HstK, a cyanobacterial protein with both a serine/threonine kinase domain and a histidine kinase domain: implication for the mechanism of signal transduction

Vincent PHALIP*, Jian-Hong LI†‡ and Cheng-Cai ZHANG†‡¹

*Unité d'Immunotechnologie et Microbiologie Moléculaire, Ecole Supérieure de Biotechnologie de Strasbourg, Université Louis Pasteur de Strasbourg, Boulevard Sébastien Brandt, F-67400 Illkirch, France, †Laboratoire de Chimie Bactérienne (UPR9043), C.N.R.S., 31 Chemin Joseph Aiguier, Marseille cedex 20, France, and ‡College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China

Two distinct families of protein kinases are involved in signal transduction: Ser, Thr and Tyr kinases, which are predominantly found among eukaryotes, and His kinases, as part of bacterial two-component signalling systems. Genetic studies in *Arabidopsis* and *Saccharomyces* have demonstrated that bacterial-type two-component systems may act upstream of Ser/Thr kinases in the same signalling pathway, but how this coupling is accomplished remains unclear. In the present study, we report the characterization of a protein kinase, HstK, from the N₂-fixing cyanobacterium *Anabaena* sp. PCC 7120, that possesses both a Ser/Thr kinase domain and a His kinase domain. Proteins with a structural architecture similar to that of HstK can be found in the eukaryote, *Schizosaccharomyces pombe*, and the bacterium,

Rhodococcus sp. M5. HstK was present in cells grown with NH₄⁺ or N₂ as the nitrogen source, but was absent in cells grown with NO₃⁻. The *hstK* gene was inactivated and the mutant phenotype was characterized. The catalytic domain of the Ser/Thr kinase of HstK functionally replaced that of Hog1p, a well-characterized protein kinase required for the response to high osmolarity in the *S. cerevisiae* heterologous system. The unusual multidomain structure of HstK suggests that a two-component system could be directly coupled to Ser/Thr kinases in the same signal transduction pathway.

Key words: *Anabaena*, mitogen-activated protein kinase, protein kinase, signal transduction, two-component system.

INTRODUCTION

In both prokaryotes and eukaryotes, signal transduction involves protein phosphorylation catalysed by protein kinases. There are two distinct families of protein kinases: Ser/Thr- and/or Tyr-specific protein kinases, predominantly found among eukaryotes [1,2], and His kinases, involved in bacterial two-component signalling pathways [3]. In eukaryotes, protein kinases often form cascades or networks of protein phosphorylation in order to co-ordinate cellular responses to a variety of signals [1,2]. In a typical bacterial two-component signalling system upon stimulation by a given signal, a receptor His kinase undergoes autophosphorylation on a conserved His residue, and then transfers the phosphate group to a conserved Asp residue of a response regulator which, in most cases, is a transcription factor [3]. Although ubiquitous mostly in bacteria, two-component systems have been found in several eukaryotes over the last few years [3–5]. In parallel, eukaryotic-type protein kinases and phosphatases have also been reported in several bacterial species [6]. Bacterial genome sequencing projects have also unveiled the existence of Ser/Thr and Tyr kinases and phosphatases in many bacterial strains [7].

Although the existence of Ser/Thr and/or Tyr kinases in prokaryotes has been established, the function of these enzymes in bacterial signal transduction remains unclear. One possibility is that they form a cascade of protein phosphorylation similar to that found in eukaryotes. Another possibility could be a coupling of Ser/Thr kinases to two-component systems. Considering the

coexistence of both two-component signalling pathways and eukaryotic-like signalling molecules in bacteria, such as in the cyanobacterium *Anabaena* sp. PCC 7120 [6], it is legitimate to ask whether an interaction between these two types of signalling molecules is possible. Indeed, such a coupling mechanism has already been suggested in several eukaryotes. Genetic studies in *Saccharomyces cerevisiae* and *Arabidopsis thaliana* have shown that two-component systems may act upstream of Ser/Thr protein kinases [4,5,8–13]. In *S. cerevisiae*, a two-component system functions upstream of a mitogen-activated protein kinase (MAPK) pathway involved in the regulation of the cellular response to high osmolarity [3,10,11,14]. This signal transduction pathway is mediated by a MAPK homologue, Hog1p. When the catalytic domain of Hog1p is deleted or mutated to an inactive form, cells become osmosensitive [12]. In *A. thaliana*, a His kinase, ETR1, acts as an ethylene sensor and transmits the signal to a Raf-related Ser/Thr kinase, CTR1 [8,9]. These studies reveal a new type of signal transduction mechanism. However, in all of these cases, how the two-component systems are coupled to Ser/Thr kinases remains obscure; in particular, it is not yet clear whether the interaction of the two-component system with Ser/Thr kinases is direct or indirect through another, yet unidentified, factor.

Anabaena sp. PCC 7120 is a filamentous strain of cyanobacteria [15]. It can grow on combined nitrogen sources, such as NO₃⁻ or NH₄⁺. In the absence of a combined nitrogen source, some regularly spaced cells along each filament are induced to differentiate into heterocysts specialized in N₂ fixation. In the present

Abbreviations used: GAF, cGMP-specific and stimulated phosphodiesterase, *Anabaena* adenylate cyclase and *E. coli* FhlA (where G is cGMP, A is adenylate and F is FhlA); GFP, green fluorescent protein; MAPK, mitogen-activated protein kinase; NCBI, National Center for Biotechnology Information; ORF, open reading frame; RbcL, large subunit of ribulose-1,5-bisphosphate carboxylase.

