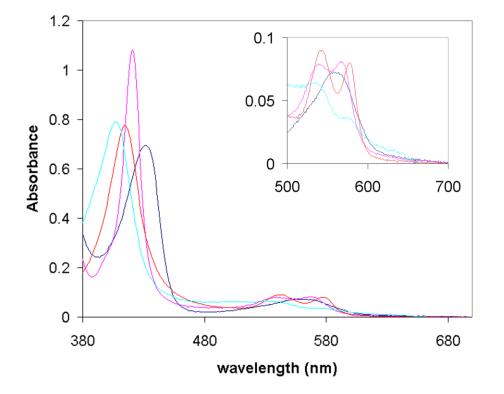
2.3 Å X-ray Structure of the Heme-Bound GAF Domain of Sensory Histidine Kinase DosT of *Mycobacterium tuberculosis*

Larissa M. Podust, Alexandra Ioanoviciu, Paul R. Ortiz de Montellano

From the Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158-2517.



SUPPLEMENTAL FIGURES

Figure S1. UV-visible spectral characterization of GAF_{DosT} . Spectra of the ferric GAF_{DosT} , *cyan*; ferrous GAF_{DosT} , *dark blue*; ferrous GAF_{DosT} -CO complex, *magenta*; and ferrous GAF_{DosT} -O₂ complex, *red*, are shown. The insert shows an enlargement of the visible region of the spectra.

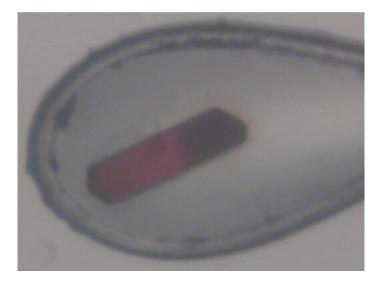


Figure S2. GAF_{DosT} x-ray induced photoreduction. A snapshot of the GAF_{DosT} crystal after collection of the two-wavelength anomalous diffraction data set is shown. The irradiated part of the crystal on the right is of a much darker red color.

Fig. S3 (page1)

50

DosT_M.tuberculosis(053473)

DosT_M.tuberculosis(053473) M.smegmatis(ABK70610) M.vanbalenii(ABM12229) M.gilvum(ABF46316) DosS_M.ulcerans(ABL04774) DosS_M.bovis(BCG_CAL73144) DosS_M.tuberculosis(NP217648) N.farcinica(BAD57741) Rhodococcus_RHA1(ABG91885) T.fusca(AA257079) S.coelicolor(CAB53427) S.tropica(ABF54044) Nocardioides_JS614(ABL30469)

DosT_M.tuberculosis(053473)

DosT_M.tuberculosis(053473) M.smegmatis(ABK70610) M.vanbaalenii(ABM12229) M.gilvum(ABP46316) DosS_M.ulcerans(ABL04774) DosS_M.bovis(BCG_CAL73144) DosS_M.tuberculosis(NP217648) N.farcinica(BAD57741) Rhodococcus_RHA1(ABG91885) T.fusca(AAZ57079) S.ccelicolor(CAB53427) S.tropica(ABF54044) Nocardioides_JS614(ABL80469)

DosT_M.tuberculosis(053473)

DosT_M.tuberculosis(053473) M.smegmatis(ABK70610) M.vanbaalenii(ABM12229) M.gilvum(ABP46316) DosS_M.ulcerans(ABL04774) DosS_M.tuberculosis(NP217648) N.farcinica(BAD57741) Rhodcocccus_RHA1(ABG91885) T.fusca(AAZ57079) S.ccelicolor(CAB53427) S.tropica(ABP54044) Nocardioides_JS614(ABL80469)

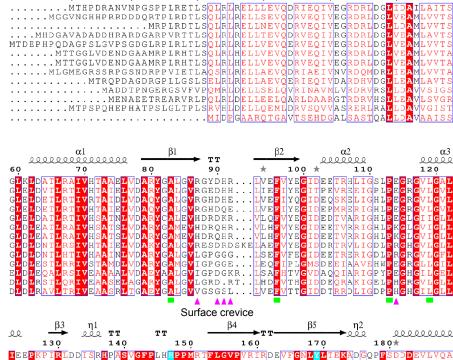
DosT M.tuberculosis(053473)

