Supplementary Material

Structure of the entire cytoplasmic portion of a sensor histidine kinase protein

Alberto Marina, Carey D. Waldburger and Wayne A. Hendrickson

Materials and methods

Cloning and mutagenesis. The predicted cytoplasmic domain of *Thermotoga maritima* ORF TM0853 (residues 232-489) was cloned from the *T. maritima* genomic DNA into the pET24b vector (Novagen) using PCR primers that contain the *Nco*I and *Blp*I restriction sites. The resulting plasmid, pHK853CD, has an initiating ATG codon replacing to codon 232 of *HK853*. The full-length ORF TM0853 was cloned in similar fashion into the *NdeI* and *BlpI* restriction sites of the pET28b vector (Novagen). The resulting plasmid, pHK853-His, fuses an N-terminal 6-histidine tag separated by a thrombin digestion site to the initial HK853 methionine residue.

pHK853 is a pBR322-derived plasmid used to express the full length HK853 protein for *in vivo* phosphotransfer assays. This plasmid was constructed by cloning the 2.0 kb *Nde*I-*Ssp*I fragment from pHK853-His into the 3.5 kb *Nde*I-*Nru*I backbone of pLQ3[N] (Regelmann *et al*, 2002). In the resulting plasmid *phoQ* is removed and the *lacUV5* promoter is fused to the *HK853-His* gene. pLQ3[N] (Regelmann *et al*, 2002) is a pBR322-derived plasmid in which *phoQ* is under control of the *lacUV5* promoter and pEnvZ (Hsing and Silhavy, 1997) is a pBR322-derived plasmid containing *envZ*. pNL2 (Lesley and Waldburger, 2001) is a low-copy-number pSC101-derived plasmid that carries the PhoP-activated *phoN-lacZ* fusion (Waldburger and Sauer, 1996).

Site-directed mutagenesis of pHK853CD was carried out using the PCR-based QuikChange system (Stratagen) and primers summarized in Table SI.. The methionine substitutions were introduced to increase selenium anomalous scattering signals since HK853-CD possesses only a single internal methionine residue, besides the new N-terminal methionine. All constructed DNA vectors were confirmed by DNA sequencing (Columbia University Biopolymer Facility, New York; and Servicio de Secuenciación del Laboratorio de Estudios Genéticos, Fundación de Investigación HCUV, Valencia) .

TM0853 full-length localization in E. coli.

E. coli strain BL21-codon plus(DE3)- RIL (Stratagen) transformed with pHK853-His was grown at 20, 25, 30 or 37 ºC in LB media supplemented with 100 μg/ml ampicillin and 33 μg/ml chloramphenicol. His6-HK853 expression was induced at an OD600 of ~0.5 by the addition of 1mM IPTG for up to 3, 4, 5 and 6 hours for 37, 30, 25 and 20 ºC cultures, respectively. Cells were harvested, washed with buffer S (50 mM sodium phosphate pH 8.0, NaCl 0.5 M and 1 mM phenylmethylsulfonyl fluoride) and lysed by sonication. Cellular debris and possible inclusion bodies were eliminated as described (Knaust and Nordlund, 2001). The soluble and membrane fractions were separated by ultracentrifugation at 200000 g for 1h, and the membrane fraction was brought to equal volume of soluble fraction with buffer S using a homogenizer. The level of His-HK853 in each fraction was quantified by pull-down assay as follows. Equal volumes (100 μl) of the soluble and membrane fractions for each temperature were diluted to 500 μl with buffer S and 10 μl of Ni-Chelating Sepharose beads (Amersham Biosciences) equilibrated in buffer S were added. The mixture was incubated 3 h at 4°C with rotation, pulled down by centrifugation and washed four times

with buffer S containing 25 mM imidazole. The proteins bound to the resin were eluted with buffer S containing 0.5 M imidazol and separated by 12 % SDS-PAGE. The gels were stained with Sypro Ruby (Molecular Probes), read with a Fluoro Image Analyzer FLA-5000 (Fuji) and HK853 was quantified using the Multi Gauge 2.0 software.