¹ To whom correspondence should be addressed (e-mail cczhang@ibsm.cnrs-mrs.fr).

The nucleotide sequence data reported have been submitted to the GenBank® under the accession number AF230361.

study, we report the isolation and characterization in this cyanobacterium of a gene encoding a protein with an N-terminal Ser/Thr kinase domain and a C-terminal His kinase domain. Such an unusual arrangement in a single molecule suggests a mechanism by which a two-component system may be coupled to a eukaryotic-type protein kinase in signal transduction.

EXPERIMENTAL

Strains and growth conditions

Anabaena sp. PCC 7120 was grown in BG11 medium as described previously [16,17], and *Escherichia coli* in Luria–Bertani broth [18]. *E. coli* strains used for conjugation were reported previously [19]. When required, 100 µg/ml spectinomycin was used for the selection of *E. coli* transformants and 5 µg/ml for the selection of conjugants in *Anabaena* sp. PCC 7120. The yeast strain *S. cerevisiae* BMA41 (MAT α ura3::lade2::1Δtrp1his3::11leu2::3,112) was used for genetic analysis. Yeast strains were grown as described previously [20] at 30 °C on SD medium [0.67% (w/v) bacto-yeast nitrogen base, 2% (w/v) glucose and 50 µg/ml Trp] or YEPD medium [1% (w/v) yeast extract, 2% (w/v) Bactopeptone and 2% (w/v) dextrose]. For phenotypic testing, 1×10^4 or 1×10^3 cells, grown on SD or YEPD, were washed with water and spotted on to the same medium supplemented with 0.9 M NaCl or 1.5 M sorbitol.

Gene cloning, sequencing and inactivation

Techniques for the screening of gene libraries and DNA sequencing have been described previously [17]. For the inactivation of *pknB*, a 2.8 kb *Hind*III subfragment (Figure 1) was cloned into the *Hind*III site of pBluescript SK[−] (Stratagene), giving rise to the plasmid, p3C5. p3C5 was partially digested by *Eco*RV, and the Ω fragment, bearing spectinomycin resistance [21], was inserted into one of the *Eco*RV sites within the coding region of *pknB* (Figure 1). The resulting plasmid was called p3C5 Ω . The insert of p3C5 Ω was isolated as a *Pst*I/*Xho*I fragment and inserted into the integrative vector pRL271 [19]. This final construct, pRL3C5 Ω , was mobilized by conjugation into *Anabaena* sp. PCC 7120 as described previously [19].

For the inactivation of *hstK*, a 4.8 kb *Hind*III fragment (Figure 1) was subcloned into the *Hind*III site of pBluescript SK[−]. The

resulting plasmid, p3C3, was partially digested by *Xmn*I, into which an *Eco*RV fragment bearing the green fluorescent protein (GFP) [22] and an *Sma*I fragment bearing Ω [21] was inserted, giving rise to the plasmid p3CGfp Ω . The GFP and Ω fragments were amplified by PCR from pBAD-GFPuv (Clontech) and pHP45 Ω [21] respectively. The insert of p3CGfp Ω was cloned as a *Pst*I/*Xho*I fragment into the integrative plasmid, pRL271, and conjugated into *Anabaena* sp. PCC 7120 as described previously [19].

Expression of recombinant proteins in *E. coli* and Western blotting

The coding sequences of PknB and those of the two kinase domains of HstK were cloned after PCR amplification using the high-fidelity Vent DNA polymerase (New England Biolabs). Each amplified fragment was digested by *Nde*I and *Bam*HI and inserted into the *Nde*I/*Bam*HI-digested pET15b (Novagen). Inserts of the resulting plasmids, pET15b(HstK-S/T), pET15b(HstK-His) and pET15b(PknB), were sequenced to ascertain the accuracy of the DNA sequence. Recombinant proteins were then overexpressed in *E. coli* BL21(DE3) [23].

Protein labelling by [³²P]P_i was carried out as described previously [24]. Crude protein extracts from *Anabaena* sp. PCC 7120 were prepared as described previously [25]. Polyclonal antibodies against GFP were purchased from Molecular Probes. Western blotting was carried out using an enhanced chemiluminescence detection kit (ECL[®]; Amersham Pharmacia Biotech) as described previously [25] except that the concentration of Tween 20 was 0.4% (v/v).

