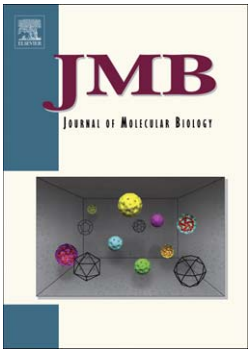


Structural insights into the HWE histidine kinase family: The Brucella blue light-activated histidine kinase domain

Jimena Rinaldi, Mehrnoosh Arrar, Gabriela Sycz, María Laura Cerutti, Paula M. Berguer, Gastón Paris, Darío Ariel Estrín, Marcelo Adrián Martí, Sebastián Klinke, Fernando Alberto Goldbaum

PII: S0022-2836(16)00081-4
DOI: doi: 10.1016/j.jmb.2016.01.026
Reference: YJMBI 64980

To appear in: *Journal of Molecular Biology*

Received date: 17 December 2015
Revised date: 22 January 2016
Accepted date: 27 January 2016

Please cite this article as: Rinaldi, J., Arrar, M., Sycz, G., Cerutti, M.L., Berguer, P.M., Paris, G., Estrín, D.A., Martí, M.A., Klinke, S. & Goldbaum, F.A., Structural insights into the HWE histidine kinase family: The Brucella blue light-activated histidine kinase domain, *Journal of Molecular Biology* (2016), doi: 10.1016/j.jmb.2016.01.026

Structural insights into the HWE histidine kinase family: The *Brucella* blue light-activated histidine kinase domain

Jimena Rinaldi[1], Mehrnoosh Arrar[2], Gabriela Sycz[1], María Laura Cerutti[1,3], Paula M. Berguer[1], Gastón Paris[1+], Darío Ariel Estrín[2], Marcelo Adrián Martí[4], Sebastián Klinke[1,3], Fernando Alberto Goldbaum[1,3]

[1]Fundación Instituto Leloir, IIBBA–CONICET, Avenida Patricias Argentinas 435, C1405BWE, Buenos Aires, Argentina

[2]Departamento de Química Inorgánica, Analítica y Química Física e INQUIMAE-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires , Ciudad Universitaria, Pabellón 2, C1428EHA, Buenos Aires, Argentina

[3]Plataforma Argentina de Biología Estructural y Metabolómica PLABEM, Avenida Patricias Argentinas 435, C1405BWE, Buenos Aires, Argentina

[4]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires , Ciudad Universitaria, Pabellón 2, C1428EHA, Buenos Aires, Argentina

[+]Present address: Centro de Investigaciones en Bionanociencias, CIBION-CONICET, Godoy Cruz 2390, C1425FQD, Buenos Aires, Argentina

Contact information

Fernando Alberto Goldbaum
fgoldbaum@leloir.org.ar
Fundación Instituto Leloir
Av. Patricias Argentinas 435
C1405BWE, Buenos Aires, Argentina
Phone:+054-11-5238-7500, ext 2304
Fax: +054-11-5238-7501

Abstract

In response to light, as part of a two-component system, the Brucella blue light-activated histidine kinase (LOV-HK) increases its autophosphorylation, modulating the virulence of this microorganism. The Brucella histidine kinase (HK) domain belongs to the HWE family, for which there is no structural information. The HWE family is exclusively present in proteobacteria and usually coupled to a wide diversity of light sensor domains. This work reports the crystal structure of the Brucella HK domain, which presents two different dimeric assemblies in the asymmetric unit: one similar to the already described canonical parallel homodimers (C) and the other, an antiparallel non-canonical (NC) dimer, each with distinct relative subdomain orientations and dimerization interfaces. Contrary to these crystallographic structures and unlike other HKs, in solution the Brucella HK domain is monomeric and still active, showing an astonishing instability of the dimeric interface. Despite this instability, using cross-linking experiments, we show that the C dimer is the functionally relevant species. Mutational analysis demonstrates that the autophosphorylation activity occurs *in cis*. The different relative subdomain orientations observed for the NC and C states highlight the large conformational flexibility of the HK domain. Through the analysis of these alternative conformations by means of molecular dynamics simulations, we also propose a catalytic mechanism for Brucella LOV-HK.

Keywords

Signal transduction; two-component system; X-ray crystallography; molecular dynamics simulations; autophosphorylation mechanism

Abbreviations

LOV-HK, blue light-activated histidine kinase; HK, histidine kinase; C, canonical; NC, non-canonical; SHK, sensor histidine kinase; DHp, dimerization/histidine phosphoacceptor subdomain; CA, catalytic and ATP-binding subdomain; LOV, blue-light sensor; ACP, AMP-PCP; S-SAD, sulfur single-wavelength anomalous dispersion method; SEC-SLS, static light scattering assay coupled to size-exclusion chromatography; MW, molecular weight; DLS, dynamic light scattering.

Highlights

- The crystal structure of the *Brucella* HK domain reveals two distinct dimers.
- In solution the *Brucella* HK domain is an active monomer.
- The canonical dimer is the functionally relevant species.
- The autophosphorylation occurs *in cis*.
- A catalytic mechanism for *Brucella* LOV-HK is proposed.

Introduction

Bacteria make use of two-component systems in order to sense and adapt to changing environmental conditions. The signaling process starts by signal detection, which results in the autophosphorylation of a conserved histidine residue in the sensor histidine kinase (SHK). Subsequent transfer of the phosphoryl group from the histidine of the histidine kinase (HK) to a conserved aspartate residue of its cognate response regulator propagates the signal and triggers the cellular response [1, 2]. SHKs are modular proteins that typically assemble as homodimers [3], although a monomeric SHK has also been recently identified [4]. SHKs bear at least a sensor domain connected to a HK domain, whereas they can increase in complexity with multiple domains on the same polypeptide chain presenting considerable architectural variety across the superfamily. There are several well established models of HK that have been thoroughly studied, namely the dimeric HK853 from *Thermotoga maritima* [5, 6], EnvZ from *Escherichia coli* [7, 8], DesK from *Bacilus subtilis* [9, 10], CpxA from *E. coli* [11] and the monomeric EL346 from *Erythrobacter litoralis* [4].

The HK domain is comprised of the dimerization/histidine phosphoacceptor subdomain (DHp) and the catalytic and ATP-binding subdomain (CA), which are connected by a flexible linker. The DHp subdomain, with two helices (α1 and α2) connected by a hairpin loop, forms a homodimeric four-helix bundle. The catalytic histidine residue is located on the solvent-exposed side of the α1 helix. The CA subdomain is a highly conserved α/β sandwich, called the Bergerat fold, with three α helices packed against five antiparallel β strands. The nucleotide binds between two helices and is held by a loop known as the ATP lid. The HK superfamily includes several conserved motifs: the H box corresponding

to the region flanking the phosphorylatable histidine residue and the N, G1, G2, G3 and F boxes, which comprise conserved nucleotide-binding sequences [12]. Autophosphorylation occurs when the ATP molecule bound to the CA phosphorylates the histidine residue in the DHp. The available structural evidence shows that the relative orientations of these subdomains greatly impact the proximity of the bound ATP to the histidine phosphoacceptor, however the mechanistic details of the activation of the HK domain as an autokinase remain unclear.

HK dimerization implies that the directionality of the autophosphorylation reaction can proceed either in a *cis* (intramolecular) or in a *trans* (intermolecular) manner [5, 13, 14]. Biochemical data have confirmed the existence of both *cis* and *trans* mechanisms in different HKs [5, 10, 13, 15]. It has been postulated and proven for several HKs that the *cis vs. trans* autophosphorylation propensity stems from the intrinsic handedness of the hairpin loop between the DHp helices (α1-α2 loop) [16, 17]. If this loop turns right (looking from the N-terminus to the C-terminus of α1), the catalytic domain of one monomer is believed to be closer to the catalytic histidine residue of the dimeric partner, and autophosphorylation proceeds *in trans*. In the alternative case, the loop turns left and the catalytic domain is closer to its own histidine residue, and autophosphorylation occurs *in cis*.

The blue light-activated SHK (LOV-HK) discovered in *Brucella abortus*, the causative agent of brucellosis, was identified to be crucial in its pathogenesis [18, 19] and is involved in the general stress response system of this microorganism [20]. *Brucella* LOV-HK comprises an N-terminal blue-light sensor (LOV) domain, followed by a central PAS domain and a C-terminal HK domain (Figure 1A).

The HK domain from *Brucella* LOV-HK belongs to a bioinformatically identified family called HWE [21]. There is no structural information available for this family. The HWE HK domains are present in LOV-HKs and bacteriophytochromes, among other proteins. Members of this family differ from most others by the absence of a recognizable F box and the presence of several uniquely conserved residues, including a histidine residue in the N box and a Trp-X-Glu sequence in the G1 box, which was used to define the family. A tyrosine residue can also be found at the position of the histidine residue in the N box, as it is the case for the HK domain from *Brucella* LOV-HK (see Figure S5).

In order to gain insight into the structure of the HWE family and to elucidate the mechanism of light activation, in this work we performed an in depth structural and functional study of the HK domain from *Brucella* LOV-HK by means of X-ray crystallography, molecular modeling and in-solution biophysical and biochemical approaches. *Brucella* HK forms a canonical homodimer with unusual instability at the dimer interface, presenting an *in cis* autophosphorylating activity.

Results

The crystal structure of the HK domain from *Brucella* LOV-HK shows two different dimeric assemblies.