Protein purification. For isolation of HK853-CD, *E. coli* strain BL21 containing pHK853CD was grown in 2L of LB broth at 37 °C. Induction was performed at an OD_{600} of ~0.5 with 1mM isopropyl-β-D-thiogalactopyranoside for up to 3 h. Cells were harvested, lysed by sonication, and the debris was removed by centrifugation. The supernatant was heated to 70°C for 10 min and denatured material was removed by centrifugation at 20000 g for 20 min. HK853-CD was purified from the soluble fraction by ammonium sulfate fractionation, ion-exchange chromatography (MonoQ) and size exclusion chromatography (Superdex75), concentrated with Ultrafree (Milipore) to ~ 20 mg/ml, and stored at -70 °C. The typical yield was 15-20 mg of purified protein per liter of bacterial culture.

Mutant HK853CD forms were expressed in 1l of LB broth and purified following the previous protocol. The I370M/V373M HK853-CD double mutant selenomethionine-containing protein was produced by a nonauxotrophic protocol (Doublie, 1997) and purified as above. Complete selenomethionine replacement was confirmed by mass spectrometry.

Autokinase assays. The autokinase activity of HK853-CD and its mutants (2-5 μM) were measured in 25 mM Tris pH 8.0, 100 mM KCl, and 2 mM $MgCl_2$ at 37 ºC. Reactions were initiated upon addition of [γ-^{32}P] ATP (0.1 mM, 0.1 mCi ml^{-1}). At specific time intervals, aliquots were removed and quenched by addition of SDS-PAGE

sample buffer containing 50 mM ethylenediaminetetraacetic acid (EDTA). The samples were then subjected to gel electrophoresis. Phosphorylated protein was read using a Fluoro Image Analyzer FLA-5000 (Fuji) and quantified using the Multi Gauge 2.0 software. Although the autokinase reaction does not strictly follow Michaelis-Menten steady state kinetics because phosphorylation of protein effectively reduces the concentration of the enzyme-substrate complex, we believe that the initial velocities calculated are still useful in comparing mutant and wild type variants.

β-galactosidase assays. Phosphotransfer to PhoP or OmpR was assayed *in vivo* by measuring β-galactosidase activity from a phospho-PhoP-activated *phoN-lacZ* fusion or a phospho-OmpR-activated *ompC-lacZ* fusion, respectively. Strains used were *E. coli* CSH26ΔQ [K12 *ara* Δ*(lac-pro) thi PhoQ*] (Waldburger and Sauer, 1996) containing the pNL2 reporter plasmid (*phoN-lacZ*) (Lesley and Waldburger, 2001) for assessing phosphotransfer to PhoP and *E. coli* 5137 [MC4100 Φ(*ompC'-lacZ*$^{+}$) *envZ*::Tn10] (Slauch and Silhavy, 1989) to assess phosphotransfer to OmpR. β-galactosidase activities of CSH26ΔQ and 5137 *E. coli* strains transformed with pHK853, pLQ3[N], pEnvZ, or pBR322 were determined as described (Lesley and Waldburger, 2003). For each evaluation, β-galactosidase assays were performed in triplicate from cultures derived from two independent colonies. The error bars reflect the standard deviations of measurements on the six independents samples for data.

Crystallization and data collection. Crystals of the HK853-CD in 4mM AMPPNP were grown at 20 ºC by vapor diffusion in hanging drops containing equal volumes of protein solution (10 mg/ml) and reservoir solution (100 mM cacodylate (pH 6.5), 1.25 M Li_2SO_4 and 50-200 mM ammonium acetate). Crystals were in space group $C222_1$ with

unit cell dimensions of a= 79.3 Å, b= 162.1 Å, c= 42.5 Å. For cryo-experiments, crystals were flash frozen in mother liquor plus 7.5 % ethylene glycol and 15 % sucrose. Crystals of selenomethionyl I370M/V373M HK853-CD in AMPPNP grew under similar conditions. All X-ray datasets were collected at 100K at the NSLS Beamline X4A on a CCD Quantum 4 detector. The diffraction data were processed and reduced with Denzo and Scalepack (Otwinowski and Minor, 1997). Statistics for all used data sets are given in Table II.

