Signal amplification and transduction in phytochrome photosensors

Heikki Takala[1,2*], Alexander Björling[2,*], Oskar Berntsson[2], Heli Lehtivuori[1], Stephan Niebling[2], Maria Hoernke[2], Irina Kosheleva[4], Robert Henning[4], Andreas Menzel[3], Janne A. Ihalainen[1], Sebastian Westenhoff[2]

[*]these authors contributed equally

Corresponding author: janne.ihalainen@jyu.fi, westenho@chem.gu.se

1. *Nanoscience Center, Department of Biological and Environmental Science, University of Jyväskylä, 40014 Jyväskylä, Finland*
2. *Department of Chemistry and Molecular Biology, University of Gothenburg, 40530 Gothenburg, Sweden*
3. *Paul Scherrer Institut, 5232 Villigen PSI, Switzerland*
4. *Center for Advanced Radiation Sources, The University of Chicago, IL 60637, USA*

Sensory proteins must relay structural signals from the sensory site over large distances to regulatory output domains. Phytochromes are a major family of red-light sensing kinases that control diverse cellular functions in plants, bacteria, and fungi.[1-9] Bacterial phytochromes consist of a photosensory core and a C-terminal regulatory domain.[10,11] Structures of photosensory cores are reported in the resting state[12-18] and conformational responses to light activation have been proposed in the vicinity of the chromophore.[19-23] However, the structure of the signalling state and the mechanism of downstream signal relay through the photosensory core remain elusive. Here, we report crystal and solution structures of the resting and active states of the photosensory core of the bacteriophytochrome from *Deinococcus radiodurans*. The structures reveal an open and closed form of the dimeric protein for the signalling and resting state, respectively. This nanometre scale rearrangement is controlled by refolding of an evolutionarily conserved "tongue", which is in contact with the chromophore. The findings reveal an unusual mechanism where atomic scale conformational changes around the chromophore are first amplified into an Ångström scale distance change in the tongue, and further grow into a nanometre scale conformational signal. The structural mechanism is a blueprint for understanding how the sensor proteins connect to the cellular signalling network.

The most common domain architecture of the photosensory core of phytochromes in bacteria, plants and fungi is PAS-GAF-PHY (Per/Arndt/Sim-cGMP phosphodiesterase/adenyl cyclase/FhlA-phytochromespecific).[10] We use the bacteriophytochrome from *D. radiodurans* to investigate the structural dynamics of photoswitching in these domains. Fig. 1a shows that the PAS-GAF-PHY domains share the prototypical response of phytochromes to red/far-red illumination. Its biliverdin chromophore relaxes into the Pr (red-absorbing) state when kept in the dark or after exposure to far-red light. The majority of molecules transform into Pfr (far-red absorbing) state after red light exposure. The dark relaxation from Pfr to Pr occurs in a matter of days, and the optical response of PAS-GAF-PHY is very similar to the full-length protein.[24] For the PAS-GAF construct, which lacks the PHY domain, the absorption spectra (Fig. 1a) and dark reversion within minutes indicate an incomplete photocycle.[24]

Time-resolved solution X-ray scattering reports on structural changes in protein reactions.[25,26] We used it here to characterize the structural dynamics of the phytochrome photocycle. In a first experiment, the phytochrome fragments were repeatedly switched between the Pr and Pfr states using laser flashes of 10 ms duration at 671 nm (Pr → Pfr) and 750 nm (Pfr → Pr). X-ray scattering intensities were recorded in between the laser flashes for approximately 1 s and exposures after far-red flashes were subtracted from those of exposures after red flashes (see Supplementary Information for details). These difference scattering curves, $\Delta S(q)$, encode the structural change of the protein during the Pr → Pfr transition. The recorded difference scattering from the complete photosensory core (PAS-GAF-PHY) shows a large oscillating signal at low angles $q < 2$ nm^{-1} (Fig. 1b, red line), indicating large structural changes on the nanometre scale. Oscillations at higher q, which report on structural changes on smaller length scales, were also observed (Extended Data Figure 1e), but are not discussed further here. The laser-induced ΔS of the PAS-GAF-PHY construct was reproduced by a standard Small-Angle X-ray Scattering (SAXS) experiment with

pre-illuminated samples (Fig. 1b, blue line). Time-resolved difference X-ray scattering data covering the micro- and millisecond time scales also reproduce these features when photoconversion is complete (delay time: 30 ms, Fig. 1b, black line). These data further establish that the discussed low-*q* features grow in with a time-constant of 4.3 ms (Fig. 1b, Extended Data Figure 1d), which coincides with the formation of the Pfr state as measured by optical absorption of the chromophore (1.0 ms, Extended Data Figure 1d).

When performing the same experiments on the PHY-less photosensory core (PAS-GAF), the oscillatory signal at low-*q* is absent (Fig. 1c). The observed oscillations in the PAS-GAF domains are much smaller in amplitude and confined to higher *q* ranges (Fig. 1c and Extended Data Figure 1e), indicating rearrangements of amino acid side chains and smaller shifts of secondary structural elements.[26] These data show that quaternary structural changes occur only when the PHY domain is present. In the wild-type protein the PHY domain connects the chromophore binding domains PAS/GAF with the output kinase, making it likely that this movement is functionally relevant.

In order to examine the nature of these conformational changes and to find out how they arise, we crystallised the photosensory core in the dark (referred to as *dark*) and under periodic illumination at 655 nm (referred to as *illuminated*), which transforms a large fraction of the protein molecules into the Pfr state (Fig. 1a and Extended Data for details). Absorption spectra of the *dark* crystals indicate the Pr state, while the *illuminated* crystals are Pfr-enriched (Extended Data Figure 3c). Omit maps indicate a change in the conformation of the chromophore between the two crystal forms (Extended Data Figure 3b). Although the biliverdin conformation cannot be fit unambiguously to the electron density in the *illuminated* structure, the remainder of the electron density is homogeneous (Extended Data Figure 2b).

