# Amino acid identity at one position within the α1 helix of both the histidine kinase and the response regulator of the WalRK and PhoPR two-component systems plays a crucial role in the specificity of phosphotransfer

Inga Jende,[1] Kottayil I. Varughese[2] and Kevin M. Devine[1]

[1]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland

[2]University of Arkansas for Medical Sciences, Department of Physiology and Biophysics, 4301 West Markham, Little Rock, AR 72205, USA

Correspondence

Kevin M. Devine

kdevine@tcd.ie

Kottayil I. Varughese

KIVarughese@uams.edu

Two-component systems usually function as cognate pairs, thereby ensuring an appropriate response to the detected signal. The ability to exclusively phosphorylate a partner protein, often in the presence of many competing homologous substrates, demonstrates a high level of specificity that must derive from the interacting surfaces of the two-component system. Here, we identify positions within the histidine kinases and response regulators of the WalRK and PhoPR two-component systems of *Bacillus subtilis* that make a major contribution to the specificity of phosphotransfer. Changing the identity of the amino acid at position 11 within the α1 helix of WalK and at position 17 within the α1 helix of PhoP altered discrimination and allowed phosphotransfer to occur with the non-cognate partner. Changing amino acids at additional positions of the WalK kinase increased phosphotransfer, while changes at additional positions in PhoP only had an effect in the presence of the change at position 17. The importance of amino acid identity at these two positions is supported by the fact that the amino acid combinations of Ile and Ser in WalRK, and Leu and Gly in PhoPR, are very highly conserved among orthologues, while modelling indicates that these amino acid pairs are juxtaposed in the WalRK and PhoPR complexes.

## INTRODUCTION

Two-component systems (TCSs) are widely used by bacteria to adapt to the prevailing environmental conditions. The prototypical TCS is composed of a sensor histidine kinase and its cognate response regulator (RR), which is often a transcription factor. Three activities have been ascribed to histidine kinases: (i) autophosphorylation of an invariant His residue, (ii) phosphorylation of the RR and (iii) phosphorylated RR (RR~P) phosphatase activity, which is found in many sensor kinases (for reviews see Hoch & Silhavy, 1995; Stock *et al.*, 2000; Gao & Stock, 2009). Thus, TCSs function by the sensor kinase autophosphorylating in response to a specific signal(s) and then transferring the phosphoryl group to a conserved Asp on its cognate RR; the latter reaction involves interaction

between conserved protein domains. The RR~P has an altered activity, usually increased DNA-binding affinity, to effect changes in gene expression. In the absence of the signal, RR~P phosphatase activity of the sensor kinase often predominates, ensuring that the regulon is not activated inappropriately by non-cognate kinases or by acetyl phosphate.

There are 34 TCSs in *Bacillus subtilis*, and each responds to a specific environmental or nutritional signal(s) (Kunst *et al.*, 1997; Fabret *et al.*, 1999; Hoch, 2000). The majority of TCSs function as cognate pairs, ensuring that the cellular response is appropriate to the detected signal. The observed specificity among TCSs implies a high level of discrimination among members of the histidine kinase and RR families: each sensor kinase and RR can recognize their partners from among 34 candidates present in the cell. Since the interacting transmitter and receiver domains of histidine kinases and RRs, respectively, show a high degree of conservation, discerning how this specificity is achieved presents an interesting challenge (Hoch & Varughese, 2001). Insight into the nature of histidine kinase and RR

interacting surfaces, and of how phosphotransfer is effected, has been established by crystallographic and mutational analysis of many TCSs (Madhusudan et al., 1996; Tzeng & Hoch, 1997; Varughese et al., 1998; Jiang et al., 1999; Zapf et al., 2000; Skerker et al., 2008; Casino et al., 2009; Yamada et al., 2009; Schug et al., 2009; for reviews see Varughese, 2002; Mukhopadhyay & Varughese, 2005; Gao & Stock, 2009). For illustrative purposes, discussion here is confined to the Spo0B : Spo0F complex of B. subtilis that serves as a prototype for such interactions. In this system, Spo0F is a truncated RR composed of a receiver domain only, while Spo0B is a phosphotransferase that contains the transmitter domain typical of histidine kinases (Hoch & Varughese, 2001). The interacting surface of the RR (Spo0F) comprises the α1 helix and the five β–α loops (the connecting regions between the β strands and α helices) within which three amino acid subsets are recognized: (i) a globally conserved set of amino acids that form a hydrophobic patch; (ii) a set of amino acids that is conserved within each RR family and is proposed to mediate discrimination at this level; and (iii) a set of amino acids that varies among members of a family and is proposed to confer specificity on its RRs. The five β–α loops surround the active-site pocket (Asp10, Asp11, Asp54, Thr82 and Lys104), where the Asp54 residue is phosphorylated. The interacting surface of the transmitter domain of histidine kinases (Spo0B) comprises a four-helix bundle formed from two protomers. Nearly half of the 24 aa that comprise this surface are located on the α1 helix, which is C-terminal to the active site His (Zapf et al., 2000). Bioinformatic and covariance analysis approaches have also identified amino acids that might contribute to the specificity of interaction between cognate TCSs (Zapf et al., 2000; Skerker et al., 2008; Szurmant et al., 2008; Weigt et al., 2008). Reassuringly, many of the covariant positions have mapped to the interacting surfaces determined from the Spo0F : Spo0B co-crystal structure. The importance of these positions in histidine kinases has been experimentally verified by Skerker et al. (2008), who succeeded in rewiring the specificity of EnvZ. By changing EnvZ domains, subdomains and clusters of specific amino acids to those found in other histidine kinases, Skerker et al. (2008) showed that the hybrid EnvZ* could phosphorylate the partner RR of the histidine kinase from which the swapped domains were derived. In some cases, this was achieved by changing as few as three specific amino acids within the α1 helix (e.g. EnvZ*–RstA), while other cases showed more complexity (e.g. EnvZ*–CpxA, PhoR, AtoS), and required additional changes within the loop region located between the α1 and α2 helices to change specificity (Skerker et al., 2008). Together, these studies have provided insight into the features of TCSs that contribute to the specificity of cognate pair interaction.

The WalRK and PhoPR TCSs control cell wall metabolism and the response to phosphate limitation, respectively, in B. subtilis (Hulett, 1996; Sun et al., 1996; Bisicchia et al., 2007, 2010). Although closely related phylogenetically,

significant differences in amino acid composition occur within the interacting surfaces of the TCSs, and these might be expected to contribute to the specificity of cognate partner recognition. For example, amino acids Arg, Thr, Arg and Tyr are located at positions +3, +8, +10 and +12, respectively, within the α1 helix of WalK (numbered relative to the active site His, 0), whereas PhoR has Lys, Ser, Lys and Phe, respectively, at these positions. Thus, while the nature of the amino acid at each position is similar, the sizes of the side chains in WalK are larger than those in PhoR, suggesting that steric hindrance may play a role in specificity in these TCSs (Howell et al., 2006). Unusually, some cross-talk between WalRK and PhoPR has been shown in vivo and in vitro: PhoR can phosphorylate both PhoP and the non-cognate WalR (but at a lower level than it phosphorylates PhoP), while WalK can phosphorylate its cognate WalR RR only (Howell et al., 2003, 2006). The fact that hybrid RRs WalR′–′PhoP and PhoP′–′WalR are functional suggests that these systems are sufficiently robust to test the specificity of cognate protein interactions by changing amino acids between the histidine kinases and RRs. We focused our analysis on five positions within the α1 helices of WalK and PhoR histidine kinases, and on four positions in the WalR and PhoP RRs, at which significant amino acid differences have been seen. The amino acids of WalK and PhoP at these positions were changed compared with those found at the same positions of PhoR and WalR, respectively, and the extent to which phosphorylation of the non-cognate and cognate proteins was gained and lost was then determined. We show here that the specificity of phosphorylation in the WalRK and PhoPR TCSs can be altered by changing the identity of amino acids at some of these positions in WalK and PhoP, with the identity of the amino acid at two positions (one within the histidine kinase and the second within the RR) being especially important. However, the rewired proteins phosphorylated the non-cognate partner at a low rate, indicating that additional features to those examined in this study are required for specific protein recognition and phosphotransfer in these systems.