Engineering of *HOG1* replacement mutants in *S. cerevisiae*

A *HOG1* deletion mutant was constructed in which the region that encodes the catalytic domain of Hog1p was replaced by *TRP1* (phosphoribosylanthranilate isomerase prototrophic selectable marker gene). The two following PCR primers were designed; HOG-D1, 5'-ACAAAGGGAAAACAGGGAAAAC-TCAACTATCGTATATAATAATGGGCCAAGAGGGAG-GGC-3' [the initiation codon of *HOG1* (ATG) is shown in bold], and HOG-D2, 5'-CAAGTCAGTGATGATCGAAAATTGGT-GAACATGATCTTTCCCTTAAATAAATACTACTC-3'. These two primers were used to amplify *TRP1*, and the PCR product was flanked by a short region of *HOG1* (corresponding to the underlined letters in the two primers above). The PCR product was then used to transform *S. cerevisiae* BMA41 for gene replacement. For the construction of other mutants in *S. cerevisiae*, the *HOG1* promoter was PCR-amplified using the P-HOG-H and P-HOG-B primers (see below), and genomic DNA of *S. cerevisiae* was used as a template. The PCR product was tagged with *Hind*III and *Bam*HI sites and cloned into the corresponding sites of the vector pFL34 [26]. DNA fragments, corresponding to the Ser/Thr kinase catalytic domain of HstK or PknB, were cloned downstream of the promoter using the *Bam*HI/*Eco*RI sites to give pHOG-HST and pHOG-BST respectively. A region including the promoter, the catalytic domain coding region and the selection marker, *URA3*, was amplified using the AccuTaq LA DNA polymerase system (Sigma). The sense PCR primer used was Rep-Hog and the anti-sense primer was Rep-Hog-HST for pHOG-HST or Rep-Hog-BST for pHOG-BST. The fragment obtained was then used to transform *S. cerevisiae* BMA41 using the lithium acetate method [27]. The sequences of the PCR primers were: P-HOG-H, CCCAAGCTT-ACCTCAAAGCGCTTCG; P-HOG-B, GCGGATCCTATTATACGATAGTTG; Rep-Hog, AAAGCGTTCTTTATCT-

Figure 1 Organization of the *hstK-pknB* genomic region and strategy for *hstK* inactivation

Deduced ORFs (*orf1*, *hstK* and *pknB*) are shown under the restriction map, with an arrowhead indicating the direction of transcription. The black bar represents the Ser/Thr kinase catalytic domain, and the cross-hatched bar the His kinase catalytic domain. The location of GFP (*gfp*) and the Ω fragment are shown. Only restriction sites relevant to the text are indicated.

CCAAGACTTTGCCCTGTATATAATTCGGTAATCTC-
CGA; Rep-Hog-HST, GTCAGTGATGATCGAAAATTGG-
TGAACATGATCTTTCCCTGGGTACGTCACGATAA;
Rep-Hog-BST, GTCAGTGATGATCGAAAATTGGTGAA-
CATGATCTTTCCCGGGTAGTACTCCTGTTA.

RESULTS

Cloning of a novel Ser/Thr kinase gene, *pknB*, from *Anabaena* sp. PCC 7120

We have previously reported [17,25,28,29] evidence showing the existence of a family of eukaryotic-type protein kinases in *Anabaena* sp. PCC 7120, and four of these genes have already been characterized. Using a PCR-generated probe, PCR-PKN4 [17], another member of this gene family, *pknB*, was cloned by colony hybridization from a gene library of *Anabaena* sp. PCC 7120. One positive clone, with an insert of approx. 19.5 kb, was further analysed. The *pknB* gene was identified on a 2.8 kb *Hind*III fragment from the insert (Figure 1) and proposed to encode a protein, PknB, of 461 amino acid residues with a molecular mass of 51804 Da. PknB possesses an N-terminal region of 260 residues similar to the catalytic domains of known Ser/Thr kinases, with the 11 subdomains conserved to various degrees. This domain is 40% identical with the putative kinase catalytic domain of PknA [17] from the same strain (results not shown). The C-terminal region of PknB has no significant sequence similarity to any proteins in the National Center for Biotechnology Information (NCBI) protein database.

hstK, located upstream of *pknB*, could encode a protein with a multidomain structure

During the sequencing process, a partial open reading frame (ORF) was found 325 bp upstream of *pknB*. The deduced peptide showed similarity to His kinases of bacterial two-component signal transduction pathways [3]. This discovery prompted us to sequence further upstream. A large ORF, with a coding region spanning 5.3 kb and a coding capacity of 1777 amino acid residues, was identified. The deduced protein had a calculated molecular mass of 200724 Da and possessed a structure unusual among signalling molecules described to date. The amino acid sequence of residues 1–250 was similar to the catalytic domains of the eukaryotic-type Ser/Thr kinases, whereas residues 1517–1777 resembled the His kinase domains found in bacterial two-component systems (Figure 2). Thus this protein was designated as HstK (for His and Ser/Thr kinase). The putative Ser/Thr kinase domain of HstK was more related to that of STE7p [6,30] from *S. cerevisiae* (with 27% identity) than to the other known eukaryotic-type protein kinases from *Anabaena* sp. PCC 7120 (with approx. 20% identity). The putative His kinase domain was approx. 23% identical with other His kinase domains, such as TlcP or ResE from *Bacillus subtilis* [31,32]. The His residue, which is the target of autophosphorylation [3], was also conserved (Figure 2).

Two other proteins from the NCBI protein database show a structure similar to HstK

The unusual structure of HstK suggests a new signalling scheme in which two-component systems and Ser/Thr and/or Tyr kinases are intimately linked. To our knowledge, HstK was the first case of such a 'hybrid' protein when it was cloned and sequenced in our laboratory [6]. Recently, however, a search in the NCBI protein database allowed the identification of two other proteins with multidomain structures similar to that of