Our dark crystal structure was modelled from data up to 3.80 Å resolution (Fig. 2a). The observed domain arrangement is similar to previously reported structures of bacterial PAS-GAF and PAS-GAF-PHY protein fragments,[12-15,17] but the dimer arrangement differs to some of these structures.[13,17] The *illuminated* structure, modelled against data cut at 3.24 Å, shows dramatic differences compared to the *dark* structure (Fig. 2a). Firstly, the dimer adopts an open Y-like shape of the PHY domains. This is in contrast to all reported resting state structures, where the dimer is closed (Fig. 2a).[12,13,17,20] Secondly, the tongue of the PHY domains (residues 446-477) appears as an α-helix and a loop in the *illuminated* model and is a β-sheet in the *dark* model (Fig. 2b and Extended Data Figure 2). These two folds have been observed separately in prototypical cyanobacterial[13,17] and non-canonical bacterial phytochromes,[12,20] respectively (Extended Data Figure 4). Our structures establish that the refolding of the tongue occurs within the same bacterial phytochrome, and suggest that this refolding is associated with opening of the dimer. The change in fold causes the length of the tongue to vary between the *dark* and *illuminated* structures (Fig. 2b), which is key for understanding the role of the tongue in this signal relay.

Before detailing this mechanism, it is important to test how well the crystal models represent the solution structures of the Pr and Pfr states. To this end, we performed structural refinement against the solution difference X-ray scattering data (Fig. 1), which is a very sensitive indicator of conformational change. It has the advantage over absolute scattering profiles that experimental error and uncertainties arising from the choice of method for scattering calculations largely cancel out (Extended Data 5).

We generated candidate solution structures by running two sets of molecular dynamics (MD) simulations starting from the *dark* and *illuminated* crystal structures (see Supplementary Information for details). These are referred to as Pr and Pfr trajectories since they aim to model solution scattering data representing these states. Snapshots were recorded every 50 ps and the solution

X-ray scattering was calculated from each. All pairs of Pr/Pfr candidate structures were then compared to the static difference X-ray scattering (Fig. 1b, Extended Data Fig. 7, see Supplementary Information for details). We selected the 100 Pr/Pfr pairs with best agreement to the data and consider the participating structures to be representative of the protein in solution. Considering all pairwise differences between these structures, a consistent set of 747 curves was generated and evaluated (Fig. 3a). The agreement of the model and experiment is excellent, and striking improvement is made compared to the crystal structures.

In order to verify the choice of solution structures based on difference scattering, these structures were cross-validated against absolute, population-corrected SAXS data as described in Supplementary Information (see also Extended Data Figure 5). Fig. 3b shows that the agreement with absolute SAXS data is strongly correlated with the separation of the PHY domains, and that the solution structures, proposed purely on difference scattering, cluster in the low-error parts of these correlations. The agreement of the solution structures to the solution scattering is significantly improved compared to the crystal structures as can be seen in Fig. 3b. Representative structural models for Pr and Pfr states in solution are depicted in Fig. 4.

The solution-structural models show that the PHY domains of opposing monomers come fairly close to each other in the Pr state but move apart by ~3 nm during the Pr → Pfr transition (Fig. 4). The Y shape assumed by the solution structures in the Pfr state is in qualitative agreement with a low-resolution envelope of a *Rps. Palustris* phytochrome determined by SAXS.[27] Our crystal structures capture the essence of this quaternary conformational change, albeit with differences in the amplitude and details of the motion. These differences may be caused by crystal contacts (Extended Data Figure 6). Nevertheless, the qualitative agreement with the solution structures lends strong support to that the *dark* and *illuminated* crystal structures reflect the relevant structural rearrangement between Pr and Pfr states. Thus we consider the refolding of the

tongue observed in the crystal structures to be a direct consequence of illumination. From the crystal and solution structures the following mechanism for signal transduction through the photosensory core emerges. It is known that the rotation of the biliverdin D-ring causes atomic rearrangements in the chromophore binding pocket, including displacement of Asp207 and Tyr263.[19-22] Our crystal structures suggest that this controls the fold of the tongue of the PHY domain. Importantly, the tongue is shortened as a result of the refolding and the distance between the GAF and PHY domain is reduced by 2.5 Å (Fig. 2b) upon red-light illumination. As a consequence of this, and demonstrated by our Pr and Pfr solution structures, the dimer opens up between the PHY domains by several nanometres (Fig. 4).

Three factors are essential for the proposed structural mechanism. Firstly, the PHY and GAF/PAS domains, which are connected by the tongue, are known to be very rigid internally such that they cannot deform to absorb the length variation of the tongue.[11] Secondly, the tongue and its junctions to PHY and GAF have to be rigid, especially in Pr where the PHY and GAF domains are pushed away from each other. In our *dark* structures the direction of the tongue is rigidly fixed with respect to the GAF domains by the conserved interaction Arg466 to Asp207/Tyr263 and additional backbone hydrogen bonding (Fig. 2b).[13] In the *illuminated* structure the tongue binds to GAF through interaction of Ser468 with Asp207/Tyr263,[12] and through hydrophobic interactions, among others by Tyr472 (Extended Data Fig. 4b).[17] The tongue is also locked close to the PHY domain by additional backbone hydrogen bonding. Thirdly, the PAS/GAF and PHY domains are connected by the unusually long scaffolding helix, which is necessary to redirect the shortening of the tongue into bending of the monomer. Indeed, our simulations reveal a hinge in the scaffolding helix (Fig. 4b). All phytochrome structures published to date support these three requirements.[12-18]

Phytochromes support different modes of signal output. In bacterial and fungal phytochromes output is through a C-terminal histidine kinase domain which

autophosphorylates and further phosphotransfers to a response regulator.[3,7] In plant phytochromes two additional PAS domains are included in the C-terminal regulatory region and a more complex pattern of functions has to be controlled, such as serine/threonine kinase activity,[4] and affinity to interaction partners.[9,10] In all cases, the output activity is likely controlled by a structural change in the photosensory core. The photosensory core and the key amino acid sequence in the tongue region ^{465}PRxSF^{469} are highly conserved over the whole phytochrome superfamily (Extended Data Figure 2c). We suggest here a mechanism, in which the shortening of the PHY-tongue pulls together (Pfr) or pushes apart (Pr) the PHY and PAS/GAF domains. The scaffolding helix redirects this force so that the monomers bend strongly. For the isolated photosensory unit in this study, this results in dramatic opening of the dimer (Fig. 2 and Fig. 4). It is conceivable that full-length phytochromes modify the distance of output domains in a similar manner. Alternatively, previously buried patches of the protein could become accessible to interaction partners, or monomers could rearrange with respect to each other. It will be intriguing to see the emergence of these mechanisms, which become deductible with experimental approaches similar to the one presented here.

