.2 mol/L to neutralize the system and mimic a physiological ionic strength.

The solvated ClCry4 structure was first pre-equilibrated for 5 ns and then for 10 ns with a time step of 1 fs. The two pre-equilibrations differed in the constraints used. In the first one, the entire protein and FAD were harmonically constraint to the atomic positions taken from the crystal structure. The second equilibration left the backbone atoms constrained while the side-chain atoms were free to move. The system was then further equilibrated with a 2 fs integration time step for a total of 355 ns, constraining the hydrogen atoms to be rigidly bonded to the heavier atoms.[37] The FAD cofactor was assumed to be in a fully oxidized form during the equilibration, i.e., the protein was in the DS.

In total, three replica simulations were performed following the protocol as visualized in Fig. 1. After the equilibration, two pairs of production simulations were initialized. Each pair consists of (i) the continuation of the DS-simulation and (ii) a branched-off RPD configuration of the protein. These simulations are denoted as DS1/RPD1 and DS2/RPD2,

respectively. A third simulation was initialized at the time instance 1 μs of DS1, at which an additional RPD state (RPD3) was forked off, and the DS was prolongated (DS3). Every simulation was carried out for 1 μs each. Throughout the summary, the results corresponding to the different simulations can be identified by their names (i.e., DS1) and their color as indicated in Fig. 1. The simulation pairs will be henceforth called simulation set 1, 2, and 3.

All simulations were analyzed based on averaged any-to-any distance matrices. Additionally, set 1 was analyzed employing PCA. A detailed account of the used analyses, including the workflow and relevant equations, is given in the SM.

Results

The changes of distances for all residues, assessed based on their centers of mass, were calculated as

$$a_{ij} = \left\langle \left\| \vec{r}_i^{DS} - \vec{r}_j^{DS} \right\| \right\rangle - \left\langle \left\| \vec{r}_i^{RPD} - \vec{r}_j^{RPD} \right\| \right\rangle, \tag{1}$$

with $\vec{r}_j^{DS,RPD}$ being the position of the center of mass of residue j in either the DS or the RPD state of ClCry4. $\langle \cdot \rangle$ denotes the average value over the trajectory under consideration. The obtained distance differences were sorted into an $N \times N$-distance matrix. The information in the matrix can be visualized using a matrix plot, where each entry value is assigned a color as shown for DS1 and RPD1 in Fig. 2A. To show a more explicit image of the conformational changes per residue, the average over the absolute values for each column in the matrix was calculated and is visualized in Fig. 2B. The plot in Fig. 2B thus quantifies the average displacement of the average movement of the RPD1 residues relative to the ones in the DS1. In Fig. 2A, one can see two marked crossing stripes, corresponding to the residues 40-50 and residues 220-245. These stripes highlight the regions, which show the most prominent structural rearrangements in the RPD1 state of ClCry4 relative to the
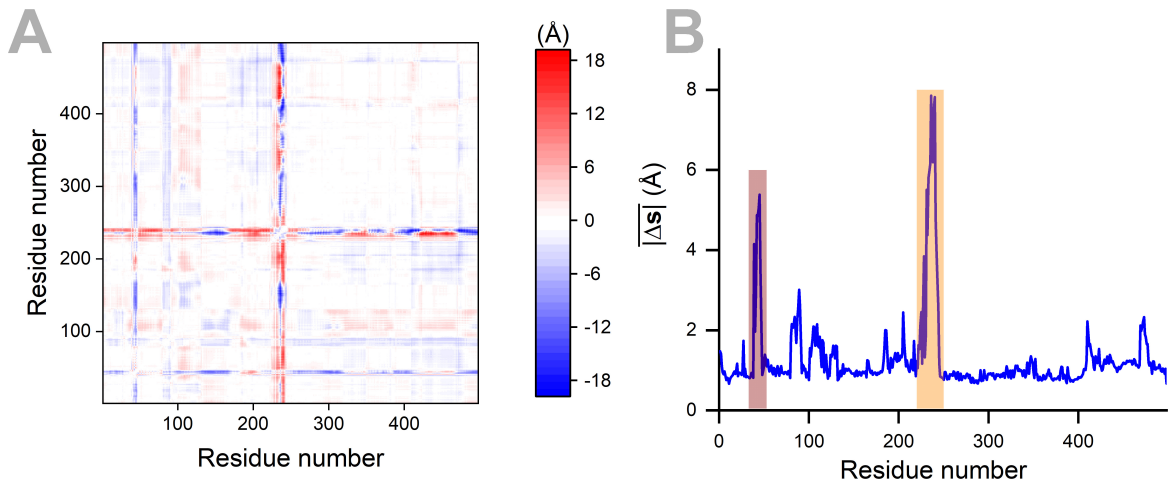
Figure 2: Panel A: A matrix plot of the distance difference matrix defined in Eq. (1) is shown. The colors indicate the difference between the matrix' elements averaged over the last 625 ns of DS1 and RPD1. One can identify the stripes assigned to residues 40-50 and residues 220-245, which exhibit a noticeable distance change upon ClCry4 changing its state from the DS1 to the RPD1 state. The results from Panel A can be further elaborated by taking the absolute sum over each column of the matrix in A and dividing by the number of residues ($\overline{|\Delta\mathbf{s}|}$). This analysis yields the average distance change per residue once the redox state of ClCry4 changes. The representation in Panel B shows which residues in the RPD1 state move considerably compared to the DS1 simulation. The regions highlighted in red and orange refer to the most mobile areas, associated with residues 40-50 and residues 220-245, respectively.

DS1 structure. The representation in Fig. 2B features the mobile regions from the matrix plot even more pronouncedly. The comparison of the results with the data obtained in an earlier study[7] suggests that the present $\overline{|\Delta s|}$ values associated with the versatile regions in the protein are larger by a factor 3. One also does not notice any large fluctuations in the C-terminal part of the protein. To get a comprehensive picture of the changes happening with the introduction of the radical pair inside ClCry4, the distance matrices were also computed for the replica simulations 2 and 3. The resulting averaged columns $\overline{|\Delta s|}$ are shown in Fig. 3.

