

## Supplemental Data

### Coordination and redox state-dependent structural changes of the heme-based oxygen sensor *AfGcHK* associated with intraprotein signal transduction

Martin Stranova, Petr Man, Tereza Skálová, Petr Kolenko, Jan Blaha, Veronika Fojtikova, Vaclav Martinek, Jan Dohnálek, Alzbeta Lengalova, Michal Rosůlek, Toru Shimizu and Markéta Martínková

#### Table of Contents

[1] Crystal structure of *AfGcHK* globin domain

[2] Structural comparison of the isolated globin domains of *AfGcHK* to other GCSs

**Table S1.** Autoxidation rates of the wild type full-length *AfGcHK* and mutants at Tyr45 and Leu68 obtained in the present study.

**Figure S1.** Schematic representation of the *AfGcHK* two-component signal transduction system.

**Figure S2.** Changes in absorption spectra (left) and absorption intensity at 575-580 nm (right) associated with the autoxidation of the Fe(II)-O<sub>2</sub> complex to the Fe(III) complex of the full-length *AfGcHK*.

**Figure S3.** Absorption spectra of the Fe(III) (black), Fe(II) (red), Fe(II)-CO (blue) and Fe(II)-O<sub>2</sub> (green) complexes of the full length *AfGcHK* wild type [A], Y45F [B], L68F [C], L68G [D], L68N [E], L68V [F].

**Figure S4.** Electron density maps supporting the existence of two alternative positions in chain G (heme, protein side chains and protein main chain).

**Figure S5.** HDX-MS protection plots showing the differences in deuteration for the full-length Fe(III)-OH-bound (active) and Fe(II)-bound (inactive) forms of *AfGcHK*.

**Figure S6.** Domain structures of *AfGcHK* showing active (A) and inactive (B) states.

**Figure S7.** An isolated kinase domain (specifically, a truncated form of *AfGcHK* lacking the N-terminal globin domain is removed) exhibits no autophosphorylation activity, suggesting that the globin domain is required for autophosphorylation of the kinase domain.

**Figure S8.** HDX-MS results demonstrating the differences between single isolated domains alone in solution and the mixture of the isolated globin and kinase domains.

## [1] Crystal structure of AfGcHK globin domain

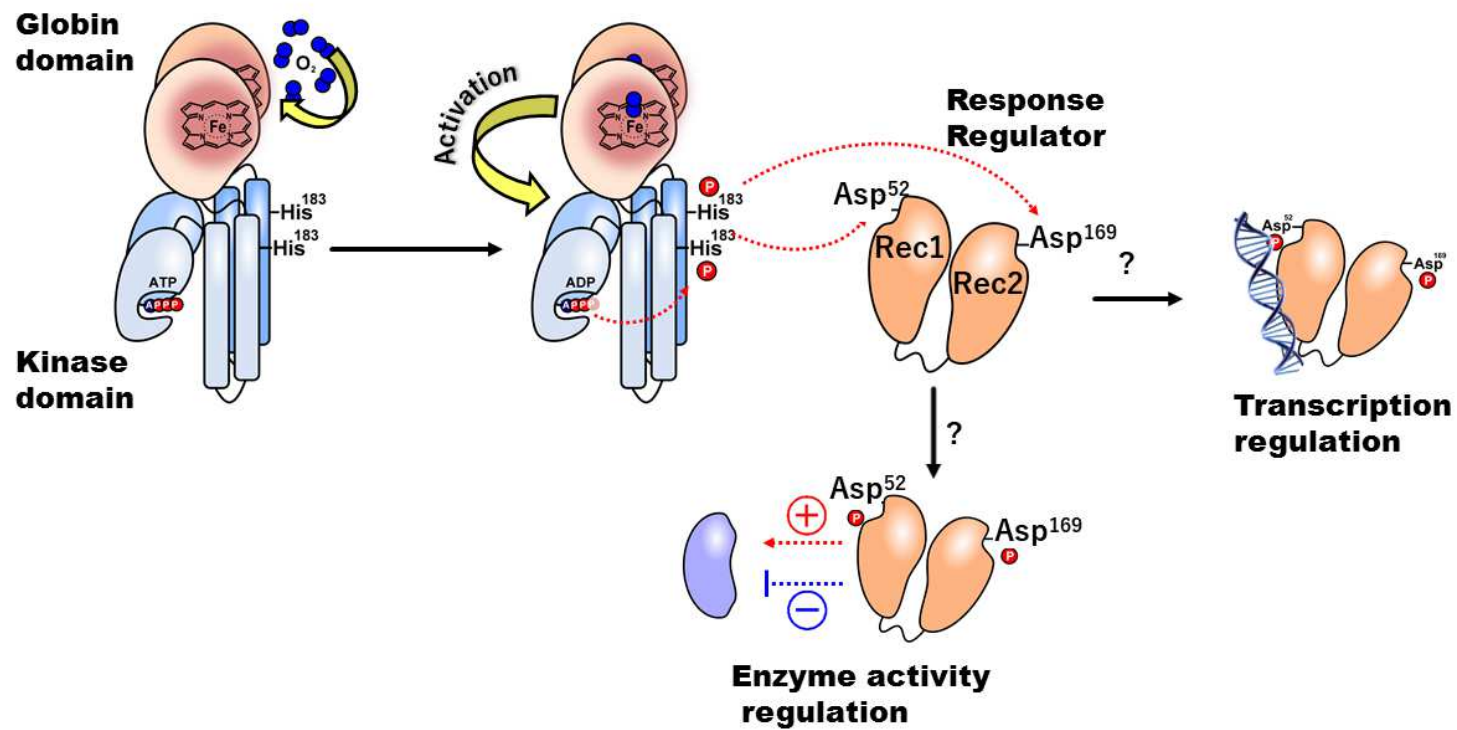
The two presented crystal structures were selected from a larger series of diffraction data sets for the crystallized globin domain of AfGcHK. Diffraction data sets were collected from 11 crystals: 3 non-soaked and 8 that were soaked in sodium dithionite for 10 - 65 min (for the longer soaks, the crystals were re-soaked at 15 min intervals with freshly diluted sodium dithionite). For the non-soaked crystals, the data set with the highest resolution was selected (with the other two providing the same structure). In the soaked crystals, the presence of alternatives in chain G was basically the same in all cases, but less defined with decreasing data resolution. Experiments with longer soaks always yielded two heme group positions; there was no case in which the second heme position was exclusively present. There was a weak trend towards an increase in the occupancy of the second iron position (as determined by anomalous occupancy) as the soak length increased, but this was accompanied by a gradual reduction in the resolution of the structure. The structure presented in the Results section is the one with the highest resolution, and was obtained after a relatively short 15 min soak; this structure was deemed most suitable for reliably building alternative conformations.