A

```

Atpk1  -----QIMRTLGTS---FGRVHLVRSVHNGRYIAKVLKQKQVVKMKQVEHTNDR
Fused  -----MNRVAVSSLVQGS---FGCVKATRKDDSKVVAIKVSK-RGRATKELNLRREC
HstK   -MITLPHNYQIETLQEGV---KTNVYHAISIPDSKLVKIKLT-EYPELKDIALLKHEY
SpAC   SGLGEFQFQPIRTQSNFSSLLSACLVRVAVHVTERRKVIKFSQK---TFKLENYEFLRLQL
BpdS   -----MRFDTVKILRDSF---GITTTLVVSGEDREKRVVLRRLDL-GRTWLMSRRWLEDEV

Atpk1  RMLKLVHEHFFLIRMWGTQDARNIFMVMNDYIEGGLFSLLRKSQRFPNPAKFAAEVIL
Fused  DIQARLKHPHVIEIESFESKTDLFVTEFALMD-LHRYLSYNGAMGEEPARRVTGHVLS
HstK   ELIKMLDIPGVKAGHFEKYNNOFALVEEEDGTSLHKVQIENKIGLLDFCKIG-IQITQ
SpAC   SSHPNGRYAIAPAYILLNLTETGALTYDDPGNILDWLGNNPDLKLFKFAFGVSY
BpdS   EAIRRRLSHVITAPTIVHGGPDHIDLVRPFIEGMDVREWFASHPTRSFABQLRIMATLPH

Atpk1  ALEYLHAHNIYRDLKPENILLDRNGH-----IKITDFGFAKEVQTVTWT---
Fused  ALYVYLSNRILHRDLKPNVLLDKNMH-----AKLCDPGLARNMTGLTHVLTLS-
HstK   ALGELHRKYI IHKDIKPNQNI IVNLETHQ-----VKIIDFSSISLLFQERAKLNLP
SpAC   VLLCFLEKKIVHGEIRLDTFHYDLNAPTHAKLLTGSSVSPIRPTLSLSLNMKRLYQVQN-
BpdS   ALACLHLRLGIAGGGVTPANIVLTENSE-----LVLLDAGVTRVQLAANDQNP--

Atpk1  --LCGTFDYIAPEVIT--TKFYNKSVDMWSLGVLIYENLAGY-----
Fused  --IKQTPLYMAPELLA--DEPYDHADMMWSLGCIAYESMAGQ-----
HstK   NLLBGTIAVMSPEQTGRNRSIDVTDYPSLGTFTYENLTGQ-----
SpAC   --ICHKLQFFSPEQIGNVGRPLDSRSIDYSLGILFVYLITKQ-----
BpdS   --TTSIGADRGDPSRN--NGTGFSADLFGAGVLLESVAQANHTASVLRRTQPAAGDLT

Atpk1  -----TPFYDTTPMKTYEKILQCKVVPYPYHPDWDLLSKLITADLTIRIGNL
Fused  -----PPFCASSILHLVKMKHKEDVWKPSTLTCECRSFLQGLLEKDPGLRISWT
HstK   -----LFFIVTDFMELVHCHIAKQFISITNQLIPIEVEVSIIMKLKSK---
SpAC   -----YPWGGQS--MRIVQSIHMRQFSPVLRPRRPAFPAL-----
BpdS   DLSRQFGLMGVFDQLRPLVTKLLDPVVAQRYQTABEALGLEALMATGQDGTGPKGLLPR

```

B

```

*
DctB   HEINQPVATIR-----AYADNARTFLHRGQITVTA-----E
FixL   HELNQPLSAIA-----NYSHGCTRLRLDMDDAVATRI-----RE
HstK   HEINNPVSFTGHNSHIDNYANDLLTLDIYRSIPEATPEIEEFLENIDVDFIKEDMPK
SpAC   HELRTPFSSEFY-----GMLSLSTDKLNEEQ-----YD
BpdS   HDEFGPLFAAA-----MSG---IGILQRSTDPVER-----E

DctB   NMESIAELTERVGAITDELRFARKGHFAAGP-----TAMKEVVEGALMLLSRFAGRMD
FixL   ALEEVAQSILRAGQIIKHLREFVTKGETEKAP-----EDIRKLVESAAALVGSREQGV
HstK   TSSMKIGTQRIREIVLTLRNFSLRLDEADMKP-----VDIHGIESLTLILQSRLOAKPG
SpAC   IVSTAKQSCTSLVQIIDDLLNFSLEKSGMKLEPKVDFVEENIADCIBLVPSLSKSPV
BpdS   TATEVREILRAGITVARSVAWSLPSGLDLGLG-----LACSIHYVEDFQARFPIRVD

DctB   AIRIDLFP-----DGLQALGNRIQLVQLINLLQNALEAGDS-----ENGAIQVR
FixL   RIVFEYLP-----GAEMVLVDRIQVQVILNLMRAIEAMRHV-----DRELTTIR
HstK   KPAIEIKNYAELPKVECYTQGLNQVFNILNNAIDSLKLNREHNLEIKTDSATIAIR
SpAC   QSVYDIYPN--VPALLAGDSAKLRQVITNLLGMSVKFTTEG-----HILLRCMAIDEE
BpdS   LTRIGNIP-----ALPPAVATAFRIVQEAITNVARHSGAR-----EGSVMLVG

DctB   --CKAAPGGIALTVDNPGFGIAADVREELTFPNTSK-----EDGLGLGLAIS-KEIVS
FixL   T-MPADPGEVAVVVEDTGCQIPEEAVAGQLKPPVTTK-----ASQMGIGLSIS-KRIVE
HstK   T-EVVNPDVLAISIKDNLGMSDSVRQRIIDFPFTTKP-----IGQGTGLGLSISYQIVVD
SpAC   INAEENQCKLRFIEDTQIGLKEQQLKLLNFFQVQDGTSTTRYGSGGLGLSIC-LQICK
BpdS   -----SADSLRVVVEDNAGAFDVLAGE-----RKS-----LGLVGV-QERAR

DctB   DYGGTIEVESGFGSG-TTFTVNLKKA-----
FixL   AHGGEMTVSKNEAGGATPRFTLPAYLDERIVAND
HstK   KHRGSIECISTFGQGAEFIIQIPCRQK
SpAC   IMDGIDGVQSVYEGSTTFWVHVLQNRVTSKLSQK
BpdS   LIGARLFVSESPNGGTTINVEVPIKL-----