The comparison of the three replica simulations suggests a notable similarity in simulation sets 1 and 2 for residues 220-245, while the second versatile region found in the first set
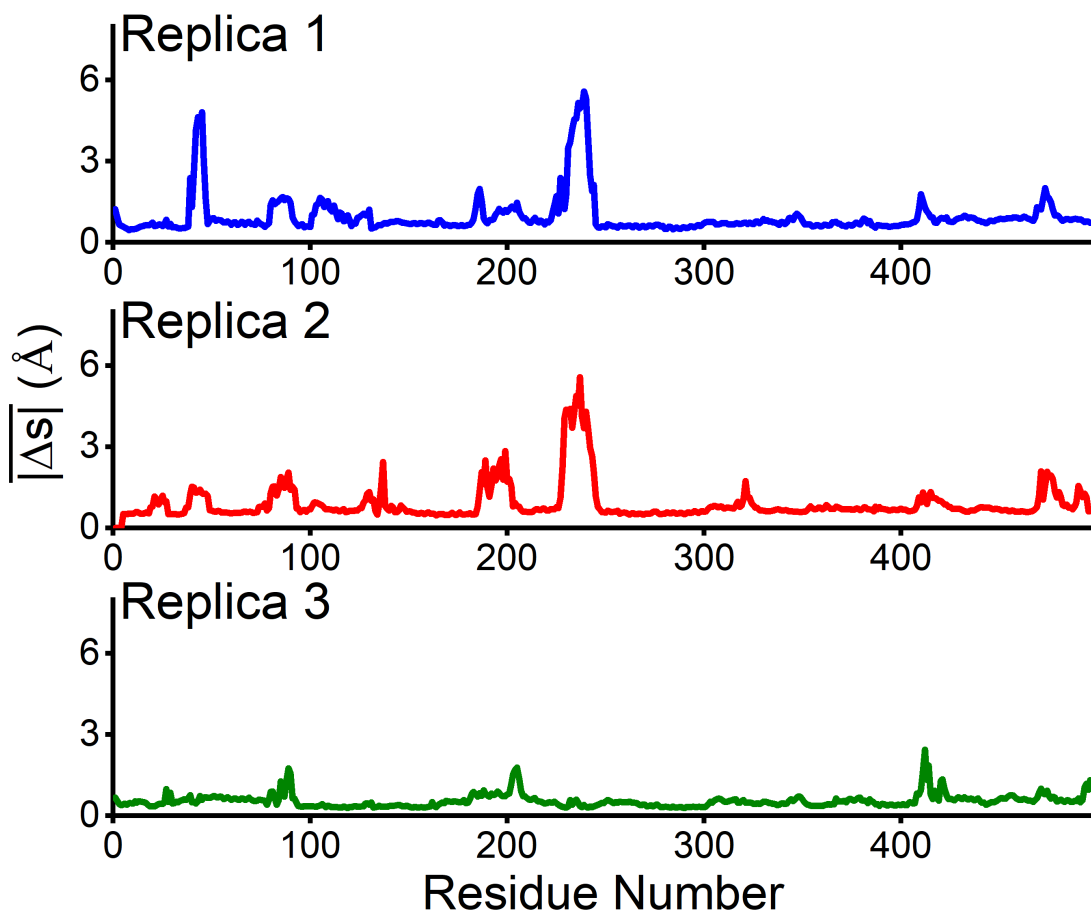
Figure 3: Comparison of the interresidue distance $\overline{|\Delta s|}$ computed for three sets of replica simulations for ClCry4. One notices that the results for set 1 and 2 show similar differences for residues 220-245. The fluctuations for simulation set 3 show only limited reorganizations during the studied time frame.

(residues 40-50) is not prominently flexible in the other replicas. Furthermore, the distance matrix analysis shows no large-scale conformational changes for the two simulations of DS3 and RPD3.

In order to identify if the changes in the distance matrix are the result of a single or many reorganization modes in ClCry4, a PCA approach was employed for simulation set 1. Judging from the distance matrix analysis in Fig. 3, one expects a similar outcome for simulation sets 1 and 2. In contrast, no significant rearrangements are expected in the case of replica simulation 3. The PCA was carried out exclusively for simulation 1. A detailed

9

Table 1: The first eight eigenvalues λ_i of the covariance matrix C defined in Eq. (S3) are shown for the combined $2 \cdot 625$ ns of DS1 and RPD1 together with the percentage of the total variance that they express.

	Eigenvalue Å^2	Percentage to the variance	Accumulated percentage
1	1508.43	65.9	65.9
2	69.03	3.0	68.9
3	58.89	2.6	71.5
4	40.44	1.8	73.3
5	33.66	1.5	74.8
6	28.56	1.3	76.1
7	25.64	1.1	77.3
8	23.49	1.0	78.3

account and an explanation of variables are given in the SM. In short, the $3N-$dimensional vector $\vec{R}$ of the combined residue locations is constructed, where $N = 497$ is the number of residues in the structure. One such vector $\vec{R}$ is obtained for each state of ClCry4 (DS1 and RPD1) and for each of the $\mathcal{N} = 125.000$ snapshots that make up the $2 \cdot 625$ ns of DS1 and RPD1, used for the analysis. Note that we use the 625 ns of the DS1/RPD1 trajectories for the analysis as both structures are deemed sufficiently equilibrated for that interval (see SM). The covariance matrix C (Eq. (S3)) was constructed and its eigenvalues were calculated.