## [2] Structural comparison of the isolated globin domains of AfGcHK to other GCSs

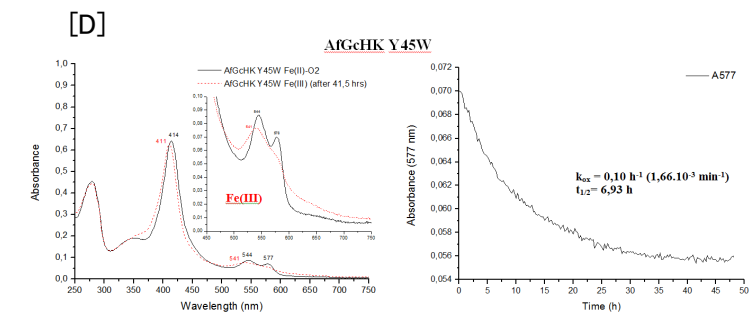
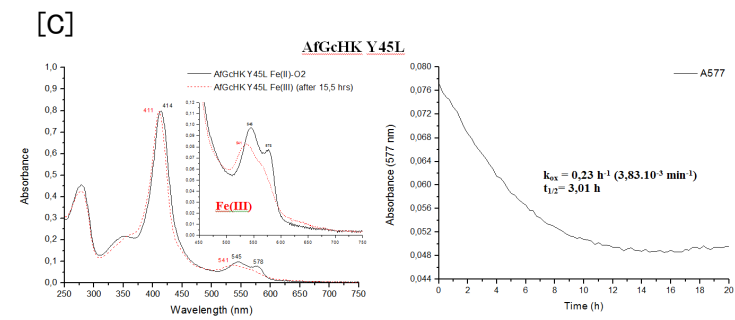
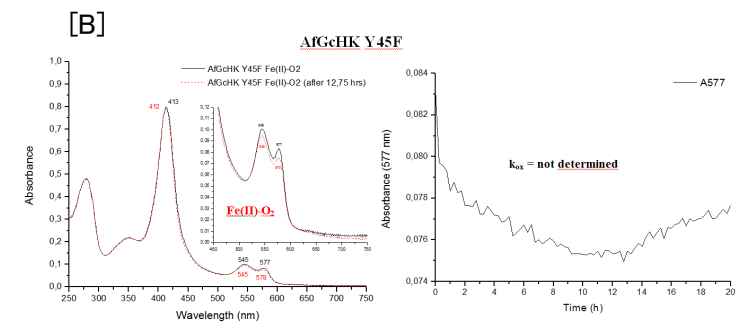
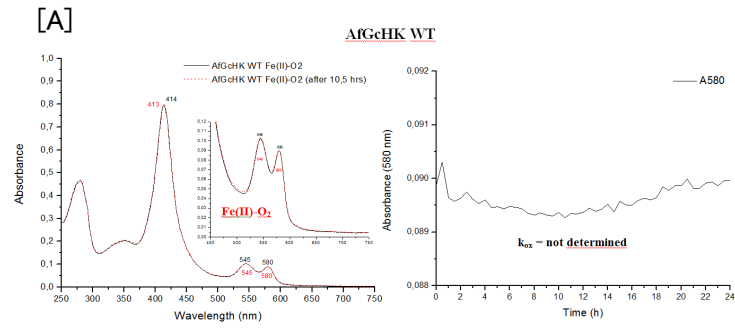
The AfGcHK globin domain structure closely resembles the structures of the globin domains from some other heme-based oxygen sensors. Closely related globin structures deposited in the PDB are the isolated globin domain of GsGCS from *Geobacter sulfurreducens* (17), PDB code 2w31; 35% sequence identity and 1.03 Å RMSD for C<sup>α</sup> atoms), the isolated globin domain of HemAT from *Bacillus subtilis* (18), PDB code 1or6; 22% sequence identity and 1.5 Å RMSD for C<sup>α</sup> atoms), and the isolated globin domain of YddV from *Escherichia coli* (16), PDB code 4zva; 21.5% sequence identity and 4.0 Å RMSD for C<sup>α</sup> atoms). The most structurally similar known protein is the isolated globin domain of GsGCS (17). However, it should be noted that the observed arrangement of this domain in the vicinity of the heme GsGCS may be artificial, because its bis-histidyl (His66 and His93) 6-coordinate heme (17) has not been observed for any other heme-based oxygen sensor. Accordingly, while AvGReg was initially reported to have a bis-histidyl 6-coordinated heme [Thijs et. al. (2007) Characterization of a globin-coupled oxygen sensor with a gene-regulating function. *J. Biol. Chem.* **282**, 37325–40], a corrected structure with a 5-coordinated heme complex has been deposited in the PDB under accession code 4UII. Compared to the BsHemAT globin domain, the isolated AfGcHK globin domain has a significantly shorter helix H3 and a longer helix H4 (18), and its H6 and H7 helices are shifted by approximately 4 Å. In addition, helix H1 from the AfGcHK globin domain has no analogous structure in the isolated globin domain of YddV (16), and helix H2 of the AfGcHK globin domain is split into the two separate helices with different orientations in the YddV ortholog.

