Published in final edited form as:

Nat Chem Biol. 2019 May ; 15(5): 510–518. doi:10.1038/s41589-019-0265-y.

Structure of McsB, a protein kinase for regulated arginine phosphorylation

Marcin J. Suskiewicz¹, Bence Hajdusits¹, Rebecca Beveridge¹, Alexander Heuck¹, Lam Dai Vu^{1,2}, Robert Kurzbauer¹, Katja Hauer³, Vanessa Thoeny³, Klaus Rumpel³, Karl Mechtler^{1,4}, Anton Meinhart¹, and Tim Clausen^{1,5,*}

¹Research Institute of Molecular Pathology (IMP), Vienna BioCenter (VBC), Vienna, Austria

³Boehringer Ingelheim RCV GmbH & Co KG, Vienna, Austria

⁴Institute of Molecular Biotechnology of the Austrian Academy of Science (IMBA), Vienna BioCenter (VBC), Vienna, Austria

⁵Medical University of Vienna, Vienna BioCenter (VBC), Vienna, Austria

Abstract

Protein phosphorylation regulates key processes in all organisms. In Gram-positive bacteria, protein arginine phosphorylation plays a central role in protein quality control by regulating transcription factors as well as marking aberrant proteins for degradation. Here, we report structural, biochemical, and *in vivo* data of the responsible kinase, McsB, the founding member of an arginine-specific class of protein kinases. McsB differs in structure and mechanism from protein kinases that act on serine, threonine, and tyrosine residues and instead has a catalytic domain related to phosphagen kinases (PhKs), metabolic enzymes that phosphorylate small guanidino compounds. In McsB, the PhK-like phosphotransferase domain is structurally adapted to targeting protein substrates and is accompanied by a novel phosphoarginine (pArg)-binding domain that allosterically controls protein kinase activity. The identification of distinct pArg reader domains in this study points to a remarkably complex signalling system, thus challenging simplistic views of bacterial protein phosphorylation.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence and Lead Contact: Tim Clausen, tim.clausen@imp.ac.at.

²Present address: VIB/UGent, Ghent, Belgium

Data Availability

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession codes 6FH1 (McsB), 6FH2 (McsB·AMP-PN), 6FH3 (McsB·pArg), 6FH4 (1-75 CtsR·pArg). All other source data are included in the paper or provided upon request.

Author Contributions

M.J.S and T.C. designed and performed experiments, analysed data, and wrote the paper with input from all authors; B.H., A.H., L.D.V., R.K., K.H., V.T., K.R., and A.M. helped with biochemical and structural analyses; R.B. and K.M. with mass spectrometric measurements.

Competing Interests

The authors declare no competing interests.

Introduction

Protein phosphorylation is a ubiquitous regulatory mechanism that affects almost every cellular process^{1,2}. Phosphoryl groups, typically originating from adenosine triphosphate (ATP), are attached to amino acid side chains on protein substrates by protein kinases and removed by protein phosphatases, resulting in reversible regulation on a shorter time scale than allowed by transcriptional responses^{3,4}. While phosphorylation can directly alter the activity of a modified protein, it often exerts its effects by creating docking sites for phospho-binding domains and thus initiating new protein–protein interactions^{5,6}. Phosphorylation-based signal transduction systems are therefore best thought of as tripartite toolkits composed of writers (protein kinases), erasers (protein phosphatases), and phospho-binding domains that serve as downstream readers⁷. However, this nomenclature has only rarely been used to describe protein phosphorylation cascades in bacteria, where phospho-binding domains are scarce⁸ and phosphorylation is thought to mainly directly control the function of specific transcription factors, as exemplified by the two-component regulatory system⁹.

With hundreds of phosphoarginine (pArg) sites detected in *Bacillus subtilis*^{10–13} and *Staphylococcus aureus*^{14,15}, protein arginine phosphorylation of Gram-positive bacteria is emerging as a novel, major example of a phosphorylation-based signalling system with established links to stress response and pathogenicity^{16–19}. Arginine side chains are phosphorylated by McsB and dephosphorylated by YwlE, both of which can distinguish (phospho)arginine from other amino acid residues^{20,21}. Although the structure of McsB has remained elusive, sequence analysis demonstrated that it is unrelated to canonical protein kinases and instead shows weak homology to phosphagen kinases (PhKs), metabolic enzymes exemplified by human creatine kinases and insect arginine kinases^{22–24}. In the species in which it is found, McsB does not have any further sequence homologues and appears to be the only protein arginine kinase, as knocking out its gene abolishes all detectable pArg sites in *B. subtilis*¹¹.

McsB appears to coordinate the stress response in two ways. First, it phosphorylates arginine residues within the DNA-binding domains of CtsR and HrcA, thus directly inhibiting two major negative regulators of the heat-shock response^{11,12,20}. Relieving the transcriptional repression imposed by CtsR and HrcA induces expression of McsB itself, as well as of the core protein quality control components including the protease ClpCP and the chaperone DnaK¹⁹. Second, McsB functions as a ‘degradation labeller’, marking both regulatory proteins and aberrant polypeptide molecules for degradation¹². The phosphorylated substrates are recognised by a pArg-specific reader domain within the ClpCP proteolytic complex. The importance of the pArg-dependent degradation pathway for the bacterial heat-shock response is confirmed by the impaired survival at elevated temperatures of a *B. subtilis* strain harbouring pArg binding-deficient ClpC¹². Conceptually, it can be considered a simple bacterial version of the eukaryotic ubiquitin-proteasome system, with the McsB kinase playing a similar role to that of the E3 ubiquitin ligases.

The present study reveals that the protein arginine phosphorylation reaction is catalysed by the PhK-like catalytic domain of McsB, which is structurally adapted to targeting protein

substrates. Moreover, McsB possesses a pArg-binding domain that allows pArg-carrying proteins to allosterically enhance its kinase activity. With the writer McsB, the eraser YwIE, and the pArg reader domains of McsB, CtsR, and ClpC now all structurally characterised, the pArg signalling toolkit of Gram-positive bacteria is a phospho-signalling system come of age.

Results

Overall structure of the protein arginine kinase McsB

To exert its biological functions as a transcriptional regulator and a degradation labeller, McsB catalyses the transfer of the ATP γ -phosphoryl onto the arginine side chain of a protein substrate (Fig. 1a). For the structural characterisation of this distinct protein kinase activity, we determined the crystal structure of McsB from *Geobacillus stearothermophilus* in the apo state and in complex with an adenosine diphosphate (ADP) analogue, AMP-PN (Fig. 1b–d, Supplementary Table 1). Crystals were only obtained after deleting the C-terminal residues 356–363, which are predicted to be disordered and are dispensable for *in vitro* activity (Supplementary Fig. 1a–c). Additionally, we used the YwIE protein arginine phosphatase to dephosphorylate pArg sites that accumulated on McsB during expression due to autophosphorylation (Supplementary Fig. 1d). As seen at high resolution, the McsB dimer adopts a flat ‘domino tile’ shape (dimensions $120 \times 40 \times 30 \text{ \AA}$) with the two active sites opening on the same side (Fig. 1c). Individual subunits are composed of the N-terminal catalytic, ATP:guanido phosphotransferase domain (PD, residues 1–263) and the C-terminal dimerisation domain (DD, residues 264–355), which are linearly organised in a PD-DD-DD*-PD* manner (the asterisk denotes the partner protomer). The PD comprises an antiparallel, 9-stranded β -sheet surrounded by seven α -helices (Supplementary Fig. 2) and contains at its centre the active site composed of adjacent binding pockets for Mg-ATP and the guanidino group of the targeted arginine side chain (Fig. 1d). In our nucleotide-bound structure, the Mg-ATP-binding site shows a clear electron density for AMP-PN (Supplementary Fig. 3a). The DD is a 4-helix bundle that is tightly associated with the PD.