Table 1 shows the first eight eigenvalues of the covariance matrix and their contribution to the total variance. Note that these first eight eigenvalues and therefore the corresponding eigenvectors (principal components) explain almost 80% of the total variance of the movement in the DS1 and RPD1 state; in other words, 0.5% of all eigenvectors are sufficient to explain 80% of the variance of the ClCry4 activation dynamics. Even more clear cut and already nominally visible is the first eigenvalue λ_1, which alone describes 66% of the structural motion variability in ClCry4 and is clearly set apart from the subsequent eigenvalues. One can visualize the conformational changes described by each principal component by calculating the root-square value for each residue as

$$RS_{ij} = \sqrt{\left(e_{j(3i-2)}^2 + e_{j(3i-1)}^2 + e_{j(3i)}^2\right) \cdot \lambda_j}. \tag{2}$$

Here $e_{j(\xi)}$ is the ξth entry in the principal component $\vec{\Lambda}_j$ (i.e $\vec{\Lambda}_1$ is the eigenvector corresponding to the largest eigenvalue λ_1). The index i in Eq. (2) labels the residues in ClCry4. The conformational changes were described through RS values weighted by their respective eigenvalues to distinguish the contributions and their importance to the total variance; the results are visualized in Fig. 4. From bottom to top, the eight most dominant principal components are illustrated through the corresponding RS values that follow from Eq. (2). The dominant influence of the first principal component on the residues 40-50 and 220-245 is is clearly visible. Some conformational change can be seen in the existing parts of the C-terminal tail of ClCry4, but it is not as distinct as the changes in other aforementioned regions of the protein.

Figure 5 visualizes the relationship of the DS1 and the RPD1 state towards the respective principal component by showing the similarityof the deviation of the structures from the average contribution and the principal components for each state of ClCry4 for each snapshot taken from the MD simulations. To compare this similarity, the vectors describing the deviation of ClCry4 residues from the average position are constructed and denoted by $\vec{\delta}^{DS}$ and $\vec{\delta}^{RPD}$, see Eq. (S8). Next, the dot product is computed between each principal component and the obtained $\vec{\delta}^{DS}$ and $\vec{\delta}^{RPD}$ vectors. Expressing the δ in terms of the principal components, the components along particular PCs reflect the similarity between a given configuration and this PC. Results show that the first principal component exhibits a definite difference in the DS1 and RPD1 state throughout the whole simulation (Fig. 5A), whereas this is not the case for the other principal components. Only in the second principal component, some differences between the DS1 and RPD1 states can be distinguished (Fig. 5B), but even there, one cannot judge the direction of the conformational change,
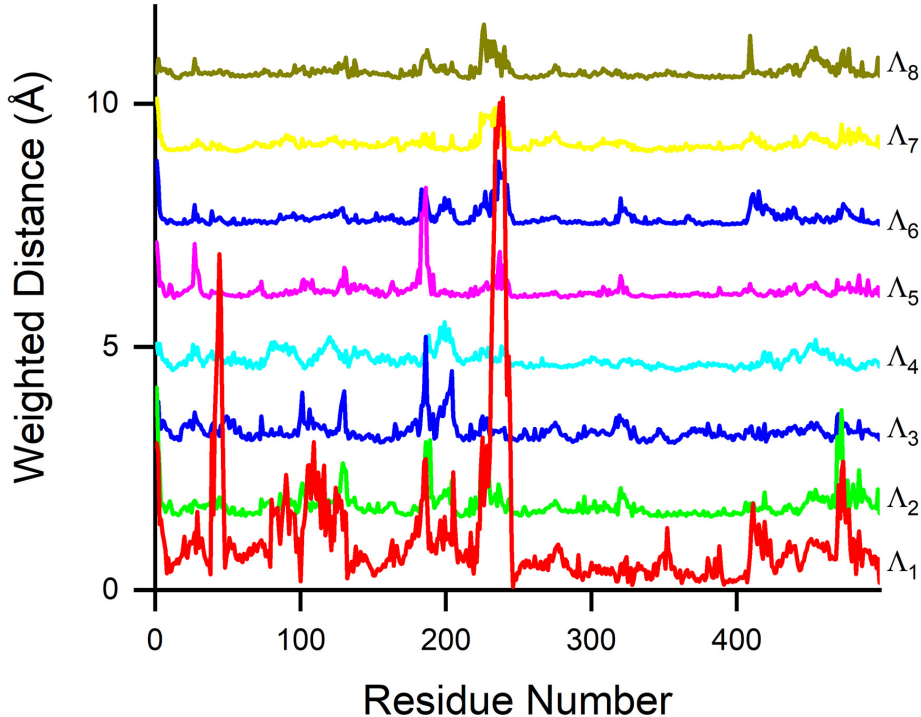
Figure 4: The root square displacement for each residue has been calculated and weighted by the eigenvalue for the first eight most dominant principal components, as described in Eq. (2). The root square displacement associated with the jth principal component has been translated by $(j-1)$ along the weighted distance axis for better visibility. The analysis shows a similar result to the distance matrix seen in Fig. 2 for the most dominant principal component $\vec{\Lambda}_1$, where the prominent peaks rapidly decline in subsequent principal components. One notes the prominent rearrangements in residues 40-50 and residues 220-245 to be mostly, if not exclusively, expressed through the very first principal component.

i.e., the principal component reflects general variability but no clear distinction of DS1 and RPD1 state, and thus, activation dynamics. Starting from principal component three, the similarity of the displacements with the principal components (as expressed in Eq. (S9) and Eq. (S10)) becomes quickly indistinguishable. This reveals that the movement contains only minor components along the corresponding principal component.