## METHODS

**Bacterial strains and growth conditions.** The bacterial strains and plasmids used in this study are listed in Supplementary Table S1, available with the online version of this paper. All strains were grown aerobically in Luria–Bertani (LB) medium (Miller, 1972) at 37 °C. Strain TG1 was used for general cloning in Escherichia coli. Antibiotics were added, as appropriate, at the following concentrations: ampicillin, 100 μg ml$^{-1}$; chloramphenicol, 3 μg ml$^{-1}$; neomycin, 7 μg ml$^{-1}$; spectinomycin, 100 μg ml$^{-1}$; erythromycin, 3 or 150 μg ml$^{-1}$. Xylose was added to the medium at concentrations specified in the text, and X-Gal was added to LB agar plates at 100 μg ml$^{-1}$.

**Plasmid and strain construction.** Oligonucleotides used in strain construction are listed in Supplementary Table S2. To construct strains in which mutated WalK kinases were placed under the control of the inducible P$_{xyl}$ promoter at the thrC locus, a 1140 bp DNA fragment, containing an optimized ribosome-binding site and the 5′

end of *walK*, was amplified using primers P_IJ100F and P_IJ100R, and cloned into a filled-in *Eco*RI site of pXT (Derre *et al.*, 2000) to generate plasmid pIJ100. DNA fragments encoding the remainder of the *walK* gene with the desired mutations were generated using eight primer pairs: primers P_IJ101F–P_IJ108F were each paired with P_IJ101R. The fragments were then individually cloned into *Eco*RI/*Sma*I-digested pBluescript KS− (Stratagene), excised by digestion with *Eco*RI/*Sph*I, and individually cloned into the *Eco*RI/*Sph*I-digested pIJ100, resulting in plasmids pIJ101–pIJ108. To generate plasmids pIJ112 and pIJ114, two 310 bp fragments were generated using primer pairs P_IJ112F/P_IJ112R and P_IJ114F/P_IJ112R. The fragments were subcloned into *Sma*I-digested pBluescript KS−, excised by digestion with *Eco*RI/*Bgl*II, and cloned into *Eco*RI/*Bgl*II-digested pIJ108, thereby generating plasmids pIJ112 and pIJ114. A gene encoding a hybrid WalK′–′PhoR kinase (containing the WalK sensing domain, transmembrane helices and PAS domain fused to the phosphotransfer and autokinase domains of PhoR) was generated by overlapping PCR. Two DNA fragments, generated using primer pairs P_IJ110F/P_IJ111R and P_IJ111F/P_IJ110R, were mixed in equal amounts, annealed, and used as a template for a third PCR using primers P_IJ110F and P_IJ110R. This fragment was digested with *Bam*HI/*Hin*dIII, and cloned into similarly digested pXT, to generate plasmid pIJ115. To construct strains in which mutated PhoP RRs were placed under the control of the inducible P*xyl* promoter at the *thrC* locus, a DNA fragment containing an optimized ribosome-binding site, and a full-length wild-type *phoP*, with added restriction sites *Nde*I 5′ and *Xho*I 3′, was generated using primers P_IJ140F/P_IJ140R, and cloned into the *Bam*HI and *Hin*dIII site of pBluescript KS− to generate plasmid pKS140. Using primers P_IJ141F/P_IJ141R, pKS140 was used as a template to generate a DNA fragment encoding *phoP* with amino acid 107 mutated, and it was cloned into the *Bam*HI and *Hin*dIII site of pBluescript KS− to generate plasmid pKS141. Plasmids pKS140 and pKS141 were used as templates for amplification of *phoP* genes with the desired combination of mutations, using P_IJ141F–P_IJ155F and P_IJ141R–P_IJ155R. The PCR-generated plasmids pKS141–pKS155 were digested with *Dpn*I, and sequenced. Plasmids pKS140–pKS155 were digested with *Bam*HI and *Hin*dIII to excise the mutated *phoP* genes which were cloned into the respective sites of pXT, resulting in plasmids pIJ140–pIJ155. To construct a transcriptional *phoA*–*lacZ* fusion, the *phoA* promoter region was amplified using primers P_IJ99F and P_IJ99R, digested with *Eco*RI and *Bam*HI, and cloned into a similarly digested plasmid pDG268, to generate plasmid pIJ99. All plasmid constructs were verified by sequencing. Strains IJW99 (*trpC2 amyE::*P*phoA*–*lacZ* Cm^R), IJR99 (*trpC2* Δ*phoR amyE::*P*phoA*–*lacZ* Cm^R) and IJPR99 (*trpC2* Δ*phoPR amyE::*P*phoA*–*lacZ* Cm^R) were constructed by transforming strains 168 (*trpC2*), AH057 (*trpC2* Δ*phoR*) and AH024 (*trpC2* Δ*phoPR*), respectively, with linearized pIJ99. The series of strains IJW101–IJW115 and IJW140–IJW155 were generated by transforming strain IJW99 with each of the pIJ101–pIJ115 and pIJ140–pIJ155 plasmids linearized with *Sma*I. Transforming strain IJR99 with plasmids pIJ101–pIJ115 and pIJ140–pIJ155 linearized with *Sma*I generated the strain series IJR101–IJR115 and IJR140–IJR155. The IJPR101–IJPR107 series of strains was generated by transforming the corresponding IJW strain with chromosomal DNA from strain AH024, and selecting for erythromycin resistance.