Enzymatic mechanism of protein arginine phosphorylation

To delineate the molecular mechanism of the protein arginine kinase reaction, we investigated the evolutionary relationship between McsB and PhKs. The PD of McsB shows low sequence homology (~25–30% identity) to the ‘large’, catalytic domain of PhKs that act on free arginine and creatine^{22,23}. The catalytic mechanism of PhKs primarily depends on the precise stereoelectronic alignment of the reacting γ -phosphoryl and guanidino moieties by a set of properly arranged positively and negatively charged residues^{25–28}. The key residues include five arginine residues coordinating the β - and γ -phosphates of ATP and two glutamate residues and one cysteine residue positioning the guanidino group of the arginine substrate for the phosphoryl transfer^{25,26,28,29}. Superposition of the nucleotide-bound catalytic domains of McsB and arginine kinase (AK) from the crab *Limulus polythemus*²⁶ showed structural conservation of all eight catalytic residues (Fig. 2a,b). To test their involvement in McsB-mediated protein arginine phosphorylation, we performed a mutational analysis of a subset of them (Arg208, Glu122, Glu213, and Cys168), using β -casein as a model substrate in a quantitative, radioactive kinase assay (Fig. 2c). In addition,

we applied a 'lag-time' assay that was based on the serendipitous observation that leaky ectopic expression of wild-type (WT) McsB results in a marked growth defect in *Escherichia coli*. Namely, McsB-expressing cells have an acutely extended lag-time, suggesting that the protein arginine kinase activity is toxic in this species (Fig. 2d). When analysing the active-site mutations, we consistently observed a normalised growth in the *E. coli* lag-time assay and a loss of β -casein phosphorylation in the radioactive kinase assay, thus revealing their impaired kinase function (Fig. 2c,d). In the lag-time assay, the mutations reduced the toxic effect of the McsB kinase without preventing its leaky expression (Supplementary Fig. 4a,b). To confirm that the demonstrated lack of activity is caused by an incomplete active site rather than compromised protein stability, we validated the structural integrity of each mutant in thermofluor experiments (Supplementary Fig. 5a,b). To further corroborate the mutational analysis data, we developed an *in vivo* assay monitoring the activity of endogenous McsB in *Bacillus subtilis*. Since McsB is required for targeting proteins to pArg-dependent degradation, we probed the kinase function by monitoring the stability of its major substrate, CtsR12,30. In fact, as predicted by our structural analysis, mutating the catalytic Glu122 (E122A) of the endogenous *mcsB* stabilised CtsR to an extent comparable to that observed for the *mcsB* deletion (Fig. 2e). According to these data, residues implicated in positioning ATP and the arginine side-chain in PhKs are essential for the protein phosphorylation activity of McsB *in vivo*.

We next examined the nucleotide-binding properties of the protein arginine kinase. When superposing apo and AMP-PN-bound McsB subunits, we noticed that, as in PhKs31, the catalytically competent active site arrangement in McsB is only achieved upon nucleotide binding, which triggers closure of the loop at residues 210–220 (the 'lid') (Fig. 2f; electron density maps in Supplementary Fig. 3b,c). This conformational change brings Glu213, a lid residue that is part of the conserved catalytic set, close to the predicted reaction site. The importance of the accurate positioning of Glu213 is highlighted by the inactivating effect of the E213D mutation, which changes the length but not the chemical properties of the side chain (Fig. 2c,d). To further test the relevance of the nucleotide-dependent lid opening and closing, we mutated Tyr211 and His92, two side chains that form different interactions in the open and closed states and thus seem to be involved in the conformational switch (Fig. 2f). Replacing either of these residues by alanine resulted in a loss of detectable *in vitro* activity and a normalised growth of *E. coli* cells in the lag-time assay (Fig. 2c,d). The impairment of activity was further confirmed *in vivo* upon introducing a mutation equivalent to Y211A into endogenous *mcsB* in *B. subtilis*, which led to robust CtsR stabilisation (Fig. 2e). Together, the *in vitro* and *in vivo* results strongly suggest that, despite a large evolutionary distance, McsB and PhKs utilise the same set of catalytic residues to phosphorylate guanidino-containing compounds. The conservation even extends to a defined open-closed conformational switch that yields a functional active site once nucleotide is bound. Indeed, the McsB/PhK active site is so far the only known catalytic architecture that can catalyse the challenging chemical reaction of guanidino group phosphorylation.

Specificity for protein arginine residues

The observed high degree of structural homology between the active sites of McsB and PhKs raises the question of how these two enzyme classes can modify substrates of very

different size and complexity (proteins vs. small guanidino compounds), thus contributing to distinct biological processes (protein quality control vs. energetic homeostasis). Therefore, we next turned our attention to structural differences between McsB and PhKs. In each case, the catalytic domain is accompanied by a different auxiliary domain. Whereas McsB has the C-terminal DD, PhKs possess an N-terminal ‘small’ domain (Supplementary Fig. 3a,b). In PhKs, it is the small domain that defines the specificity towards cognate small molecules^{25,32,33} and that forms, together with an extended loop-helix-loop motif protruding from the large domain, a dynamic gate to the substrate-binding site³⁴. Interestingly, these two regions are different in McsB (Supplementary Fig. 3c). The small domain is absent altogether, while the loop-helix-loop segment (residues 92–105, encompassing helices $\alpha 2$ and $\alpha 3$) is shorter by 10 residues. These adaptations remove a major part of the active site boundary, creating sufficient space to accommodate polypeptide substrates (Supplementary Fig. 3d). In agreement with this model, an McsB fusion protein that has the small domain of AK as an additional N-terminal steric block is inactive in the radioactive kinase assay (Supplementary Fig. 6e). It is noteworthy that, while WT McsB does not show any measurable free arginine kinase activity, it is able to phosphorylate short CtsR-derived peptides, as demonstrated using intact mass spectrometry (Supplementary Fig. 7a–c and ref. 20). This suggests that McsB specifically recognizes arginine residues present in a peptide context. As the protein arginine residues targeted by McsB *in vivo* do not share an obvious consensus motif¹¹, the interactions with native substrates are expected to be formed primarily with the main chain. However, the sequence of neighbouring side chains is likely to have some effect on recruitment efficiency. Since the area around the active site is strongly negatively charged (Supplementary Fig. 7d), we hypothesised that McsB might disfavour, through electrostatic repulsion, substrate motifs that contain glutamate or aspartate residues near the arginine phospho-acceptor. Indeed, in the peptide Ac-KRGGGGYIKIIV, substituting one or three lysine residues with glutamate residues decreased phosphorylation to undetectable levels (Supplementary Fig. 7c). Although this electrostatic filtering is expected to play a role in substrate selection *in vivo*, it does not seem to completely exclude motifs with negatively charged residues, as these can be found among sites identified in bacteria^{10–15}. Of note, the negatively charged patch bordering the active site entrance may also impact nucleotide binding, as ATP and ADP had relatively weak binding affinities to McsB, reflected in dissociation constant (K_D) values of $262 \pm 27 \mu\text{M}$ and $208 \pm 60 \mu\text{M}$, respectively (Supplementary Fig. 8a,b). While the exact mechanism of substrate engagement and Michaelis complex formation by McsB remains to be resolved, we conclude that the specificity of McsB for certain proteins and peptides and against isolated arginine is determined by the accessibility of the active site, interactions with the peptide backbone, and electrostatic effects.

Dimerisation and putative negative cooperativity

Our crystal structures reveal a dimeric arrangement of McsB (Fig. 1c,d). The dimer interface runs along helix $\alpha 8$ formed by residues 265–281 and comprises polar interactions between highly conserved residues, such as the Arg273–Glu288* salt bridge and the Phe272–Arg282* cation– π interaction (Fig. 3a). The conservation of the interface as well as the large buried surface area of $1505 \text{ \AA}^2$ suggest a stable association, a conclusion that is at variance with a published report describing McsB as a monomer³⁵. Since the previous

characterisation was based on behaviour in size-exclusion chromatography (SEC), which provides an indirect estimate of molecular weight (MW), we sought to obtain a more accurate measurement of MW by using native mass spectrometry (MS) (Fig. 3b and Supplementary Fig. 9) and SEC coupled to multi-angle light scattering (MALS) (Fig. 3c). Both methods unequivocally identified the dimer as the predominant oligomeric form of McsB in solution, with tiny amounts of higher oligomers seen in native MS experiments. Moreover, consistent with the interface observed in the crystal structures, dimerisation was largely abolished by a charge-reversal mutation of Arg273 (R273E). The narrow charge state distribution in the native MS spectrum of McsB-R273E indicates that the monomeric form of McsB is still well-folded³⁶, as further confirmed by thermofluor and circular dichroism (CD) measurements (Supplementary Fig. 5a–e). We subsequently tested a series of DD–DD* interface mutants in SEC-MALS (Fig. 3d and Supplementary Fig. 5f,g) and in the radioactive kinase assay (Fig. 3e). While individual mutations might cause additional effects, the overall tendency for the kinase activity to decline with decreasing dimerization (down to 5–10% of the WT for R273E and R282E mutants) clearly indicates that the association between the two protomers contributes to the activity. To test the implications of this observation *in vivo*, we introduced a mutation equivalent to R273E into endogenous *mcsB* in *B. subtilis* and monitored CtsR degradation. Although CtsR was stabilised (Fig. 2e), this effect was less pronounced than it would be predicted from the severe kinase activity impairment seen *in vitro*. We presume that the McsB reaction is not rate-limiting for CtsR degradation and that other factors may affect the oligomeric state or activity of McsB in the cell. Taken together, our analysis shows that dimerisation, while not strictly necessary for McsB activity, is nonetheless key to achieving its full catalytic potential.

Unlike the apo McsB structure, which shows an approximately symmetric dimer, the AMP-PN-bound state is asymmetric, being composed of one liganded and one unliganded subunit (Fig. 1d). The lid is found to be in the closed conformation in the AMP-PN-bound protomer, but in the open conformation in the empty protomer. The asymmetry of this state further extends to protein dynamics, with the empty subunit having increased thermal motion (B) factors relative to its nucleotide-bound partner subunit and lacking electron density for three substantial loops (Fig. 3f). Although asymmetry in crystal structures may reflect packing constraints and should be interpreted with caution, the described features of the AMP-PN-bound McsB might hint at negative cooperativity or, more generally, allosteric communication between the two active sites. If such communication were to promote catalysis, this would also explain why dimerisation is important for McsB activity despite being dispensable for stability. In that respect, the observed asymmetry in occupancy and B-factors is in line with a mechanism recently proposed for a dimeric model enzyme, fluoroacetate dehalogenase, where the catalysis in one liganded protomer is facilitated by an increase in disorder in the other empty protomer, which counterweighs entropy losses suffered upon substrate binding³⁷. However, whether a similar allosteric mechanism holds true for McsB awaits further experimental investigation.