METHODS SUMMARY

The PAS-GAF-PHY and PAS-GAF domains from *Deinococcus radiodurans* were expressed in the *Escherichia coli* strain BL21 (DE3) and purified by affinity and size-exclusion chromatography. Crystallographic data was collected at beamline ID23-1 of the ESRF. Time-resolved X-ray scattering with millisecond time resolution were recorded at beamline cSAXS of the Swiss Light Source.[28] SAXS measurements were performed at beamline BM29 of the European Synchrotron Radiation Facility (ESRF). Time-resolved X-ray scattering data in the micro- and millisecond ranges were collected at beamline

ID-14-B, BioCARS, of the Advanced Photon Source at Argonne National Laboratory. Molecular dynamics simulations (GROMACS 4.5.5)[29] were used to generate trial solution structures and theoretical scattering curves were evaluated using Zernike expansion as implemented in SASTBX.[30] The structures were scored against experimental scattering data as detailed in Supplementary Information.

AUTHOR CONTRIBUTIONS

JI, SW, AB, and HT conceived the experiments, HT prepared the samples, crystallized the protein, solved the structure, and measured crystal spectra, AB, HT, SW, JI, SN, HL, MH, AM, RH, and IK performed the solution X-ray scattering experiments, HL performed the spectroscopic measurements in solution, AB performed the MD simulations, AB, HT, OB, and SN analysed data, SW and JI supervised all parts of the project, and SW, JI, HT, and AB wrote the paper with input from all authors.

ACKNOWLEDGEMENTS

We acknowledge beamline access at BM29 and ID23-1 at the ESRF, cSAXS at the Swiss Light Source, whereby funding was received from the European Community's FP7 grant agreement n.°312284 (CALIPSO), and BioCARS (supported by National Institutes of Health, National Institute of General Medical Sciences grant 1R24GM111072 and in part through collaboration with Philip Anfinrud) at the Advanced Photon Source, Argonne National Laboratory. K. Forest and R. Vierstra are acknowledged for the donation of the plasmids. Jari Ylänne and his group are acknowledged for advice on crystallography, and Heikki Häkkänen and Alli Liukkonen for their assistance. We thank Maria-Andrea Mroginski for kindly providing force-field parameter files for the

biliverdin. We acknowledge Vladimir Chukharev and Nikolai V. Tkachenko at the Department of Chemistry and Bioengineering, Tampere University of Technology for the facilities for transient absorption spectroscopy measurements. We also acknowledge grants to SW by the Foundation of Strategic Research, Sweden, and the Swedish and European Research Councils, support to JI by Finnish Academy grant 138063, and to HT by the Finnish Cultural Foundation grant 0131067.

DATA DEPOSITION:

Coordinates and structure factor amplitudes have been deposited in the RCSB Protein Data Bank, www.pdb.org (PDB ID codes 4O01 and 4OP1). The solution structures are available as source data.

COMPETING FINANCIAL INTEREST

The authors declare no competing financial interest.

SUPPLEMENTARY INFORMATION

Supplementary materials and methods, Source Data, including solution structures, graphical data in ASCII format, sequence alignment, and force-field parameters for the biliverdin chromophore.

FIGURE LEGENDS

Figure 1: Time-resolved solution X-ray scattering of the PAS-GAF and PAS-GAF-PHY fragments from *D. radiodurans*

a, Absorption spectra of protein fragments after illumination with far-red (780 nm) and red (655 nm) light, labelled Pr and Pfr respectively. **b** and **c**, Solution X-ray scattering data from the PAS-GAF-PHY and PAS-GAF fragments shown on the same scale. Time-resolved data (black, BioCARS), direct static data collected by laser-induced population switching (red, cSAXS), and indirect static data from a standard SAXS experiment with pre-illuminated samples (blue, BM29) is shown. ΔS is the difference in scattered X-ray intensity caused by illumination at 671 nm. $q = 4\pi/\lambda \cdot \sin\theta$ at wavelength λ and scattering angle 2θ.

Figure 2: Dark and illuminated crystal structures of PAS-GAF-PHY from *D. radiodurans*

a, Crystal structures of the PAS-GAF-PHY dimer in the *dark* and *illuminated* forms. The tongue of the PHY domain (green) changes fold and the dimer opens up in the illuminated state. The biliverdin chromophore is shown in orange. **b**, Fold and interactions of the PHY tongue. The β-sheet (*dark*) coordinates to Asp207 and Tyr263 via Arg466, whereas the α-helix (*illuminated*) coordinates via Ser468, both of which are part of the conserved ^{465}PRxSF^{469} motif. The named residues are shown as sticks. The β-sheet (*dark*) is further held by hydrogen bonding interactions between the amide groups of Ala450, Gly452 and Arg202. The change in PHY domain conformation leads to a shortening of the tongue by 2.5 Å as measured between GAF (Arg202) and PHY (Tyr479) domains (arrows). A backbone interaction close to the PHY domain between Leu445 and Tyr479, shared by both crystal structures, is also indicated. The green dashed lines indicate regions that are not modelled or not shown for clarity.

Figure 3: Refinement of solution structures against difference SAXS data

a, Calculated differences X-ray scattering between the proposed solution structures (Pfr-Pr, grey) agree with experimental data (black). As a safeguard against overfitting, the grey curves show all 747 differences between solution structures, not just the 100 best curves on which the pairs were selected (See text for details). The agreement with experiment is dramatically improved for the solution structures, compared to the difference scattering calculated from the two crystal forms (green). **b**, Validation of the obtained solution structures (squares) against absolute X-ray scattering. All MD snapshots (dots) and crystal structures (diamonds) are scored against absolute SAXS data (BM29) after correction for the mixed Pr/Pfr populations (see Extended Data Figure 5d). The two shades of red and blue correspond to different simulation conditions as detailed in Supplementary Information. The PHY domain separation is measured as the distance between the centers of mass of the two C-terminal helices (residues 484-503).