**Protein production and purification.** DNA fragments encoding the intracellular domain of each mutated WalK kinase were generated by PCR using primers P_WalK5′ and P_WalK3′, and the appropriate pIJ plasmids as templates digested with *Nde*I and *Bam*HI. The fragments were inserted into similarly digested pET21b (Novagen), and transformed into *E. coli* strain BL21(DE3), to generate the pET plasmid series shown in Supplementary Table S1. The members of the pET plasmid series encoding the mutated PhoP* proteins were generated by digestion of pKS140–pKS155 with *Nde*I and *Xho*I, and the resulting DNA fragments were cloned into the respective restriction sites of pET21b (Novagen), and transformed into *E. coli* strain BL21(DE3), as shown in Supplementary Table S1. All plasmid constructs were verified by sequencing. The *E. coli* strains were grown in LB medium containing 75 μg ampicillin ml^−1, at 37 °C, with shaking at 200 r.p.m. Proteins were purified essentially according to the manufacturer's instructions: briefly, protein production was induced by addition of 1 mM IPTG to a culture at OD$_{600}$ 0.6–0.8, and growth was continued at 28 °C. Cells were harvested by centrifugation at 10 000 *g* 3 h post-induction, resuspended in ice-cold lysis buffer [50 mM NaH$_2$PO$_4$, 300 mM NaCl, 20 mM imidazole, 0.1 %, v/v, Tween) and lysed by addition of lysozyme (1 mg ml^−1, containing 1 %, v/v, protease inhibitor) with sonication. Cell debris and unbroken cells were removed by centrifugation at 22 000 *g* for 30 min at 4 °C, and the supernatant was bound to Ni$^{2+}$-NTA His.Bind resins (Novagen) for 2 h, and then washed four times in wash buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 40 mM imidazole) containing 0.1 % v/v Tween, and then four times in wash buffer without Tween. Proteins were eluted in elution buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl and 250 mM imidazole containing 1 %, v/v, protease inhibitor). Eluates were dialysed overnight against 50 % storage buffer (20 mM Tris/HCl, pH 8.0, 300 mM NaCl, 50 %, v/v, glycerol), and stored at −20 °C.

**Protein phosphorylation assays.** Phosphorylation reactions were carried out at room temperature in a final volume of 30 μl in phosphorylation buffer (100 mM Tris/HCl, pH 8.0, 200 mM KCl, 4 mM MgCl$_2$, 0.5 mM DTT, 0.1 mM EDTA, 3.5 %, v/v, glycerol). For a qualitative assessment of the ability of WalK* to phosphorylate WalR and PhoP, reactions contained 14 μM of each RR, and 9 μM of each kinase. The reaction was started by addition of 3 μl ATP mixture [1.0 μM ATP plus 5 μCi (185 kBq) [γ-$^{32}$P]ATP]. This ratio of radioactive to cold ATP increases the phosphorylation signal. Aliquots (5 μl) were extracted at the indicated times and added to an equal volume of 2× SDS loading buffer, and heated to 37 °C to stop the reaction. The samples were analysed by SDS-PAGE on 10 % acrylamide gels (Laemmli, 1970), and the transfer reactions were detected by autoradiography.

To estimate initial rates of phosphorylation, reactions were carried out as described above, with the following modifications: (i) the final concentration of cold ATP was increased to 200 μM to be non-limiting in the reaction; (ii) the ′WalK and ′PhoR kinases were added to a final concentration of 2.5 μM; and (iii) for ′WalK/WalR, ′WalK*/WalR and ′WalK/PhoP* reactions, RRs were added to a final concentration of 25 μM, while for ′PhoR/PhoP and ′PhoR/PhoP*, RRs were added to a final concentration of 2.5 μM. Aliquots (5 μl) were extracted at the indicated times and added to an equal volume of 2× SDS loading buffer, and heated to 37 °C to stop the reaction. Samples were separated on a 10 % acrylamide gel, which was exposed to a phosphoscreen. Phosphotransfer reactions were detected and quantified using a Fujifilm phosphoimager with Multi Gauge Version 2.0.

**Determination of *β*-galactosidase activity.** *β*-Galactosidase activity was determined by using the method described by Msadek *et al.* (1990). Values presented are the means of three determinations.

# RESULTS

## Significant amino acid differences within the interacting surfaces of the WalRK and PhoPR TCSs

Because the WalRK and PhoPR TCSs are closely related phylogenetically, we reasoned that amino acid differences

at positions located within their interacting surfaces would be important for the specificity of cognate protein interaction. The α1 helix of histidine kinases contains the active site His residue, and plays a major role in interactions with the RR. In RRs, the α1 helix plus the loops connecting helices and β strands make up the interaction surface. A comparison of the amino acid composition of the α1 helices of the WalK and PhoR histidine kinases, and of the anchor (positions 11, 14, 17, 55, 82, 83, 104 and 105), catalytic (positions 9, 10, 53, 81 and 103) and putative specificity (positions 13, 20, 84, 106 and 107) regions of the WalR and PhoP RRs, is shown in Table 1 (Hoch & Varughese, 2001; Mukhopadhyay & Varughese, 2005). There are five positions within the α1 helices of WalK and PhoR histidine kinases at which significant differences occur (Table 1, marked in bold type). Of the five positions in RRs termed specificity residues, only three (positions 13, 20 and 107) differ between WalR and PhoP (Table 1, marked in bold type). There is a further interesting difference among the proposed anchor amino acids (a term denoting an involvement in initial protein binding) at position 17 (Ile in WalR, Leu in PhoP) of the RRs. The amino acids at these nine positions, five in WalK and four in PhoP, were changed to those at the same positions of PhoR and WalR, respectively. The amino acid or combination of amino acids changed in each version of the mutated WalK* and PhoP* proteins are shown in Table 1, with the changed amino acids in bold type.

## System for detection of PhoP phosphorylation *in vivo*

An *in vivo* system to detect activation of PhoP, by monitoring production of β-galactosidase from a P*phoA*–*lacZ* transcriptional fusion, was established (Fig. 1). This allowed us to detect phosphorylation of wild-type PhoP by mutated versions of WalK (*walK** in Fig. 1a), and to detect phosphorylation of mutated PhoP (*phoP**) by wild-type WalK (Fig. 1b). The genes encoding the mutated WalK* and PhoP* proteins were located in single copy at the *thr* locus, and were expressed from a xylose-inducible promoter. The activity of each mutated WalK* protein was tested in strains IJR101–IJR115, which have the endogenous *phoR* gene deleted (Fig. 1a). To investigate activation of mutated PhoP* proteins by wild-type WalK, the mutated alleles of *phoP* were inserted into the chromosome of a strain with *phoPR* deleted, to generate the series of strains IJPR140–IJPR155 (Fig. 1b). The P*phoA*–*lacZ* transcriptional fusion was chosen as a reporter of PhoP phosphorylation because expression of the *phoA* promoter is totally dependent on the activated phosphorylated PhoP (PhoP~P) RR that is generated under phosphate-limiting conditions only. The medium used in this study (LB) is phosphate replete, and thus the P*phoA*–*lacZ* fusion is not expressed (Hulett, 1996). We verified that the P*phoA*–*lacZ* fusion was not expressed during growth in LB in strains lacking a xylose-inducible mutated WalK*

**Table 1.** Amino acid composition of the interacting surfaces of WalRK and PhoPR, and the changes made in each mutated protein

The amino acid changes made in each protein are given in bold type. Numbers are shown vertically, and indicate the amino acid position within the α1 helix of WalK numbered according to the active site histidine (0) and their actual position within the PhoP protein.