The C-terminal DD of McsB contains a pArg-binding pocket

We next focused our analysis on the structure of the DD, a domain that appeared unique to McsB. Surprisingly, its architecture closely resembles that of the C-terminal part of the CtsR

transcription factor (Fig. 4a). While the overall sequence identity is low (~15%), structure-guided sequence alignments uncovered a shared RDXXR motif (Supplementary Fig. 10a), which forms a surface pocket composed of a positively and a negatively charged area. Since this rare bipolar arrangement is reminiscent of the recently identified pArg recognition sites of ClpC12, we hypothesised that the DD might bind pArg residues. Indeed, upon soaking McsB crystals with the pArg amino acid, we observed additional electron density in each pocket of the dimer (Fig. 4b, Supplementary Table 1). We were also able to obtain a co-crystal structure of the homologous C-terminal domain of CtsR with pArg, validating the discovered phospho-receptor site (Fig. 4c, Supplementary Table 1). While the folds of the pArg-binding domains of McsB/CtsR and ClpC are unrelated, they appear to have evolutionarily converged on a similar mode of binding, recognising the guanidino group via a carboxylic amino acid residue and the phosphoryl group via appropriately spaced positively charged residues and hydrogen-bond donors (Fig. 4d and Supplementary Fig. 10b). The affinity of McsB for free pArg measured in solution is low ($K_D = 850 \pm 400 \mu\text{M}$), but we could confirm that the detected binding is mediated by the identified pocket, as it disappeared in the R338A/D339A mutant affecting the RDXXR motif (Supplementary Fig. 8c). The native ligand is expected to be a pArg-carrying protein, which might form additional interactions and thus bind more tightly. Accordingly, the pArg-binding pockets in McsB and CtsR add to the reader domain of ClpC, suggesting a complex pArg signalling system operative in Gram-positive bacteria.

Putative allosteric control of McsB by pArg polypeptides

Strikingly, the structure of the pArg-bound McsB dimer, although devoid of any nucleotide, shows an asymmetric conformation that is similar to that for the AMP-PN-bound McsB, with one closed and one open active site and an unequal distribution of B-factors (Fig. 4e and Supplementary Fig. 11). According to these structural data, ligand binding to the pArg pocket appears to be communicated to the active site. To test whether pArg binding is involved in allosteric modulation of McsB activity, we monitored β -casein phosphorylation at low ATP concentration, where the reaction is slow and can be resolved in detail by our radioactive kinase assay (Fig. 5a,b). For WT McsB, the reaction velocity is initially low and only increases after a considerable delay, suggesting that the enzyme is activated by a product that must accumulate above a certain threshold. To test whether the product that activates McsB is the arginine-phosphorylated form of β -casein, we analysed the kinetics of the pArg binding-deficient R338A/D339A mutant. The mutant, although forming a stable dimer (Supplementary Fig. 5f,g), exhibited slow kinetics comparable to the initial velocity of the WT McsB. Strikingly, adding exogenous pArg amino acid to the reaction led to instantaneous activation of the WT McsB but not of the pArg binding-deficient mutant. These data clearly indicate that the engagement of the pArg pocket is necessary and sufficient to exert stimulation (Fig. 5). Moreover, it should be noted that our model substrate β -casein naturally contains five exposed phosphoserine residues³⁸. Given a delayed boost phase in the absence of the pArg activator (Fig. 5a), *O*-phosphorylations are apparently incapable of activating McsB, as would also be expected from the distinct shape and electrostatics of the phospho-binding site.

After seeing the stimulatory effect of free pArg, we next moved to a phospho-peptide as a more relevant effector mimicking interactions with pArg-containing proteins. For this purpose, we used a lysozyme-derived peptide (AWRNRCKGTDVQAWIRGCRL) that could be quantitatively phosphorylated by McsB (Supplementary Fig. 7e). This pArg peptide was very efficient in stimulating the kinase activity of McsB (Fig. 5d). By comparing the initial velocity at various pArg peptide concentrations, we determined the concentration needed to achieve half-maximal activation (EC_{50}) to be $21 \pm 9 \mu\text{M}$, with as little as $5 \mu\text{M}$ exerting a significant, >2-fold increase in rate. The stimulatory efficiency at such low concentration would be expected for a biologically relevant effector molecule. Taken together, our data suggest that the identified pArg-binding pocket is involved in the allosteric control of the enzymatic activity of the McsB kinase, providing regulatory access to pArg-containing effector and substrate proteins.

Discussion

The present study provides the first structural picture of a protein arginine kinase while simultaneously completing the structural and biochemical analysis of the core pArg writer-eraser-reader toolkit of Gram-positive bacteria. We show that McsB catalyses protein arginine phosphorylation using a domain that is also found in PhKs, animal metabolic enzymes that buffer ATP levels by catalysing reversible phosphorylation of guanidino compounds. The core active site is remarkably conserved between McsB and PhKs, despite the enormous evolutionary distance separating bacteria from animals. This is consistent with previous suggestions that guanidino group phosphorylation is catalysed primarily by near-perfect alignment of the reacting compounds^{25,26,39}. Such alignment can only be achieved by a complete set of appropriately positioned residues, leaving little room for evolutionary plasticity at the heart of the enzyme.

Because McsB represents a novel class among protein kinases, we compared it to classic eukaryotic protein kinases that act on serine, threonine, and tyrosine residues (Fig. 6a). In addition to the particular reliance of McsB on substrate alignment, the crucial difference between the two kinase types lies in the distinct mechanism of activating the reacting residue for phosphotransfer. Whereas the eukaryotic protein kinases typically interact with the substrate's hydroxyl moiety through only one aspartate residue^{40,41}, McsB – like PhKs – employs three negatively charged residues to position and activate the arginine guanidino group for nucleophilic attack on the ATP γ -phosphorus atom.

While the catalytic core of McsB and PhKs is highly conserved, the two enzyme classes have evolutionarily diverged in another respect. Our comparative analysis demonstrates how structural adaptations in the surroundings of the active site allow McsB and PhKs to target completely different classes of guanidino compounds and, therefore, to serve distinct roles in the cell. Whereas PhKs have selectivity filters that accept only specific small energy-storage compounds, the active site of McsB is accessible to, and indeed actively selects, arginine residues that are part of a polypeptide chain. The McsB–PhK case thus illustrates how new enzymatic and biological functions evolve through preservation of the catalytic properties with concomitant alterations in substrate specificity.

Our study also underscores the role of auxiliary domains in enzyme regulation. By mediating both dimerisation and recognition of pArg residues, the DD domain present in McsB but absent from PhKs turns the bacterial protein arginine kinase into an allosteric enzyme, allowing a precise tuning of its catalytic activity (Fig. 6b). Such regulation seems particularly important for a cellular factor that is central to both the heat shock response (which involves coordinated launching of protective mechanisms) and protein degradation (where selectivity and timing are key). Specifically, our data suggest that the pArg-binding pocket in the DD mediates the stimulation of the kinase by its products, pArg-carrying proteins or peptides. It is tempting to speculate that *in vivo* this mechanism places McsB under two control circuits acting on different time scales: while the protein levels of the kinase are under transcriptional control relying on the McsB-mediated inhibition of the repressor CtsR, the enzyme activity directly reacts to pArg-containing proteins by the identified allosteric mechanism. The two interlinked, positive feedback-loops could together ensure induction of the stress response in a fast but robust manner, as suggested for other feedback systems⁴². Additionally, the uncovered pArg pocket could mediate recruitment of already-phosphorylated substrates for further rounds of phosphorylation. This product stimulation could be important, for example, for the efficient marking of large protein aggregates, the preferred *in vivo* substrates of McsB, as targets for degradation.

The pArg-binding DD of McsB has a homologue in the heat-shock repressor CtsR, and both domains add to the evolutionarily unrelated pArg-binding pocket of ClpC. These new findings underscore the importance of pArg as a signal that can trigger new protein–protein interactions. Indeed, if phospho amino-acid residues are well-suited for specific recognition by a cognate protein module in general⁴³, then the pArg residue, with its zwitterionic nature, extended shape, and the ability to both donate and accept multiple hydrogen bonds seems tailor-made for this purpose. The relative complexity of the writer-eraser-reader pArg toolkit (Fig. 6c) challenges the established perception that prokaryotic phospho-signalling systems, unlike their eukaryotic counterparts, are generally restricted to simple linear on/off pathways with little role for phosphorylation-dependent protein–protein interactions. This notion was previously suggested by the scarcity of bacterial homologues of known phospho-binding domains, which, except for sporadic cases of horizontal gene transfer, seem limited to phosphothreonine-specific FHA domains⁸. The identified pArg-binding domains, however, extend the spectrum of known bacterial phospho-binding domains by providing first examples which go beyond modules known from eukaryotes. Given the important role of the pArg signalling system in virulence of human pathogens such as *S. aureus*^{15,17}, our structural insights into the writer, eraser, and readers of this modification may facilitate the development of badly needed new antimicrobial interventions.