Figure 4: Proposed solution structures of the bacterial phytochrome from *D. radiodurans*

Representative solution structures for the Pr and Pfr states of the photosensory core, identified from solution X-ray scattering experiments. Nine Pfr and ten Pr structures are presented viewed along (**a**) and perpendicular (**b**) to the dimer symmetry axis. The long scaffolding helix is highlighted in blue, the PHY tongue in green, and the biliverdin chromophore in orange. The PHY domain separation differs by about 3 nm between the Pr and Pfr structures as shown in Figure 3b. The hinge region at the scaffolding helix (residues Val318 and Lys319) in Pfr is indicated with black arrowheads.

REFERENCES

1 Butler, W. L., Norris, K. H., Siegelman, H. W. & Hendricks, S. B. Detection, Assay, and Preliminary Purification of the Pigment Controlling Photoresponsive Development of Plants. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1703-1708, doi:DOI 10.1073/pnas.45.12.1703 (1959).

2 Kehoe, D. M. & Grossman, A. R. Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. *Science* **273**, 1409-1412, doi:DOI 10.1126/science.273.5280.1409 (1996).

3 Yeh, K. C., Wu, S. H., Murphy, J. T. & Lagarias, J. C. A cyanobacterial phytochrome two-component light sensory system. *Science* **277**, 1505-1508, doi:DOI 10.1126/science.277.5331.1505 (1997).

4 Yeh, K.-C. & Lagarias, J. C. Eukaryotic phytochromes: Light-regulated serine/threonine protein kinases with histidine kinase ancestry. *Proceedings of the National Academy of Sciences* **95**, 13976-13981, doi:10.1073/pnas.95.23.13976 (1998).

5 Jiang, Z. Y. *et al.* Bacterial photoreceptor with similarity to photoactive yellow protein and plant phytochromes. *Science* **285**, 406-409, doi:DOI 10.1126/science.285.5426.406 (1999).

6 Hughes, J. *et al.* A prokaryotic phytochrome. *Nature* **386**, 663-663, doi:Doi 10.1038/386663a0 (1997).

7 Bhoo, S. H., Davis, S. J., Walker, J., Karniol, B. & Vierstra, R. D. Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature* **414**, 776-779, doi:Doi 10.1038/414776a (2001).

8 Blumenstein, A. *et al.* The Aspergillus nidulans phytochrome FphA represses sexual development in red light. *Curr Biol* **15**, 1833-1838, doi:DOI 10.1016/j.cub.2005.08.061 (2005).

9 Ni, M., Tepperman, J. M. & Quail, P. H. Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature* **400**, 781-784 (1999).

10 Rockwell, N. C., Su, Y. S. & Lagarias, J. C. Phytochrome structure and signaling mechanisms. *Annual Review of Plant Biology* **57**, 837-858, doi:DOI 10.1146/annurev.arplant.56.032604.144208 (2006).

11 Moglich, A., Yang, X. J., Ayers, R. A. & Moffat, K. Structure and Function of Plant Photoreceptors. *Annu Rev Plant Biol* **61**, 21-47, doi:DOI 10.1146/annurev-arplant-042809-112259 (2010).

12 Yang, X., Kuk, J. & Moffat, K. Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: Photoconversion and signal transduction. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14715-14720, doi:DOI 10.1073/pnas.0806718105 (2008).

13 Essen, L. O., Mailliet, J. & Hughes, J. The structure of a complete phytochrome sensory module in the Pr ground state. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14709-14714, doi:DOI 10.1073/pnas.0806477105 (2008).

14 Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325-331, doi:Doi 10.1038/Nature04118 (2005).

15 Li, H., Zhang, J. R., Vierstra, R. D. & Li, H. L. Quaternary organization of a phytochrome dimer as revealed by cryoelectron microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10872-10877, doi:DOI 10.1073/pnas.1001908107 (2010).

16 Cornilescu, G., Ulijasz, A. T., Cornilescu, C. C., Markley, J. L. & Vierstra, R. D. Solution Structure of a Cyanobacterial Phytochrome GAF Domain in the Red-Light-Absorbing Ground State. *J. Mol. Biol.* **383**, 403-413, doi:DOI 10.1016/j.jmb.2008.08.034 (2008).

17 Anders, K., Daminelli-Widany, G., Mroginski, M. A., von Stetten, D. & Essen, L. O. Structure of the Cyanobacterial Phytochrome 2 Photosensor Implies a Tryptophan Switch for Phytochrome Signaling. *Journal of Biological Chemistry* **288**, 35714-35725, doi:DOI 10.1074/jbc.M113.510461 (2013).
18 Narikawa, R. *et al.* Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 918-923, doi:DOI 10.1073/pnas.1212098110 (2013).
19 Yang, X. J., Kuk, J. & Moffat, K. Conformational differences between the Pfr and Pr states in Pseudomonas aeruginosa bacteriophytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15639-15644, doi:DOI 10.1073/pnas.0902178106 (2009).
20 Yang, X. J., Ren, Z., Kuk, J. & Moffat, K. Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature* **479**, 428-U190, doi:Doi 10.1038/Nature10506 (2011).
21 Song, C. *et al.* Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3842-3847, doi:DOI 10.1073/pnas.1013377108 (2011).
22 Burgie, E. S., Walker, J. M., Phillips, G. N. & Vierstra, R. D. A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes. *Structure* **21**, 88-97, doi:DOI 10.1016/j.str.2012.11.001 (2013).
23 Ulijasz, A. T. *et al.* Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature* **463**, 250-U143, doi:Doi 10.1038/Nature08671 (2010).
24 Wagner, J. R. *et al.* Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *Journal of Biological Chemistry* **283**, 12212-12226, doi:DOI 10.1074/jbc.M709355200 (2008).
25 Cammarata, M. *et al.* Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nature Methods* **5**, 881-886, doi:10.1038/nmeth.1255 (2008).
26 Andersson, M. *et al.* Structural Dynamics of Light-Driven Proton Pumps. *Structure* **17**, 1265-1275, doi:10.1016/j.str.2009.07.007 (2009).
27 Evans, K., Grossmann, J. G., Fordham-Skelton, A. P. & Papiz, M. Z. Small-angle x-ray scattering reveals the solution structure of a bacteriophytochrome in the catalytically active Pr state. *J. Mol. Biol.* **364**, 655-666, doi:DOI 10.1016/j.jmb.2006.09.045 (2006).
28 Westenhoff, S. *et al.* Rapid readout detector captures protein time-resolved WAXS. *Nature Methods* **7**, 775-776 (2010).
29 Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845-854 (2013).
30 Liu, H. G., Hexemer, A. & Zwart, P. H. The Small Angle Scattering ToolBox (SASTBX): an open-source software for biomolecular small-angle scattering. *J Appl Crystallogr* **45**, 587-593, doi:Doi 10.1107/S0021889812015786 (2012).
10.1074/jbc.M611824200 (2007).