| Protein | Amino acid |
|---|---|
| **WalK and PhoR histidine kinases** | |
| | 0 1 2 3 4 5 6 7 8 9 1 1 1 1 1 1 1 1 |
| | _ _ _ _ _ _ _ _ _ _ 0 1 2 3 4 5 6 7 |
| WalK | A N S H E L **R** T P L T **T M R S Y** L E A L A |
| PhoR | A N S H E L **K** T P I T **S I K G F** T E T L L |
| | |
| WalK108 (WT) | A N S H E L R T P L T T M R S Y L E A L A |
| WalK101 | A N S **A** E L R T P L T T M R S Y L E A L A |
| WalK102 | A N S H E L **K** T P L T T M R S Y L E A L A |
| WalK103 | A N S H E L R T P L T **S** M R S Y L E A L A |
| WalK104 | A N S H E L R T P L T T M **K** S Y L E A L A |
| WalK105 | A N S H E L R T P L T T M R **G** Y L E A L A |
| WalK106 | A N S H E L R T P L T T M R S **F** L E A L A |
| WalK112 | A N S H E L R T P L T T M R **G F** L E A L A |
| WalK114 | A N S H E L **K** T P L T **S** M R **G F** L E A L A |
| WalK107 | A N S H E L **K** T P L T **S** M **K G F** L E A L A |
| WalK'–'PhoR | A N S H E L K T P I T S I K G F T E T L L |
| **WalR and PhoP RRs** | |
| | 9 1 1 1 1 1 2 5 5 8 8 8 8 1 1 1 1 |
| | 0 1 3 4 7 0 3 5 1 2 3 4 0 0 0 0 |
| | _ _ _ _ _ _ _ _ _ _ _ _ 3 4 5 6 7 |
| WalR | D D E **P** I **I F** D M T A K D K P F S **T** |
| PhoP | D D I **S** I **L Y** D M T A K D K P F S **P** |
| | |
| PhoP140 | D D I S I L Y D M T A K D K P F S P |
| PhoP141 | D D I S I L Y D M T A K D K P F S **T** |
| PhoP142 | D D I **P** I L Y D M T A K D K P F S P |
| PhoP143 | D D I S I **I** Y D M T A K D K P F S P |
| PhoP144 | D D I S I L **F** D M T A K D K P F S P |
| PhoP145 | D D I **P** I **I** Y D M T A K D K P F S P |
| PhoP146 | D D I **P** I **I F** D M T A K D K P F S P |
| PhoP147 | D D I S I **I F** D M T A K D K P F S P |
| PhoP148 | D D I **P** I **I F** D M T A K D K P F S P |
| PhoP149 | D D I **P** I L Y D M T A K D K P F S **T** |
| PhoP150 | D D I S I **I** Y D M T A K D K P F S **T** |
| PhoP151 | D D I S I L **F** D M T A K D K P F S **T** |
| PhoP152 | D D I **P** I **I** Y D M T A K D K P F S **T** |
| PhoP153 | D D I **P** I L **F** D M T A K D K P F S **T** |
| PhoP154 | D D I S I **I F** D M T A K D K P F S **T** |
| PhoP155 | D D I **P** I **I F** D M T A K D K P F S **T** |

or PhoP* protein (data not shown). Thus, any β-galactosidase production in our test strains must have been generated by the introduced mutated WalK* or PhoP* proteins. The absence of detectable β-galactosidase production in strain IJR99 (P*phoA*–*lacZ* Δ*phoR*) was especially important for this study: it demonstrated that PhoP was not activated in the absence of its cognate PhoR kinase under the growth conditions used. Because the

phosphatase activity of sensor kinases often predominates in the absence of an activating signal (LB medium being phosphate replete), deletion of a histidine kinase (here PhoR) can sometimes result in inappropriate activation of its cognate RR (here PhoP), either by cross-talk with non-cognate histidine kinases or by high-energy phosphate intermediates, such as acetyl phosphate (McCleary et al., 1993; McCleary & Stock, 1994). We verified that such inappropriate PhoP activation did not occur in the IJR99 (ΔphoR) strain under our culture conditions. Therefore, detection of β-galactosidase production was a reflection of the generation of PhoP~P in vivo by the introduced WalK* or PhoP* proteins. It is also important to note that, because WalRK is essential, all strains necessarily have a single copy of wild-type walRK at the normal locus (Fig. 1).

## Activation of PhoP by mutated WalK* proteins in vivo

Five positions at which notable differences in amino acid identity occurred within the α1 helices of WalK and PhoR were chosen for investigation (Table 1): there were conservative changes at positions 3 (Arg→Lys), 8 (Thr→Ser) and 10 (Arg→Lys), and two non-conservative changes at positions 11 (Ser→Gly) and 12 (Tyr→Phe). To determine the contribution of these amino acids to the specificity of phosphorylation observed among the PhoP/PhoR and WalR/WalK cognate pairs, we replaced the amino acids at these positions of WalK with the amino acids at the same positions of PhoR, singly and in combination, and analysed whether the mutated WalK* proteins could activate PhoP in vivo, as detected by expression of the $P_{phoA}$–lacZ reporter fusion. The ability of each WalK* kinase to phosphorylate PhoP in the IJR (ΔphoR) series of strains was determined by monitoring $P_{phoA}$–lacZ expression qualitatively by streaking the strains onto LB agar plates containing X-Gal, and quantitatively by measurement of β-galactosidase accumulation [values given are 2 h after the cessation of exponential growth ($T_2$)] in broth cultures (Fig. 2). As controls, we generated

strains that produced the wild-type WalK (strain IJR108) and WalK101 (strain IJR101), which has the active site His changed to Ala, rendering it inactive. A background level of $P_{phoA}$–lacZ production was observed in each strain, and this showed that PhoP was not activated by wild-type WalK, nor was it phosphorylated by other sensor kinases or by high-energy intermediates, such as acetyl phosphate, even in the absence of the cognate PhoR (Fig. 2). Of the five strains producing WalK* with a single amino acid change in the α1 helix, background levels of β-galactosidase production were found in four of the strains (containing amino acid changes Arg3Lys, Thr8Ser, Arg10Lys and Tyr12Phe) when grown on LB agar plates or in liquid medium (Fig. 2). However strain IJR105 (ΔphoR), carrying the Ser11Gly change, was exceptional: this strain was always light blue when grown on agar plates and showed significant accumulation of β-galactosidase activity (215 units) during growth in liquid culture (Fig. 2). Importantly, when WalK105 (Ser11Gly) was expressed in strain IJPR105 (ΔphoPR), then β-galactosidase activity was reduced to background levels, showing that activity in strain IJR105 was dependent on PhoP activation (data not shown). A high level of activity was also observed in the strains producing WalK* with 2 (IJR112), 4 (IJR114) and 5 (IJR107) aa changes within the α1 helix. Activity was also observed in strain IJR115 that produced a WalK′–′PhoR hybrid kinase – in this strain, the interacting surface was contributed entirely by the cognate PhoR kinase. Three observations were noteworthy: (i) while strain IJR106 (Tyr12Phe) was white on agar plates, producing background levels of β-galactosidase, and strain IJR105 (Ser11Gly) was light blue on plates, with significant accumulation of β-galactosidase (215 units), changing both amino acids [strain IJR112 (Ser11Gly, Tyr12Phe)] resulted in increased β-galactosidase activity (567 units), indicating a synergistic effect of these two amino acids on PhoP phosphorylation; (ii) activation of PhoP was further increased by making 4 (strain IJR114) and 5 (strain IJR107) aa changes in WalK, with the accumulation of 796 and 1124 units β-galactosidase activity, respectively; and (iii)
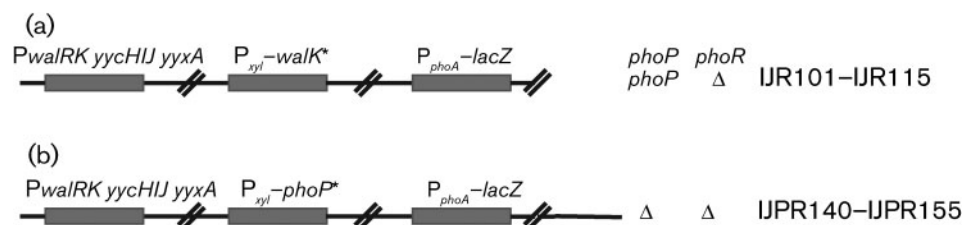