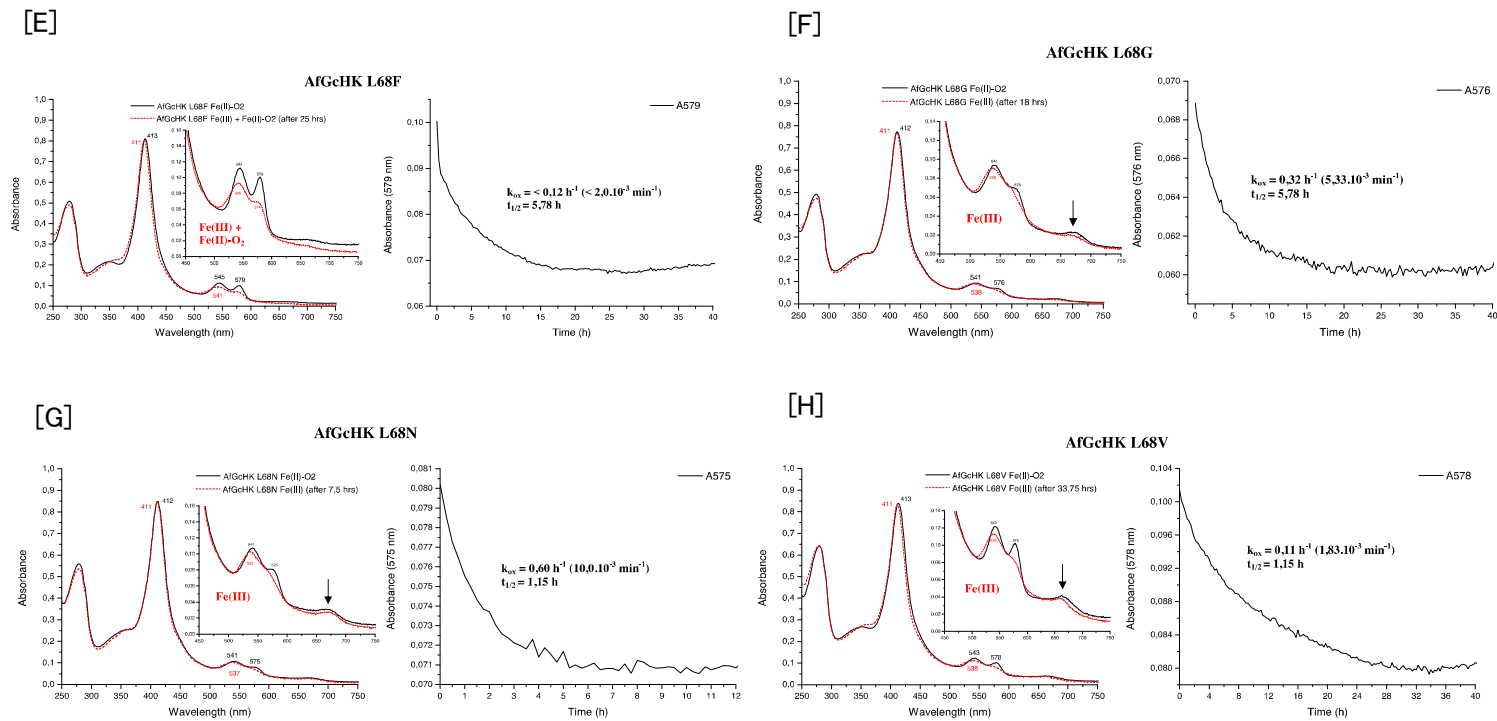
**Table S1.** Autoxidation rates of the wild type full-length *AfGcHK* and mutants at Tyr45 and Leu68 obtained in the present study. Autoxidation rates are also shown for corresponding mutants of other GCSs: YddV (21), HemAT-Bs (22, 23), and sperm whale myoglobin (24).