We designed a HK construct that is truncated immediately after the PAS domain, comprising residues 266 to 489 (Figure 1A). The protein was crystallized in complex with the ATP analogue AMP-PCP (ACP) and Mg^{2+} and its structure was solved at 2.51 Å resolution by means of the sulfur single-wavelength anomalous dispersion method (S-

SAD) in the $P2_1$ space group [22]. The final model was refined to R = 0.202 and R_{free} = 0.242 with very good stereochemistry (Table 1). Ramachandran statistics show that over 95% of the residues lie in the most favored region of the plot. Four polypeptide chains were found in the asymmetric unit of the crystal, forming two types of dimers (Figure S1 and Figure 1B and C). Residues 480-489 from chain A, 266-277, 307-309, 412-413 and 480-489 from chain B, 266-308 and 480-489 from chain C and 266-308, 433-435 and 478-489 from chain D could not be located in the electron density and were excluded from the crystallographic model. Additionally, several residues show weak or missing electron density for their side chains and were therefore modeled only up to their beta carbon atoms. One ACP molecule was found for each chain, whereas Mg^{2+} was found only associated with chains A and B.

The dimer formed by chains A and B corresponds to a canonical (C) HK dimer, consisting of the parallel four helix bundle known as the DHp comprising α1 and α2, and the CA, which consists of α3-α5 and β1-β7 (Figure 1B). Additionally, chain A includes an extra three-residue long β strand (β-1), which corresponds to the C-terminus of the PAS domain. Chain B lacks electron density for strands β-1 and the N terminus of α1 (266-278). The other dimeric conformation, which is composed of chains C and D, will be referred to as the “non-canonical” (NC) dimer and consists of an antiparallel dimer. The β-1 strand and the whole α1 helix (residues 266-308) are missing, presumably due to an enhanced flexibility in the protein at the α1- α2 loop. All helices (α2-α5) are involved in dimerization. The α2 helices from each monomer are located adjacent to each other in an antiparallel orientation (Figure 1C).

The monomers belonging to the same dimer (chains A-B and chains C-D) are nearly identical to each other, with C^{α} r.m.s.d. values of 0.45 and 0.70 Å, respectively. For this

reason from now on chains A and B will be referred to as “canonical monomers” and chains C and D will be called “non-canonical monomers”.

The CA subdomains from the C and NC monomers are very similar - C^{α} r.m.s.d. values range from 0.99 to 1.18 Å- (Figure S2), whereas the relative orientations of the CA with respect to the DHp differ significantly (Figure 2A). First, the long axis of the CA is nearly parallel to α2 in the C monomer and approximately orthogonal in the NC monomer. Second, the CA embraces either the outer or inner side of the α2 helix, in the C or NC monomer, respectively. We use the term “inner” to describe the side of the α2 helix that faces the four-helix bundle in the C dimer, and similarly describe the opposite side of the helix as “outer”. This difference is driven by the fact that the α2-α3 loop, which hinges the CA to the principal helical axis, adopts two alternative conformations. The alternation of the α2-α3 loop configurations corresponds to different torsion angles adopted by the residues Glu335 and Trp337 in the C and NC forms as observed in the Ramachandran plot (Figure 2B).

The HK domain is a monomer in solution and presents autokinase activity.

In order to determine the oligomeric state of the HK domain in solution, we performed a static light scattering assay coupled to size-exclusion chromatography (SEC-SLS). The protein eluted as a single and symmetric peak with a molecular weight (MW) of 26 kDa, which is compatible with the expected value for the monomer. This experiment showed that contrary to the two dimeric assemblies observed in the crystal structure and those of other reported HKs, the HK domain from *Brucella* LOV-HK is monomeric in solution (Figure 3A). However, a marginal tendency to dimerize is observed in static and dynamic light scattering (DLS) experiments. SEC-SLS experiments with increasing amounts of

protein show a displacement of the unique peak to smaller elution volumes with larger average MW (Figure S3A). The hydrodynamic diameter of the HK domain estimated by DLS corresponds to an intermediate value between the one expected for the monomer and the one expected for the C dimer and increases with the protein concentration (Figure S3B). In order to have a measured value of the dimer a stabilized C dimer (mutant I286C, see below) was included in the experiment.

According to the dimerization interface area and the nature of the contacts analyzed by the PISA server (see interface area in Table S1) both the C and the NC dimers may exist in solution. However, a comparative analysis indicates that both dimerization interfaces may be less stable than those of other dimeric HKs (see ΔG^{diss} in Table S1). In agreement, the C dimer also appears to be less stable compared to other HKs in molecular dynamics simulations (see discussion).

We analyzed whether the HK domain was able to present autokinase activity. For this purpose, we performed an activity assay using radio-labeled ATP. The results show that the HK domain is active (Figure 3B). The results allow two possible and not exclusive interpretations: the monomer is active or the dimer is active but not stable in solution.

Taken together, these experiments demonstrate that the isolated HK domain is a monomer in solution, with a slight tendency to dimerize and presents autokinase activity.

Cysteine cross-linking experiments show that the C dimer can be stabilized and is active in solution.

In order to evaluate the biological relevance of the two dimeric species observed in the crystal structure, we engineered disulfide bonds exclusively at the dimerization interface of either the C (mutants I286C, V282C and A287C-L333C) or the NC (mutants L322C-

L325C, L441C and F437C-H329C) dimers and evaluated their ability to cross-link, as well as their autophosphorylation activity (Figure 4). According to their retention volumes in SEC experiments and their far UV CD spectra, all mutants fold as the wild type protein (Figure S4). The dimer formation in each case was evaluated by SDS-PAGE with and without the addition of DTT as a reducing agent. SEC-SLS experiments were also performed, in the absence of DTT.

For the three mutants designed to cross-link the C dimer, the main peak in the SEC-SLS experiments corresponds to the dimer (51-56 kDa). Accordingly, a band at the height of the dimer (52 kDa) is observed in the SDS-PAGE gels in the absence of DTT. This band presents autokinase activity. Even in the absence of DTT, some activity at the height of the monomer is present for the I286C and V282C mutants.

On the contrary, for the NC cross-linked mutants the main peak corresponds to a monomer-dimer equilibrium displaced towards the monomer (29-31 kDa) in SEC-SLS experiments. Residual or no band is observed at the height of the dimer in the SDS-PAGE gels. Activity at the height of the monomer was detected for the L322C-L325C and L441C mutants.

Taken together, the cross-linking analysis suggests that a functional HK can be stabilized by specifically cross-linking the C dimer. Additionally, the existence of activity at the height of the monomer in the SDS-PAGE gels strongly suggests that the monomer is active.

The C dimer is functionally relevant.

We noted that the activity of the three C cross-linked dimers, I286C, V282C and A287C–L333C was different. The Ile286 and the Val282 mutations are located in the α1 helix N-terminal to the histidine phosphoacceptor. In these two cases a single mutation to cysteine

was enough to generate a cross-linked dimer as both Ile286 and Val282 are adjacent to the same position in the partner subunit. The third mutant was designed based on the proximity of Ala287 in the α1 helix to Leu333 in the α2-α3 loop. In order to compare the autophosphorylation activity of the three mutants, we isolated the three cross-linked dimers and the wild-type HK domain by gel filtration and measured their activities (Figure 5). We found that one of the disulfide-linked mutants, I286C, was actually more active than the wild-type HK, underlining the biological relevance of the C dimer. The activity of the isolated V282C dimer is lower than the wild-type, followed by the A287C–L333C mutant with the lowest autokinase activity. The diminished activity of the A287C–L333C dimer was anticipated since a disulfide bridge in this position would restrict the flexibility of the α2-α3 loop. The difference in activity between the I286C and V282C dimer, however, was not expected since these residues are just one helical turn apart, both at the dimeric interface of the C dimer.

The fact that a cross-linked C dimer with higher activity than the monomer was identified is a strong evidence of the biological relevance of the C dimer.

The crystallographic C dimer: insights into the HWE HK family

Having established the biological relevance of the crystallographic C dimer we focused on the analysis of its structure.

In the DHp (Figure 6B), formed by α1 and α2 and their connecting loop, most of the interactions at the dimer interface comprise the inner face of the α1 helix contacting either the α1 helix (Val282–Val282´, Ile286–Ile286´, Phe290–Phe290´) or the α2 helix (Phe290–His329´, Lys291–Asp330´) of its partner. The residue Phe318, which is located in the α2 helix, interacts with the same residue from the neighboring molecule. Most of these

positions present a high degree of conservation within the LOV-PAS-HWE_HK proteins (Figure S5). Despite the structural conservation of the four-helix bundle among several well studied HKs, their sequences strongly differ within the superfamily (Figure S6). The α1-α2 loop is 10 residues long (304-313) and left-handed (Figure 6B). The length of this loop, but not its sequence, is conserved within the LOV-PAS-HWE_HK proteins (Figure S5).

The overall structure of the CA is highly conserved, consisting of an α/β sandwich with its three helices (α3, α4 and α5) packed against one of the faces of the β-sheet, which consists of five antiparallel β strands (β2, β4, β5, β6 and β7, see Figure 1B). The nucleotide binding pocket is very well conserved (Figure 6C). In this sense, Asn388 and Lys391 from the N box contact the β-phosphate of the ACP molecule, while Glu384 also from the N box contacts the Mg^{2+} ion. As already described [21] the F box is absent. In *Brucella* HK and in all HWE family the ATP lid, which comprises the F and the G2 boxes, is shorter in comparison with those of the HisKA family (Figure S5, Figure S6 and Figure S7). It is worth noting that DesK and *E. litoralis* LOV-HK also lack the F box (Figure S6). In *Brucella* HK, the phosphate groups bind residues Arg434, Arg435, Gly436, Gly438 and Gln439. Glu423 and Phe458 from the G1 and G3 boxes, respectively, bind the adenine moiety of the ACP molecule. Additionally, Tyr392 interacts with the adenine ring and the ribose moiety of ACP (Figure 6C).