The rapid decline in the contribution of higher principal components to ClCry4 dynamics leads to the interpretation that the movement in both regions identified in the ClCry4
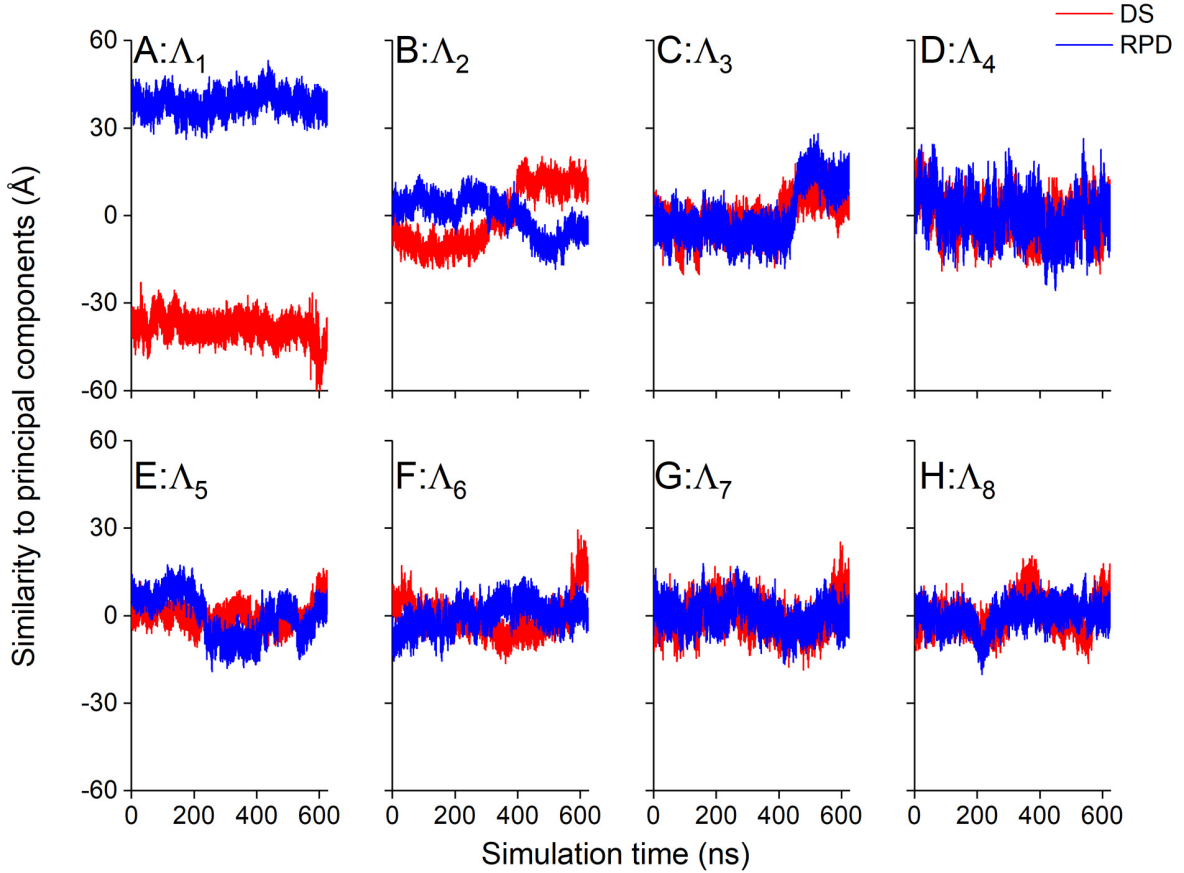
Figure 5: The plot shows the similarity between the principal components and the average configuration for the two states of ClCry4. The similarity is calculated as the dot product of the vectors $\vec{\delta}^{DS}$ (red) and $\vec{\delta}^{RPD}$ (blue) with the first eight principal components as defined as $\alpha^{DS}(k)$ and $\alpha^{RPD}(k)$ in Eq. (S9) and Eq. (S10). Panel A shows the similarity to the first principal component with a clear distinction between the DS1 and the RPD1 states. The subsequent principal components (panels B-H) do not show a clear difference between the states and the change can be less expressed by the principal components; already the second principal component exhibits a mean close to zero.

structure, namely residues 40-50 and residues 220-245, can be attributed to just one principal component, which links the two versatile regions of ClCry4 and their importance to the conformational change in the protein. Furthermore, a visual inspection reveals that the two regions are located on a straight line passing through the FAD cofactor in the center of the protein, see Fig. 6.

13

The very unambiguous difference between the DS1 and the RPD1 states of ClCry4 raises the question as to how long the RPD1 protein structure takes to divert from the DS1 structure. To this end, the principal component analysis was repeated starting from the time instance of 80 ns before the RPD1 state was initiated until 300 ns after this event. The vectors $\vec{\delta}^{DS}$ and $\vec{\delta}^{RPD}$ as defined in Eq. (S8) generated for this simulation time window were then compared with the principal components obtained earlier. A rapid decrease of the RPD1 state similarity with its parent state is manifested after about 100 ns once the radical pair is formed, see Fig. S3. The analysis again demonstrates clearly that this change is only observed in the first principal component, while the subsequent principal components do not show significant differences between the DS1 and the RPD1 state.