Nucleotide hydrolysis analysis. The nucleotide content of crystals and the AMPPNP hydrolysis capacity of HK853-CD were analyzed by FPLC as described (Lima *et al*, 1997). Adenosine nucleotides run as standards showed retention times that decreased stepwise from AMP to ADP to ATP with AMPPNP at a position intermediate between ADP and ATP in keeping with its reduced negative charge relative to ATP.

A number of crystals (~10) were washed several times in 250 µl of crystallization reservoir buffer and dissolved in 200 µl of 9 M urea. Protein was eliminated by filtration using microcon 3K (Amicon). The filtrate nucleotide showed a retention time between that of ADP and AMP and shifted from ADP by an amount analogous with the shift of AMPPNP from ATP. This suggests that the nucleotide retained in the crystal is the hydrolysis product ADPβN.

Nucleotide hydrolysis reactions were conducted by incubation of 50 µg of HK853-CD in 20 µl of 1.25 M Li_2SO_4, 100 mM ammonium acetate and 100 mM cacodylate pH 6.5 or Tris-HCl pH 8.0 with Mg^{2+}ATP or AMPPNP 4mM as substrates at 20 ºC for 24 h. Controls in absence of protein were carrying out in parallel. The reaction was stopped by adding 180 µl of 10 M urea and filtrate nucleotides were analyzed as above. Nucleotide retention times after incubation of ATP with and without enzyme

were identical with those for the ADP and ATP standards, but the products from incubations of AMPPNP were the same in both cases and identical with that of the putative ADPβN product extracted from the crystals.

Structural determination and refinement. The structure was solved at 2.1 Å resolution using data from a four-wavelength MAD experiment on a single crystal of the selenomethionyl I370M/V373M HK853-CD double mutant. Coordinates for all four possible selenium sites were found by SOLVE (Terwilliger and Berendzen, 1999) giving a mean figure of merit <m>=0.55 and Z-score=12.09 and refined with SHARP (de la Fortelle and Brigogne, 1997). Initial MAD phases to 2.1 Å spacings were improved and extended to 1.9 Å by density modification using histogram matching and solvent flattening against the native HK853-CD data set as implemented in program DM (CCP4, 1994). The quality of phases enabled WarpNtrace (Perrakis *et al*, 1999) automatically to build and refined 182/257 visible residues including side chain positions. The structure was completed after cycles of interactive manual building with O (Jones and Kjeldgaard, 1997) and computational refinement with CNS (Brunger *et al*, 1998). The final model consists of 240 residues out of a total of 257 residues, five with double conformation (residues 291,293,440,441 and 442), a hydrolyzed molecule of AMPPNP, 179 water molecules and one sulfate ion**.** Stereochemistry checks indicate that the refined model is in quite good agreement with expectations for models within this resolution range (Laskowski, 1993). The refinement statistics are summarized in Table I. Atomic coordinates and structure factors have been deposited into the Protein Data Bank with accession code 2C2A.

Structure and sequence comparisons. The least-squares superimpositions were calculated using LSQMAN (Kleywegt, 1996). The overall best fits between the structures were determined using the 'Brute' option in the program with a cut-off distance criterion of 3.5 Å and a minimum fragment length of three consecutive residues. The coordinates were taken from the Protein Data Bank with entry codes: CheA-KD, 1B3Q (Bilwes *et al*, 1999) PhoQ-KD, 1IDO (Marina *et al*, 2001) and NtrB-KD, 1R62 (Song *et al,* 2004);. Sequence alignments among these proteins were based on structural superimposition. The multiple alignments of HK853 with the histidine kinase DHp and CA domains were taken from Pfam database [http://www.pfam.wustl.edu] with accession number PF02895 and PF02518, respectively.