The conserved Trp421 residue, after which the HWE family was named, is located at the β5 strand of *Brucella* HK and interacts with Phe382, Leu385 and Ala386 from the α4 helix. Other HKs present a valine residue at this position (Figure S6), which interacts in a similar way with the α4 helix. This position in the β5 strand seems to have an essential role in the stabilization of the Bergerat fold. In *Brucella* HK, the distance between the β-strand and

the α4 helix is longer than in other HKs, due to the presence of Trp421. Two residues ahead, Glu423, after which the HWE family was also named, interacts with the amine group of the nucleotide adenine ring. The same interaction is observed in other HKs, which have an aspartate residue at this position (Figure S6). The substitutions Val to Trp and Asp to Glu could be correlated, since a longer β5-α4 distance implies also longer distance between the acidic residue and the nucleotide.

Apart from the defined boxes, the LOV-PAS-HWE_HK proteins have two extra conserved motifs located between the H and the N boxes, which we name the R box, containing an invariable arginine residue (Arg321 in *Brucella* HK) and the G0 box. Interestingly, Arg321 from the α2 helix of the DHp contacts Glu384, Tyr383 as well as the γ-phosphate from ACP, closing the binding pocket (Figure 6C).

The C dimer was trapped in the crystal in an inactive conformation, since the distance between the phosphorylatable histidine residue and the γ-phosphate of ACP is 27 Å (Figure 6A). The following interactions between the DHp and the CA stabilize the inactive conformation: Gln303–Arg440, Arg289–Asp372, Asn292–Arg373 and the already mentioned contacts of Arg321 (Figure 6D). In order to evaluate the role of the interactions of Arg321 in the stabilization of the inactive conformation we performed an activity assay on the R321A mutant. We expected an increase of activity or no effect, if the Arg321 interactions are essential or not for the stabilization of the inactive form, respectively. Contrary to our expectation, the R321A has no autophosphorylation activity (Figure S8).

In summary, despite the low sequence similarity of *Brucella* HK with other HK structures present in the PDB, its overall structure is conserved. The structural hallmarks of the HWE family are (a) the instability of the DHp dimerization interface, (b) the shorter ATP lid, which corresponds to the shorter G2 box and the lack of the F box and (c) the presence of the R

box, with a conserved arginine residue in the α2 helix anchoring the ATP in the inactive state.

Autophosphorylation occurs *in cis* in the dimer.

To further probe the activation mechanism of the HK domain, we studied whether autophosphorylation occurred *in cis* or *trans*. Two evidences put on to the *cis*-autophosphorylation: the fact that the monomer is active and the left-handedness of the α1-α2 loop. We evaluated the reconstitution of activity from the mix of two non-active mutants: H288A, lacking the phosphorylatable histidine residue and N388A, which is unable to bind ATP [6, 23, 24]. We expected an onset of activity for a *trans*-autophosphorylation mechanism, as a subunit able to bind the ATP molecule can transfer the phosphate group to the histidine residue of the other subunit. No activity was observed, indicating that the autophosphorylation occurs exclusively *in cis* (data not shown). To confirm this result, based on a classic approach [5, 13, 17] we generated heterodimers formed by a long (L, the HK domain plus a 23 kDa tag, 49kDa) and a short (S, 26 kDa) version of the HK domain to distinguish which subunit within the heterodimer was phosphorylated. Additionally, the short version has the N388A mutation, which impairs its autophosphorylation capacity. Finally, both versions have the I286C mutation in order to control the dimer formation with a reducing agent. The purified long and short HK domains were mixed in equimolar proportions, incubated with DTT and then extensively dialyzed before the phosphorylation assay (Figure 7). Both the long homodimer (98 kDa) and the heterodimer (75 kDa) are expected to be active and the short homodimer (52 kDa) inactive, regardless of the directionality of the autophosphorylation. After adding DTT in the loading buffer, however, the radio-labeled phosphate must be found on the short version if

the phosphotransfer occurred *in trans*, and on the long version if *in cis*. Since only the long monomer was phosphorylated, the results clearly show that the phosphotransfer exclusively occurs *in cis*.

The *cis* activation is compatible with our observation of the active monomer. These results suggest that the dimerization of the HK domain through the DHp subdomain enhances its autokinase activity by specifically influencing the dynamics of the CA.

A catalytic mechanism is proposed through molecular modeling and simulations.

Having established the functional relevance of the C dimer, we considered whether the artefactual NC crystallographic conformation could provide structural insight into the activation of the HK domain. We were able to construct two models of the active monomeric HK, one of which is compatible with the C dimer interface.

The different relative subdomain orientations observed for the NC and C conformations highlight the large conformational flexibility of the HK domain (Figure 2A). As mentioned above, the region responsible for this behavior is the α2-α3 loop (residues 333 to 341), which hinges the CA to the DHp, more specifically the residues Glu335 and Trp337 (Figure 2B). By only transforming Glu335 from its C to its NC conformation, the CA moves from the outer to the inner side of the α2 helix, and with the C-to-NC transformation of Trp337, the CA changes its orientation with respect to the α2 helix (Figure 2C). Importantly, the conformational changes in this loop determine the proximity of the bound ATP to the histidine phosphoacceptor. Therefore, we reasoned that its motion could allow access to the active conformation.

We performed rigid-body rotations of the CA about Trp337, sweeping its dihedral space and monitoring the resulting His288-ATP distance. In this sweep of the Trp337 dihedral

space, we obtained a potentially active conformation, with the ATP γ-phosphate within 7 Å of the histidine phosphoacceptor, and the CA perpendicular to the principal helical axis (Figure 8A, left). Consistent with the *in cis* activity of the HK, this model illustrates an intramolecular conformational change of a single monomer of the crystallographic C dimer. Interestingly, the modeled active C monomer is essentially a mirror image of the NC monomer. The key difference between the two conformations is that the CA wraps around the outer side of the α2 in the C conformation instead of around the inner side of the α2 in the NC conformation. Indeed, we modeled the missing α1 helix of the crystallographic NC monomer, using its crystal neighbor α2 as a template, and the ATP γ-phosphate was again positioned within 7 Å of the histidine phosphoacceptor, resulting in an alternative active monomeric conformation (Figure 8A, right). Thus, in the monomeric species the CA domain may wrap around either the outer or inner side of the α2 and approach the His288 phosphoacceptor.

To evaluate both potentially active conformations, we performed a total of 0.5 microseconds of classical molecular dynamics simulations (MD), in five independent trials. We considered only the monomeric active models because the crystallographic C dimer was not particularly stable in MD simulations (Figure S9), supporting the tendency to dissociate, observed in solution. For both active models we specifically monitored their stability (Figure S9) as well as the distance between the γ-phosphate and the histidine phosphoacceptor (Figure 8B). To our surprise, we found that the CA in both cases was able to further approach the histidine phosphoacceptor, exploring potential catalytic conformations with reactive distances of 4 Å. We note that the reactive distance in a comparable simulation of the histidine kinase CpxA, crystallized in the active conformation is 3.9 ± 0.9 Å.

Taken together, the MD simulations show that, in the monomeric HK domain, the phosphotransfer may occur by the CA wrapping around either side of the α2 helix. We should point out that the modeled active NC monomer, with the CA wrapped around the inner side of the α2 helix would disrupt the formation of the four-helix bundle at the canonical dimeric interface, whereas the transformation about Trp337 that brought the bound ATP molecule within 7 Å of the phosphoacceptor in the modeled active C monomer is permitted in the canonical dimer context. Based on this model we propose a catalytic mechanism (Figure 9 and see discussion below).

Discussion

Here we report the crystal structure of the HK domain from *Brucella* LOV-HK, the first determined structure from the HWE family of histidine kinases. Two different dimeric assemblies were observed in the crystal: the parallel C dimer and an antiparallel NC dimer, each of which capture different CA-DHp relative orientations. Contrary to these crystallographic structures and unlike other HKs, we found that the *Brucella* HK domain is monomeric in solution. Using cross-linking mutagenesis we demonstrated that the C dimer interface is functionally relevant and when stabilized, enhances its *in cis* activity. By means of MD and exploring the different CA-DHp relative orientations, we also propose a catalytic mechanism.

The crystal structure of *Brucella* HK, together with in-solution biophysical assays unveiled as a structural hallmark of the HWE family the instability of the DHp. In agreement, we note that MD simulations of the HK C dimer revealed a great deal of plasticity in the 10-15 N-terminal residues of the α1 helices at the dimer interface (Figure S9). In contrast, the

DHp of another truncated HK domain of *T. maritima* HK853 retain a stable dimeric interface on simulations of the same timescale (data not shown). Contrary to the general understanding of the structure of HKs, the DHp dimerization interface in *Brucella* LOV-HK evidently does not suffice to preserve the dimeric assembly in solution. Because the canonical dimeric interface could be stabilized using cross-linking, and, once stabilized, could even augment the autokinase activity, we believe that the C dimer is the biologically relevant species, being the monomer an artifact derived from working with the isolated HK domain. We envision that the LOV and PAS domains play important roles in stabilizing the dimeric HK. Moreover, since the residues at the dimeric interface are highly conserved throughout the HWE family with shared architecture, we expect the role in the dimerization of the LOV and the PAS domains to be a conserved feature. Supporting this idea, two HKs coupled to a PAS domain, VicK from *Streptococcus mutans* [25] and a LOV domain, the chimeric YF1 [26] were reported, where both the LOV and PAS domains play an important role in dimerization. On the contrary, another LOV-HK, EL346, has been recently identified as the first full-length monomeric HK [4]. This difference in the oligomeric state could be due to the divergence in domain architecture, as *Brucella* LOV-HK has a central PAS domain between its LOV and HK domains. Future structural studies of full-length HKs from the HWE family are needed to understand the roles of LOV and PAS domains in promoting dimerization.