Extended Data Figure Legends

Extended Data Figure 1 | Difference scattering and kinetics. a-c, Singular Value Decomposition ($\Delta S(q,t) = U \cdot S \cdot V^T$) of time-resolved solution scattering data from

PAS-GAF-PHY. Two components suffice to describe the data, the final product ($n = 1$) and a transient low-q depression ($n = 2$). **a**, The first three basis spectra (1st, 2nd and 3rd columns of $U \cdot S$), and original X-ray scattering data (black) with reconstruction based on the first two singular values (all columns of $U \cdot S \cdot V^T$, red) are shown. **b**, Relative amplitudes of the two first basis spectra (1st and 2nd columns of V). Lines are guides for the eye. **c**, Singular values (diagonal elements of S). **d,** The rise of the Pfr product state as measured by direct integration of difference scattering curves ($< \Delta S(q)^2 >_{1.2<=q<=2.5}$) and by absorption spectroscopy (Abs. at 754 nm). Solid curve is the function $(1 - e^{-t/\tau})$. These data establish that the structural change occurs just after the Pfr state is formed in the chromophore. Note the positive signal in the absorption curve with very small amplitude at >3 ms, which appears to decay while the structural signal rises. This could be because the absorption properties depend weakly on the large-scale rearrangement. **e**, Direct static difference data from Figure 1, amplified by q^2 to reveal wide-angle oscillations.

Extended Data Figure 2 | Light-induced changes in the secondary structure of the evolutionally conserved PHY tongue. a, Secondary structure and topology of the *Deinococcus radiodurans* PAS-GAF-PHY construct. The structural elements in our crystal structures are very similar to other published phytochrome structures [12,13]. The PHY tongue region (box), however, was found to refold upon illumination. The five-stranded ß-sheet core of the GAF domain is extended by a small sixth β-strand (called 2') that interacts with the PHY tongue (see Fig. 2, Extended Data Figure 4). The mini-sheet structure at the knot region [14] is not included in the graph. The figure is adapted from [12]. **b**, Omit map of the PHY tongue in the dark crystal form (upper panel) and the illuminated crystal form (lower panel). In the dark crystal form, the omit map density (blue) supports the built ß-turn secondary structure (orange sticks), even though most of the side chains are poorly resolved. In the illuminated crystal form, the omit map (blue) clearly reveals the density of a helix with its bulky side chains (orange). The omit maps were calculated by repeating molecular replacement and a refinement step (see Supplementary Information) with a structure where the PHY tongue was removed. All electron density maps are contoured at sigma level 3.0. **c**, Sequence alignment of part of the GAF domain and of PHY loop region. The conserved ^{207}DIP^{209} motif in the GAF domain and ^{465}PRxSF^{469} motif in the PHY tongue are marked by asterisks (*). Five

representatives from eubacterial (BphP), cyanobacterial (Cph), higher plant, fungi (Fph) and PAS-less phytochromes are shown. Colour coding: dark grey - identical residues; light grey - similar residues. The amino acid sequences were aligned with Clustal Omega using the default settings.[31] Sequences include *Deinococcus radiodurans* BphP, *Pseudomonas syringae* pv. tomato T1 BphP, *Rhodopseudomonas palustris* TIE-1 BphP3, *Pseudomonas aeruginosa* PAO1 BphP, *Agrobacterium fabrum* str. *C58* BphP1 (Agp1), *Synechocystis* sp. PCC6803 Syn-Cph1, *Microcystis aeruginosa* NIES-843, *Nodularia spumigena* CCY9414, *Cyanothece* sp. PCC 7822, *Anabaena variabilis* ATCC 29413, *Physcomitrella patens* Phy1, Zea mays PhyB1, *Populus trichocarpa* PhyA, *Selaginella martensii* Phy1, *Arabidopsis thaliana* PhyA, *Synechococcus* OSB' SyB-Cph1, *Synechococcus* OSA SyA-Cph1, *Nostoc punctiforme* PCC73102, *Lyngbya* sp. PCC 8106, *Anabaena variabilis* ATCC 29413.

Extended Data Figure 3 | Biliverdin structure and spectra in crystals, and crystal fluorescence. a, Photographs of the crystals under cryogenic conditions at the beamline ID23-1. **b**, Biliverdin omit maps of the dark (upper left) and illuminated (upper right) form support the existence of the modelled biliverdin conformations (yellow and orange). Comparison of the electron density around the biliverdin with published structures (lower panels). In the dark form, the electron density indicates a conformation similar to the published Pr structures[13,14,32,33], including a *D. radiodurans* structure (2O9C, cyan).[32] Therefore we modelled the biliverdin as a 15*Z*a isomer. In the illuminated form, however, the electron density supports neither the biliverdin as determined in the Pfr structure of *Pa*BphP (3NHQ, red),[20] nor as determined in the Pr structure of *D. radiodurans* (2O9C, cyan).[32] Therefore the rotation of the biliverdin D-ring cannot be reliably determined and is modelled with both possibilities (15*Z*a, and 15*E*a, orange and yellow in upper right panel). Omit maps were calculated as in Extended Data Figure 2 and contoured at a sigma level of 3.0. **c**, Representative absorption spectra of the dark (black) and illuminated (grey) crystals, recorded at 123K. Note that the terms “illuminated” and “dark” refer here to the crystallization conditions (see Supplementary Information for details). The dark crystal spectrum resembles the Pr spectrum in solution (Fig. 1a). Illumination with red light in the crystallization drops at ambient temperature led to a slight increase of far-red absorption and disintegration of the crystals (data not shown). Dark crystals were unaffected by far-red illumination. The