Discussion and Conclusions

A visual inspection of the mobile region for ClCry4 standing out for simulation sets 1 and 2 clarifies the origin of the observed conformational change. Panels A and B in Fig. 6 show the two versatile regions identified in replica simulation 1. Panel C in Fig. 6 shows the phosphate-binding loop in the context of the protein and panel D in Fig. 6 isolates the loop. The reorganization of the phosphate-binding loop might be interpreted as an opening of a molecular gate towards the FAD inside ClCry4.[11] Our results show a clearly visible difference for the spatial organization of that region in the case of replica simulation 1. Figure 7 shows the spatial reorganization of the phosphate-binding loop in the case of simulation sets 2 and 3 in comparison with the one from simulation set 1. Panels A and B in Fig. 7 show a similar opening movement of the phosphate-binding loop for simulation sets 1 and 2, even though the gate opens in a slightly different direction. Panels C and D in Fig. 7, on the other hand, illustrate the absence of significant rearrangements in the RPD3 state compared to the DS3. Furthermore, one notices a small difference in the spatial reorganization of the phosphate-
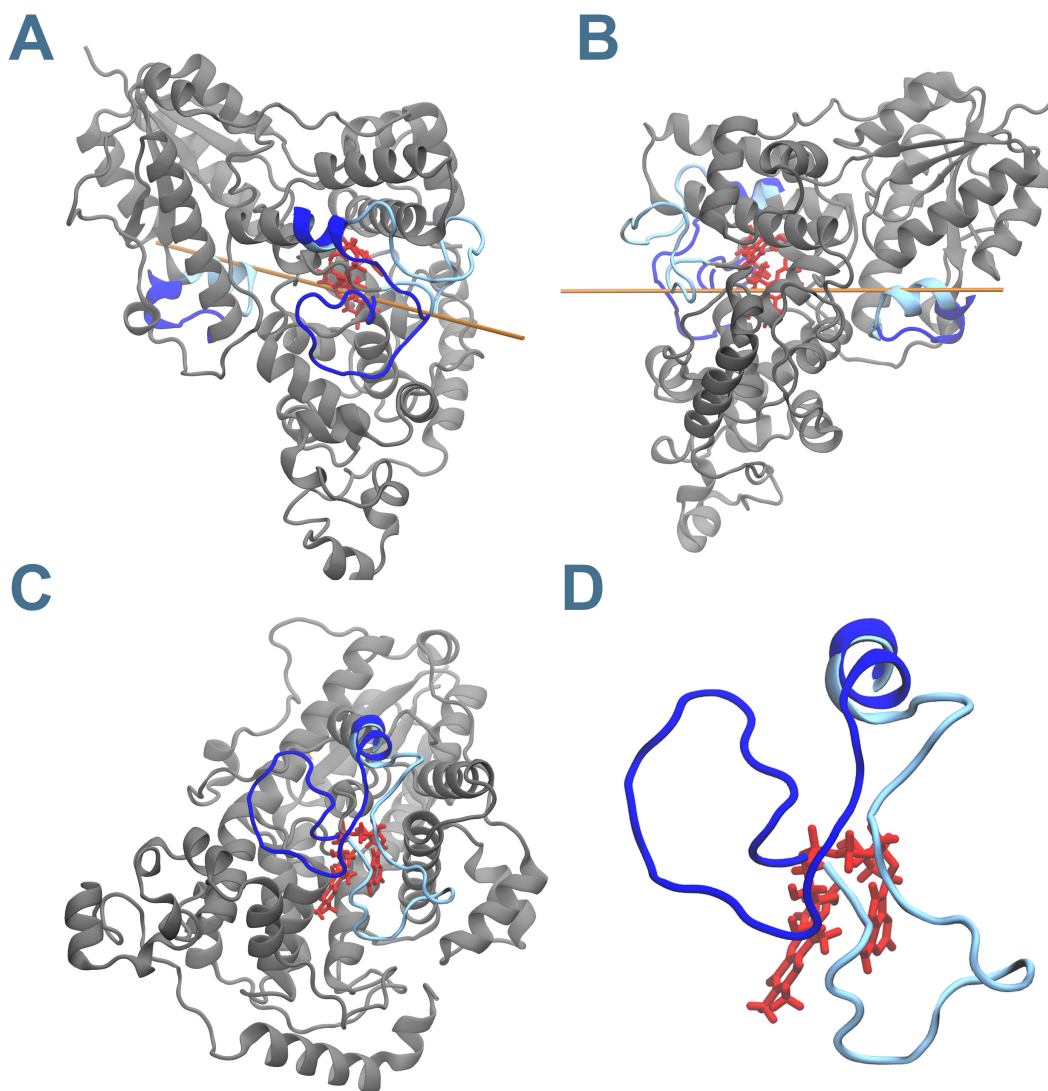
Figure 6: ClCry4 is shown with a focus on its versatile regions. The RPD1 is shown in blue, while the DS1 is highlighted in light blue. FAD is shown in red. A visual inspection of the flexible regions in ClCry4 shows that residues 40-50 and 220-245 are located on a line that passes through the FAD on opposite sites of the protein structure, as indicated in panels A and B by the orange line. Panel C and D focus on the phosphate-binding loop as it shows to be a region of primary reorganization. Panel D shows the opening of the gate to the FAD through the movement in the RPD1 simulation.

binding loop in DS3/RPD3 as compared to the DS1. This change might explain why in the case of simulation sets 1 and 2, one observes the opening movement during the simulation time frame, while it is not present in the case of replica 3 simulation. All renderings are

taken from the last snapshot of the respective simulation trajectory.

Our results show a clear difference between the DS and RPD state for the two simulation sets 1 and 2 of ClCry4 in residues 220-245 and additionally, a clear difference in residues 40-50 in case of the first simulation set. This result is also described by the first principal component for simulation set 1, which is more distinct for ClCry4 compared to the previous study[7] for ErCry4a, in which the same variance was covered by the first three principal components. The phosphate-binding loop in ClCry4 consists of the same residues as in the ErCry4a except for two exceptions (92% sequence homology), which might indicate that the movement in ErCry4a is expected to resemble the conformational changes in ClCry4. The mobile region around residues 80-135 in ErCry4a is not clearly distinguished for ClCry4, even though it exhibited movement expressed by the second principal component for ErCry4a. Additionally, the earlier investigation found the phosphate-binding loop to be versatile in the RPD relative to the DS in the employed distance matrix analysis, but the dominant principal component did not markedly feature for this particular movement. Here, the study for ClCry4 yields a more clearcut picture in that the phosphate-binding loop predominantly changes its conformation if the RPD state is formed. Our observations are in line with the experimental study by Zoltowski et al.[10] that on proteolysis assays under light and dark conditions, suggests that the photoactivation of ClCry4 could be coupled to order-disorder transitions in the phosphate binding loop.