α4

DosT_M.tuberculosis (053473) M.smegmatis (ABK70610) M.vanbaalenii (ABM12229) M.gilvum (ABP46316) DosS_M.ulcerans (ABL04774) DosS_M.tuberculosis (NP217648) N.farcinica (BAD57741) Rhodcocccus_RHA1 (ABG91885) T.fusca (AAZ57079) S.coelicolor (CAB53427) S.tropica (ABP54044) Nocardioides_JS614 (ABL80469)

DosT_M.tuberculosis(053473)

DosT_M.tuberculosis(053473) M.smegmatis(ABK70610) M.vanbaalenii(ABM12229) M.gilvum(ABP46316) DosS_M.ulcerans(ABL04774) DosS_M.tuberculosis(NF217648) N.farcinica(BAD57741) Rhodococcus_RHA1(ABG91885) T.fusca(AAZ57079) S.ccolicolor(CAB53427) S.tropica(ABF54044) Nocardioides_JS614(ABL80469)





Proximal heme-binding motif

10

20

30

4 0

1	2000000	2000	тт				
19	p	200	210	220	230	240	250
	LAAAAGIA	VDNARI	LFEESRT <mark>R</mark> EA <mark>M</mark>	IEATRDIGTOML	AGADPAMVE	RLIAEEALTL	MAGAATLVAVPL
	LAAAAGIA	AVD <mark>NAR</mark> I	LYEESQA <mark>r</mark> qa <mark>w</mark>	IAATRDIGTQLL	SGTDPATVE	RLVAAEALTL	TGADGTLVAVPA
- 6	LAAA <mark>AG</mark> IA	AIE <mark>NAR</mark> I	LYQQSKT <mark>R</mark> RS <mark>≬</mark>	IEATRDIGTEML	SGADPSKVE	RLVADQSRLL	SGAQSTLVAVRS
				IEATRDIGTELL			
				IGATRDIATELL			
				IEATRDIATELL			
				IEATRDIATELL			
				LEATQQVATRLL			
				LEATREIATELL			
				LDASDEITTRLL			
				LQVNAEITHTLM			
				LAAAAEITTLLL			
	LASTAGEN	/ID <mark>NAR</mark> /	AYGLSER <mark>R</mark> RQ M	LEASAELTEMLQ	PPIEL <mark>G</mark> RAI	SQVARSARSM	SGAVAASVLRLD

	260	270	280	290	300	310
				•	•	RRFDR.LDLA.VDGPVE
	DPDASAAE	ELVIVEV	AGAVPAEVEAS	AIPVQDNAIGQ	AFRDRAP	RRLDV.LDGPGL
	DLDELDGRVD	ELVVVAI	AGD.GP <mark>G</mark> VSLS	AIPTHGSAIGE	VFSTKTP	VRFDT.LELE.P.RHLA
	DPDLPVGEVA	ELTVAAS	AGE.TI <mark>A</mark> AAHD	PIPVAGTAIGD	AFVHRTP	GMTDR.ADIG.I.GATV
	DEDVPASEVA	QLLVIET	VGNAVAAAEGC	TIAVAGTALAE	VLLDSAP	RQVDK.IAVEDVDELGD
	DEDMPAADVG	ELLVIET	VGSAVA <mark>S</mark> TVGR	TIPVAGAVLRE	VFVNGIP	RRVDR.VDLEGLDELAD
)	DEDMPAADVG	ELLVIET	VGSAVA <mark>S</mark> IVGR	TIPVAGAVLRE	VFVNGIP	RRVDR.VDLEGLDELAD
	DPDVPPEDLT	ELVVVAA	AGADADTLTGT	RLPMDETHTGV	AFRDGRP	LAVDT.ADAPSFATSLE
	DPDIPSDEVT	ELVVTAS	AGTVSD <mark>R</mark> IIGR	TIPVDKSTSGE	AFRERTP	LRVTALAFDPGFETGTR
	RDEEEL	VVRIADG	PDAS <mark>V</mark> VRGR	STPVEGSLTGR	AFRGADPIIT.	ELARESATSPTLLHGLD
	SGTDTL	AVELAVG	HEADAWRGI	VLPVEGTLIGQ	AF VQRAPVHSA:	DVCRDSRVSAGPPRFEG
	PGADTF	TVDVVDG	AGEQAAALVGT	VLPAADTSFGA	AVIHGRH	DRVADLAYAAPWPALFD
	NGTPVHTTAV	DPADADR	VGGALELLGDE	V <mark>W</mark> PRGD <mark>V</mark> SPVG	RTIGGLE	