```

Figure 2 Sequence comparison of the Ser/Thr kinase and the His kinase domains of HstK, BpdS and SpAC27E2.09

(A) Ser/Thr kinase domains. For known Ser/Thr kinases, Atpk1 (NCBI protein database accession number P42818) is from *A. thaliana* and Fused (NCBI protein database accession number L34782) is from *Drosophila melanogaster*. (B) His kinase domains. For known His kinases, DctB and FixL are both from *Rhizobium meliloti* [31,32]. SpAC, SpAC27E2.09. Bold characters represent identical residues in three out of five aligned sequences. Gaps were introduced to optimize the alignment. The conserved His residues, which are the site for autophosphorylation, are indicated by an asterisk.

HstK (Figure 2). One of these, BpdS, was described in the Gram-positive actinomycete *Rhodococcus* sp. M5 [33]. The function of *bpdS* itself is unknown. The other protein, SpAC27E2.09 (GenBank® accession number Z98978), was found during the genome sequencing project of the yeast *S. pombe*.

Figure 3 Protein kinase activity

(A) Protein labelling in *E. coli* in the presence of [32 P] P_i . *E. coli* cells bearing a plasmid overexpressing each polypeptide were grown to an D_{600} of 0.5. Isopropyl β -D-thiogalactoside ('IPTG', 0.5 mM) was added to induce protein overexpression in the presence of [32 P] P_i for 2 h at 37 °C. Total proteins were isolated and separated by SDS/PAGE, transferred on to a nylon membrane and exposed to X-ray film. Lane 1, pET15b; lane 2, pET15b(HstK-S/T); lane 3, pET15b(HstK-His); lane 4, pET15b(PknB). Molecular-mass markers (in kDa) are shown on the left. (B) Wild-type *S. cerevisiae* BMA41 and three mutant constructs, Δ HOG1 (the catalytic domain of Hog1p is deleted), Δ HOG1-BST (the Ser/Thr catalytic domain of Hog1p replaced with that of PknB) and Δ HOG1-HST (Ser/Thr kinase catalytic domain of Hog1p replaced with that of HstK), were grown on plates with SD medium either alone (SD) or containing 1.5 M sorbitol (SD + Sorbitol) or 0.9 M NaCl (SD + NaCl). For each strain, 1×10^4 (on the left) and 1×10^5 (on the right) cells were spotted on to each plate and grown for 2 days.

All three proteins, HstK, BpdS and SpAC27E2.09, possess an N-terminal Ser/Thr kinase domain with features common to Ser/Thr and/or Tyr kinases ([1,2] and Figure 2A), and a C-terminal His kinase domain similar to those in bacterial two-component His kinases ([3] and Figure 2B). The three Ser/Thr kinase domains shared 17.4–24% sequence identity, and the three His kinase domains were 18.7–29% identical. All three proteins were large in size with 1777, 2310 and 1576 residues for HstK, SpAC27E2.09 and BpdS respectively. Although overall sequence similarity in the central domains of these proteins was not found, two regions of 34 and 110 residues (corresponding to 452–565 and 841–875 of HstK respectively) showed significant sequence conservation (results not shown). In addition, HstK (residues 1319–1467) and SpAC27E2.09 (residues 1457–1602), but not BpdS, contained a GAF [cGMP-specific and stimulated phosphodiesterase, *Anabaena* adenylate cyclase and *E. coli* FliA

(where G is cGMP, A is adenylate and F is FliA)] domain found in many prokaryotic and eukaryotic signalling proteins [34].

Functional analysis of PknB and HstK in *E. coli* and *S. cerevisiae* heterologous systems

PknB as well as the Ser/Thr and His kinase domains of HstK were overexpressed in *E. coli*. The overexpression of the full-length HstK in *E. coli* was unsuccessful, presumably because of its large size. During the course of induction in *E. coli*, all three recombinant polypeptides were radiolabelled in the presence of [32 P] P_i (Figure 3A). These results suggest that PknB and the two distinct protein kinase domains of HstK have autophosphorylating protein kinase activities, a feature shared by most protein kinases [35].

Although all of the His kinases tested in bacteria were capable of autophosphorylating, Ser/Thr kinases could also be substrates of other protein kinases. To further demonstrate that HstK and PknB functioned as Ser/Thr kinases, we tested the ability of their respective Ser/Thr kinase domains to replace the catalytic domain of Hog1p, a well characterized protein kinase in the osmolarity-response pathway in *S. cerevisiae* [12,14]. Clear sequence similarity between Hog1p and HstK was identified in subdomains I–IX of the catalytic domains [1,2]. This region corresponds to amino acid residues 1–208 of HstK and residues 1–215 of Hog1p. Three yeast strains were constructed: Δ HOG1, in which the coding region of residues 1–215 of Hog1p was replaced by *TRP1*; Δ HOG1-HST, in which the coding region of residues 1–215 of Hog1p was changed to the corresponding 1–208 residues of HstK; and Δ HOG1-BST, in which the same region of *HOG1* was replaced with the corresponding region of PknB. The growth of these mutants and that of wild-type *S. cerevisiae* BMA41 were tested using SD and YEPD growth media, in both liquid and solid culture. As shown in Figure 3(B), the Δ HOG1 mutant was unable to grow under conditions of high osmolarity in the presence of either 0.9 M NaCl or 1.5 M sorbitol; findings consistent with a published observation [12]. The ability to adapt to high osmolarity pressure was restored in both Δ HOG1-HST and Δ HOG1-BST mutants as they both grew normally in the presence of either NaCl or sorbitol when compared with the wild-type.