**Fig. 1.** Schematic representation of the in vivo system developed to identify amino acids important for the specificity of phosphotransfer in the WalRK and PhoPR TCSs. Genes or operons are represented by filled rectangular boxes. The locations of these operons at separate chromosomal loci are indicated by double slanted lines. Different walK alleles encoding the documented amino acid changes are indicated by walK*. (a) The series of strains IJR101–IJR115 carrying walK* alleles under control of the $P_{xyl}$ promoter are indicated. Δ, phoR gene deleted. (b) The IJPR series of strains carrying the mutated phoP* genes under the control of the $P_{xyl}$ promoter, with the endogenous phoP and phoR genes deleted (Δ).

unexpectedly, activation of PhoP in strain IJR115, which produced the WalK′–′PhoR hybrid kinase with a perfect PhoP interacting surface, was lower than that observed in IJR107 (WalK107 with 5 aa changes). This latter result might have been due to strain within the molecule because of its hybrid nature or due to formation of heterodimers within the cell. These results show that the identity of the amino acid at position 11 (Ser in WalK, and Gly in PhoR) is critically important for the specificity of phosphorylation – changing the amino acid at this position alone allowed WalK to activate PhoP, albeit at a low level. Changing the amino acid at position 11 in combination with changes at other positions (that alone had little effect) resulted in an increased ability of WalK* to phosphorylate PhoP, indicating synergistic effects of the amino acids at these positions.

## Phosphorylation of WalR and PhoP by mutated WalK* kinases *in vitro*

Above, we showed that changing the amino acids at five positions within the WalK α1 helix to those found at the same positions of the PhoR α1 helix conferred on WalK* the ability to phosphorylate the non-cognate PhoP RR *in vivo*. However, due to the essential nature of WalRK, it was not possible to determine *in vivo* how these amino acid replacements in WalK affected phosphorylation of its cognate partner WalR. Therefore, to establish the effect of the amino acid changes on recognition of WalR by WalK*, and to confirm that the mutated WalK* kinases could indeed phosphorylate PhoP, as shown *in vivo*, we performed *in vitro* phosphorylation reactions, selecting those histidine kinases that showed PhoP activation *in vivo* (i.e. ′WalK105, ′WalK112, ′WalK114 and ′WalK107) (Fig. 3). Initial experiments were performed with a low level of cold ATP to maximize the detection of radioactive phosphoryl groups in RR proteins (Fig. 3a). The truncated wild-type ′WalK protein rapidly phosphorylated WalR, as expected, with high levels of phosphorylated WalR (WalR~P) observed within 1 min, but phosphorylation of PhoP was not detected after 30 min of reaction; these

results were consistent with previous *in vivo* and *in vitro* studies (Howell *et al.*, 2003, 2006). Similarly ′PhoR kinase phosphorylated PhoP with high efficiency (high levels of PhoP~P were observed within 1 min), but some phosphorylation of WalR was also observed after 5 min of reaction; again these results were consistent with a previous study (Howell *et al.*, 2006). ′WalK105 kinase (Ser11Gly) phosphorylated PhoP to a significant level (PhoP~P was first detected at a low level, at 2 min of reaction), and this was accompanied by a concomitant reduction in WalR phosphorylation, supporting the results obtained *in vivo* (Fig. 2). ′WalK112 kinase (Ser11Gly, Tyr12Phe) had a slightly increased ability to phosphorylate PhoP when compared with ′WalK105, but, surprisingly, its ability to phosphorylate the cognate WalR was not as impaired as that observed for ′WalK105 (Fig. 3a). The ability of ′WalK114 (4 aa changes) to phosphorylate PhoP was somewhat increased when compared with the ′WalK112 protein (2 aa changes), but its ability to phosphorylate the cognate WalR was reduced, and the level was similar to that observed with ′WalK105. The highest level of PhoP phosphorylation was observed with ′WalK107 that contained 5 aa changes, although, interestingly, ′WalK107 retained a significant ability to phosphorylate its cognate WalR RR under these conditions. To assess the efficiency of phosphotransfer from ′WalK107 (with all 5 aa substitutions) to WalR and PhoP, reactions were performed with high non-limiting levels of cold ATP, which allowed an estimation of the initial rates of phosphorylation (Fig. 3b). A high level of RR~P was observed within 10 s of starting the reaction when WalR and PhoP were incubated with their cognate ′WalK and ′PhoR kinases, respectively. Changing amino acids at five positions (′WalK107) led to a dramatic loss of ability to phosphorylate WalR, while the rate of PhoP phosphorylation was also low, but clearly detectable. These data confirmed the importance of the amino acid at position 11 of the WalK and PhoR kinases, and supported the results observed *in vivo* (see strain IJR105 in Fig. 2) – changing the amino acid at this position alone conferred on WalK the ability to phosphorylate PhoP, albeit at a low rate. Changing up to 5 aa gradually
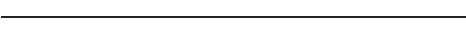
| Growth on LB agar containing X-Gal | Strain | β-Gal activity | | |
|---|---|---|---|---|
| | | IJW | IJR | IPR |
| | IJR108 | ~7 | ~6 | ~5 |
| | IJR101 | ~5 | ~5 | ~5 |
| | IJR102 | ~5 | ~5 | ~5 |
| | IJR103 | ~5 | ~5 | ~5 |
| | IJR104 | ~5 | ~5 | ~5 |
| | IJR105 | ~5 | 215 | ~5 |
| | IJR106 | ~4 | 8 | ~5 |
| | IJR112 | ~7 | 567 | ~5 |
| | IJR114 | 45 | 796 | ~5 |
| | IJR107 | ~5 | 1124 | ~5 |
| | IJR115 | 21 | 451 | ~3 |