Online Methods

Cloning for recombinant expression in *E. coli*

Truncated *mcsB* (356-363) from *G. stearothermophilus* NCA 26 (ATCC 12980), *ctsR* (1-75) from *Bacillus subtilis*, and genes encoding arginine kinase (AK) from *Limulus polyphemus* and fusion proteins AK^{small}-McsB (residues 1-118 of AK and 24-363 of McsB) and AK^{small}-McsB^{PD} (residues 1-118 of AK and 24-264 of McsB) were each cloned into a

pET-21a vector, yielding respective constructs: pET21-mcsB(1-356)-His₆, pET21-ctsR(76-154)-His₆, pET21-argk-His₆, pET21-argk(1-118)-mcsB(24-363)-His₆, pET21-argk(1-118)-mcsB(24-264)-His₆. For truncated *clpA* (1-146) from *Escherichia coli* and full-length *ywlE* and *mcsB* from *G. stearothermophilus*, published pET-21a-derived plasmids were used: pET21-clpA(1-146)-His₆12, pET21-ywlE-His₆21, pET21-mcsB-His₆20. Point mutants of full-length *mcsB* were generated using a standard site-directed mutagenesis protocol. All primers used are listed in Supplementary Table 3.

Protein expression and purification

Proteins were expressed in *E. coli* BL21 (DE3) strain grown in LB containing 50 µg/ml ampicillin. Expression was induced with 0.5 mM IPTG and allowed to proceed for either 12 h at 18 °C (356-363 McsB) or 3 h at 37 °C (other proteins). All proteins were purified on an ÄKTA FPLC system (GE Healthcare Life Sciences). YwlE was purified as described previously²¹. Purification of other proteins was performed with Ni-nitrilotriacetic (Ni-NTA) affinity chromatography in 10 mM Tris, pH 7.5, 150 mM NaCl and increasing imidazole as the first step. For all proteins except for 356-363 McsB and 1-75 CtsR, NiNTA was followed by a Superdex200 size-exclusion column in 10 mM Tris, pH 7.5, 50 mM NaCl.

356-363 McsB after NiNTA was purified by buffer exchange to 20 mM Tris, pH 8, 50 mM NaCl; dephosphorylation with YwlE (400 µM McsB was mixed with 700 µM YwlE) at room temperature for 1 h; and two chromatography steps for YwlE removal: a Superdex200 size-exclusion column in 50 mM Tris, pH 8, 50 mM NaCl and a Resource Q anion exchange chromatography column in 50 mM Tris, pH 8, 50 mM NaCl and a 50 to 500 mM NaCl gradient. 1-75 CtsR after NiNTA was purified by a Superdex200 size-exclusion column in 6 mM Na/K phosphate, pH 7.25, 1.35 mM KCl, 68.5 mM NaCl. Proteins were concentrated using Vivaspin devices (Sartorius). Concentration was determined using Bradford protein assay (Bio-Rad) and/or absorbance at 280 nm.

Dot-blot analysis

McsB (E122A, WT, or 356-363, in each case at 400 µM) was incubated either alone or with YwlE (at 700 µM) for 1 h at room temperature in 20 mM Tris, pH 8, 50 mM NaCl.

356-363 McsB after NiNTA (without dephosphorylation) was used. 2 µl of each sample were pipetted onto nitrocellulose membrane. The membrane was blocked with 3% bovine serum albumin (BSA) in TBS (50 mM Tris, pH 7.5, 150 mM NaCl) and immunoblotted with the 2 µg/ml human F_{(ab')₂} anti-pArg primary antibody²¹ followed by 1:7,000 dilution of horseradish peroxidase (HRP)-coupled anti-human IgG F_{(ab')₂} antibody (AbD Serotec, now Bio-Rad). The signal was visualised using the Clarity enhanced chemiluminescence (ECL) reagent (Bio-Rad).

Crystallisation, data collection, structure solution, refinement, and analysis

Monoclinic 356-363 McsB crystals were obtained within 3 days at 4 °C by sitting drop vapour diffusion upon mixing 100 nl of the optimised Morpheus screen (Molecular Dimensions) solution (25% v/v ethylene glycol, 75 mM MES-imidazole, pH 6.2, and 58 mM of each of the following: Na formate, NH₄ acetate, Na citrate, Na K tartrate, Na oxamate) with 200 nl of protein solution containing 480 µM 356-363 McsB in 50 mM Tris, pH 8, 240 mM NaCl supplemented with 1 mM Tris(2-carboxyethyl)phosphine (TCEP).

Soaking was performed by incubation with ligand (2 mM pArg, Sigma Aldrich, or 10 mM MgCl₂ and 10 mM AMP-PNP, Jena Bioscience) dissolved in the reservoir solution supplemented with 1 mM TCEP. McsB crystals were cryo-protected in 40% v/v PEG 400 and 60% reservoir solution supplemented with, for the soaked crystals, the matching ligand. Electron density analysis demonstrates that AMP-PNP hydrolysed in the slightly acidic crystallisation condition to an ADP analogue, AMP-PN.

Orthorhombic 1-75 CtsR·pArg crystals formed within 3 days at 19 °C by the same technique upon mixing 100 nl of 10% v/v polypropylene glycol P 400 with 200 nl of protein solution containing 2.4 mM 1-75 CtsR in 6 mM Na/K phosphate, pH 7.25, 1.35 mM KCl, 68.5 mM NaCl, 5 mM pArg. CtsR crystals were cryo-protected in 40% PEG 400, 6 mM pArg, 6 mM Tris, pH 7.5, 90 mM NaCl.

Diffraction data were collected at 100 K at beamlines: ID23-1 (McsB) and ID23-2 (McsB·AMP-PN) at ESRF, Grenoble; P13 at DESY, Hamburg44 (McsB·pArg); and at an in-house Rigaku X-ray generator (1-75 CtsR·pArg). The wavelengths used were 0.979, 0.873, 0.976 Å, and 1.542 Å, respectively. All data were integrated and scaled using XDS45. For McsB, an initial low-confidence molecular replacement solution obtained with Phaser46 using a fragment of PDB entry 1M15 as a search model was improved by iterative manual and automatic rebuilding with O47, Coot48, and ARP/wARP49. Ligand-bound McsB structures were solved by molecular replacement in Phaser using the final apo model, followed by rebuilding and ligand placing in Coot. The 1-75 CtsR·pArg structure was solved by molecular replacement in Phaser using a fragment of PDB entry 3H0D as a search model and further rebuilding and ligand placing in Coot. All structures were refined with CNS50 and PHENIX51 to good Ramachandran statistics with 0.15, 0, 0.29, and 0% outliers for McsB, McsB·AMP-PN, McsB·pArg, and 1-75 CtsR·pArg, respectively. Data collection and refinement statistics are provided in Supplementary Table 1. PyMOL (Schrodinger) was used for figure preparation.

Sequence and structure alignments

Sequence alignment of C-terminal regions of McsBs from various *Bacilli* was performed using MAFFT 752. Structure-guided sequence alignment of C-terminal regions of selected McsB and CtsR proteins from *Bacilli* was performed using PROMALS3D53. Structure superpositions were performed in PyMOL (Schrodinger) and RMSD calculations with SuperPose version 1.054.

Radioactive protein kinase assay

All experiments except for those shown in Fig. 5c and d consisted of a single step in which 55 µM β-casein (Sigma Aldrich) was incubated with 0.5 µM McsB and either 200 µM (for experiment illustrated in Fig. 5a and b), or 12.5 mM ATP (for all other experiments) spiked with [γ -³²P]ATP from Perkin Elmer for the indicated time periods at 37 °C in 55 mM Tris, pH 7.5, 55 mM NaCl, 2% glycerol, 10 mM β-mercaptoethanol, and 12.5 mM MgCl₂. In the experiment shown in Fig. 5a and b, 10 mM phosphoarginine (Sigma Aldrich) was added to indicated samples. The reaction was stopped by adding denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) sample buffer. After resolving the samples

on a 15% SDS-PAGE gel, radioactive signal was visualised using Phosphorimager (GE Healthcare Life Sciences). Bands from Phosphorimager scans were quantified using ImageJ (NIH). To average results obtained from different repeats of the same radioactive kinase assay experiment, the volumes of all bands were expressed as fractions of the volume of the band for the final time point of the WT sample (in the experiment shown in Fig. 5, specifically: WT sample without pArg) in the same scan, which was treated as 1 arbitrary unit (arb. unit).