References

Bilwes AM, Alex LA, Crane BR, Simon MI (1999) Structure of CheA, a signal-transducing histidine kinase. *Cell* **96**: 131-141

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**: 905-921

Collaborative Computational Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallog*. **D 50**: 760-763

de la Fortelle E, Bricogne G (1997) Maximum-likelihood heavy-atom parameter refinement for the Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction methods. *Methods Enzymol* **276:** 472-494

Doublie S (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* **276**:523-530

Hsing W, Silhavy TJ (1997) Function of conserved histidine-243 in phosphatase activity of EnvZ, the sensor for porin osmoregulation in *Escherichia coli. J Bacteriol* **179**: 3729-3735

Jones TA, Kjeldgaard MO (1997) Electron-Density Map Interpretation. *Methods Enzymol* **277**: 173-208

Kleywegt GJ (1996) Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallog D***52**:842-857

Knaust RK, Nordlund P (2001) Screening for soluble expression of recombinant proteins in a 96-well format. *Anal Biochem* **297:** 79-85

Laskowski RA, Moss DS, Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**: 1049-1067

Lesley JA, Waldburger CD (2001) Comparison of the *Pseudomonas aeruginosa* and *Escherichia coli* PhoQ sensor domains: evidence for distinct mechanisms of signal detection. *J Biol Chem* **276**: 30827-30833

Lima CD, Klein MG, Hendrickson WA (1997) Structure-based analysis of catalysis and substrate definition in the HIT protein family. *Science* **278**:286-290

Marina A, Mott C, Auyzenber A, Hendrickson WA, Waldburger CD (2001) Structural and mutational analysis of PhoQ histidine kinase catalytic domain. *J Biol Chem* **276**: 41182-41190

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **277**: 307-326

Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**: 458-463

Regelmann AG, Lesley JA, Mott C, Stokes L, Waldburger CD (2002) Mutational analysis of the Escherichia coli PhoQ sensor kinase: differences with the *Salmonella enterica* serovar Typhimurium PhoQ protein and in the mechanism of Mg2+ and Ca2+ sensing. *J Bacteriol* **184**: 5468-5478

Slauch JM, Silhavy TJ (1989) Genetic analysis of the switch that controls porin gene expression in Escherichia coli K-12. *J Mol Biol* 210:281-292.

Song Y, Peisach D, Pioszak AA, Xu Z, Ninfa AJ (2004) Crystal structure of the C-terminal domain of the two-component system transmitter protein nitrogen regulator II (NRII; NtrB), regulator of nitrogen assimilation in Escherichia coli. *Biochemistry* **43**: 6670-6678

Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D* **55**: 849-861

Waldburger CD, Sauer RT (1996) Signal detection by the PhoQ sensor-transmitter. Characterization of the sensor domain and a response-impaired mutant that identifies ligand-binding determinants. *J Biol Chem* **271**: 26630-26636