In order to investigate the effects of the conformational changes for the protein structure, simulation set 1 was analyzed in terms of its solvent-accessible surface area (SASA) per residue and possible phosphorylation sites. Using the software VMD[36] and a probe sphere radius of 1.4 Å, the SASA was calculated for every protein residue, including FAD. The values were normalized, i.e., divided by, the nominal SASA values for each residue as tabulated earlier.[38] The values of SASA computed for the FAD cofactor have not been normalized. The FAD exhibits the greatest change in SASA, climbing from 36.44 Å^2 to 60.38 Å^2, while
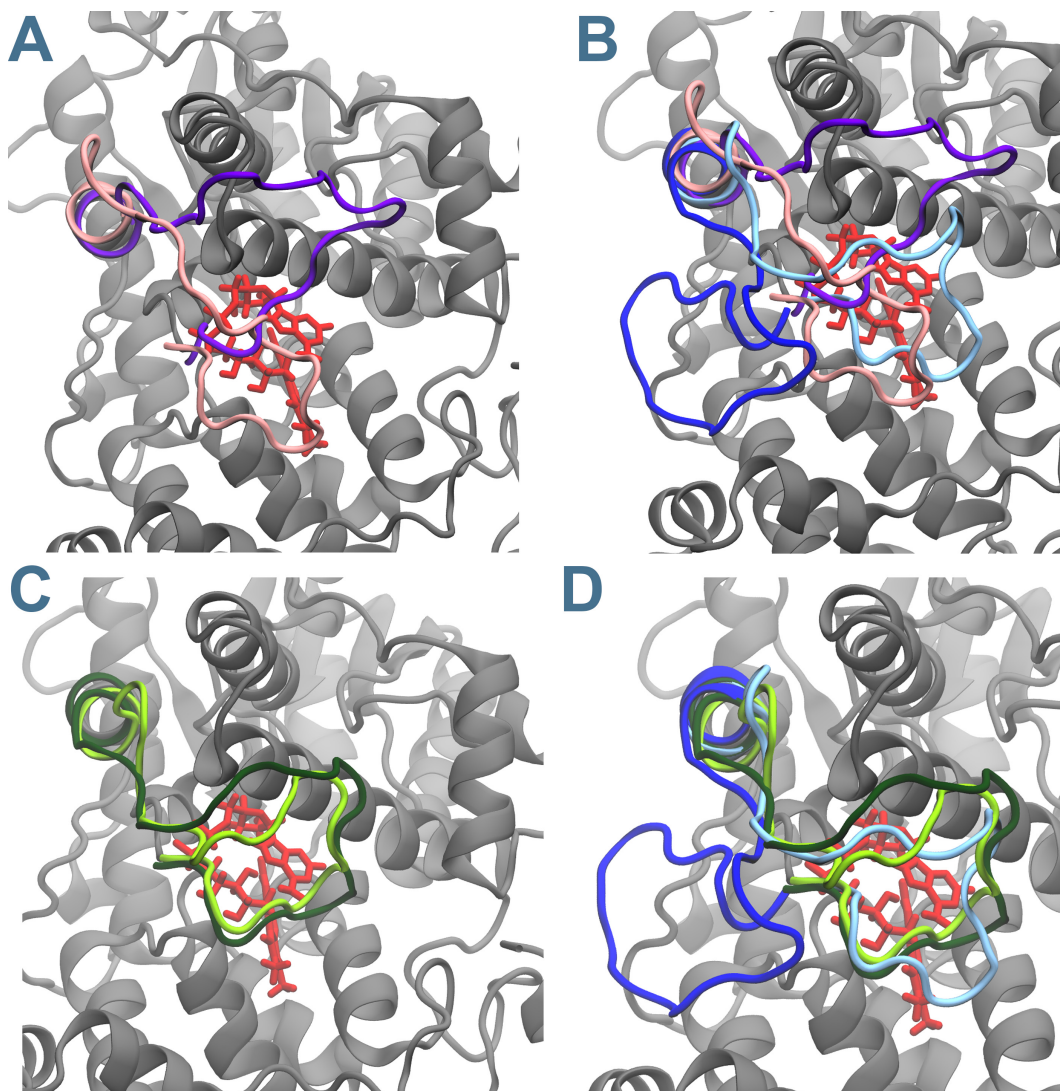
Figure 7: Visual inspection of all replica simulations is shown. FAD is colored red. Panel A shows the confirmation of DS2 (mauve) and RPD2 (violet). One notices a similar opening movement as in RPD1, but in a slightly different direction. Panel B compares the structures of the phosphate-binding loop with simulations DS1 (light blue) and RPD1 (blue). One notices that the organization of the phosphate-binding loop in the dark states is relatively similar while both RPD states swing open in slightly different directions. Panel C shows DS3 (lime) and RPD3 (dark green). Here the phosphate-binding loop is stronger stabilized, not showing marked reorganization movements upon ionization during the simulated time interval. Panel D compares the structures once more to the reorganisation motif exhibited in simulation set 1. One notices that the RPD3 will have to pass through the DS1 conformation to evolve to the RPD1 state. All renderings have been generated from the last snapshot of the respective simulation trajectory.

still being located at its original binding pocket. A visualization of the FAD cofactor's SASA is given in Fig. S4A-B. This analysis reveals that the FAD cofactor can be approached by water molecules after the RPD1 conformational changes through an opening provided by the phosphate-binding loop rearrangement. Looking at the other residues, one identifies that Ala45 and Pro239 exhibit the greatest increase in their SASA values, and Leu242 exhibits the greatest decrease, measured relative to their nominal values. Note that both, Pro239 and Leu242, are located in the phosphate-binding loop, which showed the greatest conformational change. The changes in SASA values for all residues of ClCry4 are compiled in Table S1.

Serine is the most common phosphorylation site of proteins. Phosphorylation can lead to activation or deactivation, functioning as a regulating agent, which is also well-established for plant cryptochromes, for which the phosphorylation was found to be magnetosensitive.[39] Serines could also be responsible for mediating direct interaction of a protein with membranes and lipid head groups. Such interactions may be specifically important to describe intercellular cryptochrome interaction networks responsible for signalling.[40] Netphos 3.1[41] discovered 17 potential serine phosphorylation sites in the whole protein structure of ClCry4. Ser44, Ser206, and Ser250 are located either within or very close to the identified moving regions. Table 2 summarizes the changes in SASA for serines identified as most likely linase targets by Netphos.

Table 2 shows that Ser278 exhibits the greatest difference in SASA among all the studied serines. A visual inspection of that residue reveals that the phosphate-binding loop takes a major role here, even though Ser278 itself remains rather static; in the DS1, the phosphate-binding loop almost completely blocks Ser278, whereas it is freed up in the RPD1 state. Ser44 is influenced more directly, being located within the mobile region around residues 40-50. A visual inspection here reveals that the rotation of an α-helix in that region of ClCry4 rotates Ser44 to be tucked away, which is aided by the formation of an α-helix involving

Table 2: The 17 serine residues of ClCry4 identified by Netphos[41] with the corresponding change in SASA upon protein activation from the DS1 to the RPD1 state. Specifically highlighted are serines 44, 206 and 250 for their closeness to the moving regions. Notably, Ser44 becomes less surface exposed, whereas Ser206 and Ser250 seem to be far enough away from the phosphate binding loop to not be noticeably influenced. Ser278's SASA increases drastically.