Consistent with recent evidence indicating that the directionality of the autophosphorylation reaction is determined by the handedness of the α1-α2 loop, here we demonstrated that the autophosphorylation reaction occurs intramolecularly, or *in cis* in the left-handed *Brucella* HK. A *cis* autophosphorylation allows the monomeric activity observed although it may not be natural.

Stabilization of the canonical dimer interface achieved in the cross-linking studies limits the conformational freedom of the CA domain, in particular prohibiting the formation of a NC active structure. All three of the engineered cross-linked C dimers bear autokinase activity, and one in particular resulted in an increase in activity. This enhanced activity suggests that the formation of the four-helix bundle is critical in directing the dynamics of the CA domains, even if the resulting autophosphorylation reaction occurs intramolecularly. In the context of thermodynamic coupling, the I286C mutation can be thought of as a way to stabilize the on-state of the DHp, which in turn promotes the formation of the active complex.

The C dimer was trapped in the crystal in an inactive conformation. A conserved arginine residue, in the R box (Arg321) binds the conserved glutamate (Glu384) and tyrosine (Tyr383) residues in the binding pocket and the γ- phosphate of the nucleotide (Figure 5C). These interactions seem to stabilize the inactive conformation. The same interactions are present in EL346, in which the mutation R175A resulted in an increase of activity [4]. On the contrary, the replacement of Arg321 to alanine here impaired autokinase activity. The loss of activity of the R321A mutant may be a consequence of a diminished ATP affinity, taking into account the interaction between the ATP analogue and Arg321 (Figure 6C). Recently, two structures of the ATP-bound Michaelis complex primed for autophosphorylation were reported [11, 17]. The autophosphorylation Michaelis complex is defined by a close approach between the CA and DHp; about 10–15 Å in the Michaelis complex *vs.* 20–30 Å in other states. Additionally, these autophosphorylation complexes, and even a third putative active structure [25], have been characterized by an angular preference of the CA with respect to the DHp [3]. Interestingly, the modeled active canonical conformation presents a similar inter-subdomain orientation (Figure S11). In the

two Michaelis complex structures, the carboxylic side chain of the acidic residue which follows the histidine phosphoacceptor binds to the N^{δ} atom of the histidine residue (see in Figure S4 residues Glu261 and Asp244 of HK853 and EnvZ, respectively). The acidic residue is proposed as a general base in the reaction, as it might abstract a proton from the histidine to induce its appropriate tautomeric form for the nucleophilic attack of N^{ϵ} atom to the ATP γ-phosphate. This hypothesis is supported by the fact that the mutation of the corresponding acidic residue in several HisKA HKs has a similar deleterious effect in the kinase reaction [17, 27, 28]. At the same time, its side chain forms a hydrogen bond to an asparagine residue of the N box (see in Figure S4 residues Asn376 and Asn343 in HK853 and EnvZ, respectively), helping the close positioning of the ATP and the histidine phosphoacceptor. In the HWE and HisKA2 families an arginine residue (Arg289 for *Brucella* HK), instead of the conserved acidic residue found in the HisKA family, follows the histidine phosphoacceptor. The mutation of the basic arginine residue following the histidine phosphoacceptor in EL346 impaired its autophosphorylation activity [4]. In the canonical catalytic complex observed in MD simulations (Figure 9), Arg289 forms a salt bridge with Glu384, which is located at the same position as the asparagine residue of EnvZ and HK853, in the N box. Both Arg289 and Glu384 are conserved within the HWE family (Figure S5). As already mentioned, this glutamate residue interacts with Arg321 in the inactive conformation, presumably switching from Arg321 in the inactive state to Arg289 in the active state. The general base in the reaction in the HWE family could be a glutamate residue located near the histidine phosphoacceptor, which is conserved in the LOV-PAS-HWE_HK family (Glu285 for *Brucella* HK, Figure 9).

The structure reported here is the first one from the HK HWE family, providing structural information about the family, particularly about the flexibility of the dimeric interface of DHp. This work also proposes an autophosphorylation mechanism for the HWE HK families.

Materials and Methods

Site-directed mutagenesis

All point mutant versions of the HK domain were achieved by site-directed mutagenesis using Q5 High-Fidelity DNA Polymerase (New England Biolabs), specific primers and a pET24d recombinant vector containing residues 266 to 489 from *Brucella abortus* LOV-HK and a C-terminal 6xHis tag [22]. The final constructs were checked by DNA sequencing.

Protein purification

Escherichia coli BL21(DE3) competent cells (Stratagene, La Jolla, California, USA) were transformed with the expression plasmids mentioned above. Precultures were grown overnight in 5 ml of LB medium added with 25 μg ml^{-1} kanamycin at 37 ºC with agitation (200 rev min^{-1}), and then diluted to 500 ml and grown to an absorbance (at 600 nm) of 0.6. At this point, isopropyl β-D-1-thiogalactopyranoside was added to a final concentration of 0.5 mM and the cultures were further incubated overnight at 20 ºC with agitation (200 rev min^{-1}). The purification for protein crystallization is described elsewhere [22]. The bacteria were centrifuged at 5000g for 10 min at 4 ºC. Pellets were resuspended and sonicated in a solution containing 50 mM sodium phosphate, 0.5 M sodium chloride, 20 mM imidazole, 1 mM DTT and 1 mM PMSF (pH 7.5, buffer A) and then centrifuged at 160000g in a

Beckman Coulter L7-65 ultracentrifuge (Brea, California, USA) for 60 min at 4 ºC. The supernatant was filtered through a 0.45 µm membrane and loaded onto a HisTrap HP column (all columns were from GE Healthcare, Little Chalfont, England). Elution was performed with a linear gradient of buffer B containing 50 mM Na phosphate, 0.5 M sodium chloride, 0.5 M imidazole, 1 mM DTT and 1 mM PMSF (pH 7.5). A major peak was observed. The appropriate fractions were pooled and dialyzed extensively overnight at 4 ºC against buffer C (50 mM MES, 0.25 M sodium chloride and 1 mM PMSF, pH 6.5). The protein was aliquoted and stored at -70 ºC at a concentration of approximately 1 mg/ml. The quality of the final preparation was checked by SDS–PAGE (15% gel) and UV spectrophotometry. For the experiment of Figure 8 the dimers of I286C, V282C and A287C-L333C mutants and the wild type version of the HK domain were further purified by Superdex-75 in a buffer containing 50 mM MES pH 6.5 and 250 mM sodium chloride. The concentration of the protein was estimated using the absorbance at 280 nm and the molar extinction coefficient calculated from the sequence using the ProtParam tool from the Expasy server (http://web.expasy.org/protparam/).

Crystallization, X-ray data collection and structure resolution

Detailed information about the crystallization of the HK domain, together with diffraction data collection and processing, phasing, model building and refinement, is provided elsewhere [22]. Briefly, a total of three HK crystals were used for highly redundant sulfur anomalous data collection (λ = 1.80 Å) at the PROXIMA 1 beamline at the SOLEIL Synchrotron (France), taking into advantage the natural content of cysteine and methionine residues of the protein (11 S sites in 242 residues). After successful S-SAD

phasing and partial model building, native data (λ = 0.98 Å) from a fourth HK crystal diffracting to 2.51 Å resolution were used for the refinement of the model. The final coordinates were subjected to validation with MolProbity [29] and with the validation module implemented in Coot [30]. Detailed statistics on the refinement process are found in Table 1.

Coordinates representation

Coordinates representation were performed in Pymol (http://www.pymol.org) and Visual Molecular Dynamics (VMD) [31] software.

Crystallographic coordinate analysis

The PDBePISA server [32] was used to evaluate the dimerization interfaces (http://www.ebi.ac.uk/pdbe/pisa)

Size exclusion chromatography - Static light scattering measurements

The average MW of wt and mutant HK domains in solution were determined on a Precision Detectors PD2010 90º light scattering instrument tandemly connected to a high-performance liquid chromatography and a LKB 2142 differential refractometer and to a 486 Absorbance Detector (Waters), set at 280 nm. The columns used were Superdex 75 GL 10/300 (GE Healthcare). 500 µL of each purified protein were injected into the column, and the chromatographic runs were performed with a buffer containing 50 mM MES pH 6.5, 250 mM sodium chloride under isocratic conditions at a flow rate of 0.4 mL/min at 20º C. The concentration of the injected samples was ~40 µM, except for the experiments of Figure S3A. The MW of each sample was calculated relating its 90° and RI signals and

comparison of this value with the one obtained for BSA (MW: 66.5 kDa) as a standard using the software Discovery32. The reported MW values are an average between the values relating RI and UV with scattering.