Fig. S3 (page2)

380

370

DosT_M.tuberculosis(053473)

320

330

340

DosT M.tuberculosis(053473) M. smegmatis (ABK70610) M. vanbaalenii (ABM12229) M. gilvum (ABP46316) DosS_M.ulcerans(ABL04774) DosS_M.bovis(BCG_CAL73144) DosS_M.tuberculosis(NP217648) N.farcinica (BAD57741) Rhodococcus_RHA1 (ABG91885) T.fusca(AAZ57079) S.coelicolor(CAB53427) S.tropica(ABP54044) Nocardioides_JS614 (ABL80469)

DosT M.tuberculosis(053473)

DosT_M.tuberculosis(053473) M.smegmatis(AEK70610) M.vanbaalenii(AEM12229) M.gilvum(AEP46316) DosS_M.ulcerans (ABL04774) DosS_M.bovis (BCG_CAL73144) DosS_M.tuberculosis (NP217648) N.farcinica (BAD57741) Rhodococcus_RHA1 (ABG91885) T.fusca (AAZ57079) S.ccelicolor(CAB53427) S.tropica(ABF54044) Nocardioides_JS614(ABL80469)

DosT_M.tuberculosis(053473)

DosT_M.tuberculosis(053473) M.smegmatis(ABK70610) M.vanbaalenii(ABK70610) M.gilvum(ABP46316) Doss_M.ulcerans(ABL04774) Doss_M.bovis(BCG_CAL73144) Doss_M.BOVIS(BCG_cH/S144) Doss_M.tuberculosis(NP21768) N.farcinica(BAD57741) Rhodococcus_RHA1(ABG91885) T.fusca(AAZ57079) S.coelicolor(CAB53427) S.tropica (ABP54044) Nocardioides_JS614 (ABL80469)

DosT_M.tuberculosis(053473)

DosT M.tuberculosis(053473) M. smegmatis (ABK70610) M.vanbaalenii (ABM12229) M.gilvum (ABP46316) DosS_M.ulcerans(ABL04774) DosS_M.bovis(BCG_CAL73144) DosS_M.tuberculosis(NP217648) N.farcinica (BAD57741) Rhodococcus_RHA1 (ABG91885) T.fusca(AAZ57079) S.coelicolor(CAB53427) S.tropica (ABP54044) Nocardioides_JS614 (ABL80469)

DosT_M.tuberculosis(053473)

DosT_M.tuberculosis(053473)							
M.smegmatis (ABK70610)							
M.vanbaalenii(ABM12229)				•	•		
M.gilvum(ABP46316)							
DosS_M.ulcerans(ABL04774)	•			•			
DosS_M.bovis(BCG_CAL73144)							
DosS_M.tuberculosis(NP217648)							
N.farcinica (BAD57741)							
Rhodococcus_RHA1 (ABG91885)				•			
T .fusca (AA Z57079)							
S.coelicolor(CAB53427)	ΕC	GG	31	R	S	A	R
S.tropica(ABP54044)							
Nocardioides_JS614 (ABL80469)				•			