Residue number	SASA DS ($\mathring{A}^2$)	SASA RPD ($\mathring{A}^2$)	Absolute Difference ($\mathring{A}^2$)
28	0.22	0.08	0.14
44	**0.79**	**0.42**	**0.38**
59	0.00	0.00	0.00
132	0.20	0.11	0.09
146	0.46	0.51	0.05
159	0.68	0.71	0.03
181	0.15	0.25	0.10
204	0.95	0.75	0.20
206	**0.26**	**0.25**	**0.01**
250	**0.01**	**0.06**	**0.04**
259	0.03	0.07	0.04
268	0.35	0.25	0.10
278	0.07	0.46	0.39
371	0.16	0.03	0.12
389	0.01	0.07	0.06
449	0.49	0.55	0.06
475	0.62	0.52	0.11

residues 194-199.

Lastly, the distance between Trp369 (W_D) to FAD has been calculated for the DS1 and RPD1 simulations as given by their respective geometric center. The residue deserves such a detailed treatment, as it changes its charge during ClCry4 activation process. We find that Trp369 on average moves 0.61 Å towards the FAD in the process of protein activation. Such a difference in distance may lead to a siginificant change in electron transfer rates and could potentially be important for ClCry4 functionality.[42] These structural changes are visualized in Fig.S2.

We want to summarize that employing multiple extensive simulations of ClCry4 in its

two pertinent states, DS and RPD, one particular region was identified, which moved significantly in two simulation pairs upon protein activation. The first simulation also showed some peculiar movement on the opposite side of the protein, which was not reproduced by its replica simulation. In the first simulation, it was shown that these changes could be attributed solely to the first principal component describing the conformational dynamics of the protein. Furthermore, it was demonstrated that these dynamics allow for a possibility for solvent to reach the FAD inside ClCry4 and expose several possible phosphorylation sites while FAD remains bound. Finally, in the third replica simulation, we found that the DS can reach a more stable conformation of the phosphate-binding loop, in which the movement induced by the RPD3 state appears to require a significantly longer time for activation.

Acknowledgement

by the North-German Supercomputing Alliance (HLRN).

Supporting Information Available

Supplementary Material

- Data on protein stability after equilibration and during production simulations.

- Detailed description of methods employed in the analysis:

  - Guide to calculating the distance matrix;

  - Guide to do the PCA;

- Additional supporting visualizations:

  - Visualization of the Trp369-FAD distance in DS1 and RPD1 states.

  - Similarity of the DS1 and RPD1 states upon the DS1→RPD1 transition;

  - Table of SASA for all residues;

  - Visualization of FAD's SASA and SASA changes for DS1 and RPD1 states.

References

(1) Xu, J.; Jarocha, L. E.; Zollitsch, T.; Konowalczyk, M.; Henbest, K. B.; Richert, S.; Golesworthy, M. J.; Schmidt, J.; Déjean, V.; Sowood, D. J. C. et al. Magnetic sensitivity of cryptochrome 4 from a migratory songbird. *Nature* **2021**, *594*, 535–540.

(2) Hore, P. J.; Mouritsen, H. The Radical-Pair Mechanism of Magnetoreception. *Annu. Rev. Biophys.* **2016**, *45*, 299–344.

(3) Zapka, M.; Heyers, D.; Hein, C. M.; Engels, S.; Schneider, N. L.; Hans, J.; Weiler, S.; Dreyer, D.; Kishkinev, D.; Wild, J. M. et al. Visual but not Trigeminal Mediation of Magnetic Compass Information in a Migratory Bird. *Nature* **2009**, *461*, 1274–1277.

(4) Kobylkov, D.; Schwarze, S.; Michalik, B.; Winklhofer, M.; Mouritsen, H.; Heyers, D. A newly identified trigeminal brain pathway in a night-migratory bird could be dedicated to transmitting magnetic map information. *Proc. R. Soc. B Biol. Sci.* **2020**, *287*.

(5) Engels, S.; Schneider, N. L.; Lefeldt, N.; Hein, C. M.; Zapka, M.; Michalik, A.; Elbers, D.; Kittel, A.; Hore, P. J.; Mouritsen, H. Anthropogenic electromagnetic noise disrupts magnetic compass orientation in a migratory bird. *Nature* **2014**, *509*, 353–356.

(6) Wu, H.; Scholten, A.; Einwich, A.; Mouritsen, H.; Koch, K. W. Protein-protein interaction of the putative magnetoreceptor cryptochrome 4 expressed in the avian retina. *Sci. Rep.* **2020**, *10*, 7364.

(7) Kattnig, D. R.; Nielsen, C.; Solov'yov, I. A. Molecular Dynamics Simulations Disclose Early Stages of the Photo-Activation of Cryptochrome 4. *New J. Phys.* **2018**, *20*, 083018.

(8) Günther, A.; Einwich, A.; Sjulstok, E.; Feederle, R.; Bolte, P.; Koch, K. W.; Solov'yov, I. A.; Mouritsen, H. Double-Cone Localization and Seasonal Expression Pattern Suggest a Role in Magnetoreception for European Robin Cryptochrome 4. *Curr. Biol.* **2018**, *28*, 211–223.e4.

(9) Einwich, A.; Dedek, K.; Seth, P. K.; Laubinger, S.; Mouritsen, H. A Novel Isoform of Cryptochrome 4 (Cry4b) is Expressed in the Retina of a Night-Migratory Songbird. *Sci. Reps.* **2020**, *4*, 15794.

(10) Zoltowski, B. D.; Chelliah, Y.; Wickramaratne, A.; Jarocha, L.; Karki, N.; Xu, W.; Mouritsen, H.; Hore, P. J.; Hibbs, R. E.; Green, C. B. et al. Chemical and Structural

Analysis of a Photoactive Vertebrate Cryptochrome from Pigeon. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 19449–19457.