Autokinase assays

For all experiments, wild-type and mutant versions of the HK domain were incubated at 37°C at a concentration of 25 µM (192 pmol) in a buffer containing 100 mM Tris-HCl pH 8, 5 mM $MgCl_2$, 100 µM [γ-32P] ATP 1 µCi/nmol (PerkinElmer Life Sciences). After incubation for 30 min or at the indicated times, the reactions were stopped with Laemmli sample buffer. Samples were separated on a 15% SDS-PAGE gel, exposed to a Storage Phosphor Screen (GE Healthcare), and scanned by a Storm 840 Molecular Imager (GE Healthcare). For *cis-trans* autophosphorylation experiments, the purified long and short HK domains were mixed in equimolar proportions, incubated with 100 mM DTT and then extensively dialyzed in a buffer containing 50 mM MES pH 6.5, 250 mM sodium chloride and 1mM PMSF before the phosphorylation assay. PageRuler Prestained Protein Ladder (Thermo Scientific) and Low Molecular Weight Marker (Amersham™ GE Healthcare) were used as MW standards. For several experiments, the intensity of each band signal was estimated using the ImageQuant 5.2 program (Molecular Dynamics).

Construction of active structural models

Modeled active C monomer (Figure S10A). Beginning with chain A of the canonical crystal structure, we performed a sweep of the backbone dihedral space of Trp337. The dihedral sweep was performed on a square grid ranging from -180 to 180 degrees and points placed at 10 degree intervals. Using VMD software [31], rigid-body transformations of the

catalytic subdomain were made about residue Trp337 such that its (φ, ψ) dihedral angles took values on the grid. The active conformation was chosen based on the proximity of the resolved ATP analogue to His288, which took a value of 7 Å with Trp337 at (50,-100). This structure was then refined to remove any clashes.

Modeled active NC monomer (Figure S10B). Beginning with chain C of the non-canonical structure, we used the software Modeller v 9.14 [33] to first align the sequence between the missing α1 helix and the α2 helix, and then model the missing α1 helix using the α2 helix of chain D as a template structure, resulting in the γ-phosphate of the resolved ATP analogue within 7 Å of His288.

MD simulations

Solvation and equilibration. Each model was solvated in a 10.0 Å octahedron, and chloride counterions were added to neutralize the net charge. The parameters for the protein were taken from the AMBER ff99SB force field, with the ildn modification for isoleucine, leucine, aspartate and asparagine residues [34]. The TIP3P water model was used for solvent, with the updated Lennard Jones parameters for ions [35]. The coordinates of the resolved ACP molecules were used to position the corresponding ATP molecules. ATP and Mg^{2+} parameters were taken from the Bryce group [36]. The system was heated to 300K at constant volume, with restraints on the protein that were gradually reduced from 200 to 0 kcal/mol/$Å^2$ over a period of 150ps. The Langevin thermostat was used with a collision frequency of 1.0 ps^{-1}. The SHAKE algorithm was employed to constrain bonds to non-polar hydrogens, and a 1.0 fs time-step was used during dynamics [37]. An 8.0 Å cutoff was used for non-bonded interactions. 100 Ps equilibration runs were done in the NPT ensemble, using isotropic pressure scaling and a pressure relaxation time of 2.0 ps.

Production simulations. For each system, five independent 100 ns simulations were started from different snapshots from the equilibration simulations (taken at arbitrary intervals after density of water box was equilibrated) with randomized velocities. All production runs were done in the same NVT conditions described in 'Solvation and equilibration', with the exception of a 2.0 fs time-step. All simulations were performed on GPUs using the PMEMD (cuda) module within the Amber11 simulation package [38]. Analysis of MD simulations was performed using PTRAJ from AmberTools (distances, RMSD, radius of gyration) or VMD (solvent-accessible surface area).

Coordinates deposition

The X-ray crystallographic coordinates of the HK domain from *Brucella* LOV-HK have been deposited at the Protein Data Bank, under the code **5EPV**. Residue numbering follows the Uniprot sequence for *Brucella* LOV-HK Gene **BAB2_0652** (**Q2YKK7-1**).

Acknowledgments

This work was supported by the Argentinian Ministry of Science (MINCyT) and the Argentinian Research Council (CONICET) and by the National Science Foundation Postdoctoral Research Fellowship in Biology under Grant No. DBI-1401889. We are grateful for the access to the PROXIMA 1 beamline at the SOLEIL Synchrotron, France.

References

[1] Gao R, Stock AM. Biological insights from structures of two-component proteins. Annu Rev Microbiol. 2009;63:133-54.
[2] Igo MM, Ninfa AJ, Stock JB, Silhavy TJ. Phosphorylation and dephosphorylation of a bacterial transcriptional activator by a transmembrane receptor. Genes Dev. 1989;3:1725-34.

[3] Bhate MP, Molnar KS, Goulian M, DeGrado WF. Signal Transduction in Histidine Kinases: Insights from New Structures. Structure. 2015;23:981-94.
[4] Rivera-Cancel G, Ko WH, Tomchick DR, Correa F, Gardner KH. Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. Proc Natl Acad Sci U S A. 2014;111:17839-44.
[5] Casino P, Rubio V, Marina A. Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. Cell. 2009;139:325-36.
[6] Marina A, Waldburger CD, Hendrickson WA. Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. EMBO J. 2005;24:4247-59.
[7] Ferris HU, Coles M, Lupas AN, Hartmann MD. Crystallographic snapshot of the Escherichia coli EnvZ histidine kinase in an active conformation. J Struct Biol. 2014;186:376-9.
[8] Tanaka T, Saha SK, Tomomori C, Ishima R, Liu D, Tong KI, et al. NMR structure of the histidine kinase domain of the E. coli osmosensor EnvZ. Nature. 1998;396:88-92.
[9] Albanesi D, Martin M, Trajtenberg F, Mansilla MC, Haouz A, Alzari PM, et al. Structural plasticity and catalysis regulation of a thermosensor histidine kinase. Proc Natl Acad Sci U S A. 2009;106:16185-90.
[10] Trajtenberg F, Grana M, Ruetalo N, Botti H, Buschiazzo A. Structural and enzymatic insights into the ATP binding and autophosphorylation mechanism of a sensor histidine kinase. J Biol Chem. 2010;285:24892-903.
[11] Mechaly AE, Sassoon N, Betton JM, Alzari PM. Segmental helical motions and dynamical asymmetry modulate histidine kinase autophosphorylation. PLoS Biol. 2014;12:e1001776.
[12] Kim D, Forst S. Genomic analysis of the histidine kinase family in bacteria and archaea. Microbiology. 2001;147:1197-212.
[13] Ninfa EG, Atkinson MR, Kamberov ES, Ninfa AJ. Mechanism of autophosphorylation of Escherichia coli nitrogen regulator II (NRII or NtrB): trans-phosphorylation between subunits. J Bacteriol. 1993;175:7024-32.
[14] Yang Y, Inouye M. Intermolecular complementation between two defective mutant signal-transducing receptors of Escherichia coli. Proc Natl Acad Sci U S A. 1991;88:11057-61.
[15] Cai SJ, Inouye M. Spontaneous subunit exchange and biochemical evidence for trans-autophosphorylation in a dimer of Escherichia coli histidine kinase (EnvZ). J Mol Biol. 2003;329:495-503.
[16] Ashenberg O, Keating AE, Laub MT. Helix bundle loops determine whether histidine kinases autophosphorylate in cis or in trans. J Mol Biol. 2013;425:1198-209.
[17] Casino P, Miguel-Romero L, Marina A. Visualizing autophosphorylation in histidine kinases. Nat Commun. 2014;5:3258.
[18] Gourley CR, Petersen E, Harms J, Splitter G. Decreased in vivo virulence and altered gene expression by a Brucella melitensis light-sensing histidine kinase mutant. Pathog Dis. 2015;73:1-8.
[19] Swartz TE, Tseng TS, Frederickson MA, Paris G, Comerci DJ, Rajashekara G, et al. Blue-light-activated histidine kinases: two-component sensors in bacteria. Science. 2007;317:1090-3.
[20] Sycz G, Carrica MC, Tseng TS, Bogomolni RA, Briggs WR, Goldbaum FA, et al. LOV Histidine Kinase Modulates the General Stress Response System and Affects the virB Operon Expression in Brucella abortus. PLoS One. 2015;10:e0124058.

[21] Karniol B, Vierstra RD. The HWE histidine kinases, a new family of bacterial two-component sensor kinases with potentially diverse roles in environmental signaling. J Bacteriol. 2004;186:445-53.
[22] Klinke S, Foos N, Rinaldi JJ, Paris G, Goldbaum FA, Legrand P, et al. S-SAD phasing of monoclinic histidine kinase from Brucella abortus combining data from multiple crystals and orientations: an example of data-collection strategy and a posteriori analysis of different data combinations. Acta crystallographica Section D, Biological crystallography. 2015;71:1433-43.
[23] Bilwes AM, Quezada CM, Croal LR, Crane BR, Simon MI. Nucleotide binding by the histidine kinase CheA. Nat Struct Biol. 2001;8:353-60.
[24] Hsing W, Russo FD, Bernd KK, Silhavy TJ. Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ. J Bacteriol. 1998;180:4538-46.
[25] Wang C, Sang J, Wang J, Su M, Downey JS, Wu Q, et al. Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains. PLoS Biol. 2013;11:e1001493.
[26] Diensthuber RP, Bommer M, Gleichmann T, Moglich A. Full-length structure of a sensor histidine kinase pinpoints coaxial coiled coils as signal transducers and modulators. Structure. 2013;21:1127-36.
[27] Atkinson MR, Ninfa AJ. Mutational analysis of the bacterial signal-transducing protein kinase/phosphatase nitrogen regulator II (NRII or NtrB). J Bacteriol. 1993;175:7016-23.
[28] Willett JW, Kirby JR. Genetic and biochemical dissection of a HisKA domain identifies residues required exclusively for kinase and phosphatase activities. PLoS Genet. 2012;8:e1003084.
[29] Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta crystallographica Section D, Biological crystallography. 2010;66:12-21.
[30] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta crystallographica Section D, Biological crystallography. 2010;66:486-501.
[31] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996;14:33-8, 27-8.
[32] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774-97.
[33] Webb B, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol. 2014;1137:1-15.
[34] Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins. 2010;78:1950-8.
[35] Joung IS, Cheatham TE, 3rd. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. J Phys Chem B. 2008;112:9020-41.
[36] Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. J Comput Chem. 2003;24:1016-25.
[37] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. Journal of Computational Physics. 1977;23:327-41.
[38] Case DA, Berryman JT, Betz RM, Cerutti DS, Cheatham TE, Darden I, T. A., et al. AMBER 2015. 2015.