Expression of *hstK* and characterization of *hstK* insertion mutants

RNA–DNA hybridization failed to reveal any transcripts of the *hstK* gene, suggesting that it is expressed at a low level. Thus a mutant construct was made in which GFP [22] was translationally fused to the N-terminal region of *hstK* (Figure 1). In addition, an Ω element (a spectinomycin-resistance gene cassette) was also inserted adjacent to GFP, providing a selection marker for fusion and inactivation mutants (Figure 1). Double recombinants were obtained, and the complete interruption of the *hstK* coding region by the GFP– Ω insertion was confirmed by DNA–DNA hybridization (results not shown).

Two independent conjugants, named as *hstK*-G1 and *hstK*-G2, were further analysed, and had similar phenotypes. Heterocyst development and pattern were not affected. In the presence of NO_3^- as the nitrogen source, both mutants grew normally when compared with the wild-type. However, their growth capacities were impaired when NH_4^+ or N_2 was used as a nitrogen source. Compared with the wild-type, the mutant strains displayed a growth rate 10% slower in the presence of NH_4^+ and 40% slower under N_2 -fixing conditions.

The autofluorescence of HstK–GFP in the two clones was undetectable under fluorescence microscopy irrespective of the nitrogen source in the medium, a finding consistent with a low-

Figure 4 Detection of HstK in *Anabaena* sp. PCC 7120 by GFP-epitope tagging

The wild-type strain (lanes 1–3) and one conjugant, hstK-G1 (lanes 4–6), were cultured in the presence of NO_3^- (lanes 1 and 4), under N_2 -fixing conditions (lanes 2 and 5) or in the presence of NH_4^+ (lanes 3 and 6). Total proteins were isolated, separated by SDS/PAGE, blotted on to a nylon membrane and revealed with a polyclonal antibody against GFP (A) or RbcL (B).

level expression of HstK. Expression of the HstK–GFP fusion protein was thus assayed with polyclonal antibodies against GFP. Protein extracts were prepared from cells grown under different nitrogen regimes. A signal, which was absent in wild-type cells, was detected with protein extracts from mutant cells grown on NH_4^+ or under N_2 -fixing conditions, but not with extracts from cells grown on NO_3^- (Figure 4A). The size of the protein corresponding to this signal was approx. 70 kDa, close to the theoretical value for the HstK–GFP fusion protein (66.7 kDa; sum of the N-terminal 360 residues of HstK and the 240 residues of GFP). The intensity of this signal in cells grown on NH_4^+ was approx. 5 times stronger than that in cells grown on N_2 . The amount of proteins loaded in each lane was determined using polyclonal antibodies against the large subunit of ribulose-1,5-bisphosphate carboxylase (RbcL) (Figure 4B).

pknB insertion mutants were also isolated. They grew normally under different nitrogen regimes when compared with the wild-type strain (results not shown).

DISCUSSION

Genetic and molecular studies in three eukaryotes, *A. thaliana*, *S. cerevisiae* and *Dictyostelium*, have suggested that two-component signalling systems may be coupled to Ser/Thr kinases in the same signal transduction pathway [8–13]. It was recently shown [36] that, using a yeast two-hybrid system, the His kinase, ETR1, and the Ser/Thr kinase, CTR1, from *A. thaliana* physically interacted with each other. The structure of HstK also indicates that a two-component signalling domain can be directly linked to a Ser/Thr kinase domain to form a single protein. The structural definition of the different domains in such a protein may shed new light on the mechanism of coupling between two-component pathways and Ser/Thr kinases. In addition, the finding that proteins with a structure similar to that of HstK exist in two other organisms, the eukaryote *S. pombe* and the Gram-positive bacterium, *Rhodococcus* sp. M5, suggests that such a coupling mechanism is not a rare event. In this context, it is worth noting that *bpdS* in *Rhodococcus* sp. M5 is adjacent to the gene *bpdT* encoding a response regulator, the second member of two-component pathways [33]. It would be interesting to know whether BpdT interacts with BpdS.

Both HstK of *Anabaena* PCC 7120 and SpAC27E2.09 of *S. pombe* contain a GAF domain just in front of the His kinase catalytic domain. The GAF domain, found in a variety of signalling proteins, including many sensory His kinases, can function as a cGMP-binding motif and may serve as a signal-

reception domain [34,37]. This signalling input may influence the activity of the His kinase, thereby triggering an as yet undefined signal transduction cascade.