(11) Sjulstok, E.; Solov'yov, I. A. Structural Explanations of Flavin Adenine Dinucleotide Binding in Drosophila melanogaster Cryptochrome. *J. Phys. Chem. Lett.* **2020**, *11*, 3866–3870.

(12) Kattnig, D. R.; Solov'yov, I. A.; Hore, P. J. Electron Spin Relaxation in Cryptochrome-Based Magnetoreception. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12443–12456.

(13) Solov'yov, I. A.; Domratcheva, T.; Schulten, K. Separation of Photo-Induced Radical Pair in Cryptochrome to a Functionally Critical Distance. *Sci. Rep.* **2014**, *4*, 3845.

(14) Solov'yov, I. A.; Mouritsen, H.; Schulten, K. Acuity of a Cryptochrome and Vision-Based Magnetoreception System in Birds. *Biophys. J.* **2010**, *99*, 40–49.

(15) Solov'yov, I. A.; Chandler, D. E.; Schulten, K. Magnetic Field Effects in Arabidopsis thaliana Cryptochrome-1. *Biophys. J.* **2007**, *92*, 2711–2726.

(16) Solov'yov, I. A.; Schulten, K. Reaction Kinetics and Mechanism of Magnetic Field Effects in Cryptochrome. *J. Phys. Chem. B* **2012**, *116*, 1089–1099.

(17) Sjulstok, E.; Olsen, J. M. H.; Solov'yov, I. A. Quantifying Electron Transfer Reactions in Biological Systems: What Interactions Play the Major Role? *Sci. Rep.* **2015**, *5*, 18446.

(18) Wong, S. Y.; Frederiksen, A.; Hanić, M.; Schuhmann, F.; Grüning, G.; Hore, P. J.; Solov'yov, I. A. Navigation of migratory songbirds: a quantum magnetic compass sensor. *Neuroforum* **2021**, 000010151520210005.

(19) Sjulstok, E.; Lüdemann, G.; Kubař, T.; Elstner, M.; Solov'yov, I. A. Molecular Insights into Variable Electron Transfer in Amphibian Cryptochrome. *Biophys. J.* **2018**, *114*, 2563–2572.

(20) Lüdemann, G.; Solov'yov, I. A.; Kubař, T.; Elstner, M. Solvent Driving Force Ensures Fast Formation of a Persistent and Well-Separated Radical Pair in Plant Cryptochrome. *J. Am. Chem. Soc.* **2015**, *137*, 1147–1156.

(21) Rodgers, C. T.; Hore, P. J. Chemical Magnetoreception in Birds: The Radical Pair Mechanism. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 353–360.

(22) Kattnig, D. R.; Sowa, J. K.; Solov'yov, I. A.; Hore, P. J. Electron Spin Relaxation Can Enhance the Performance of a Cryptochrome-Based Magnetic Compass Sensor. *New J. Phys.* **2016**, *18*, 063007.

(23) Nangle, S.; Xing, W.; Zheng, N. Crystal structure of mammalian cryptochrome in complex with a small molecule competitor of its ubiquitin ligase. *Cell Res.* **2013**, *23*, 1417–1419.

(24) Foloppe, N.; MacKerell, A. D. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86–104.

(25) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain $\chi1$ and $\chi2$ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(26) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348–362.

(27) Pavelites, J. J.; Gao, J.; Bash, P. A. A Molecular Mechanics Force Field for NAD+ , NADH, and the Pyrophosphate Groups of Nucleotides. *J. Comput. Chem.* **1996**, *18*, 221–239.

(28) MacKerell, A. D.; Banavali, N. K. All-Atom Empirical Force Field for Nucleic Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution. *J. Comput. Chem.* **2000**, *21*, 105–120.

(29) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; MacKerell, A. D. Impact of 20-Hydroxyl Sampling on the Conformational Properties of RNA: Update of the CHARMM All-Atom Additive Force Field for RNA. *J. Comput. Chem.* **2011**, *32*, 1929–1943.

(30) MacKerell, A. D.; Feig, M.; Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.

(31) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. et al. All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(32) Solov'yov, I. A.; Domratcheva, T.; Moughal Shahi, A. R.; Schulten, K. Decrypting Cryptochrome: Revealing the Molecular Identity of the Photoactivation Reaction. *J. Am. Chem. Soc.* **2012**, *134*, 18046–18052.

(33) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(34) Phillips, J. C.; Hardy, D. J.; Maia, J. D.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W. et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 44130.

(35) Korol, V.; Husen, P.; Sjulstok, E.; Nielsen, C.; Friis, I.; Frederiksen, A.; Salo, A. B.; Solov'yov, I. A. Introducing VIKING: A Novel Online Platform for Multiscale Modeling. *ACS Omega* **2020**, *5*, 1254–1260.

(36) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(37) Ryckaert, J.-P. P.; Ciccotti, G.; Berendsen, H. J. C.; Ciccotti+, G.; Berendsen, H. J. C.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(38) Schuhmann, F.; Korol, V.; Solov'yov, I. A. Introducing Pep McConst — A user-friendly peptide modeler for biophysical applications. *J. Comput. Chem.* **2021**, *42*, 572–580.

(39) Gao, P.; Yoo, S. H.; Lee, K. J.; Rosensweig, C.; Takahashi, J. S.; Chen, B. P.; Green, C. B. Phosphorylation of the Cryptochrome 1 C-terminal Tail Regulates Circadian Period Length. *J. Biol. Chem.* **2013**, *288*, 35277–35286.

(40) Müller, P.; Bouly, J. P. Searching for the mechanism of signalling by plant photoreceptor cryptochrome. *FEBS Lett.* **2015**, *589*, 189–192.

(41) Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and Structure-based Prediction of Eukaryotic Protein Phosphorylation Sites. *J. Mol. Biol.* **1999**, *294*, 1351–1362.

(42) Lüdemann, G.; Solov'yov, I. A.; Kubař, T.; Elstner, M. Solvent driving force ensures fast formation of a persistent and well-separated radical pair in plant cryptochrome. *J. Am. Chem. Soc.* **2015**, *137*, 1147–1156.

Graphical TOC Entry