Protein	Autoxidation rate ( $\times 10^{-3} \text{ min}^{-1}$ )	Reference
Full-length <i>AfGcHK</i> WT	$< 10^{-2}$	this work
Y45F	$< 10^{-2}$	this work
Y45L	3.8	this work
Y45W	1.7	this work
L68F	$< 2.0$	this work
L68G	5.3	this work
L68N	10.0	this work
L68V	1.8	this work
Full-length YddV WT	9.2	21
Globin domain YddV WT	7.6	21
Y43A	Fe(II)-O <sub>2</sub> not formed	21
Y43F	69	21
Y43L	Fe(II)-O <sub>2</sub> not formed	21
Y43W	120	21
L65G	$> 100$	21
L65M	32	21
L65Q	63	21
L65T	89	21
HemAT-Bs	1.0	22, 23
Sperm whale myoglobin	1.0	24

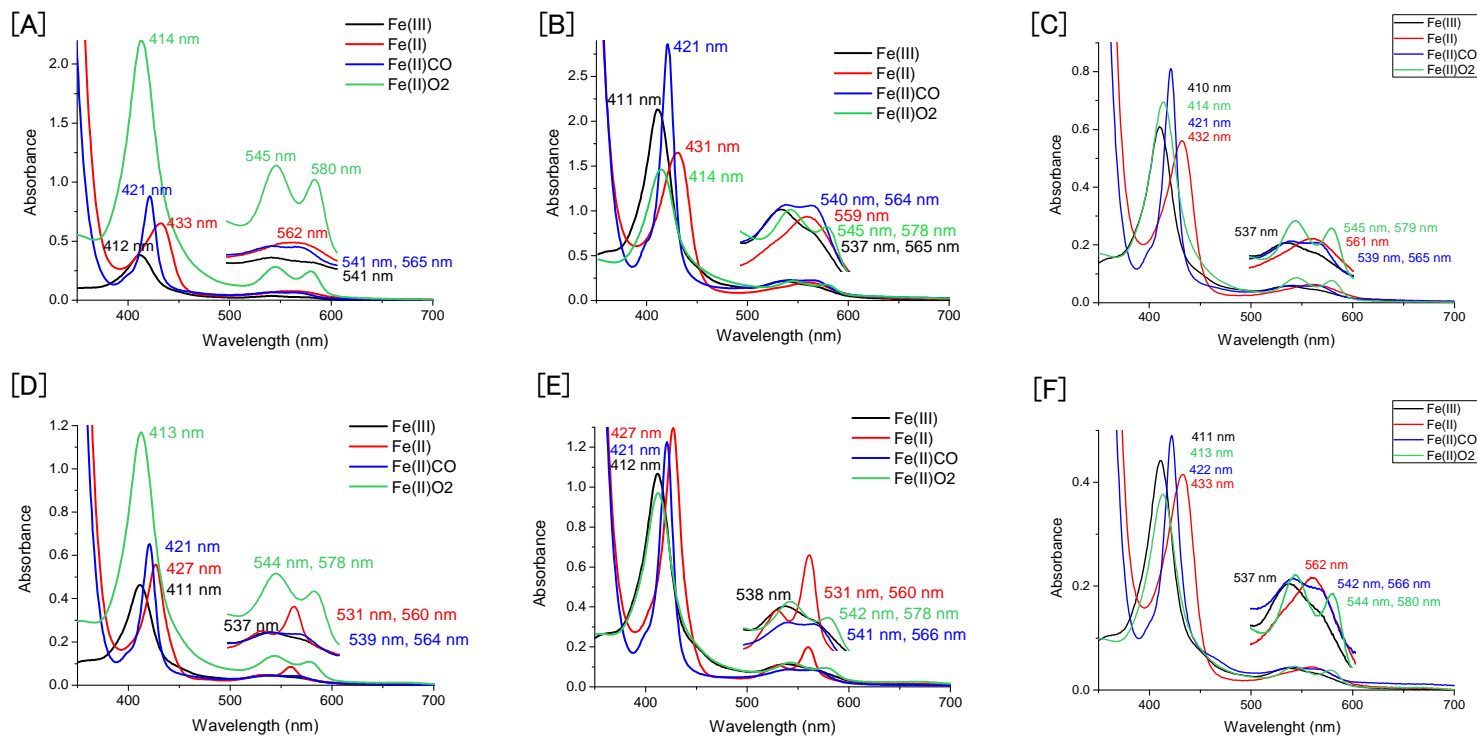


**Figure S1.** Schematic representation of the *A/GcHK* two-component signal transduction system. O<sub>2</sub> binding to the heme Fe(II) complex bound in the globin domain activates autophosphorylation at His183 in the kinase domain of *A/GcHK*. The phosphoryl group at His183 is then transferred to Asp52 or Asp169 in the cognate response regulator protein, ultimately leading to the modulation of transcription and/or enzymatic activity.