GGPALVLPLRA PGPALVLPLRA GPALLLPLRA MGPALLLPLRA AGPALLLPLRA YGPVLTLPLRA FGPALALPLRA FGPALALPLRA IGPVVLVPFGA LGPGVAVPIGS TGPAVLSPLAT	T DTVAGVLVAV TDTVAGVLVAL PDAVPGVLIVL TDAVAGVVVVL RGTVAGVVVVL GTVAGVVVVL GNAVVGLLTTM AQSVTGVLVTL PDSRRGVLILG SAEARGVVLVA	Q G S G A R P F TA E (R P E G A P P F NA E F R S V G A Q P F R S L E S Q C G P G A F T D E (S Q C G P G A F T D E (R A L G M P V D A A C R H V D A L P F T D D (K R A G R L P F P M F I R Q S G G Q E F S E E F Y S A D H D D T S E L I	2 LEMMIT GF ADQAAU 2 LEMMIA GF ADQAAU 2 LDMIA AF ADQAAQ 2 LEMMIA AF ADQAAL 2 LEMMIA AF ADQAAL 2 LEMMIA AF ADQAAL 2 LEMMIA F ADQAAL 2 LEMMIA F ADQAAU 2 LALMS GF ADQAAU 2 VQMILHAF AGHAAV 3 TEP LLVF AAQAAU 3 LIAL LGS F AAQAAU	AWRLATAQRQMREVEILIDR AWQLATSQRRMSELDVLIDR AWQLATTQRQLRELDVLIDR AWQLATSQRRMRELDVUTER AWQLATSQRRMRELDVUTER AWQLATSQRRMRELDVUTDR AUQMATQRRMRELDVUTDR ALQMATQRRMRELDVUSER ILELAEARKHAERLIVLEDR AMELAERRADAEQIALEDR AMERARGQEERELLVVLEDR ALDRAQAIEDRASLAVTSDR
390	400	410 4	120 430	440 450
DRIARDLHDHV	IQRLFAVGLTL	QGAAPRARVPAU	/RESIYSSIDDLQE	IIQEIRSAIFDLHAGPSRAT
D <mark>RIA</mark> RDLHDHV D <mark>RIA</mark> RDLHDHV	IQRLFAVGLTL IQRLFAVGLSL	QGAAPRARVPAV QGTIPRARSAEV	VRESIYSSIDDLQE VQQRIADCVDDLQA	IIQEIR <mark>SAI</mark> FDLHAGPSRAT VISEIRTAIFDLHHTAPGGT
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLAL	QGAAPRARVPAV QGTIPRARSAEV QGTIPRARAPEV	VRESIYSSIDDLQE VQQRIADCVDDLQA VRERLTDCVDDLQQ	IIQEIRSAIFDLHAGPSRAT VISEIRTAIFDLHHTAPGGT VIQEIRTAIFDLHGGHGGSN
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLAL IQRLFAVGLSL	QGAAPRARVPAV QGTIPRARSAEV QGTIPRARAPEV QGSIARARSEEV	VRESIYSSIDDLQE VQQRIADCVDDLQAV VRERLTDCVDDLQQ VQQRLSANVDDLQS	IIQEIRSAIFDLHAGPSRAT VISEIRTAIFDLHHTAPGGT VIQEIRTAIFDLHGGHGGSN VIQEIRTAIFDLHGGSSAVT
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLAL IQRLFAVGLSL IQRLFAIGLAL	QGAAPRARVPA QGTIPRARSAF QGTIPRARSAF QGSIARARSE QGSVPRTREPK	VRESIYSSIDDLQE VQQRIADCVDDLQA VRERLTDCVDDLQQ VQQRLSANVDDLQQ VQQRLSEVIDDLQG	IIQEIRSAIFDIHAGPSRAT VISEIRTAIFDIHAGPGGT VIQEIRTAIFDIHGGHGGSN VIQEIRTAIFDIHGGSSAVT VIQEIRTIFDIHGGSSQVT
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLAL IQRLFAVGLSL IQRLFAIGLAL IQRLFAIGLAL	QGAAPRARVPA QGTIPRARSAFU QGTIPRARAPEU QGSIARARSEEU QGSVPRTREPKU QGAVPHERNPEU	VRESIYSSIDDLQE VQQRIADCVDDLQA VRERLTDCVDDLQQ VQQRLSANVDDLQQ VQQRLSEVIDDLQQ VQQRLSEVIDDLQG	IIQEIRSAIFDLHAGPSRAT VISEIRTAIFDLHHTAPGGT VIQEIRTAIFDLHGGHGGSN VIQEIRTAIFDLHGGSSAVT VIQEIRTAIFDLHGSSQVT VIQEIRTIFDLHGSSQCT