spectrum of the illuminated crystals shows that a substantial proportion (>50%) of the proteins reside in Pfr state. A similar absorption was detected from solubilized crystals. The illuminated crystals could be switched to Pr-like absorption with far-red illumination. Reversely, also the Pfr-like features could be increased with red light (data not shown) with illumination at ambient temperature. Exposure with light increased the scattering background in the absorption measurements. The crystals seemed unaffected by the illumination when illuminated with red light in the crystallization drops. Although the spectral analyses of the illuminated crystals do not indicate a pure Pfr spectrum, and the biliverdin conformation cannot be fit unambiguously to the electron density, the remainder of the electron density is homogeneous (Extended Data Figure 2b). Most importantly, the tongue region of the PHY domain adopts the conformation resembling the Pfr state of *Pa*BphP (Extended Data Figure 4b)[12,19,20]. The conformations of the four monomers in an asymmetric unit are practically identical and hence we conclude that biliverdin can co-exist in both Pr and Pfr states inside this crystal form and still the protein part represents the structural aspects of the Pfr state only.

Extended Data Figure 4 | Comparison of PHY tongue interactions with published structures. a, Comparison of the dark crystal form (*green*/*dark grey*) to cyanobacterial *Cph1* in Pr state (PDB code 2VEA, *orange*/*light grey*).[13] **b**, Comparison of the illuminated crystal form (*green*/*dark grey*) to *Pa*BphP in Pfr state (PDB code 3NHQ[20], *orange*/*light grey*). In both the Pr and Pfr forms, key interactions are conserved between the phytochromes (*black dashes*), as well as the positions of three conserved tongue motives (see Extended Data Figure 2c). The residues of these three motives are indicated as: ^{451}W(G/A)G^{453}, ^{465}PRxSF^{469}, and ^{472}(W/F,Y)xE^{474},[13] with numbering from the *D. radiodurans* sequence. Trp451 was not modelled in our illuminated crystal structure, and part of the PHY tongue has been removed for clarity. Small changes in relative orientations between the difference crystal structures are observed, *e.g.* a slight tilt of helix of the Pfr tongue.

Extended Data Figure 5 | SAXS data and calculations. a, Experimental SAXS data of dark ("Pr") and pre-illuminated ("Pfr") samples. The data is merged from the concentration series (Extended Data Table 2b) and normalized on $0.4\,\mathrm{nm}^{-1} < q <$

$0.6\ \mathrm{nm}^{-1}$. **b**, Guinier plot of the low-q region, shown for all concentrations. The curves converge at low concentration (C). Inset shows the radii of gyration (Rg) calculated from the curves in **(a)** according to the Guinier approximation. Dotted lines show the q-range used for linear regression ($0.15\ \mathrm{nm}^{-2} < q^2 < 0.120\ \mathrm{nm}^{-2}$), which is such that $q \cdot R_g < 1.3$ as required.[34] Arrows indicate increasing concentration. **c,** Average difference scattering signals calculated from the solution-structural models using three methods: Crysol (default settings), SASTBX with spherical harmonic expansion (SHE, default settings), and SASTBX with Zernike polynomial expansion as described in Supplementary Information.[30] The choice of calculation method does not significantly change the predicted X-ray difference scattering. **d,** Determination of the relative Pr/Pfr populations represented by the BM29 data as described in Supplementary Information. We find that our Pr sample contained only Pr (top) whereas in the Pfr sample 64% of the protein molecules adopted the Pfr conformation (bottom). Notably the static difference scattering patterns (Fig. 1) represent, up to a scaling factor, the relation between pure Pr and Pfr populations. This is in contrast to traditional SAXS which report on population mixtures, because the Pfr state cannot be easily produced with 100% population in solution.

Extended Data Figure 6 | Packing interactions of the crystal forms. Crystal packing interactions of the (**a**) dark and (**b**) illuminated crystal forms. The dimer of an asymmetric unit is shown in red and the symmetry mates in grey. For the illuminated form, only chains A and B are shown in red. Interactions are shown from two orientations for clarity. In the dark form, crystal contacts are seen in the top regions of the PHY domains and therefore may cause artefacts in the long scaffolding helix and in the opening of the PHY domains. In the illuminated form contacts are such that the PHY domains may be pushed closer together, which is consistent with the larger separation of the PHY domains as refined from the solution X-ray scattering data. It is noteworthy that the relative orientation of the monomers in the dimer is different between all three known structures for PAS-GAF-PHY phytochromes. For *Pseudomonas aeruginosa* the dimer is parallel with variations in between different copies of the dimer in the crystallographic unit cell[12], in two cyanobacterial phytochrome an antiparallel dimer is observed[13,35], and in our Pr structure, the monomer have an angle of approximately 45º.

Extended Data Figure 7 | Solution-structural refinement. a, The distribution of PHY domain separations (R_{PP}) obtained from unbiased MD simulations (production runs 1-3). The sets *A* and *B* of solution structures suggested in Fig. 3 are indicated by crosses (with $N = 100$, $M \approx 6 \cdot 10^6$). While the Pr structures cluster in a region of high sampling, the Pfr structures lie at the edges of the PHY-PHY distribution, suggesting inadequate sampling. To remedy this, we artificially scanned the PHY domain separation in separate simulations (production runs 4 and 5) to improve sampling. **b**, The new distribution of pulled PHY domain separations in the Pfr state. The final analysis and all solution-structural conclusions drawn in this study are based on the trajectories described in **b**.
c-e, Consistency test of the structural refinement procedure. **c**, A cutoff parameter R_{cut} was introduced to reject all MD frames $R_{PP} < R_{cut}$. The resulting average over R_{PP} of the best N = 100 pairs is plotted as a function of R_{cut}. It is found that $R_{PP} \approx R_{cut}$, which indicates that the best fit to the difference X-ray scattering data is always at the highest separations available in sampling range. **d** and **e** show the dependence of the total and average error as a function of R_{cut}, respectively. It is observed that the error decreases steeply for $R_{cut} \lesssim 5$ nm, and only marginally for $R_{cut} \gtrsim 5$ nm. We therefore consider optimization in the latter range overfitting, and applied $R_{cut} = 5.0$ nm in the refinement for the solution structures. The refined structures (Fig. 4) should therefore be considered to represent lower limits on R_{PP}. **f**, Solution-structural refinement parameters and results.