PknB and the two distinct kinase catalytic domains of HstK were radiolabelled when expressed in *E. coli*. Since His kinases in bacterial signalling systems are all capable of autophosphorylation [3] and no such enzyme in bacteria has been shown to act as a substrate of other protein kinases, we thus believe that the His kinase domain of HstK can indeed function as a His kinase. In the case of PknB and the Ser/Thr kinase domain of HstK, radiolabelling in *E. coli* may be the result of either autophosphorylation or phosphorylation by an unknown protein kinase in the host. To test these hypotheses, and to avoid the problem of inclusion body formation of recombinant proteins in *E. coli*, we chose to use a genetic approach based on the well-characterized Hog signalling system in *S. cerevisiae* [14]. Hog1p belongs to the MAPK family and it is required for the response to high osmolarity by regulating the cellular level of glycerol [3,10,11]. Among the proteins in the NCBI protein database, c-Jun N-terminal kinase 1 ('JNK1'), a member of the MAPK subfamily, is most similar to the Ser/Thr kinase catalytic domain of HstK (results not shown). Deletion or point mutations in *HOG1* almost completely abolishes the ability of yeast cells to adapt to high-osmolarity pressure [12]. The osmosensitive phenotype of our ΔHOG1 mutant, in which the Ser/Thr kinase domain is deleted, is consistent with these observations. This phenotype was suppressed in the ΔHOG1 -HST strain in which the putative Ser/Thr kinase domain of HstK replaced that of Hog1p. These results not only indicate that the Ser/Thr catalytic domain of HstK is functional, but that critical residues required for the function of Hog1p are also conserved in the corresponding region of HstK. One of the possible explanations for this could be that the chimaeric protein kinase is constitutively active, enabling the cells to resist high-osmolarity pressure. Whatever the explanation, the results obtained in the *E. coli* and *S. cerevisiae* heterologous systems indicate that HstK may indeed function as a protein kinase. In addition, the expression of *hstK* is regulated in *Anabaena* sp. PCC 7120 and the inactivation mutant of *hstK* is impaired in its growth ability under certain conditions. Taken together, these results strongly suggest that HstK is a functional protein kinase. Such enzymes should allow us to understand how the two-component system and Ser/Thr kinases cross-talk in cellular signalling networks.

No proteins with a structure similar to that of HstK were identified in the entire genome of the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 [38], although this strain harbours several Ser/Thr kinases [39] and a large family of His kinases [40]. This observation suggests that HstK might be involved in a cellular activity, such as heterocyst metabolism, which is restricted to *Anabaena* sp. PCC 7120 but absent in *Synechocystis* sp. PCC 6803. The phenotypes of the mutants carrying an inactivated *hstK* support this view. As for the other eukaryotic-type protein kinases investigated in bacteria to date, the HstK protein is not essential. It appears the protein participates in nitrogen metabolism when NH_4^+ or N_2 , but not NO_3^- , is used as a nitrogen source. These results are consistent with the expression pattern of *hstK*. Indeed, the HstK–GFP fusion protein was not detected in cells grown on NO_3^- , but was present in cells grown either on NH_4^+ or under N_2 -fixing conditions.

We thank I. Kuhn and G. Brandner for technical help, Dr W.-H. Shen for providing antibodies against ribulose-1,5-bisphosphate carboxylase/oxygenase, the Fondation pour la Recherche Médicale for a fellowship to J.-H. L., the French Ministry of Research for financial support through the Programme de Recherche Fondamentale en Microbiologie et Maladies Infectieuses et Parasitaires grant, and the CNRS for an Action Thématique Incitative sur Programme et Equipe grant.