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIGLAL IQRLFAIGLAL IQRLFAIGLAL	QGAAPRARVPA QGTIPRARSAEV QGTIPRARAPEV QGSIARARSEEV QGSVPRTREPK QGAVPHERNPEV QGAVPHERNPEV	VRESIYSSIDDLQE VQQRIADCVDDLQA VRERLTDCVDDLQQ VQQRLSANVDDLQS VQQRLSEVIDDLQD VQQRLSDVVDDLQD VQQRLSDVVDDLQD	I I QE I R SAT F D L HAGP SRAT VISE I R TA I F D L HAGP SRAT VIQE I R TA I F D L HGGHGG SN VIQE I R TA I F D L HGGS SAVT VIQE I R TT I F D L HGS SQGVT VIQE I R TT I Y D L HGAS QG I T VIQE I R TT I Y D L HGAS QG I T
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIGLAL IQRLFAIGLAL IQRLFAIGLAL IQRLFAVGLSL	QGAAPRARVPA OGTIPRARSAFU QGIIPRARAPEU OGSIARARSEU QGSVPRTREPKU QGAVPHERNPEU QGAVPHERNPEU QGAVPHERNPEU QSTVSRARAPEI	TRESTYSSIDDIQE TQQRTADCVDDIQA TRERLTDCVDDIQA TQQRLSANVDDIQS TQQRLSEVIDDIQO TQQRLSEVIDDIQO TQQRLSDVVDDIQD TQQRLSDVVDDIQD TQQRLSDVVDDQS	I I QE I R SA T F D I HAGP SRAT VI SE I R TA I F D I HA TA P G T VI QE I R TA I F D I H G S S S V VI QE I R TA I F D I H G G S S V T VI QE I R TT I F D I H G S S Q S V T VI QE I R TT I Y D I H G A S Q G I T VI QE I R TT I Y D I H G A S Q G I T I Y QE I R TT I Y D I H S S S A A A
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	I QRLFAVGLTL I QRLFAVGLSL I QRLFAVGLSL I QRLFAVGLSL I QRLFAI GLAL I QRLFAI GLAL I QRLFAI GLAL I QRLFAVGLSL I QRLFAVGLSL	QGAAPRARVPA QGTIPRARSAEV QGTIPRARSAEV QGSIARASEEV QGSVPRTREPK QGAVPHERNPEV QGAVPHERNPEV QSTVSRARAPEI QSTLVSRARAPIV	TRESTYSSIDDIQE /QQRIADCVDDIQA (RERLTDCVDDIQC /QQRLSANVDDLQS /QQRLSDVVDDLQD /QQRLSDVVDDLQD /QQRLSDVVDDLQD /QQRLSDVVDDLQD /QQRLSDVVDDLQD /KTRLMDTVNDVQS /KQRLEQSIDDLHD	I I QE I R SA I F DI HAGP SRAT VI SE I R TAIF DI HAGP SRAT VI QE I R TAIF DI HGGH GGS N VI QE I R TAIF DI HGGS SAVT VI QE I R TI F DI HGSS QGVT VI QE I R TI Y DI HGAS QGIT VI QE I R TI Y DI HGAS QGIT VI QE I R SIF DI HSSSAAEA I V PI I R AIF DI HGGS GGMT
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIGLAL IQRLFAIGLAL IQRLFAIGLAL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIAMSL	QGAAPRARVPA QGTIPRARSAFV QGTIPRARSAFV QGSIARARSEV QGSVPRTREPK QGSVPRTREPK QGAVPHERNPEV QGAVPHERNPEV QSTVSRARAPEI QSTLQRAKAPII MSSVRRIEDDT	TRESTYSSIDDIQE TQQRIADCVDDIQA YRERITDCVDDIQA YQQRLSANVDDIQS YQQRLSEVIDDIQG YQQRLSDVVDDLQD YQQRLSDVVDDLQD KTRLMDTVNDVQS KKQRLEQSIDDLHD	I I QE I R SAT F D L HAGP SRAT VI SE I R TA F D L HAGP SRAT VI QE I R TA F D L HGGH GG S N VI QE I R TA F D L HGGS SAVT VI QE I R TT F D L HGS SQGVT VI QE I R TT Y D L HGS SQGVT VI QE I R TT Y D L HGS SGIT I VQE I R TT Y D L HGS SQET I VQE I R TT Y T Y D L HGS SQET I VQE I R TT Y T Y T Y T Y T Y T Y T Y T Y T Y