Extended Data Table 1 | Crystallographic data collection and refinement statistics

Extended Data Table 2 | SAXS statistics and sample details. a, Statistics of the static SAXS (BM29) data, including radii of gyration (R_g), maximum particle dimension (D_{max}), Porod volume (V_{porod}), forward scattering (I_0). The molecular weights (MM_{exp}) were estimated from the I_0 of BSA using the formula $M(sample) \approx M(BSA)*[I_0(sample)/I_0(\mathrm{BSA})]$, where $M(BSA) = 66$ kDa. Theoretical molecular weights (MM_{calc}) were calculated from the protein sequence[36]. **b,** Details of the samples used for solution scattering measurements.

Extended Data References

31 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, doi:http://www.nature.com/msb/journal/v7/n1/suppinfo/msb201175_S1.html (2011).

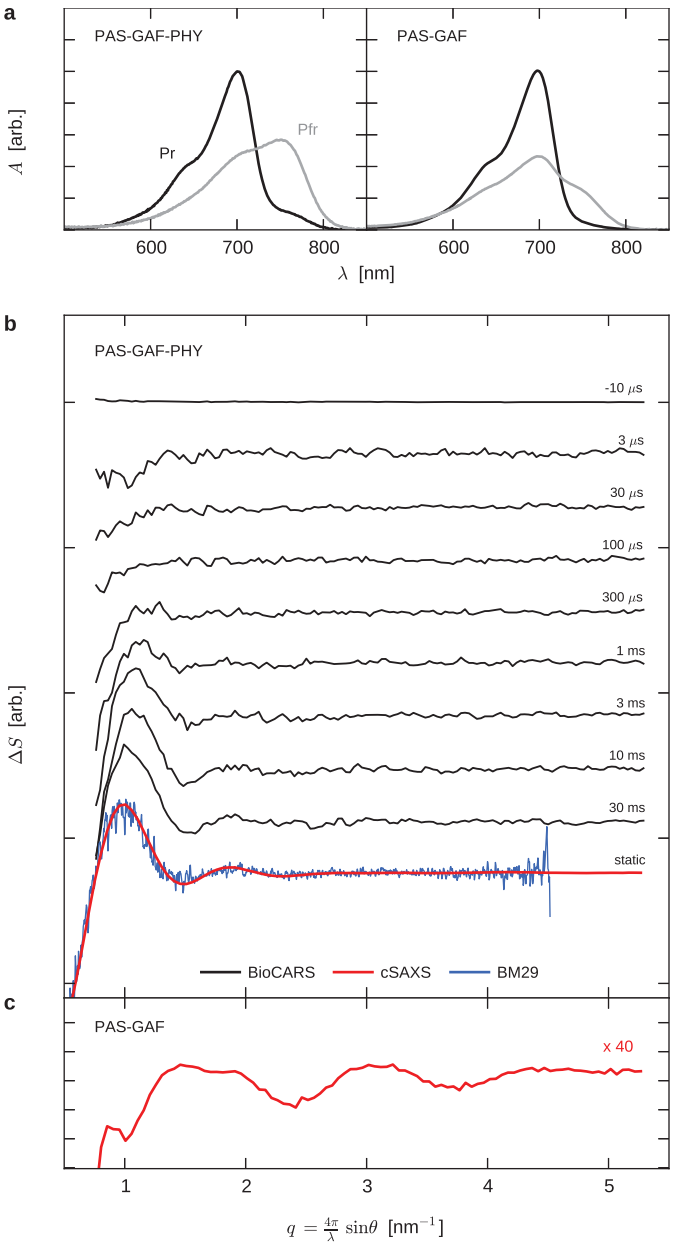
32 Wagner, J. R., Zhang, J. R., Brunzelle, J. S., Vierstra, R. D. & Forest, K. T. High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *Journal of Biological Chemistry* **282**, 12298-12309, doi:DOI 10.1074/jbc.M611824200 (2007).

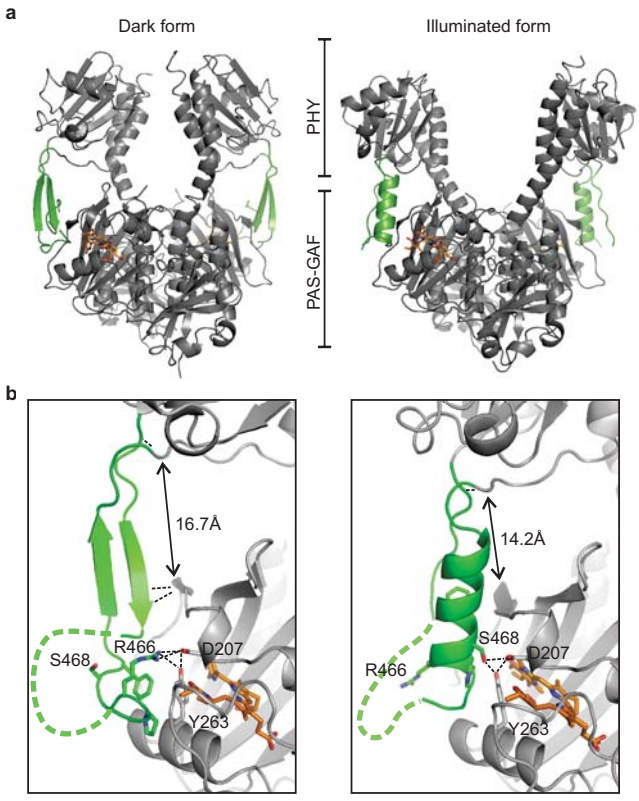
33 Mailliet, J. *et al.* Spectroscopy and a High-Resolution Crystal Structure of Tyr263 Mutants of Cyanobacterial Phytochrome Cph1. *J. Mol. Biol.* **413**, 115-127, doi:DOI 10.1016/j.jmb.2011.08.023 (2011).