REFERENCES

- 1 Hanks, S. K. and Hunter, T. (1995) The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596
- 2 Hunter, T. (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* (Cambridge, Mass.) **80**, 225–236
- 3 Stock, A. M., Robinson, V. L. and Goudreau, P. N. (2000) Two-component signal transduction. *Annu. Rev. Biochem.* **69**, 183–215
- 4 Appleby, J. L., Parkinson, J. S. and Bourret, R. B. (1996) Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. *Cell* (Cambridge, Mass.) **86**, 845–848
- 5 Wurgler-Murphy, S. M. and Saito, H. (1997) Two component signal transducers and MAPK cascades. *Trends Biochem. Sci.* **22**, 172–176
- 6 Zhang, C.-C. (1996) Bacterial signalling involving eukaryotic-type protein kinases. *Mol. Microbiol.* **20**, 9–15
- 7 Shi, L., Potts, M. and Kennelly, P. J. (1998) The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms: a family portrait. *FEMS Microbiol. Rev.* **22**, 229–252
- 8 Chang, C., Kwok, S. F., Bleecker, A. B. and Meyerowitz, E. M. (1993) *Arabidopsis* ethylene-response gene *ETR1*: similarity of product to two-component regulators. *Science* (Washington, D.C.) **262**, 539–544
- 9 Maeda, T., Wurgler-Murphy, S. M. and Saito, H. (1994) A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature* (London) **369**, 242–245
- 10 Kieber, J. J., Rothenberg, M., Roman, G., Feldmann, K. A. and Ecker, J. R. (1993) *CTR1*, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the Raf family of protein kinases. *Cell* (Cambridge, Mass.) **72**, 427–441
- 11 Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Thai, T. C. and Saito, H. (1996) Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 “two-component” osmosensor. *Cell* (Cambridge, Mass.) **86**, 865–875
- 12 Schüller, C., Brewster, J. L., Alexandre, M. R., Gustin, M. C. and Ruis, H. (1994) The HOG pathway controls osmotic regulation of transcription via the stress response element (STRE) of the *Saccharomyces cerevisiae* *CTT1* gene. *EMBO J.* **13**, 4382–4389
- 13 Thomson, P. A., Traynor, D., Cavet, G., Chang, W.-T., Harwood, A. J. and Kay, R. R. (1998) An intersection of the cAMP/PKA and two-component signal transduction systems in *Dictyostellium*. *EMBO J.* **17**, 2838–2845
- 14 Gustin, M. C., Albertyn, J., Alexander, M. and Davenport, K. (1998) MAP kinase pathways in the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **62**, 1264–1300
- 15 Wolk, C. P. (1996) Heterocyst formation. *Annu. Rev. Genet.* **30**, 59–78
- 16 Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. and Stanier, R. Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.* **111**, 1–61
- 17 Zhang, C.-C. (1993) A gene encoding a protein related to eukaryotic-type protein kinases from the filamentous heterocystous cyanobacterium *Anabaena* PCC 7120. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11840–11844
- 18 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 19 Elhai, J. and Wolk, C. P. (1988) Conjugal transfer of DNA to cyanobacteria. *Methods Enzymol.* **167**, 747–754
- 20 Phalip, V., Kuhn, I., Lemoine, Y. and Jeltsch, J.-M. (1999) Characterization of the biotin biosynthesis pathway in *Saccharomyces cerevisiae* and evidence for a cluster containing *BIO5*, a novel gene involved in vitamin uptake. *Gene* **232**, 43–51
- 21 Prentki, P. and Krisch, H. M. (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29**, 303–313
- 22 Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression. *Science* (Washington, D.C.) **263**, 802–805
- 23 Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89
- 24 Munoz-Dorado, J., Inouye, M. and Inouye, S. (1991) A gene encoding protein Ser/Thr kinase is required for normal development of *M. xanthus*, a gram-negative bacterium. *Cell* (Cambridge, Mass.) **67**, 995–1006
- 25 Zhang, C.-C. and Libs, L. (1998) Cloning and characterisation of the *pknD* gene encoding an eukaryotic type protein kinase in the cyanobacterium *Anabaena* sp. PCC 7120. *Mol. Gen. Genet.* **258**, 26–33
- 26 Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G. Y., Labouesse, M., Minvielle-Sebastia, L. and Lacroute, F. (1991) A family of low and high copy replicative, integrative and single-stranded *S. cerevisiae*/*E. coli* shuttle vectors. *Yeast* **7**, 609–615
- 27 Schiestl, R. H., Manivasakam, P., Woods, R. A. and Gietz, R. D. (1993) Introducing DNA into yeast by transformation. *Methods* **5**, 79–85
- 28 Zhang, C.-C., Friry, A. and Peng, L. (1998) Molecular and genetic analysis of two closely linked genes that encode, respectively, a protein phosphatase 1/2A/2B homolog and a protein kinase in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J. Bacteriol.* **180**, 2616–2622
- 29 Gonzalez, L., Phalip, V. and Zhang, C.-C. (2001) Characterization of PknC, a Ser/Thr kinase with broad substrate specificity from the cyanobacterium *Anabaena* sp. strain PCC 7120. *Eur. J. Biochem.* **268**, 1869–1875
- 30 Cairns, B. R., Ramer, S. W. and Kornberg, K. D. (1992) Order of action of components in the yeast pheromone response pathway revealed with a dominant allele of the STE11 kinase and the multiple phosphorylation of the STE7 kinase. *Genes Dev.* **6**, 1305–1318
- 31 Jiang, J., Gu, B. H., Albright, L. M. and Nixon, B. T. (1989) Conservation between coding and regulatory elements of *Rhizobium meliloti* and *Rhizobium leguminosarum* *dct* genes. *J. Bacteriol.* **171**, 5244–5253
- 32 David, M., Daveran, M. L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P. and Kahn, D. (1988) Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. *Cell* (Cambridge, Mass.) **54**, 671–683
- 33 Labbé, D., Garnon, J. and Lau, P. C. K. (1997) Characterization of the genes encoding a receptor-like His kinase and a cognate response regulator from a biphenyl/polychlorobiphenyl-degrading bacterium, *Rhodococcus* sp. strain M5. *J. Bacteriol.* **179**, 2772–2776
- 34 Aravind, L. and Ponting, C. P. (1997) The GAF domain, an evolutionary link between diverse phototransducing proteins. *Trends Biochem. Sci.* **22**, 458–459
- 35 Smith, J. A., Francis, S. H. and Corbin, J. D. (1993) Autophosphorylation: a salient feature of protein kinases. *Mol. Cell. Biochem.* **127–128**, 51–70
- 36 Clark, K. L., Larsen, P. B., Wang, X. and Chang, C. (1998) Association of the *Arabidopsis* CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5401–5406
- 37 Ho, Y.-S. J., Burden, L. M. and Hurley, J. H. (2000) Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *EMBO J.* **19**, 5288–5299
- 38 Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S. et al. (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**, 109–136
- 39 Zhang, C.-C., Gonzalez, L. and Phalip, V. (1998) Survey, analysis and genetic organisation of genes encoding eukaryotic-like signalling proteins on a cyanobacterial genome. *Nucleic Acids Res.* **26**, 3619–3625
- 40 Mizuno, T., Kaneko, T. and Tabata, S. (1996) Compilation of all genes encoding bacterial two-component signal transducers in the genome of the cyanobacterium, *Synechocystis* sp. strain PCC 6803. *DNA Res.* **3**, 407–414

Received 3 August 2001/17 September 2001; accepted 11 October 2001