[39] Engh RA, Huber R. Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallographica Section A. 1991;47:392-400.

Figures and figure legends

Figure 1. The crystal structure of the HK domain from *Brucella* LOV-HK shows two different dimeric assemblies. Crystal structure of the HK domain from *Brucella* LOV-HK photoreceptor. **A**. Domain architecture of *Brucella* LOV-HK. The crystal structure corresponds to the HK domain colored in yellow. The residue numbers of the domain boundaries are indicated. **B**. Canonical dimer. **C**. Non-canonical dimer. The protein is represented as ribbons. The four chains are shown in different colors, with the same color code along the article. AMP-PCP (ACP) is depicted as sticks and Mg^{2+} as a sphere. The secondary structure elements and the Dimerization and Histidine phosphotransfer (DHp) and the Catalytic and ATP binding (CA) subdomains are indicated. Secondary structure elements are also indicated.

Figure 2. Apart from the parallel- antiparallel disposition, the canonical and the non-canonical dimers differ in the relative orientation of their CA with respect to their DHp. A. Ribbon representation of the canonical (C) monomer (yellow) compared to the non-canonical (NC) monomer (light blue). The α2 helices from both monomers are superimposed. The surface representation of the CA shows their different orientations between the C and the NC monomers. The two alternative conformations of the α2-α3 loop (residues 333 to 341) are highlighted in orange (A) and blue (C) within a light oval. CA subdomains and α2-α3 loops are indicated. **B.** Ramachandran plot of the C (yellow) and

the NC (light blue) monomers, highlighting the change in the torsion angles of Glu335 and Trp337. **C.** Resulting changes in relative subdomain orientations upon transforming either Trp337or Glu335 from their C to NC conformations.

Figure 3. In solution, the HK domain is active as a monomer. **A**. Oligomeric state of the HK domain in solution. The HK domain was subjected to SEC coupled to a light-scattering instrument connected in tandem to a differential refractometer detector and an absorbance detector tuned at 280 nm. The MW estimated by the relation of scattering/RI or UV signals is plotted in the inset. **B**. Autophosphorylation activity of the HK domain. The HK domain was incubated at 37°C with radiolabeled ATP and Mg^{2+}. The phosphorylated protein was visualized by SDS-PAGE and autoradiography. The incubation time in minutes is indicated above. The MW standard values are shown on the left.

Figure 4. Cysteine cross-linking experiments show that the canonical dimer, but not the non-canonical dimer, exists and is active in solution. The following cysteine mutations were introduced in the HK domain in order to cross-link either the canonical (I286C, V282C and A287C-L333C) or the non-canonical (L322C-L325C, L441C and F437C-H329C) dimer. The mutations are mapped on the dimeric structures in the upper panels. The proteins were purified and subjected to SEC-SLS in the absence of reducing agents (middle panels). The numbers above the peaks correspond to the estimated MW values in kDa. Their activity was analyzed by the incubation with radiolabeled ATP, followed by SDS-PAGE and autoradiography (bottom panels). The samples were subjected in parallel to SDS-PAGE and stained with Coomassie Blue. The MW standard values (kDa) are shown on the left.

Figure 5. The canonical dimer is functionally relevant. A. The dimers of three canonical cysteine mutants (I286C, V282C, and A287C–L333C) and the wild-type protein were isolated by gel filtration before autophosphorylation assay. The incubation time (min) is indicated in the upper part of the gel. **B**. The band intensities were quantified and plotted *vs*. incubation time for each mutant and the wild-type protein.

Figure 6. The crystallographic canonical dimer. A. The overall canonical dimer is shown. The histidine phosphoacceptor (His288), Arg321 and the ACP molecule are shown as sticks and Mg^{2+} as a sphere. The distance between His288 and the ACP molecule (Å) is indicated in blue. Red, black and green boxes with dashed borders correspond to panels B, C and D, respectively. **B**. Dimerization contacts in the DHp. The residues involved are shown as sticks and labeled. The molecule is rotated 90° clockwise in the y axis with respect to panel A. **C**. Bottom view of the nucleotide binding site. The relevant side chains are shown as sticks and labeled. **D**. CA-DHp interface.

Figure 7. Autophosphorylation occurs *in cis* in the dimer. Two different versions of the HK domain: one long (L, the HK domain plus a 23 kDa tag, 49kDa) and other short (S, 26 kDa) were generated. The short version has the N388A which impairs the phosphorylation capacity. Both versions have the I286C mutation. The purified long and short HK domains were mixed in equimolar proportions, incubated with 100 mM DTT and then extensively dialyzed before the phosphorylation assay. The species present in the experiments and the possible autophosphorylation outcome are shown schematically at the left side. The possible phosphorylated subunits, according to a *cis*- or a *trans* directionality, are

highlighted in red. On the right side is shown the electrophoretic analysis of the autophosphorylation activity of the S, L and the mix in the presence and in the absence of DTT by SDS-PAGE and [γ-^{32}P] ATP autoradiography. The samples were subjected in parallel to SDS-PAGE and stained with Coomassie Blue. The bands corresponding to each species (LL, SL, SS, L and S) are labelled. The MW standard values (kD) are shown on the left.

Figure 8. Molecular modeling and simulations of the HK domain. **A**. Active structural models of HK monomer, based on the C (orange) and NC (blue) crystallographic conformations. Arrows highlight opposing routes of autophosphorylation. **B**. Histogram of the His288-ATP distance in 0.5-μs MD simulations of the C (orange) and NC (blue) active structural models.

Figure 9. Proposed autophosphorylating center based on the active C monomeric model. The relevant side chains are shown as sticks and labeled. Mg^{2+} ion is shown as a sphere. The distance between the γ-phosphate of the ATP molecule and the N^{δ} of His288 is indicated. Glu285 might act as a general base in the phosphotransferase reaction. The Arg289-Glu384 interaction would stabilize the active conformation.

Table 1. Refinement and validation statistics. [a]

	HK from B. abortus
Refinement	
Resolution range (Å)	48.6-2.51
Number of protein atoms	5614
Number of ligand atoms	126
Number of water molecules	-
R	0.202
R_{free}	0.242
Rms deviations from ideal values [b]	
Bond lengths (Å)	0.010
Bond angles (º)	1.2
B-factor (average, $Å^2$)	74
MolProbity validation [c]	
Clashscore	6.77
Poor rotamers (%)	10.6
Ramachandran plot	
Favoured (%)	95.8
Allowed (%)	4.1
Disallowed (%)	0.1
Deposition	
PDB code	5EPV

[a] Data collection, processing and phasing statistics are presented elsewhere [22].
[b] [39]
[c] [29]