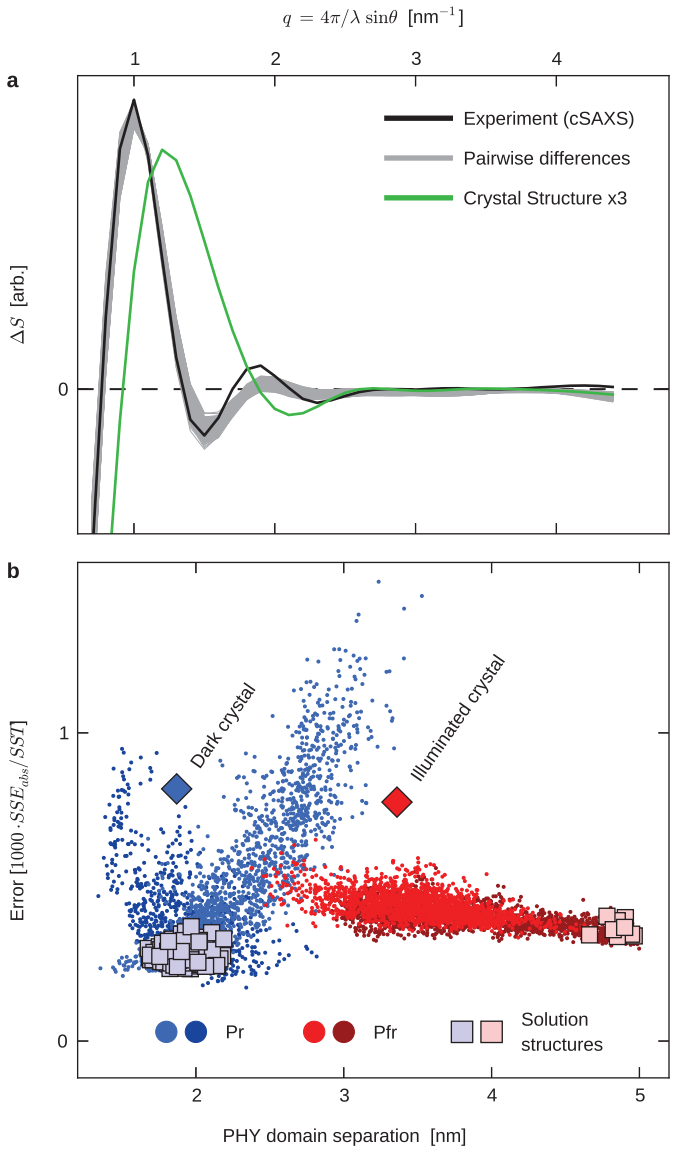
34 Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277-1282, doi:Doi 10.1107/S0021889803012779 (2003).

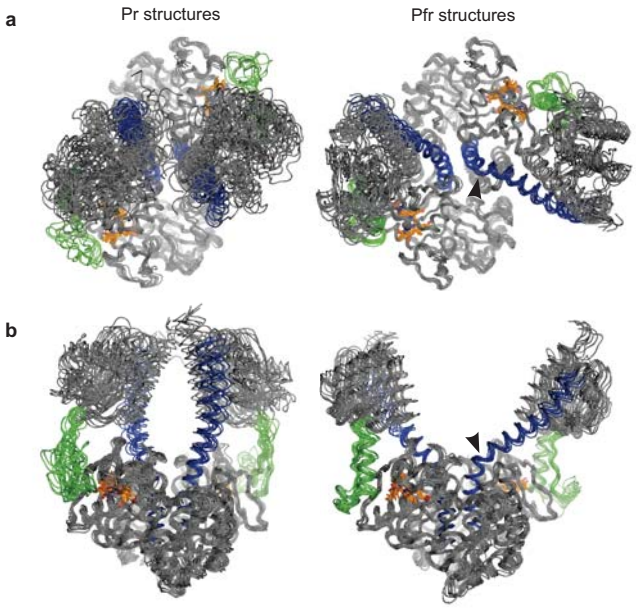
35 Anders, K., Daminelli-Widany, Mroginski, M. A., Von Stetten, D. & Essen, L. O. Structure of the cyanobacterial phytochrome 2 photosensor implies a tryptophan switch for phytochrome signaling. *The Journal of Biological Chemistry* **published online**, doi:10.1074/jbc.M113.510461 (2013).

36 Gasteiger, E. *et al.* in *The Proteomics Protocols Handbook* (ed JohnM Walker) Ch. 52, 571-607 (Humana Press, 2005).


a
PAS-GAF-PHY
PAS-GAF
Pr
Pfr
A [arb.]
600
700
800
λ [nm]
b
PAS-GAF-PHY
-10 μs
3 μs
30 μs
100 μs
300 μs
1 ms
3 ms
10 ms
30 ms
static
ΔS [arb.]
BioCARS
cSAXS
BM29
c
PAS-GAF
x 40
1
2
3
4
5
q = 4π/λ sinθ [nm⁻¹]


a
Dark form
Illuminated form
PHY
PAS-GAF
b
16.7Å
S468
R466
D207
Y263
14.2Å
S468
R466
D207
Y263


$q = 4\pi/\lambda \sin\theta$ [nm^{-1}]
1
2
3
4
a
Experiment (cSAXS)
Pairwise differences
Crystal Structure x3
ΔS [arb.]
0
b
Dark crystal
Illuminated crystal
Error [$1000 \cdot SSE_{abs}/SST$]
1
0
Pr
Pfr
Solution structures
2
3
4
5
PHY domain separation [nm]


a
Pr structures
Pfr structures
b