**Fig. 2.** Detection of PhoP activation *in vivo* in strains grown on LB agar and in liquid broth. Exponential phase cultures were slotted onto a square agar plate containing X-Gal by using the edge of a sterile microscope slide, and the plate was incubated overnight. The amino acid change(s) in the xylose-inducible WalK* copy are given in Table 1 and Supplementary Table S1. The levels of β-galactosidase accumulation (specific activity units; mean of three replicates) at T₂ of selected strains in the IJW, IJR and IJPR genetic backgrounds are shown.
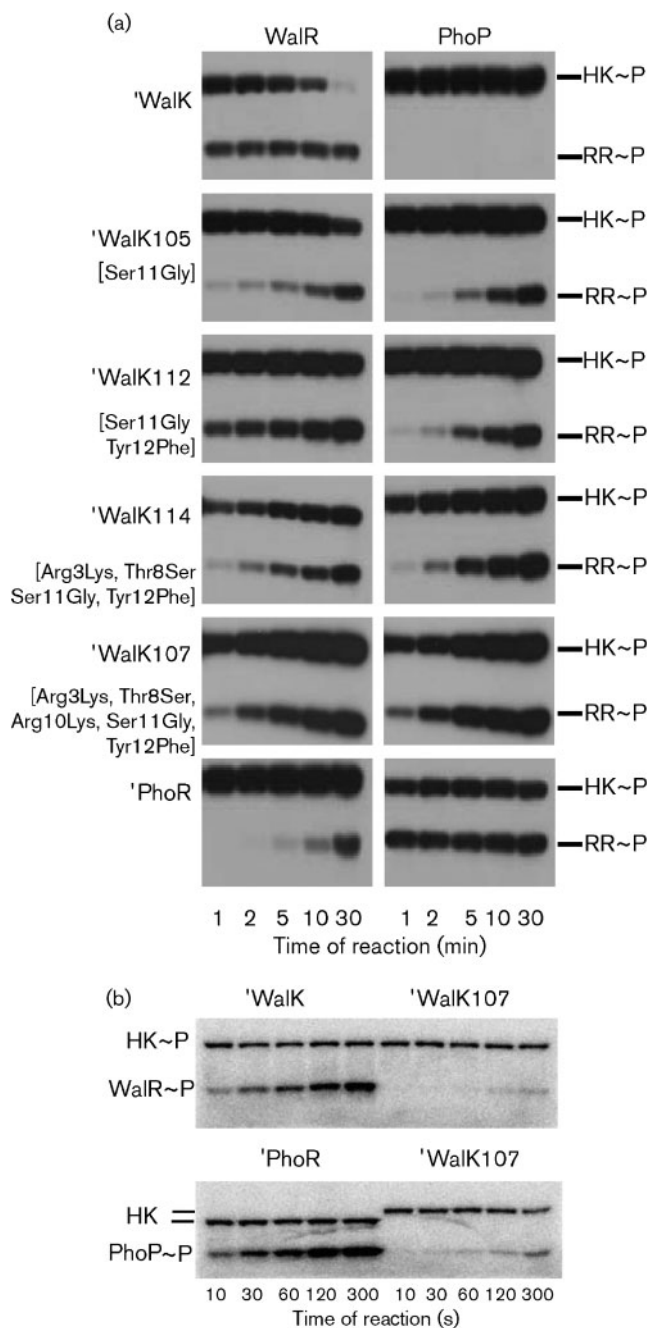
**Fig. 3.** Phosphorylation of the WalR and PhoP RRs by ′WalK, mutated ′WalK* and ′PhoR histidine kinases *in vitro*. (a) Estimation of the level of WalR and PhoP phosphorylation by wild-type and mutated kinases. HK~P, autophosphorylated histidine kinase; RR~P, phosphorylated RRs. All autoradiographs were exposed for an equal length of time. (b) Estimation of the initial rates of accumulation of phosphorylated WalR and PhoP mediated by the cognate kinases WalK and PhoR, respectively, and by mutated ′WalK107 kinase. HK and HK~P, histidine kinase and autophosphorylated histidine kinase, respectively. The phosphoimages were exposed for an equal length of time.

increased the extent to which PhoP phosphorylation by ′WalK* was achieved, although the rate was still low, suggesting that additional determinants are involved in this process. Also, the loss of the ability of ′WalK* to phosphorylate its cognate partner WalR was more erratic through successive amino acid changes.

## Activation of mutated PhoP* proteins by the wild-type WalK protein *in vivo*

A comparison of amino acid identity at the positions of WalR and PhoP that contribute to the surfaces that interact with their respective kinases is shown in Table 1. Our investigation focused on three positions within the α1 helix of these RRs: positions 13 (Pro and Ser), 17 (Ile and Leu) and 20 (Phe and Tyr), and a fourth position that resides within the β5–α5 loop region at position 107 (Thr→Pro). To assess the contribution of these amino acids to the specificity of phosphorylation, we replaced the amino acids at these positions of PhoP with those found at the same positions of WalR, both singly and in combination. The ability of wild-type WalK to phosphorylate these mutated PhoP* proteins was established *in vivo* in strains with *phoPR* deleted, using the system developed here (Fig. 1b). The results are shown in Fig. 4. There was no expression of the $P_{phoA}$–*lacZ* fusion in any strain in the absence of xylose, but expression was observed upon addition of 0.005 % xylose, confirming that accumulation of β-galactosidase was dependent on production of the introduced mutated PhoP* proteins (data not shown). The level of $P_{phoA}$–*lacZ* expression achieved by each PhoP* protein was determined qualitatively on LB agar plates containing X-Gal, and quantitatively in LB broth, as described in Methods (Fig. 4). Strains IJPR140 (expressing wild-type PhoP) and IJPR155 (expressing PhoP* with all four amino acid changes) were streaked onto each plate for comparison (Fig. 4). Strains IJPR141, IJPR142, IJPR143 and IJPR144, each of which carries a single amino acid change in the introduced PhoP* protein, showed that single amino acid changes at positions 13 (Ser13Pro), 20 (Tyr20Phe) and 107 (Pro107Thr) result in background levels of $P_{phoA}$–*lacZ* expression (Fig. 4a). However, at position 17 of PhoP, changing Leu→Ile (the amino acid at position 17 of WalR) resulted in an approximately 12-fold increase (relative to the level observed in IJPR140 carrying wild-type PhoP) in $P_{phoA}$–*lacZ* expression. The accumulation of β-galactosidase in strains with two amino acid changes in PhoP is shown in Fig. 4b. There was increased expression of $P_{phoA}$–*lacZ* (blue colonies) in three strains only (IJPR145, IJPR147 and IJPR150), and each contains a PhoP* protein with the Leu17Ile amino acid change. Expression of $P_{phoA}$–*lacZ* increased approximately 21-fold in strain IJPR145 (Ser13Pro, Leu17Ile), and approximately 8-fold in strain IJPR150 (Pro107Thr Leu17Ile). However, the increase in strain IJPR147 (Tyr20Phe Leu17Ile) was more modest. Interestingly, the increased expression in strain IJPR145 (Ser13Pro, Leu17Ile) was approximately 2-fold higher than that observed in the strain IJPR155 carrying all four amino
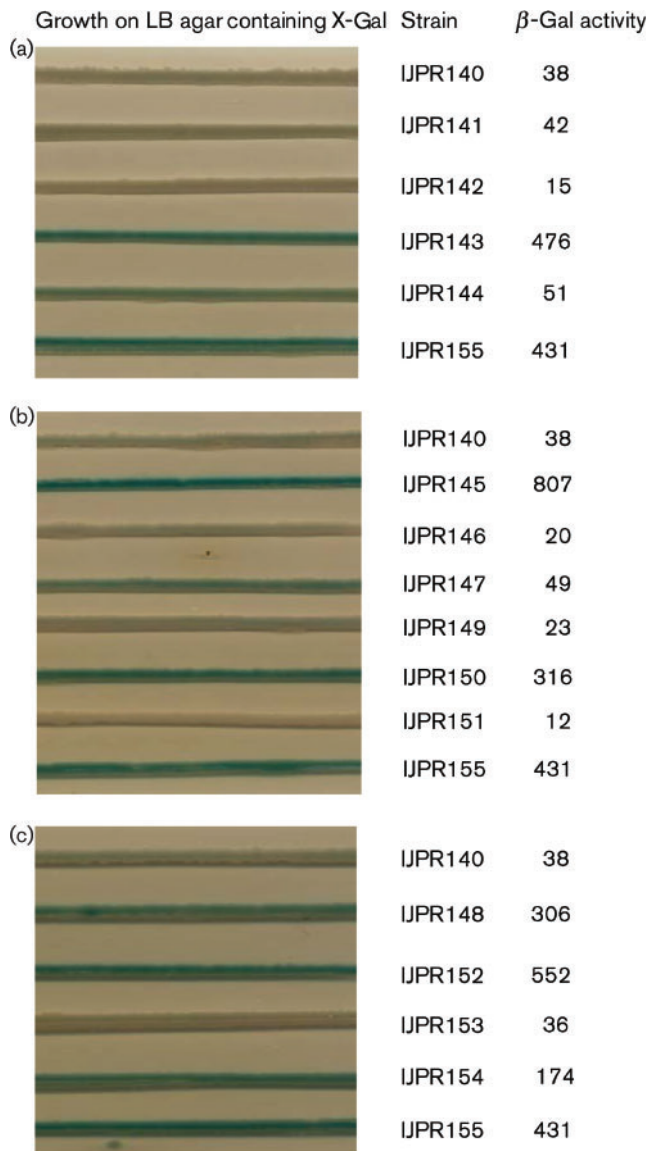
*Microbiology* **156**