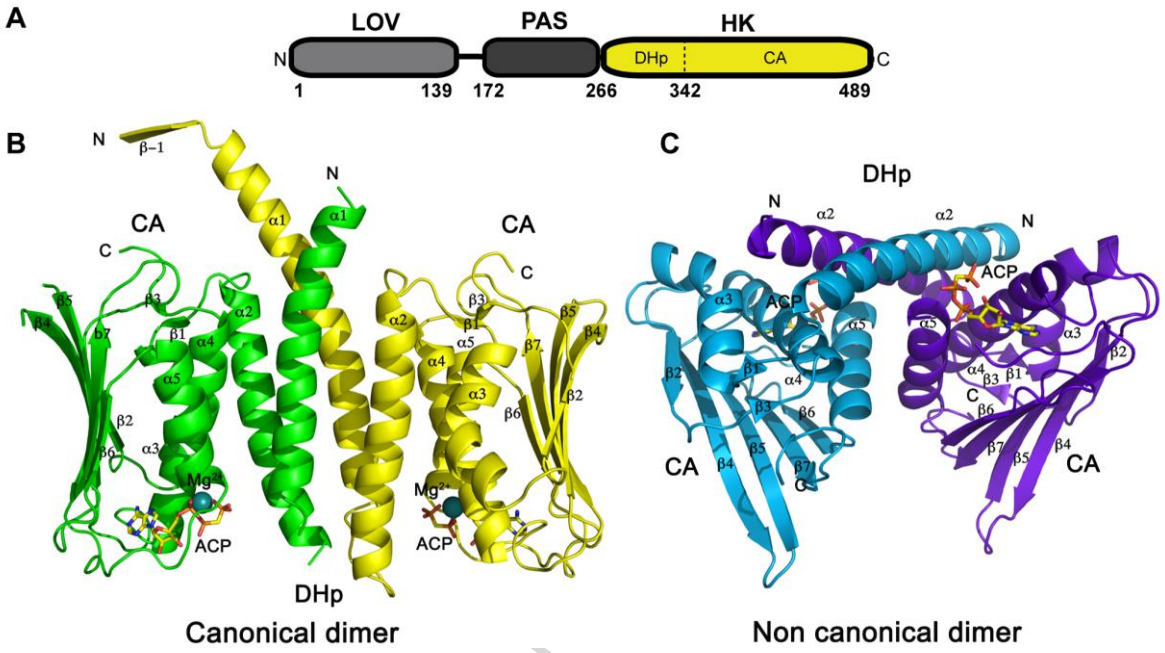


Figure 1

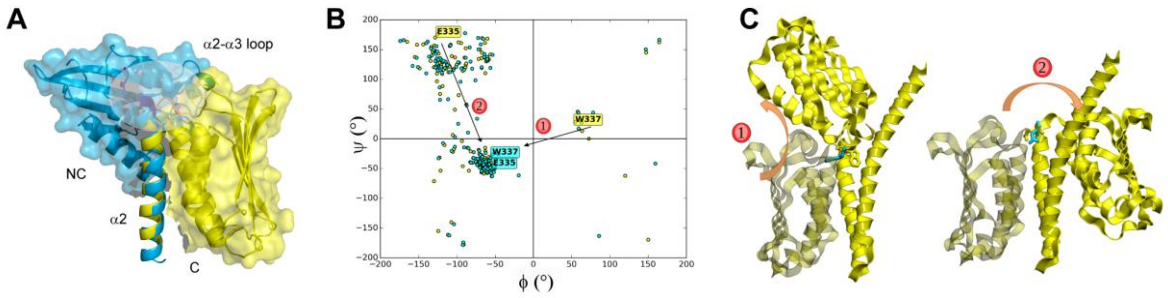


Figure 2

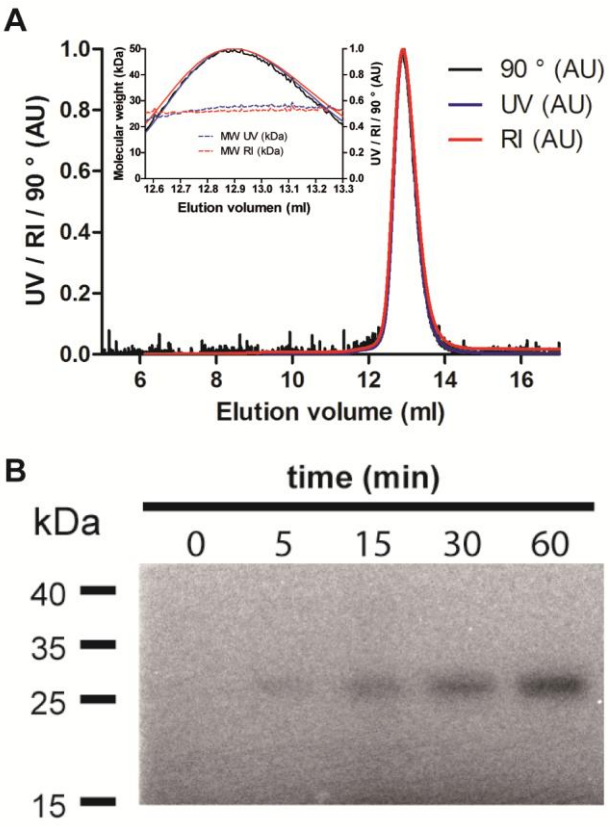


Figure 3

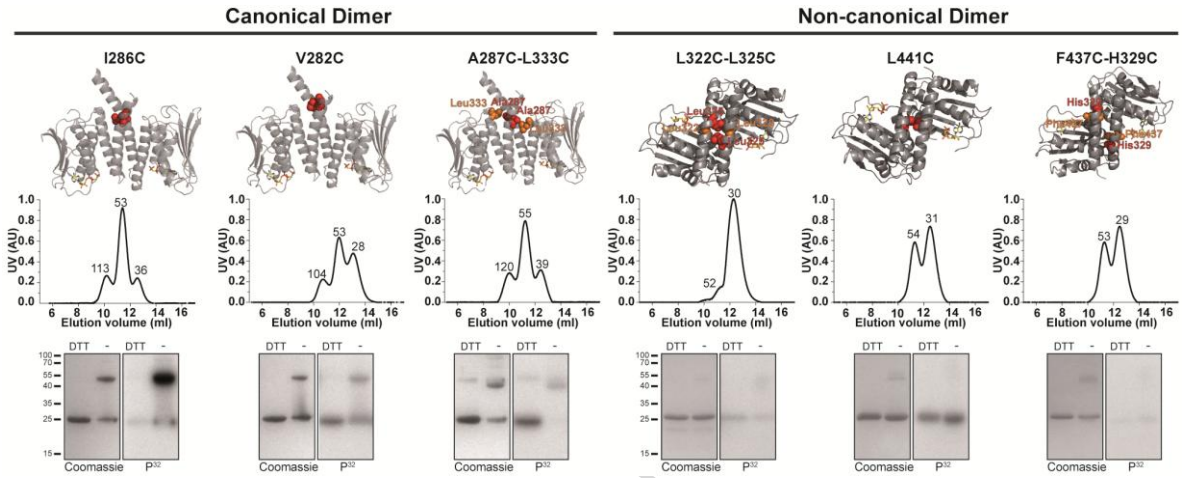


Figure 4

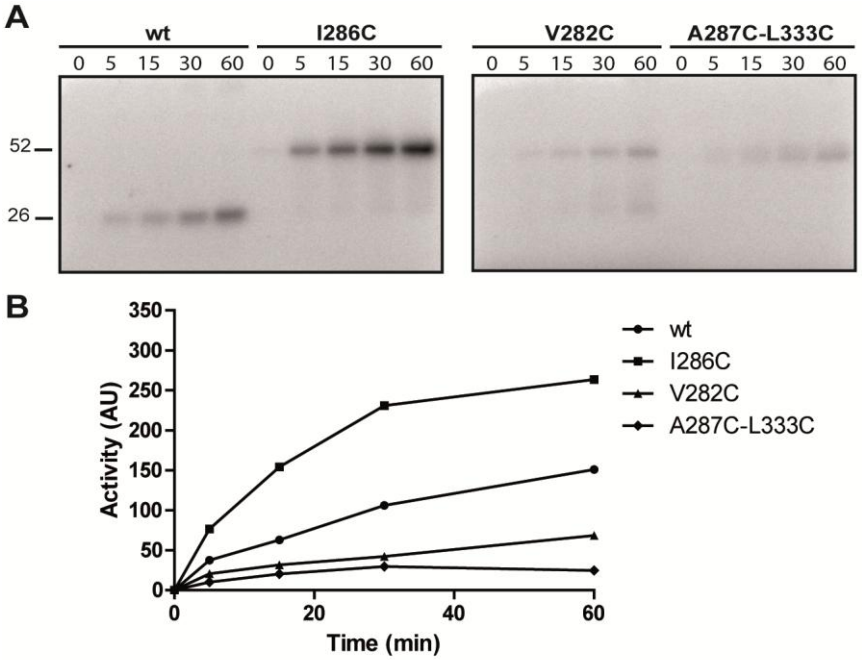


Figure 5

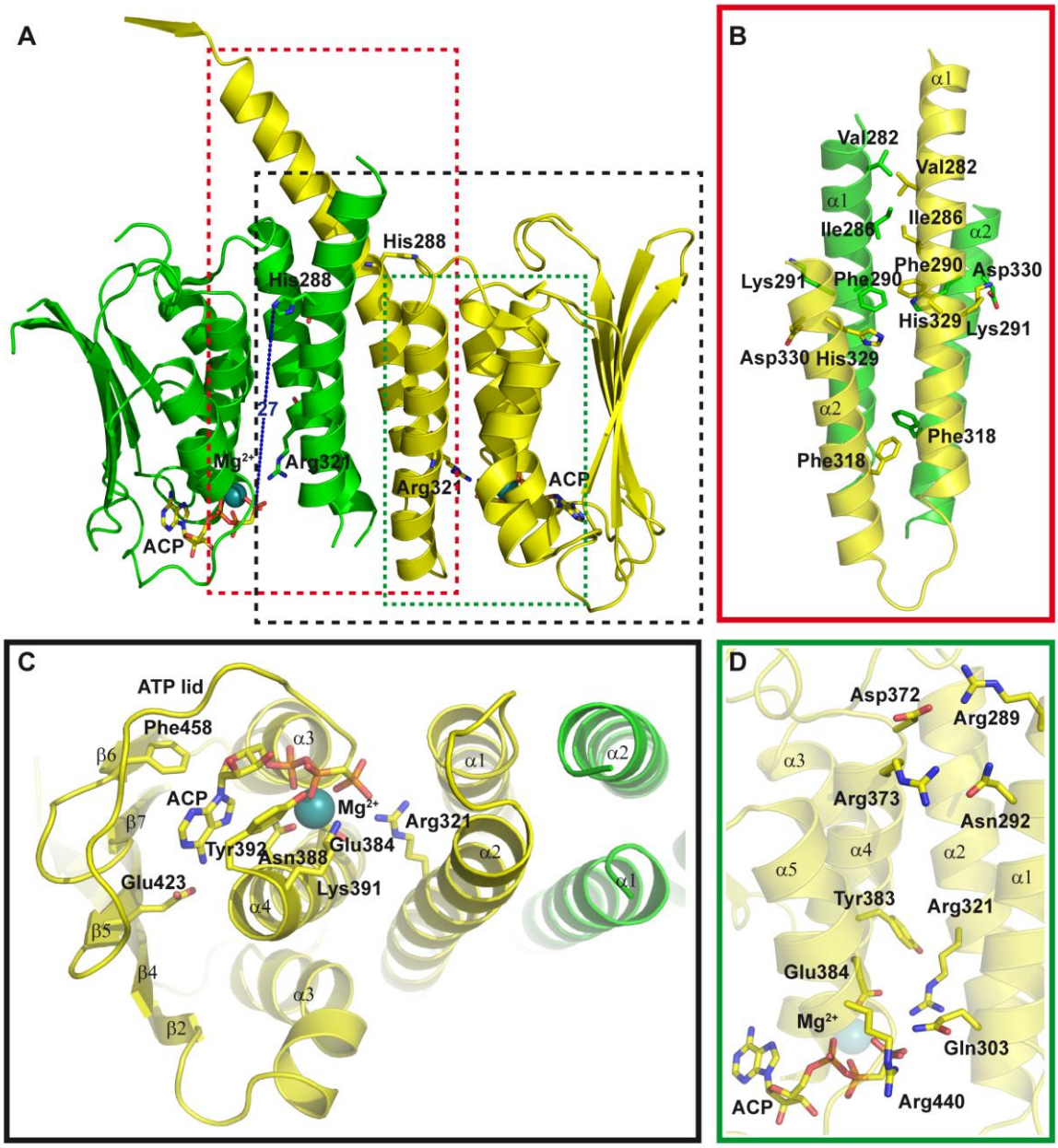


Figure 6

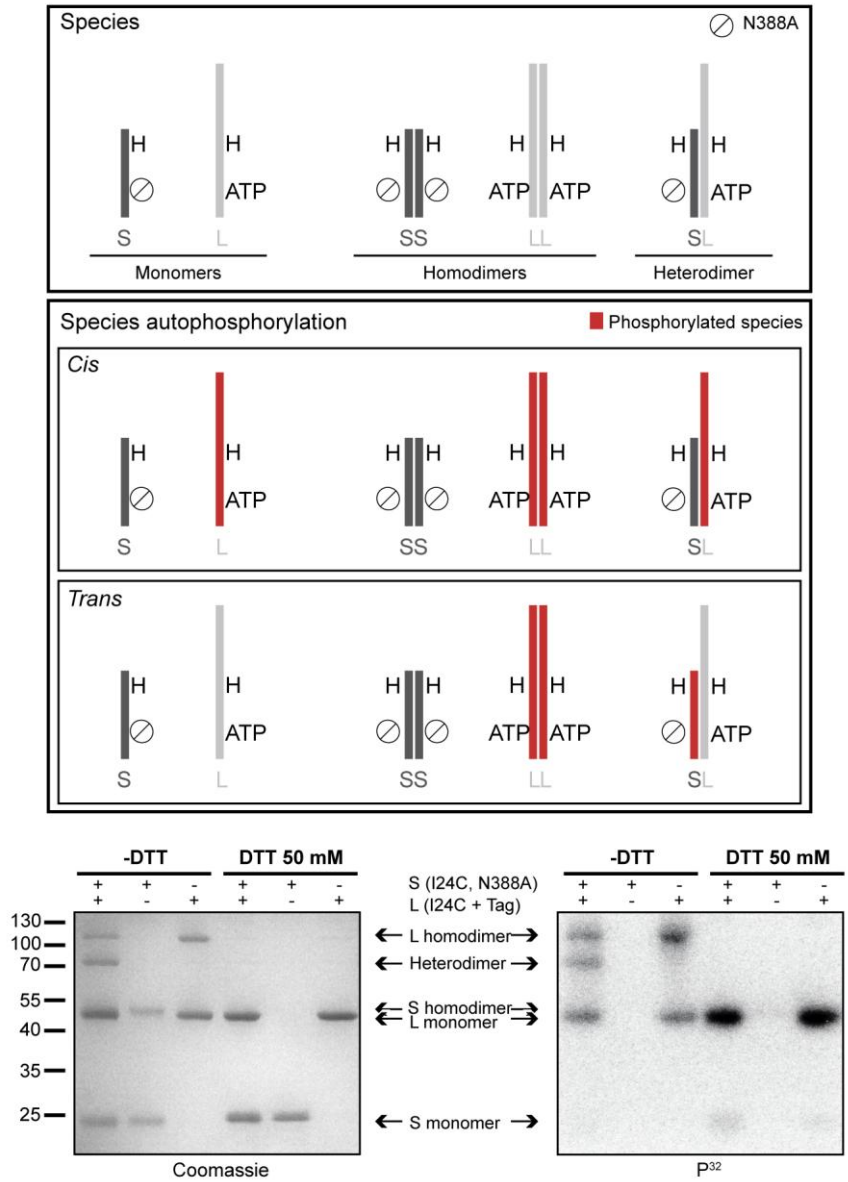


Figure 7

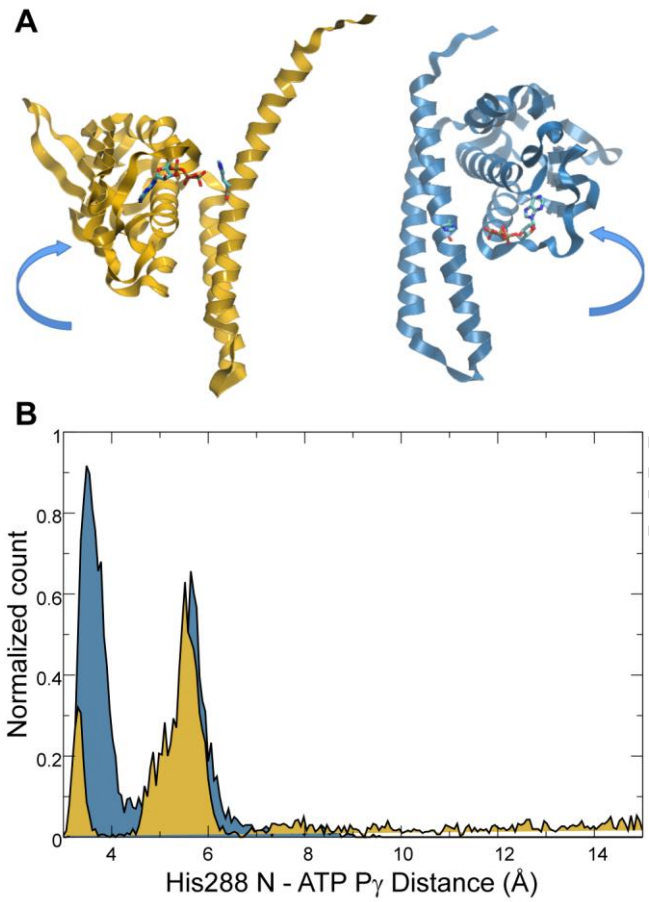


Figure 8