Table SI. Oligonucleotides used in mutant construction

Name oligo	Sequence
HK853-F312C-1	5′-CTGGACTGTTCGAGGCTTGAAAGGAAATCC-3′
HK853-F312C-2	5′-CAAGCCTCGAACAGTCCAGTAACTCATTCAGG-3′
HK853-L315A-1	5′-CTGGACTTTTCCCGGGCTGAAAGGAAATCCCTTC-3′
HK853-L315A-2	5′-GGGATTTCCTTTCAGCCCGGGAAAAGTCCAGTAACTC-3′
HK853-L315C-1	5′-CTTTTCCAGATGTGAAAGGAAATCCCTTC-3′
HK853-L315C-2	5′-CCTTTCACATCTGGAAAAGTCCAGTAACTC-3′
HK853-L315W-1	5′-CTGGACTTTTCCAGATGGGAAAGGAAATCCCTTC-3′
HK853-L315W-2	5′-GGGATTTCCTTTCCCATCTGGAAAAGTCCAGTAACTC-3′
HK853-L320P-1	5′-GGAAATCGCCGCAAATAAACAGAGAAAAAG-3′
HK853-L320P-2	5′-GTTTATTTGCGGCGATTTCCTTTCAAGCC-3′
HK853-Q321P-1	5′-GGAAATCTCTTCCAATAAACAGAGAAAAAGTTG-3′
HK853-Q321P-2	5′-CTGTTTATTGGAAGAGATTTCCTTTCAAGCC-3′
HK853-I370M/V373M-1	5′-CAAGAATGCGACAGATGATTTTGAATCTTCTCAAC-3′
HK853-I370M/V373M-2	5′-CAAAAGCATCTGTCGCATTCTTGTTGGATCGATG-3′
HK853-F428A-1	5′-GGATATTCGAACAGGCCTACAGGGTGGATTCTTCG-3′
HK853-F428A-2	5′-GAATCCACCCTGTAGGCCTGTTCGAATATCCTATCC-3′
HK853-F428E-1	5′-GGATATTCGAACAGGAGTACAGGGTGGATTCTTCG-3′
HK853-F428E-2	5′-GAATCCACCCTGTACTCCTGTTCGAATATCCTATCC-3′
HK853-F428W-1	5′-GGATATTCGAACAGTGGTACAGGGTGGATTCTTCG-3′
HK853-F428W-2	5′-GAATCCACCCTGTACCACTGTTCGAATATCCTATCC-3′
HK853-L444A-1	5′-GCCTGGAACGGGAGCCGGCCTTGCCATAACAAAAG-3′
HK853-L444A-2	5′-GTTATGGCAAGGCCGGCTCCCGTTCCAGGCACTTCG-3′
HK853-L444C-1	5′-GAACGGGATGCGGCCTTGCCATAACAAAAG-3′
HK853-L444C-2	5′-CAAGGCCGCATCCCGTTCCAGGCACTTCG-3′
HK853-L444W-1	5′-GCCTGGAACGGGATGGGGCCTTGCCATAACAAAAG-3′
HK853-L444W-2	5′-GTTATGGCAAGGCCCCATCCCGTTCCAGGCACTTCG-3′
HK853-I448A-1	5′-CTCGGCCTTGCAGCAACAAAAGAGGATCGTGG-3′
HK853-I448A-2	5′-GATCTCTTTTGTTGCTGCAAGGCCGAGTCCCG-3′
HK853-I448W-1	5′-CTCGGCCTTGCCTGGACAAAAGAGGATCGTGG-3′
HK853-I448W-2	5′-GATCTCTTTTGTCCATGCAAGGCCGAGTCCCG-3′

The underlined nucleotides introduced the mutations

Table SII. Effects of Mutation on Autophosphorylation Activity

Mutant Variant	Initial Rate (% phosphorylation / min)	Initial Rate (relative to wild type)	Equilibrium (% phosphorylation)
Type I			
WT	2.78 ± 0.19		8.0 ± 1.3
L315A	4.66 ± 0.55	1.7	14.1 ± 2.0
L315W	9.15 ± 0.68	3.3	10.7 ± 0.7
F428A	5.25 ± 0.12	1. 9	7.9 ± 0.9
F428E	7.56 ± 0.60	2.7	10.8 ± 1.3
F428W	4.52 ± 0.42	1.6	10.1 ± 0.9
L444A	3.30 ± 0.24	1.1	7.7 ± 0.5
L444W	4.23 ± 0.23	1.5	8.6 ± 0.4
I448A	16.60 ± 1.13	6.0	16.9 ± 2.5
I448W	0.05 ± 0.008	0.01	nd
Type II			
WT2.36 ± 0.10			
WT (DTT)	2.37 ± 0.13	1.0	
F312C/L444C	0.03 ± 0.005	0.01	nd
L315C/L444C	0.04 ± 0.002	0.02	nd
F312C/L444C (DTT)	5.49 ± 0.19	2.3	nd
L315C/L444C (DTT)	6.84 ± 0.41	2.9	nd
Type III			
WT2.26 ± 0.10			
L320P	> 28 *	>12.4*	8.5 ± 0.9
Q321P	5.37 ± 0.24	2.4	nd