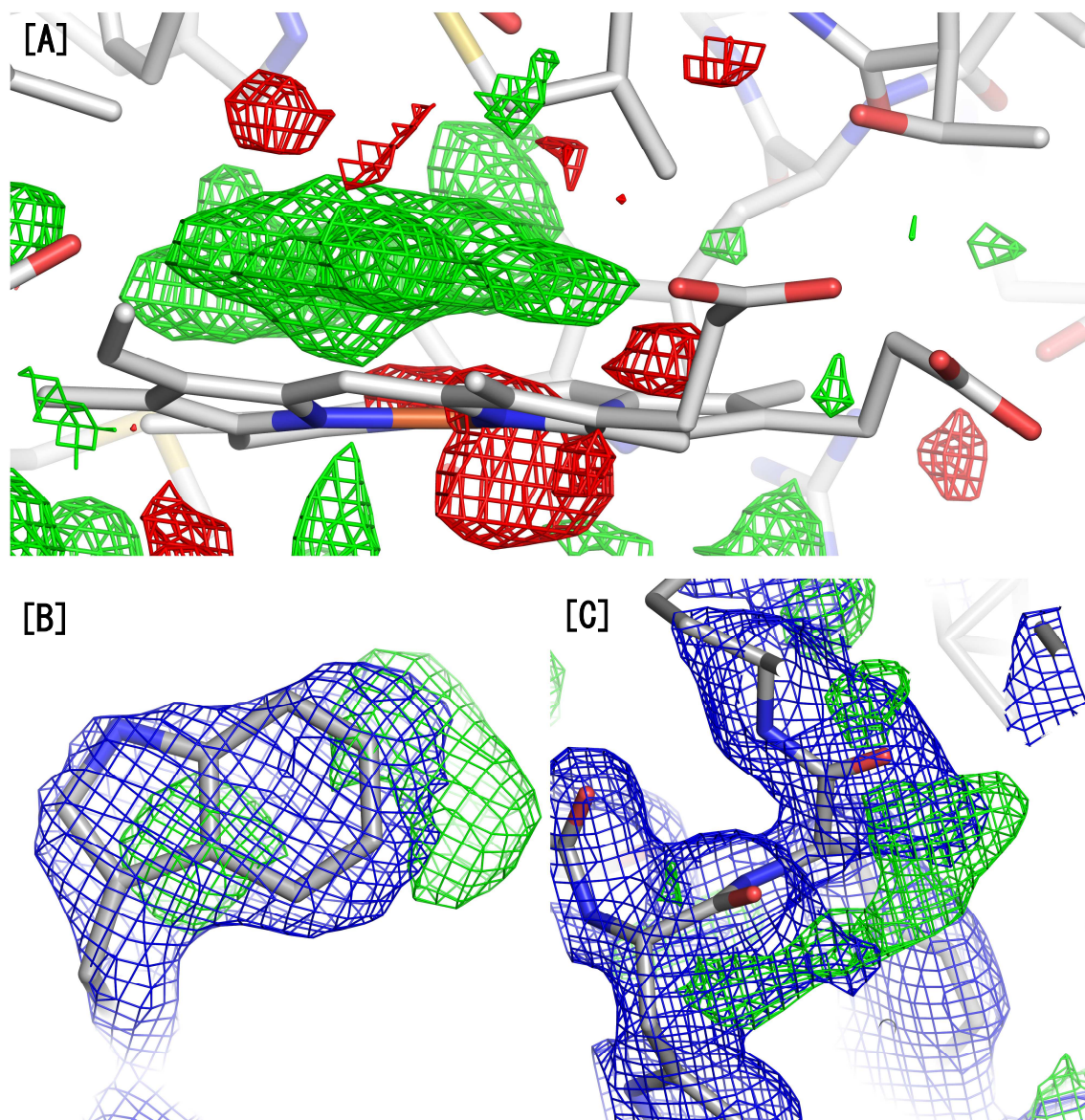




**Figure S2.** Changes in absorption spectra (left) and absorption intensity at 575-580 nm (right) associated with the autoxidation of the Fe(II)-O<sub>2</sub> complex to the Fe(III) complex of the full-length AfGcHK. [A] wild type, [B] Y45F, [C] Y45L, [D] Y45W, [E] L68F, [F] L68G, [G] L68N, [H] L68W.