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDVV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIGLAL IQRLFAIGLAL IQRLFAIGLAL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIGMSL IQRLFATGMTL	QGAAPRARVPA QGTIPRARSAFV QGTIPRARSAFV QGSIARARSEV QGSVPRTREPKV QGAVPHERNPEV QGAVPHERNPEV QSTVSRARAPEI QSTVSRARAPIV MSSVRRIEDDIV QSAGRRVQDPM	TRESTYSSIDDIQE TQQRIADCVDDIQA TRERLTDCVDDIQA TQQRLSANVDDIQS TQQRLSDVVDDIQS TQQRLSDVVDDIQS TQQRLSDVVDDIQD KTRLMDTVNDVQS TQQRLEQSIDDIHD JAQRIQQAVDDIDD SERILRAVDDIDE	I I QE I R GAT F DI HAGP SRAT VI SE I R TA I F DI HHTAP GGT VI QE I R TA I F DI HGG SGS N VI QE I R TA I F DI HGG SS AVT VI QE I R TT I F DI HGGS SQ VT VI QE I R TT I Y DI HGAS QG I T VI QE I R TT I Y DI HGAS QG I T I VQE I R HS I F DI HGS SG GT I VQE I R HS I F DI HGGS GGT I R QI R ST I F AI QHNTE GGP I KI I R ST I F GI R SR DD AV
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV	IQRLFAVGLTL IQRLFAVGLSL IQRLFAVGLSL IQRLFAVGLSL IQRLFAIGLAL IQRLFAIGLAL IQRLFAIGLAL IQRLFAVGLSL	QGAAPRARVPA OGTIPRARSAFU QGIIPRARAPEU OGSIARARSEU QGSVPRTREPKU QGAVPHERNPEU QGAVPHERNPEU QGAVPHERNPEU QSTVSRARAPEI	TRESTYSSIDDIQE TQQRTADCVDDIQA TRERLTDCVDDIQA TQQRLSANVDDIQS TQQRLSEVIDDIQO TQQRLSEVIDDIQO TQQRLSDVVDDIQD TQQRLSDVVDDIQD TQQRLSDVVDDQS	I I QE I R SA T F D I HAGP SRAT VI SE I R TA I F D I HA TA P G T VI QE I R TA I F D I H G S S S V VI QE I R TA I F D I H G G S S V T VI QE I R TT I F D I H G S S Q S V T VI QE I R TT I Y D I H G A S Q G I T VI QE I R TT I Y D I H G A S Q G I T I Y QE I R TT I Y D I H S S S A A A
DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDHV DRIARDLHDVV DRIARDLHDVV DRIARDLHDVA BRIARDLHDVA BRIARDLHDVA	I QRLF AVGLTL I QRLF AVGLSL I QRLF AVGLSL I QRLF AVGLSL I QRLF AI GLAL I QRLF AI GLAL I QRLF AI GLAL I QRLF AVGLSL I QRLF AT GMTL I QRLF AT GMTL I QRLF AT GLQL	QGAAPRARVPA QGTIPRARSAFV QGTIPRARSPEV QGSIARSEEV QGSVPRTREPKV QGSVPRTREPKV QGAVPHERNPEV QGAVPHERNPEV QSTVSRARAPEI QSTLQRAKAPIV MSSVRRIEDDTV QSAQRRVQDPM QSAVPMSTRPEI	TRESTYSSIDDIQE TQQRIADCVDDIQA TRERITDCVDDIQA TQQRLSANVDDIQS TQQRLSDVVDDIQD TQQRLSDVVDDIQD TQQRLSDVVDDIQD KTRLMDTVNDVQS TKQRLEQSIDDIHD ASERILRAVDDIDD ASERILRAVDDIDD	IIQEIRSAIFDIHAGPSRAT VISEIRTAIFDIHHTAPGGT VIQEIRTAIFDIHGGHGGSN VIQEIRTAIFDIHGGSSAVT VIQEIRTTIYDIHGSSGVT VIQEIRTTIYDIHGSSGUT VIQEIRTIYDIHGASGGT IVQEIRHSIFDIHSSAAEA IVRDIRTAIFDIHGGSGMT VRDIRTAIFDIHGGSGMT