The experiment illustrated in Fig. 5c and d was performed in two steps. In the first step, 1.1 mM of peptide (AWNRCKGTDVQAWIRGCRL), as well as a control reaction in the absence of the peptide, was incubated with 2.5 μ M McsB, 333 μ M ATP, 12.5 mM MgCl_2 , 10 mM phosphoenolpyruvate, 10 U/ml pyruvate kinase (Sigma-Aldrich) and 10 mM KCl in 50 μ l of 55 mM Tris, pH 7.5, 55 mM NaCl, 2% glycerol, 10 mM β -mercaptoethanol for 2 hours at 37 °C degrees. McsB was removed from both reactions using Spin Cup columns (Pierce) packed with 50 μ l (25 μ g) Ni-NTA Superflow beads (Qiagen). The peptide reaction was analysed with intact mass spectroscopy to confirm phosphorylation (see Supplementary Fig. 7e). In the second step, the two reactions were first mixed in an appropriate ratio depending on the desired peptide concentration and then diluted 5-fold with the reaction buffer containing fresh supply of McsB. The reactions were started with the addition of β -casein substrate spiked with [γ - ^{32}P]ATP (Perkin Elmer). The final reactions contained 0.5 μ M McsB, 50 μ M ATP spiked with hot ATP, 12.5 mM MgCl_2 , 2 mM phosphoenolpyruvate, 2 U/ml pyruvate kinase, and 2 mM KCl in reaction buffer. The use of ATP regeneration system will result in diluting hot ATP (which is not regenerated) with cold ATP (which is) over time, but this should occur in the same way in all samples regardless of peptide concentration. In the experiment shown in Fig. 5d, β -casein phosphorylation at 25 min, which is within the linear reaction phase according to Fig. 5c, was used to estimate relative reaction rate for increasing peptide^{pArg} concentrations. 3-parameter agonist vs. response curves were fitted individually to each data series using Prism version 7.04 (GraphPad) and the obtained EC_{50} values were averaged.

Bacterial culture growth conditions

All bacterial strains used (Supplementary Table 4) were cultivated either on plates with Lysogeny Broth (LB; 10 g NaCl, 5 g yeast extract, 10 g tryptone per l) solidified with 1.5% (w/v) agar or under vigorous agitation in liquid culture, in both cases, unless otherwise stated, at 37 °C. For liquid culture, we used LB or a modified version of a previously described synthetic medium⁵⁵ (50 mM Tris-HCl, pH 7.5, 15 mM $(\text{NH}_4)_2\text{SO}_4$, 8 mM MgSO_4 , 27 mM KCl, 7 mM Na citrate, 0.6 mM KH_2PO_4 , 2 mM CaCl_2 , 1 μ M FeSO_4 , 10 μ M MnSO_4 , 4.5 mM glutamate, 98 μ M tryptophan, and 0.2% (w/v) glucose).

Bacterial strains

The *B. subtilis* strains described in Supplementary Table 4 as from “this study” were generated by introducing mutations to the chromosome of the laboratory strain 168 of *B. subtilis* using the pMAD vector-based protocol⁵⁶. First, two 2-kB-long fragments of chromosomal DNA that span the 5'- and 3'-end halves of the *mcsB* gene were cloned separately into a pMAD vector, yielding plasmids pMAD-*mcsB*_{Nterm} and pMAD-

*mcsB*_{Cterm}, respectively. This was then followed by Q5 site-directed mutagenesis (New England Biolabs) of the vector best suited for a given point mutation; integration into *B. subtilis* 168; and curing of the vector. The *mcsB* deletion strain (*mcsB*) was produced by mutating the third codon of *mcsB* into a stop codon using the same technique.

Serum and antibody generation

Serum against *B. subtilis* CtsR and *B. subtilis* McsB were produced in collaboration with Gramsch laboratories by injecting recombinant protein into rabbits. Polyclonal anti-McsB antibody was produced by purifying the latter serum using an N-Hydroxysuccinimide (NHS)-coupled resin (GE Healthcare Life Sciences) charged with McsB. After washing with phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄), the antibody was eluted with 100 mM glycine, pH 2.6, pH-adjusted, and dialysed into PBS supplemented with 10% glycerol.

Lag-time assay

The experiment was performed in LB containing 50 µg/ml ampicillin at 37 °C. A set of pET-21 plasmids harbouring a control gene (encoding the fragment 1-150 of ClpA from *E. coli*), WT *mcsB*, or one of various point mutants of full-length *mcsB* were individually transformed into chemically-competent *E. coli* BL21 (DE3) cells. 5-ml cultures were grown in round-bottom falcons (Corning) for ~12 h, until the *OD*₆₀₀ of at least 1 AU. If this high density was not achieved, the growth-impairing effect of *mcsB* appeared less pronounced. The overnight cultures were each diluted into 5 ml of fresh medium to the calculated *OD*₆₀₀ of 0.0045. The *OD*₆₀₀ was then measured at regular time intervals using a small-volume spectrophotometer/fluorimeter DS-11 FX+ (DeNovix). Measured growth curves were fitted to the 3-parameter 'modified logistic equation'⁵⁷ given below by minimising squared errors between actual and predicted values using Excel Solver tool (Microsoft).

Modified logistic equation:

$$d(t) = \frac{A}{1 + e^{\frac{4\mu}{A}(\lambda - t) + 2}}$$

where $d(t)$ is the *OD*₆₀₀ at time t , A is a maximal density (allowed to be different for each curve, but between 0.9 and 2, which were minimal and maximal *OD*₆₀₀ values of test cultures after prolonged, overnight cultivation), λ is lag time (allowed to be different for each curve), and μ is maximal growth rate (forced to be the same for all fitted curves). λ was a robust parameter that did not change significantly with A or μ . We used the unpaired two-tailed Student's t -test to test the null hypothesis that the λ values measured for a mutant are not different from those for WT. To verify McsB expression levels, cell pellets from 1.5-ml cultures at *OD*₆₀₀ of 0.3 were lysed in 40 µl of B-PER solution (Thermo Fisher Scientific) containing 50 µg/ml lysozyme (Sigma Aldrich) and 1 µg/ml DNase I (Roche). The lysates were separated on a 15% SDS-PAGE gel, transferred onto a polyvinylidene difluoride (PVDF) membrane (Thermo Fisher Scientific), and immunoblotted according to a standard protocol with 1:5,000 dilution of anti-McsB antibody generated in this study or 1:1,000

penta-His tag mouse monoclonal antibody (Qiagen). The signal was visualised using 1:10,000 dilution of an HRP-coupled anti-rabbit (Sigma Aldrich) or anti-mouse (Jackson ImmunoResearch) and Clarity ECL reagent (Bio-Rad).

Thermofluor/differential scanning fluorimetry assay

Protein stability of various McsB mutants were investigated using thermofluor/differential scanning fluorimetry. The reaction mixture consisted of 0.05 mg/ml protein under study and 7.5x SYPRO orange (ThermoFisher) in 55 mM Tris-HCl pH 7.5, 55 mM NaCl, 2% v/v glycerol, 10 mM β -mercaptoethanol and 12.5 mM $MgCl_2$ in a final volume of 20 μ l. Melting curves were monitored with a Biorad CFX96 thermal cycler and fluorescence signals were obtained using the FRET channel. The thermal cycling protocol consisted of equilibration at 25 °C for 1 minute and subsequent fluorescence measurement. The next steps included heating by 1 °C followed by 1-minute equilibration and fluorescence measurement until 95 °C. Melting curve analysis was performed using a previously reported Excel (Microsoft)-based software⁵⁸.

In vivo McsB activity assay

The experiment was performed in round-bottom falcons (Corning) in synthetic medium. 2-3 middle-sized colonies of the *B. subtilis* strains 168, *mcsB*, *mcsB_{E121A}*, *mcsB_{Y210A}*, and *mcsB_{R272E}* from an overnight culture were cultivated each in 5 ml of the medium for 2 h to adapt to its composition, prior to being diluted to the OD_{600} of 0.075 in the total volume of 6 ml. At the OD_{600} of 0.5, the cultures were moved to a shaker pre-warmed to 50 °C. 1-ml samples were taken at indicated time intervals. Cells were harvested by centrifugation and the pellets were lysed in 40 μ l of B-PER solution (Thermo Fisher Scientific) supplemented with 50 μ g/ml lysozyme (Sigma Aldrich) and 1 μ g/ml DNase I (Roche). The lysates were separated on a 15% SDS-PAGE gel, transferred onto a PVDF membrane (Thermo Fisher Scientific), and immunoblotted using a standard protocol with 1:2,000 dilution of anti-CtsR rabbit serum. The signal was visualised using 1:10,000 dilution of an HRP-coupled anti-rabbit secondary antibody (Sigma Aldrich) and Clarity ECL reagent (Bio-Rad). The tested *B. subtilis* McsB residues are referred to in Fig. 2e using *G. stearothermophilus* McsB residue numbering. Bands from immunoblots were quantified using ImageJ (NIH).