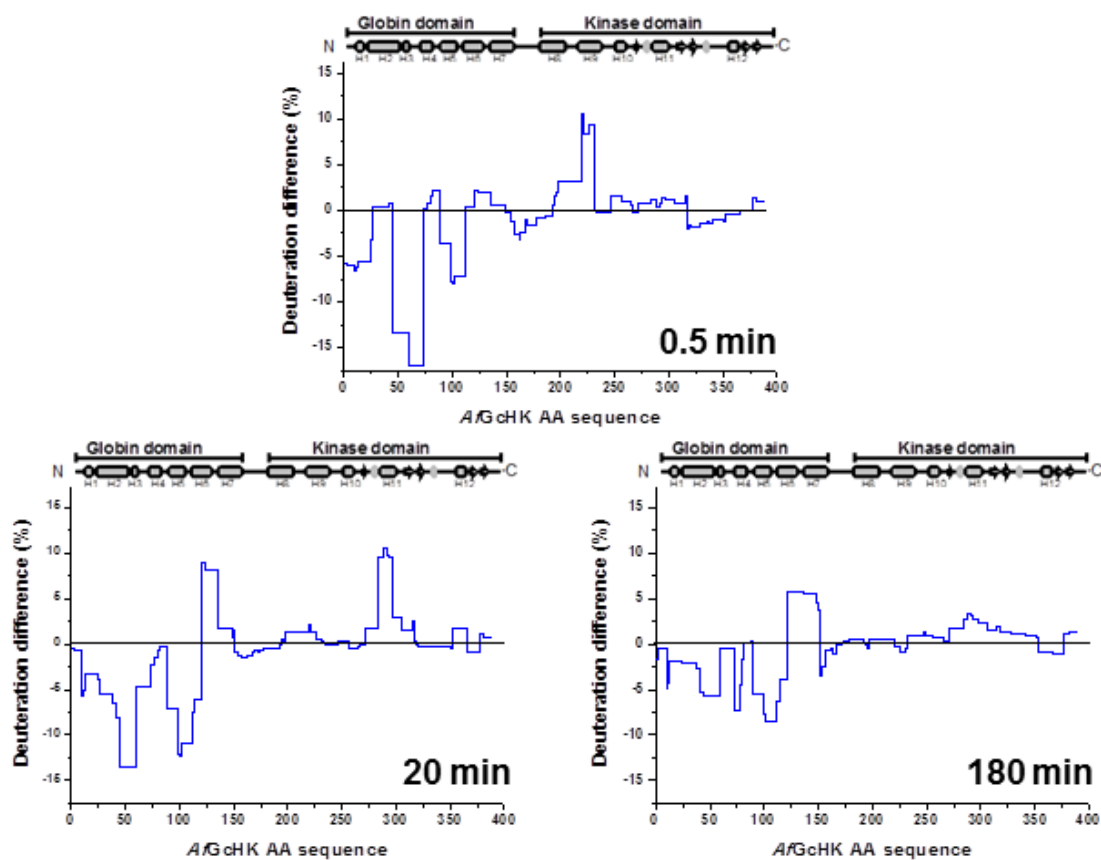


**Figure S3.** Absorption spectra of the Fe(III) (black), Fe(II) (red), Fe(II)-CO (blue) and Fe(II)-O<sub>2</sub> (green) complexes of the full length A/GCHK wild type [A], Y45F [B], L68F [C], L68G [D], L68N [E], L68V [F].