Growth on LB agar containing X-Gal | Strain | β-Gal activity

(a)

| Strain | β-Gal activity |
|--------|----------------|
| IJPR140 | 38 |
| IJPR141 | 42 |
| IJPR142 | 15 |
| IJPR143 | 476 |
| IJPR144 | 51 |
| IJPR155 | 431 |

(b)

| Strain | β-Gal activity |
|--------|----------------|
| IJPR140 | 38 |
| IJPR145 | 807 |
| IJPR146 | 20 |
| IJPR147 | 49 |
| IJPR149 | 23 |
| IJPR150 | 316 |
| IJPR151 | 12 |
| IJPR155 | 431 |

(c)

| Strain | β-Gal activity |
|--------|----------------|
| IJPR140 | 38 |
| IJPR148 | 306 |
| IJPR152 | 552 |
| IJPR153 | 36 |
| IJPR154 | 174 |
| IJPR155 | 431 |

**Fig. 4.** Detection of PhoP* activation *in vivo* in strains grown on LB agar and in liquid broth. Exponential phase cultures were slotted onto a square agar plate containing X-Gal by using the edge of a sterile microscope slide, and the plate was incubated overnight. The amino acid change(s) in the xylose-inducible PhoP* copy are given in Table 1 and Supplementary Table S1. The levels of β-galactosidase accumulation (specific activity units; mean of three replicates) at $T_2$ are shown.

acid changes (Fig. 4). Importantly, the three strains with two amino acid changes, but not the Leu17Ile change (IJPR146, IJPR149 and IJPR151), showed only background expression levels. The four strains that harboured a PhoP* protein with three amino acid changes also showed an interesting expression profile (Fig. 4c): the three strains carrying the Leu17Ile mutation showed 8-fold (IJPR148), 14-fold (IJPR152) and 4.5-fold (IJPR154) increases in P$_{phoA}$–*lacZ* expression. Significantly, strain IJPR153 car-

rying PhoP* with three amino acid changes, but not the Leu17Ile change, showed only background levels of expression. Expression of P$_{phoA}$–*lacZ* in strain IJPR155, which harboured PhoP* with all four amino acid changes, increased approximately 11.5-fold. This was a lower level of expression than that observed in strains IJPR145 and IJPR152, which had changes at only two and three positions, respectively.

These data showed that the identity of the amino acid at position 17 of the WalR and PhoP RRs plays a crucial role in discrimination by the cognate kinases: changing this amino acid alone from Leu (found in PhoP) to Ile (found in WalR) allows the PhoP (Leu17Ile) protein to be phosphorylated *in vivo*. The identity of the amino acid at the other three positions affected the level of PhoP phosphorylation, but only in the presence of the Leu17Ile change; for example, the Leu17Ile change in combination with Ser13Pro yielded 807 units β-galactosidase, in combination with P107T it yielded 316 units β-galactosidase, while with Tyr20Phe Pro107Tyr, it yielded 174 units β-galactosidase. However, changing the identity of amino acids at positions 13, 20 or 107, in any combination, did not allow PhoP* to be phosphorylated to detectable levels by WalK.

## Phosphorylation of mutated PhoP* proteins by wild-type ′WalK and ′PhoR kinases *in vitro*

A caveat to the conclusions presented above is that changing the amino acids at positions 13, 17, 20 and 107 of PhoP may relax specificity, and allow phosphorylation by cellular histidine kinases other than WalK. This possibility cannot be excluded *in vivo* because of WalRK essentiality. Therefore, to confirm that PhoP* proteins with these amino acid changes can be phosphorylated by WalK, and to assess the efficiency of phosphotransfer, the PhoP143 (Leu17Ile), PhoP150 (Leu17Ile, Pro107Thr), PhoP152 (Ser13Pro, Leu17Ile, Pro107Thr) and PhoP155 (Ser13Pro, Leu17Ile, Tyr20Phe, Pro107Thr) proteins were purified and tested for their ability to be phosphorylated by the cognate ′PhoR kinase and the non-cognate ′WalK kinase (Fig. 5). All four PhoP* proteins displayed reduced or loss of ability to be phosphorylated by the cognate ′PhoR kinase. However, each PhoP* protein was phosphorylated by the wild-type ′WalK kinase, albeit to different extents. Importantly, changing the amino acid at position 17 alone (PhoP143) resulted in a gain of ability to be phosphorylated, albeit at a low rate, by ′WalK, and this was consistent with results obtained *in vivo*. PhoP150 and PhoP155 were also phosphorylated by ′WalK, but again at a low rate, and PhoP~P levels at 20 min were comparable with those achieved with the cognate WalR regulator between 30 and 60 s (Fig. 5). Phosphorylation of PhoP152 by ′WalK was the highest level of phosphorylation achieved, and it was consistent with the results obtained *in vivo* for this protein. The level of PhoP~P phosphorylation after 2 min reaction was approximately

the same as that observed with the cognate WalR after 30 s reaction (note the different times for reactions with cognate and non-cognate protein pairs). These data showed the crucial role played by the amino acid at position 17 of WalR and PhoP in discriminating between 'WalK and 'PhoR. The amino acids at the other three positions contributed to specificity, but only when the amino acid at position 17 was changed, and this emphasized the dominant role played by the identity of the amino acid at position 17. However, although changing the amino acids at three positions of PhoP* (PhoP152) resulted in a significant gain of ability to be phosphorylated by 'WalK, the rate was lower that that observed with phosphorylation of the cognate WalR, and it indicated that additional features remain to be identified. These results also excluded the possibility that PhoP* activation observed *in vivo* (Fig. 4) was due solely to phosphorylation by cellular kinases due to a general relaxed specificity, although they did not exclude the possibility that cellular

kinases may play some more minor role *in vivo*. However, the *in vitro* studies showed that phosphorylation of PhoP* *in vivo* was primarily, if not exclusively, from WalK.