350

360

Autophosphorylation site

460	47 Q	48 Q	490	500	510
G. LRHRLDKVIDQL	A. IPALHTTVQYT	GPL.SVVDTVLA	NHAEAVLRE/	AVS <mark>N</mark> AV <mark>RH</mark> AN	ATSLAINVS
R. LRORLDEAIAOF					
R. LRQRLDEAIAQF					
R.LRQRLDAAVAAF					
R.L <mark>R</mark> QRIDAAVAAF					
R. LRQRIDAAVAQF					
R. LRQRIDAAVAQF					
PALRKRLHAVLAEM					
Q.FRRRLHEIVAET					
AWLRSRILDVVGAA					
PGLRSRAVRAVGEA					
AQLRSEIREAVEDA					
PSLRADLRQLAREY	APLLKFDPTVTTV	GP V D T A V P T E V R	EQLLPVLREA	ALS <mark>N</mark> LA <mark>RH</mark> AA	ADHAEIELA

52 Q	53 <u>0</u>	54 Q	55 0	560	57 <u>0</u>
VED.DVRVEVV <mark>D</mark> I	D <mark>G</mark> V <mark>G</mark> ISGDI.TH	ES <mark>GL</mark> RNLRQ <mark>R</mark> A	D D A <mark>G G</mark> E F T V E	ENMPTG. <mark>GT</mark> LL	RWSAPLR
VED.DLSIEVT <mark>D</mark> I	D <mark>G</mark> E <mark>G</mark> ISGPV.TH	ES <mark>GL</mark> S <mark>N</mark> LRQ <mark>R</mark> A	DQC <mark>G</mark> ELRII	[DRPGG. <mark>GT</mark> IL]	R <mark>WTAPL</mark> PD
AAD.ELIIEVA <mark>D</mark> I	D <mark>G</mark> C <mark>G</mark> MPADV.TI	ES <mark>GL</mark> T <mark>N</mark> LRQ <mark>R</mark> A	G D V <mark>G </mark> G T F T V E	ESGDAG. <mark>GT</mark> RLH	RWCAPLR
					I <mark>WRAPL</mark> P
					R <mark>WSAPL</mark> IQ
					R <mark>WSAPL</mark> SQ
					R <mark>W</mark> S <mark>APL</mark> SQ
					H <mark>W</mark> S <mark>VPL</mark> P
VYD.DLAIEVA <mark>D</mark> I	GKGFVENVVT	. S <mark>gl</mark> enlaara	REVNGHFAII)TTPGG. <mark>GT</mark> TLH	R <mark>W</mark> T <mark>APL</mark> P
					R <mark>WOVPL</mark> PRE
					VWHAPLADSPDRA
					Y <mark>W</mark> S <mark>VPL</mark> DS
VDAREVRLTVLD	NGVGL.GTLSAI	S <mark>GL</mark> RNARR <mark>R</mark> A	TTLGCSFEL	GPRQPR. <mark>GT</mark> SF	VWR <mark>VPL</mark> R

Figure S3. Heme-binding and ligand-binding motifs. Multiple sequence alignments between putative two-component sensory protein-histidine kinases from different organisms are shown. Accession numbers of proteins in Swiss-Prot/TrEMBL (http://us.expasy.org/sprot) or NCBI (http://www.ncbi.nlm.nih.gov) databases are given next to the name of the host organism (M. tuberculosis, Mycobacterium tuberculosis; M. smegmatis, Mycobacterium smegmatis; M. vanbaalenii, Mycobacterium vanbaalenii; M. gilvum, Mycobacterium gilvum, M. ulcerans, Mycobacterium ulcerans; M. bovis, Mycobacterium bovis; N. farcinica, Nocardia farcinica; T. fusca, Thermobifida fusca; S. coelicolor, Streptomyces coelicolor, S. tropica, Salinispora tropica). Secondary structure annotation and residue numbering is according to GAF_{DosT} of *M. tuberculosis*. The α helices are represented by *spirals* and β -strands by *arrows*. Alignments were performed using the MAP algorithm as implemented in the BCM Search Launcher (http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html). Residues constituting the proximal heme binding motif as deduced from the crystal structure are marked with blue triangles. Residues constituting the distal ligand-binding pocket are marked with green squares. Residues making up the surface crevice are marked with *pink triangles*. Invariant H147 and Y169 are highlighted in *cyan*. Highly conserved H392 implicated as the site of autophosphorylation is highlighted in *green*. The image was prepared by using program ESPript (http://espript.ibcp.fr/ESPript/ESPript/).