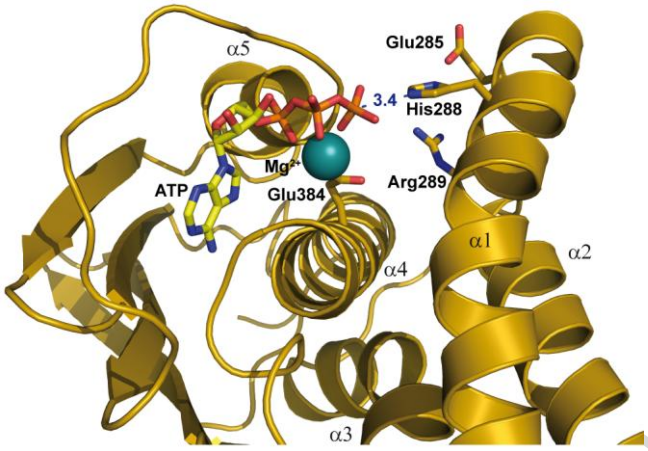
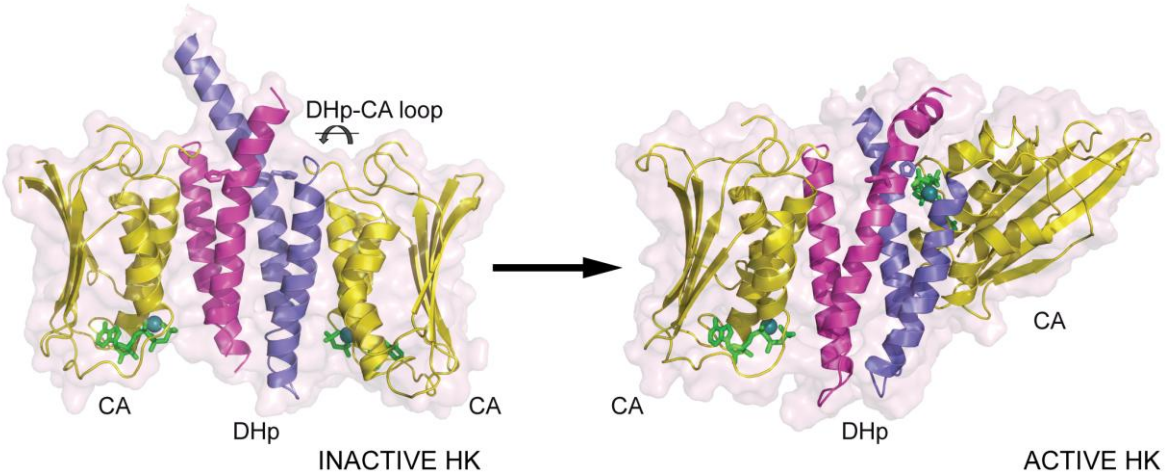


Figure 9



Graphical abstract

Highlights

- The crystal structure of the *Brucella* HK domain reveals two distinct dimers.
- In solution the *Brucella* HK domain is an active monomer.
- The canonical dimer is the functionally relevant species.
- The autophosphorylation occurs *in cis*.
- A catalytic mechanism for *Brucella* LOV-HK is proposed.