Free arginine kinase assay

We monitored free arginine kinase activity indirectly by measuring the rate of ADP generation in the presence of arginine using a coupled enzymatic reaction⁵⁹. 1.5 μ M full-length McsB or arginine kinase with or without 25 mM arginine was incubated at 25 °C with 37.5 U/ml pyruvate kinase, 42.9 U/ml lactate dehydrogenase, 0.25 mM NADH, 15 mM phosphoenolpyruvate, and 0.5 mM ATP in ATPase buffer (20 mM Hepes, pH 7.75, 100 mM NaCl, 15 mM KCl, 5 mM $MgCl_2$). Absorbance at 340 nm was recorded for 14 min using PHERAstar FS plate reader (BMG Labtech). ADP generation rate was calculated as described⁶⁰.

Arginine-phosphorylation and intact mass spectrometry of peptides

Two N-terminally acetylated CtsR-derived sequences, Ac-KRGGGGYIKIIV and Ac-YKVKSKRGGGGY, as well as two mutated versions of the first sequence (Ac-KRGGGGYIEIIV and Ac-KRGGGGYIEIIEV), were tested. 100 μ M peptide was incubated with 0.5 μ M McsB in the presence of 12.5 mM ATP at 37 °C for 2 hours in a buffer containing 55 mM Tris, pH 7.5, 55 mM NaCl, 1 mM TCEP, and 12.5 mM MgCl₂. Subsequently, McsB was removed from the reactions using Spin Cup columns (Pierce) packed with 30 μ l (15 μ g) Ni-NTA Superflow beads (Qiagen). The flow-through was collected and applied for a desalting protocol using C18 spin tips prior to intact mass spectrometry analysis. The peptides were eluted using 2 times 20 μ l 80% acetonitrile and used for intact mass spectrometry.

Intact MS of peptides was carried out on a Synapt G2Si instrument (Waters, Manchester, UK) with a nanoelectrospray ionisation (nESI) source. Mass calibration was performed by a separate infusion of NaI cluster ions. Solutions were ionised through a positive potential applied to metal-coated borosilicate capillaries (Thermo Fisher Scientific). Peptides were sprayed from 40% acetonitrile at a concentration of 10 μ M. Capillary voltage, sample cone voltage, extractor source offset and source temperature were set to 1.5 kV, 30 V, 20 V and 40 °C, respectively. Data were processed using MassLynx software version 4.1 (Waters).

Surface plasmon resonance (SPR)

SPR measurements were performed on a Biacore T200 instrument (GE Healthcare Life Sciences) at 25 °C. Full-length McsB was immobilised (11,800 resonance units) by capture-coupling to a Ni-loaded NTA Chip (GE Healthcare Life Sciences) by passing 4.8 μ M (200 μ g/mL) protein at a flowrate of 10 μ l/min in 10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% *v/v* surfactant P20 for a contact time of 300 s over the surface. Binding of ligands was probed by performing a 10-point titration (1:1 dilution steps), starting with either 2 (for ADP) or 1 (for pArg and ATP) mM ligand as the maximal concentration. For each titration point, the ligand was passed over the chip at a flow rate of 30 μ l/min in 25 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM MgCl₂, 0.005% Tween-20 for a contact time of 60 s and was subsequently washed away with buffer for a dissociation time of 120 s. Obtained sensorgrams were blank-subtracted. Dissociation constants (K_D s) were calculated using the Langmuir fit model of the BIAevaluation Software version 2.0 (GE Healthcare Life Sciences) from a plot of steady-state binding levels (obtained 5 s before the end of each injection at steady-state) against ligand concentration, assuming 1:1 binding stoichiometry.

Size-exclusion chromatography-multi-angle light scattering (SEC-MALS)

SEC-MALS analysis of 25- μ l full-length McsB samples at ~100 μ M (~4 mg/ml) was performed using a Superdex200 increase 5/150 GL column installed on an ÄKTA ETTAN LC system (GE Healthcare Life Sciences) and equilibrated in 50 mM Tris, pH 7.5, 50 mM NaCl. The flow rate of 0.33 ml/min was used. The column was connected to a miniDAWN TREOS light scattering detector (Wyatt Technologies). Average molecular mass across the central parts of main peaks was calculated using ASTRA software v6 (Wyatt Technologies).

Native mass spectrometry

Native MS experiments were carried out on a Synapt G2Si instrument (Waters, Manchester, UK) equipped with a nanoelectrospray ionisation (nESI) source. Mass calibration was performed by a separate infusion of Na iodide cluster ions. Solutions were ionised through a positive potential applied to metal-coated borosilicate capillaries (Thermo Fisher Scientific). Full-length McsB samples (5 μ M) were sprayed from 100 mM ammonium acetate, pH 7.85. During the observation of the multimeric state, the following parameters were used: capillary voltage 1.6 kV, sample cone voltage 60 V, extractor source offset 80 V, source temperature 70 °C. Data were processed using MassLynx software version 4.1 (Waters).

Circular dichroism

Raw ellipticity [θ] of 4 μ M (0.17 mg/ml) WT or R273E full-length McsB in 50 mM Na phosphate, pH 7.52, 50 mM was recorded over a range of wavelengths (260–195 nm) at increasing temperatures (20 to 90 °C, 1 °C/step, 1 °C/min heating rate) using a Chirscan-plus Circular Dichroism Spectrometer (Applied Photophysics) and Chirscan Single Cell Peltier Temperature Controller (Applied Photophysics). The temperature was monitored with a thermal probe placed in the cuvette. [θ] values for R273E McsB were scaled by a factor of 1.26 (the ratio of absorption values measured at 205 nm for WT and R273E McsB) to account for error in initial protein concentration estimation. Spectra obtained at a range of temperatures were deconvoluted using the Global 3 software (Applied Photophysics) into a linear combination of ‘folded’ and ‘unfolded’ spectra to estimate the percentage of unfolded state at each temperature and, by fitting a sigmoidal curve, the midpoint of transition temperatures (T_m).

Statistical analysis

Data analysis in X-ray crystallography, SEC-MALS, native MS, CD, and SPR, as well as EC_{50} calculations from the radioactive kinase assay measurements, was performed using standard specialised software outlined above. All additional statistical calculations were performed in Excel (Microsoft) with details, such as n values, the type of replicates and the test used, indicated in figure legends. In bar graphs, the bar represents the mean, dots – individual measurements, and error bars – standard deviation. The significance of difference between sets of measurements was assessed using two-tailed Student’s t -test (T.TEST function in Excel), either unpaired or paired, as indicated. The following significance limits were used: *, P 0.05; **, P 0.01; ***, P 0.001. Graphs were prepared in Excel and further modified in Illustrator (Adobe).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank R. Huber and all members of the Clausen group for remarks on the manuscript and discussions, J. Leodolter and M. Madalinski for support in preparing pArg-containing peptides, A. Schleiffer for help with bioinformatic analysis, A. Sedivy and P. Stolt-Bergner for assistance with CD spectroscopy measurements, N. Stanley-Wall (University of Dundee) for pMAD plasmid and advice on mutagenesis in *B. subtilis*, and staff of beamlines at ESRF (Grenoble), SLS (Villigen), and DESY (Hamburg) for excellent help during data collection.

This work was supported by a grant from the European Research Council (AdG 694978, to T.C.). The IMP is supported by Boehringer Ingelheim.