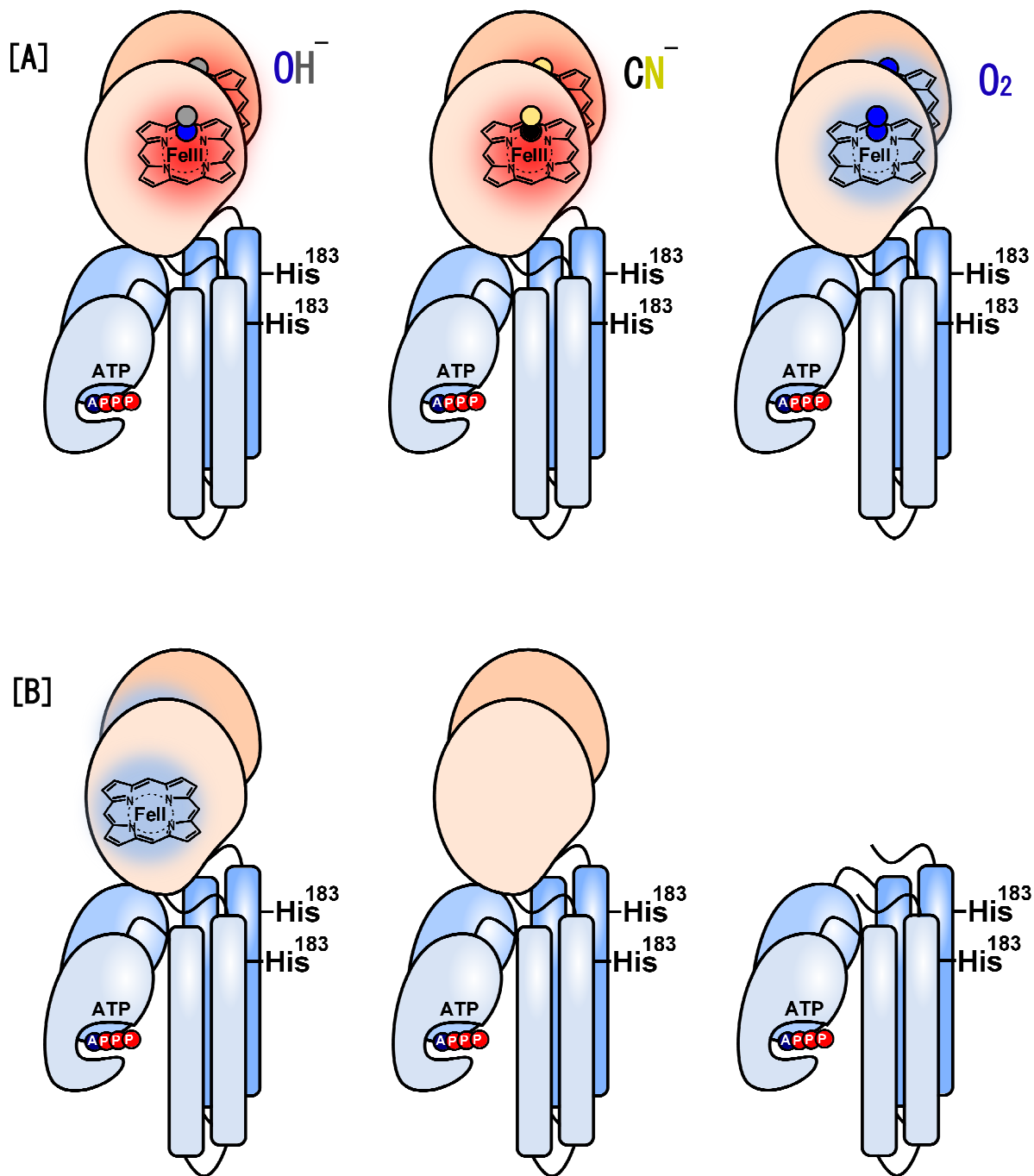


**Figure S4.** Electron density maps supporting the existence of two alternative positions in chain G (heme, protein side chains and protein main chain). [A]  $mF_o-DF_c$  map at the  $3\sigma$  level (green - positive, red - negative). The heme is built in one position with full occupancy. [B] Trp89 side chain, one position with 0.5 occupancy,  $2mF_o-DF_c$  map at  $1\sigma$  (blue) and positive  $mF_o-DF_c$  map at  $3\sigma$  (green). [C] Tyr108 main chain, one position with 0.5 occupancy,  $2mF_o-DF_c$  map at  $1\sigma$  (blue) and positive  $mF_o-DF_c$  map at  $3\sigma$  (green).



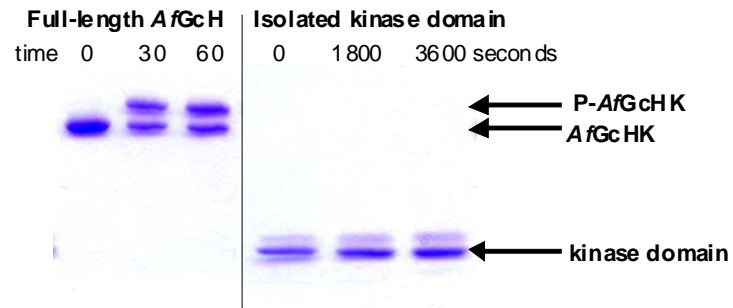


**Figure S5.** HDX-MS protection plots showing the differences in deuteration after 0.5 min (upper centre), 20 min (lower left) and 180 min (lower right) of exchange for the full-length Fe(III)-OH-bound (active) and Fe(II)-bound (inactive) forms of *AfGcHK*. The deuteration levels of the active form were subtracted from those of the inactive form and plotted against the protein's sequence (x-axis). The blue line thus represents the structural changes caused by reducing the heme iron center in the globin domain of the full-length protein.

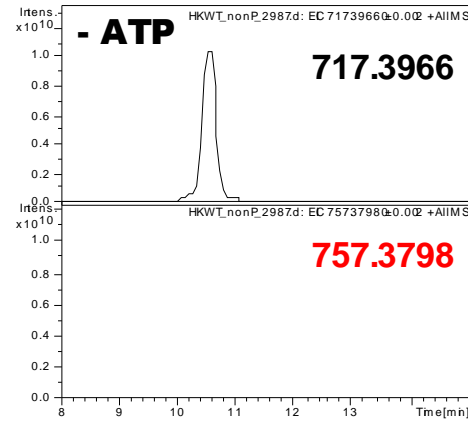


**Figure S6.** Domain structures of *A/GcHK* showing active (A) and inactive (B) states. Individual states of the globin domain (heme iron oxidation/reduction, heme iron coordination, presence or absence of heme, presence or absence of the globin domain) are associated with either the active or the inactive state (based on results from (8) and this study).

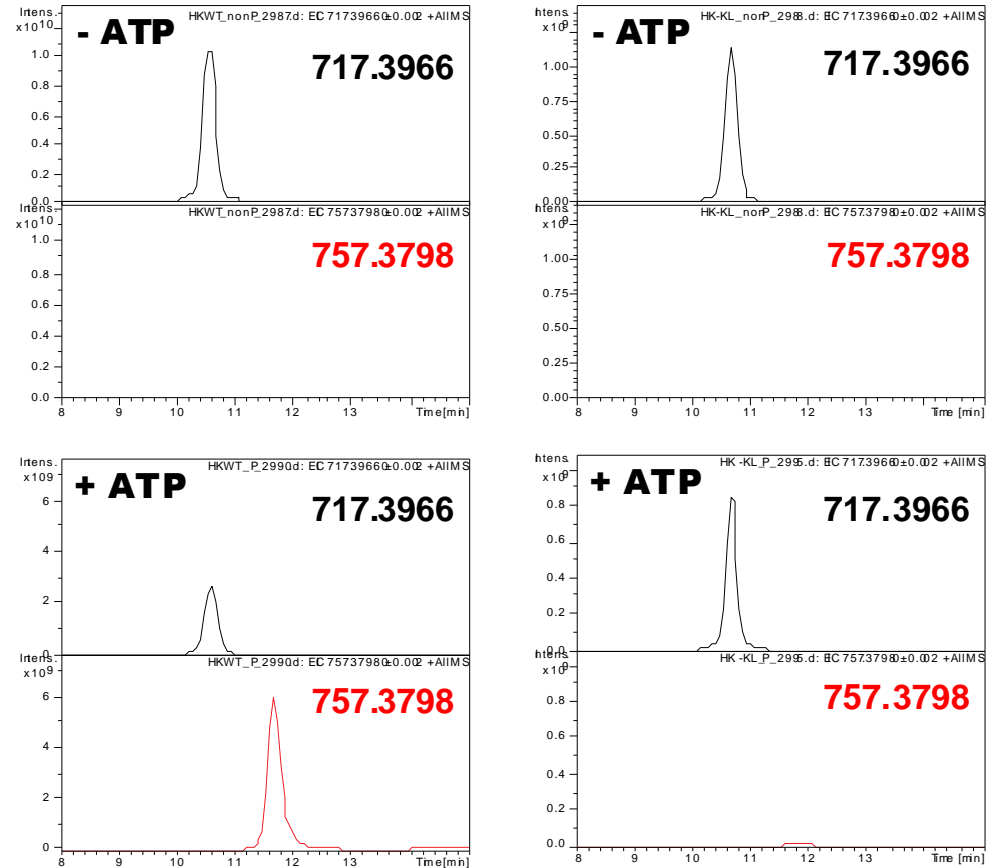
[A]