References

1. Cohen P. The origins of protein phosphorylation. *Nature cell biology*. 2002; 4:E127–130. DOI: 10.1038/ncb0502-e127 [PubMed: 11988757]
2. Hunter T. Signaling--2000 and beyond. *Cell*. 2000; 100:113–127. [PubMed: 10647936]
3. Hunter T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell*. 1995; 80:225–236. [PubMed: 7834742]
4. Eckhart W, Hutchinson MA, Hunter T. An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell*. 1979; 18(4):925–933. [PubMed: 229973]
5. Waksman G, Kominos D, Robertson SC, Pant N, Baltimore D, Birge RB, Cowburn D, Hanafusa H, Mayer BJ, Overduin M, Resh MD, et al. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature*. 1992; 358(6388): 646. [PubMed: 1379696]
6. Seet BT, Dikic I, Zhou MM, Pawson T. Reading protein modifications with interaction domains. *Nature reviews. Molecular cell biology*. 2006; 7:473–483. DOI: 10.1038/nrm1960 [PubMed: 16829979]
7. Jin J, Pawson T. Modular evolution of phosphorylation-based signalling systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 2012; 367:2540–2555. DOI: 10.1098/rstb.2012.0106 [PubMed: 22889906]
8. Pallen M, Chaudhuri R, Khan A. Bacterial FHA domains: neglected players in the phospho-threonine signalling game? *Trends in microbiology*. 2002; 10:556–563. [PubMed: 12564991]
9. Hoch JA, Varughese KI. Keeping signals straight in phosphorelay signal transduction. *Journal of bacteriology*. 2001; 183:4941–4949. [PubMed: 11489844]
10. Elsholz AK, Turgay K, Michalik S, Hessling B, Gronau K, Oertel D, Mäder U, Bernhardt J, Becher D, Hecker M, Gerth U. Global impact of protein arginine phosphorylation on the physiology of *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:7451–7456. DOI: 10.1073/pnas.1117483109 [PubMed: 22517742]
11. Schmidt A, Trentini DB, Spiess S, Fuhrmann J, Ammerer G, Mechtler K, Clausen T. Quantitative phosphoproteomics reveals the role of protein arginine phosphorylation in the bacterial stress response. *Molecular & cellular proteomics : MCP*. 2014; 13:537–550. DOI: 10.1074/mcp.M113.032292 [PubMed: 24263382]
12. Trentini DB, Suskiewicz MJ, Heuck A, Kurzbauer R, Deszcz L, Mechtler K, Clausen T. Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature*. 2016; 539:48–53. DOI: 10.1038/nature20122 [PubMed: 27749819]
13. Trentini DB, Fuhrmann J, Mechtler K, Clausen T. Chasing Phosphoarginine Proteins: Development of a Selective Enrichment Method Using a Phosphatase Trap. *Molecular & cellular proteomics : MCP*. 2014; 13:1953–1964. DOI: 10.1074/mcp.O113.035790 [PubMed: 24825175]
14. Bäsell K, Otto A, Junker S, Zühlke D, Rappen GM, Schmidt S, Hentschker C, Macek B, Ohlsen K, Hecker M, Becher D. The phosphoproteome and its physiological dynamics in *Staphylococcus aureus*. *International journal of medical microbiology : IJMM*. 2014; 304:121–132. DOI: 10.1016/j.ijmm.2013.11.020 [PubMed: 24457182]
15. Junker S, Maaß S, Otto A, Michalik S, Morgenroth F, Gerth U, Hecker M, Becher D. Spectral Library Based Analysis of Arginine Phosphorylations in *Staphylococcus aureus*. *Molecular & cellular proteomics : MCP*. 2018; 17:335–348. DOI: 10.1074/mcp.RA117.000378 [PubMed: 29183913]
16. Mijakovic I, Grangeasse C, Turgay K. Exploring the diversity of protein modifications: special bacterial phosphorylation systems. *FEMS microbiology reviews*. 2016; 40:398–417. DOI: 10.1093/femsre/fuw003 [PubMed: 26926353]
17. Wozniak DJ, Tiwari KB, Soufan R, Jayaswal RK. The mcsB gene of the clpC operon is required for stress tolerance and virulence in *Staphylococcus aureus*. *Microbiology*. 2012; 158:2568–2576. DOI: 10.1099/mic.0.060749-0 [PubMed: 22902728]

18. Singh LK, Dhasmana N, Sajid A, Kumar P, Bhaduri A, Bharadwaj M, Gandotra S, Kalia VC, Das TK, Goel AK, Pomerantsev AP. *clpC* operon regulates cell architecture and sporulation in *Bacillus anthracis*. *Environmental microbiology*. 2015; 17:855–865. DOI: 10.1111/1462-2920.12548 [PubMed: 24947607]
19. Schumann W. Regulation of bacterial heat shock stimulons. *Cell stress & chaperones*. 2016; 21:959–968. DOI: 10.1007/s12192-016-0727-z [PubMed: 27518094]
20. Fuhrmann J, Schmidt A, Spiess S, Lehner A, Turgay K, Mechtler K, Charpentier E, Clausen T. McsB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science*. 2009; 324:1323–1327. DOI: 10.1126/science.1170088 [PubMed: 19498169]
21. Fuhrmann J, Mierzwa B, Trentini DB, Spiess S, Lehner A, Charpentier E, Clausen T. Structural basis for recognizing phosphoarginine and evolving residue-specific protein phosphatases in gram-positive bacteria. *Cell reports*. 2013; 3:1832–1839. DOI: 10.1016/j.celrep.2013.05.023 [PubMed: 23770242]
22. Kruger E, Msadek T, Ohlmeier S, Hecker M. The *Bacillus subtilis* *clpC* operon encodes DNA repair and competence proteins. *Microbiology*. 1997; 143(Pt 4):1309–1316. DOI: 10.1099/00221287-143-4-1309 [PubMed: 9141693]
23. Suzuki T, Soga S, Inoue M, Uda K. Characterization and origin of bacterial arginine kinases. *International journal of biological macromolecules*. 2013; 57:273–277. DOI: 10.1016/j.ijbiomac.2013.02.023 [PubMed: 23523720]
24. Ellington WR. Evolution and physiological roles of phosphagen systems. *Annual review of physiology*. 2001; 63:289–325. DOI: 10.1146/annurev.physiol.63.1.289
25. Zhou G, Somasundaram T, Blanc E, Parthasarathy G, Ellington WR, Chapman MS. Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proceedings of the National Academy of Sciences of the United States of America*. 1998; 95:8449–8454. [PubMed: 9671698]
26. Yousef MS, Fabiola F, Gattis JL, Somasundaram T, Chapman MS. Refinement of the arginine kinase transition-state analogue complex at 1.2 Å resolution: mechanistic insights. *Acta crystallographica. Section D, Biological crystallography*. 2002; 58:2009–2017. [PubMed: 12454458]
27. Summerton JC, Evanseck JD, Chapman MS. Hyperconjugation-mediated solvent effects in phosphoanhydride bonds. *The journal of physical chemistry A*. 2012; 116:10209–10217. DOI: 10.1021/jp306607k [PubMed: 23009395]
28. Pruett PS, Azzi A, Clark SA, Yousef M, Gattis JL, Somasundaram T, Ellington WR, Chapman MS. The putative catalytic bases have, at most, an accessory role in the mechanism of arginine kinase. *The Journal of biological chemistry*. 2003; 278:26952–26957. DOI: 10.1074/jbc.M212931200 [PubMed: 12732621]
29. Gattis JL, Ruben E, Fenley MO, Ellington WR, Chapman MS. The active site cysteine of arginine kinase: structural and functional analysis of partially active mutants. *Biochemistry*. 2004; 43:8680–8689. DOI: 10.1021/bi049793i [PubMed: 15236576]
30. Kruger E, Zuhlke D, Witt E, Ludwig H, Hecker M. Clp-mediated proteolysis in Gram-positive bacteria is autoregulated by the stability of a repressor. *The EMBO journal*. 2001; 20:852–863. DOI: 10.1093/emboj/20.4.852 [PubMed: 11179229]
31. Yousef MS, Clark SA, Pruett PK, Somasundaram T, Ellington WR, Chapman MS. Induced fit in guanidino kinases—comparison of substrate-free and transition state analog structures of arginine kinase. *Protein science : a publication of the Protein Society*. 2003; 12:103–111. DOI: 10.1110/ps.0226303 [PubMed: 12493833]
32. Azzi A, Clark SA, Ellington WR, Chapman MS. The role of phosphagen specificity loops in arginine kinase. *Protein science : a publication of the Protein Society*. 2004; 13:575–585. DOI: 10.1110/ps.03428304 [PubMed: 14978299]
33. Suzuki T, Kawasaki Y, Furukohri T, Ellington WR. Evolution of phosphagen kinase. VI. Isolation, characterization and cDNA-derived amino acid sequence of lombricine kinase from the earthworm *Eisenia foetida*, and identification of a possible candidate for the guanidine substrate recognition site. *Biochimica et biophysica acta*. 1997; 1343:152–159. [PubMed: 9434106]

34. Davulcu O, Flynn PF, Chapman MS, Skalicky JJ. Intrinsic domain and loop dynamics commensurate with catalytic turnover in an induced-fit enzyme. *Structure*. 2009; 17:1356–1367. DOI: 10.1016/j.str.2009.08.014 [PubMed: 19836335]
35. Kirstein J, Zuhlke D, Gerth U, Turgay K, Hecker M. A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in *B. subtilis*. *The EMBO journal*. 2005; 24:3435–3445. DOI: 10.1038/sj.emboj.7600780 [PubMed: 16163393]
36. Beveridge R, Covill S, Pacholarz KJ, Kalapothakis JM, MacPhee CE, Barran PE. A mass-spectrometry-based framework to define the extent of disorder in proteins. *Analytical chemistry*. 2014; 86:10979–10991. DOI: 10.1021/ac5027435 [PubMed: 25353392]
37. Kim TH, Mehrabi P, Ren Z, Sljoka A, Ing C, Bezginov A, Ye L, Pomès R, Prosser RS, Pai EF. The role of dimer asymmetry and protomer dynamics in enzyme catalysis. *Science*. 2017; 355doi: 10.1126/science.aag2355
38. Dumas BR, Brignon G, Grosclaude F, Mercier JC. Structure primaire de la caséine β bovine: séquence complète. *European Journal of Biochemistry*. 1972; 25:505–14. [PubMed: 4557764]
39. Kenyon GL. Energy metabolism. Creatine kinase shapes up. *Nature*. 1996; 381:281–282. DOI: 10.1038/381281a0 [PubMed: 8692264]
40. Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*. 1991; 253:407–414. [PubMed: 1862342]
41. Endicott JA, Noble ME, Johnson LN. The structural basis for control of eukaryotic protein kinases. *Annual review of biochemistry*. 2012; 81:587–613. DOI: 10.1146/annurev-biochem-052410-090317
42. Brandman O, Ferrell JE Jr, Li R, Meyer T. Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science*. 2005; 310:496–498. DOI: 10.1126/science.1113834 [PubMed: 16239477]
43. Hunter T. Why nature chose phosphate to modify proteins. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 2012; 367:2513–2516. DOI: 10.1098/rstb.2012.0013 [PubMed: 22889903]
44. Cianci M, Bourenkov G, Pompidor G, Karpics I, Kallio J, Bento I, Roessle M, Cipriani F, Fiedler S, Schneider TR. P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *Journal of synchrotron radiation*. 2017; 24:323–332. DOI: 10.1107/S1600577516016465 [PubMed: 28009574]
45. Kabsch W. Xds. *Acta crystallographica. Section D, Biological crystallography*. 2010; 66:125–132. DOI: 10.1107/S0907444909047337 [PubMed: 20124692]
46. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *Journal of applied crystallography*. 2007; 40:658–674. DOI: 10.1107/S0021889807021206 [PubMed: 19461840]
47. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta crystallographica. Section A, Foundations of crystallography*. 1991; 47(Pt 2):110–119. [PubMed: 2025413]
48. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*. 2004; 60:2126–2132. DOI: 10.1107/S0907444904019158 [PubMed: 15572765]
49. Langer G, Cohen SX, Lamzin VS, Perrakis A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature protocols*. 2008; 3:1171–1179. DOI: 10.1038/nprot.2008.91 [PubMed: 18600222]
50. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta crystallographica. Section D, Biological crystallography*. 1998; 54:905–921. [PubMed: 9757107]
51. Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. Iterative model building, structure refinement and density modification with the