[B]

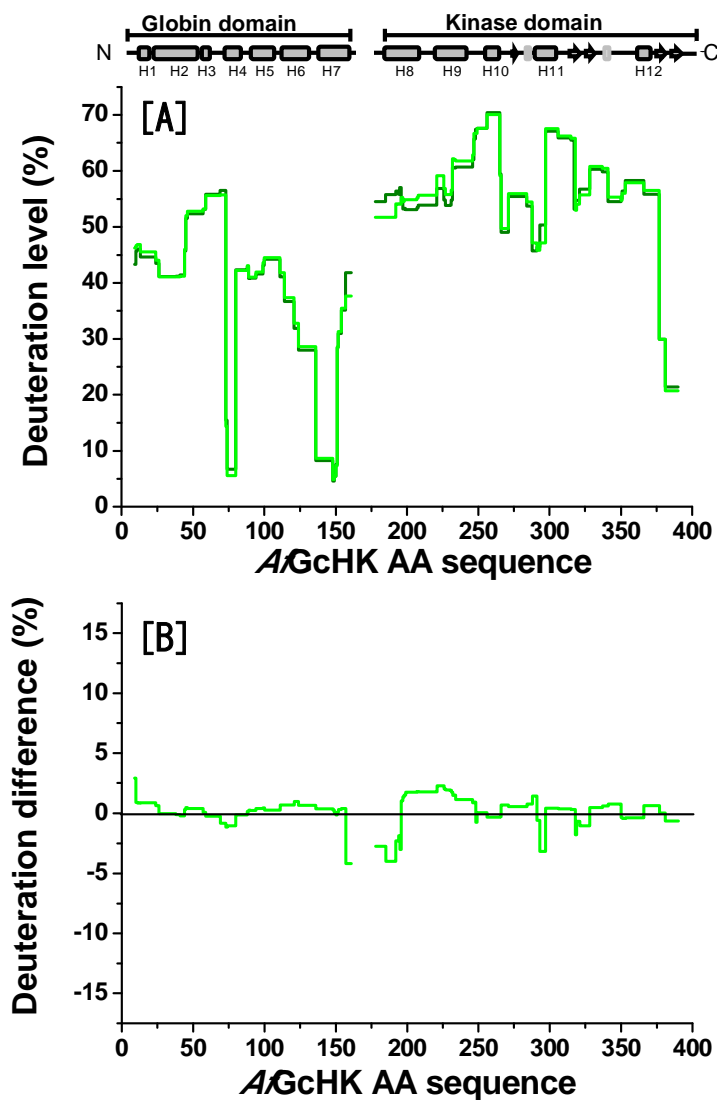


[C]



**Figure S7.** An isolated kinase domain (specifically, a truncated form of AfGcHK lacking the N-terminal heme-bound globin domain is removed) exhibits no autophosphorylation activity, suggesting that the globin domain is required for autophosphorylation of the kinase domain. [A]: Phos-tag SDS-PAGE gel

patterns demonstrate that full-length *AfGcHK* is autophosphorylated after 30 sec (left), whereas the isolated kinase domain is not autophosphorylated even incubation times of up to 1 hr (right) in the presence of 1 mM ATP. [B] and [C]: Mass spectrometric characterization of the phosphorylation status of histidine 183 in the kinase domain of full-length *AfGcHK* [B] and its isolated kinase domain [C] obtained in the absence and presence of 1 mM ATP. Extracted ion chromatograms for selected peptides covering histidine 183 in the protein's sequence – VGSIGHDLRNPLGV (178 – 191) - are shown. Black and red traces correspond to the unmodified and phosphorylated protein, respectively. The mass-to-charge ( $m/z$ ) ratios used to plot the chromatograms are indicated above each trace.



**Figure S8.** HDX-MS results demonstrating the differences between single isolated domains alone in solution and the mixture of the isolated globin and kinase domains. [A]: HDX-MS results in the form of a protection plot. The dark green lines represent the deuteration levels of the individual isolated domains alone in solution – data for the Fe(III)-OH<sup>-</sup> form of the globin domain are shown on the left, and data for the isolated kinase domain on the right. The light green lines show the results for the equimolar mixture of isolated domains. [B]: Difference in deuteration after 180 min of exchange between the individual isolated domains alone in solution and the mixture of isolated globin and kinase domains (light green lines).