- PHENIX AutoBuild wizard. Acta crystallographica. Section D, Biological crystallography. 2008; 64:61–69. DOI: 10.1107/S090744490705024X [PubMed: 18094468]
52. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in bioinformatics. 2017; doi: 10.1093/bib/bbx108
53. Pei J, Tang M, Grishin NV. PROMALS3D web server for accurate multiple protein sequence and structure alignments. Nucleic acids research. 2008; 36:W30–34. DOI: 10.1093/nar/gkn322 [PubMed: 18503087]
54. Maiti R, Van Domselaar GH, Zhang H, Wishart DS. SuperPose: a simple server for sophisticated structural superposition. Nucleic acids research. 2004; 32:W590–W594. [PubMed: 15215457]
55. Antelmann H, Engelmann S, Schmid R, Sorokin A, Lapidus A, Hecker M. Expression of a stress- and starvation-induced dps/pexB-homologous gene is controlled by the alternative sigma factor sigmaB in Bacillus subtilis. Journal of bacteriology. 1997; 179:7251–7256. [PubMed: 9393687]
56. Arnaud M, Chastanet A, Debarbouille M. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. Applied and environmental microbiology. 2004; 70:6887–6891. DOI: 10.1128/AEM.70.11.6887-6891.2004 [PubMed: 15528558]
57. Zwietering MH, Jongenburger I, Rombouts FM, van't Riet K. Modeling of the bacterial growth curve. Applied and environmental microbiology. 1990; 56:1875–1881. [PubMed: 16348228]
58. Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols. 2007; 2:2212.doi: 10.1038/nprot.2007.321 [PubMed: 17853878]
59. Norby JG. Coupled assay of Na⁺,K⁺-ATPase activity. Methods in enzymology. 1988; 156:116–119. [PubMed: 2835597]
60. Heuck A, Schitter-Sollner S, Suskiewicz MJ, Kurzbauer R, Kley J, Schleiffer A, Rombaut P, Herzog F, Clausen T. Structural basis for the disaggregase activity and regulation of Hsp104. eLife. 2016; 5doi: 10.7554/eLife.21516

Figure 1. Structure of the protein arginine kinase McsB

a, Simplified scheme of the protein arginine phosphorylation reaction catalysed by McsB. **b**, Domain organisation of *G. stearothermophilus* McsB. **c,d**, Ribbon diagram of apo (**c**) and AMP-PN-bound (**d**) McsB dimers (subunit 1 coloured as in **b**; subunit 2, white). Lid loop is thickened for clarity and its conformation (open or closed) is indicated above each subunit. The guanidino group binding site was predicted based on the AK structure (PDB:1M15).

Figure 2. Catalytic mechanism and mutational analysis of McsB

a, Structural comparison of the active sites of the nucleotide-bound McsB and AK. AMP-PN from the McsB structure and the arginine substrate from the AK structure are shown. **b**, Active site of McsB. Predicted positions of guanidino moiety and γ -phosphoryl are indicated. Observed polar interactions are shown as dashed lines. Mg^{2+} is omitted. **c–e**, Schematic presentations and results of assays used for mutational analysis: radioactive protein kinase assay (**c**), 'lag-time' assay (**d**), and *in vivo* activity assay (**e**), respectively. In **c**, the experiment was performed in the presence of 12.5 mM ATP. In **d** and **e**, error bars indicate standard deviation for 5 (for WT in **d**), 3 (for control and mutants in **d**), or 4 independent experiments (in **e**); significant differences from WT are marked (*, $P = 0.05$; **, $P = 0.01$; ***, $P = 0.001$; two-tailed Student's *t*-test, unpaired in **d**, paired in **e**). In **e**, the quantification shows the amount of CtsR after 60 min relative to the amount after 30 min for a given mutant. For the uncropped Phosphorimager scan from **c**, see Supplementary Figure 12a; for full results of **d**, see Supplementary Figure 4c; for the Coomassie staining used as a loading control in **e**, see Supplementary Figure 12b. **f**, Structural comparison of an apo and the AMP-PN-bound protomer of McsB, with a schematic presentation of lid closure shown below.

Figure 3. Dimerisation and putative negative cooperativity in McsB

a, Ribbon diagram of the dimeric interface of McsB (subunit 1, yellow; subunit 2, gray). **b**, Native MS of R273E and WT McsB. Full native MS spectrum for WT McsB is shown in Supplementary Fig. 9. **c**, SEC-MALS of R273E and WT McsB. The UV spectra (thin line) were scaled to equal height. The experimental MW measured with MALS is represented with a bold line. Full-width SDS-PAGE fragments are shown in Supplementary Figure 12c and SDS-PAGE analysis of SEC input samples in Supplementary Figure 12d. **d**, Summary of SEC-MALS results for the dimerisation-deficient mutants (raw data in Supplementary Fig. 5f,g). Vertical axis is in logarithmic scale. **e**, Radioactive kinase assay of dimerisation-deficient mutants in the presence of 12.5 mM ATP. Representative Phosphorimager scan (top) and quantification of 3 experiments with independently purified protein batches (bottom) are shown. Uncropped scan is shown in Supplementary Figure 12e. Error bars indicate standard deviation. **f**, Distribution of thermal motion factors in the apo and AMP-PN-bound McsB dimer. B factors are reflected in colouring and in cartoon putty radius.

Figure 4. The DD of McsB and the CTD of CtsR are pArg-binding domains

a, Structural comparison of the DD of McsB and the CTD of CtsR from *G.*

stearothermophilus (PDB:3H0D). The shared motif is shown in stick mode. Inset is a

fragment of a sequence alignment. **b** and **c**, pArg pockets of both McsB subunits (**b**) and of CtsR (**c**) coloured according to surface electrostatics. The $2F_o - F_c$ omit electron density of the pArg ligands is contoured at 1σ .

d, Polar interactions coordinating pArg in McsB. **e**, Distribution of thermal motion factors of the pArg-bound structure. B factors are reflected in colouring and in cartoon putty radius.

Figure 5. Allosteric stimulation of McsB by pArg binding

a, Radioactive kinase assay of WT McsB and pArg-binding-deficient mutant (R338A/D339A) $\pm$ pArg. Schematic representation of the reacting molecules (left) and Phosphorimager scan fragments from a representative experiment (right) are shown. **b**, Quantification of 3 repeats of the experiment shown in **a** performed with independently purified protein batches. Error bars indicate standard deviation. **c**, Radioactive kinase assay of WT McsB $\pm$ pArg-containing peptide (peptide^{pArg}). Sequence and nanoelectrospray spectrum of the peptide are shown in Supplementary Figure 7e. **d**, A representative Phosphorimager scan (inset) and quantification of 3 technical repeats of peptide^{pArg} titration experiments. Reactions were performed in the presence of 200 (**a**, **b**) or 50 (**c**, **d**) μM ATP. For uncropped Phosphorimager scans, see Supplementary Figure 12f–h.

Figure 6. Organisation of the McsB protein arginine kinase and the pArg signalling system
a, Schematic presentations of the active site structures of McsB and canonical protein kinases. Highlighted are residues promoting the nucleophilic attack of the incoming side chain on the ATP γ -phosphorus. **b**, Cartoon showing the overall structural organisation and key mechanistic features of McsB. Proposed allosteric pathways are indicated. **c**, Components of the 3-partite pArg signalling toolkit with writer, eraser, and readers. Ribbon diagrams of McsB, YwIE (PDB:4KK3), ClpCP (PDB entries 3PXI and 3TT6 assembled

together), and CtsR (PDB:3H0D) are shown. Functions of pArg residues recognised by the various reader proteins are indicated in italics.