## DISCUSSION

In this study, we identified nine positions within the interacting surfaces of the closely related WalRK and PhoPR TCSs that differed in amino acid composition, and tested their contributions, singly and in combination, to the specificity of phosphotransfer. A central finding is that the identity of the amino acid at two of these positions, one within the α1 helix of the WalK and PhoR histidine kinases, and the second within the α1 helix of the WalR and PhoP RRs, plays a critical role in phosphotransfer discrimination between these TCSs. When single amino acid changes were made at each of the nine positions, only the Ser11Gly change at the C-terminal region of the WalK α1 helix (WalK105), and the Leu17Ile change within the PhoP α1 helix (PhoP143), resulted in a significant change in discrimination. Specifically, at position 11 of the WalK α1 helix, changing Ser to Gly (the amino acid found at the same position of PhoR) allowed WalK* (WalK105) to phosphorylate PhoP. Likewise, at position 17 of the PhoP α1 helix, changing Leu→Ile (the amino acid found at the same position in WalR) allowed PhoP (PhoP143) to be phosphorylated by WalK. Both amino acid changes were also accompanied by a loss of ability to phosphorylate (in the case of WalK*) or be phosphorylated by (in the case of PhoP*) their cognate partners, and this emphasized the critical role of amino acid identity at these two positions in phosphotransfer specificity. However, the rate of phosphotransfer achieved by the WalK105 (Ser11Gly) and PhoP143 (Leu17Ile) proteins was low. When amino acids were changed at up to five positions in WalK (WalK107), a successive increase in the ability to phosphorylate PhoP was seen, and this showed their contribution to phosphotransfer specificity. However, the rate of PhoP phosphorylation achieved by WalK107 (with all five amino acid changes) remained relatively low when compared with the rate of PhoP phosphorylation by the cognate PhoR kinase, and this indicated that additional determinants must be required for this process.

The results obtained upon changing amino acid identity at positions within the interacting surfaces of the WalR and PhoP RRs were especially interesting. Significant differences were observed at five positions among those proposed to constitute the RR surface that interacts with their cognate kinases. Four of these positions were examined in this study (7, 13, 20 and 107, Table 1). Three of the positions (13, 20 and 107) have been proposed to contribute to specificity among this family of RRs. The other two positions (11 and 17) have been proposed to contribute to formation of a hydrophobic patch that anchors the kinase and RRs (Hoch & Varughese, 2001; Varughese, 2002). Our study shows that it is the amino acid at the anchor position (position 17 of the WalR/PhoP
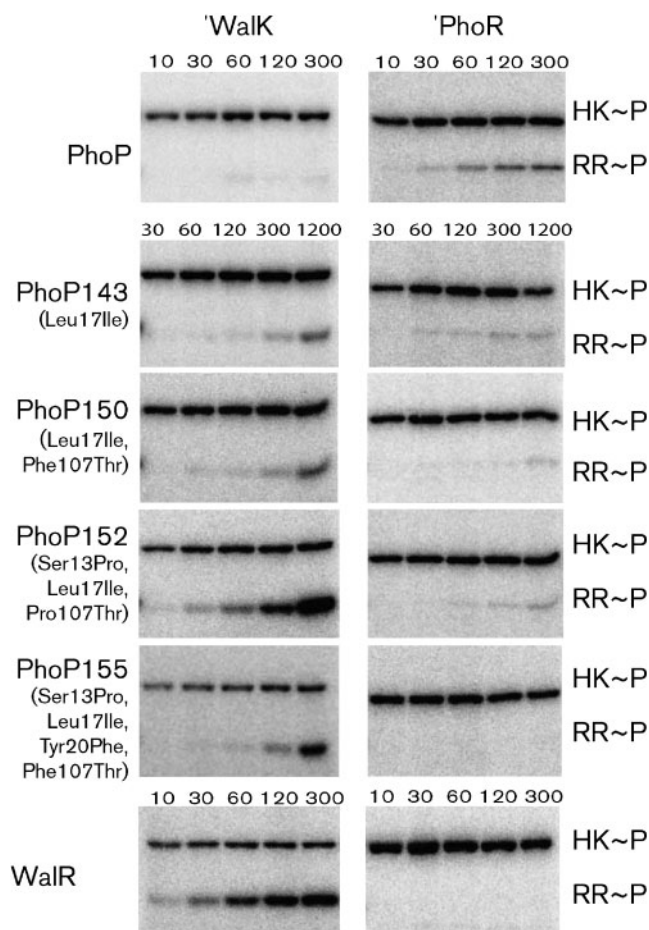


**Fig. 5.** Time course of phosphorylation of PhoP, PhoP* mutant proteins and WalR by 'WalK and 'PhoR. The positions of autophosphorylated 'WalK and 'PhoR are indicated by HK~P and the position of PhoP and PhoR mutants are indicated by RR~P. Reaction time is indicated (s).

α1 helix, corresponding to position 18 in Spo0F) that plays a central role in discrimination of phosphotransfer in the WalRK and PhoPR TCSs. This is consistent with its proposed anchoring role, a point of first contact between the cognate proteins. All combinations of amino acid changes that resulted in changed specificity, i.e. that allowed PhoP* to be phosphorylated by WalK, contained this Leu→Ile amino acid change. However, the specificity positions (13, 20 and 107) did contribute to discrimination, but only when the correct amino acid was positioned at position 17. Specifically, changing amino acid identity at positions 13, 20 and 107 of PhoP led to a loss of ability to be phosphorylated by PhoR, but there was no discernible gain of ability to be phosphorylated by WalK. However, changing the identity of amino acids at these three positions of PhoP in combination with changing Leu→Ile at position 17 did allow it to be phosphorylated by WalK to differing extents, depending on the combination of amino acids that were changed. Thus, the roles played by amino acids at specificity positions 13, 20 and 107 appear to be subservient to the role played by the amino acid at position 17, which emerged as the dominant determinant of phosphotransfer specificity in the WalRK and PhoPR TCSs. This is consistent with other models that have proposed that position 17 is the point of first contact between cognate proteins, with the specificity positions playing a role only when this first contact is made. The remaining amino acid difference (Glu in WalR, Ile in PhoP) at position 11 (another proposed anchor position) was unfortunately not included in this study, but it deserves investigation in the light of the study by Mukhopadhyay & Varughese (2005).

The importance of amino acid identity at position 11 of the kinases and position 17 of the RRs, as shown in our study, is supported by the fact that the Ile : Ser (WalRK) and the Leu : Gly (PhoPR) combinations are almost universally conserved at these positions among orthologues of these TCSs. Moreover, the Ile : Ser combination found at these positions of the WalRK proteins is unique, and can be used diagnostically to distinguish it from the closely related PhoPR TCSs. Therefore, to gain insight into the structural significance of amino acid identity at these positions, we modelled their location within the kinase–RR complex, based on the known structure of Spo0F : Spo0B (Zapf *et al.*, 2000). All nine positions examined in this study are depicted in Fig. 6: seven of these positions are shown as blue–green circles, while position 11 of the kinase (Gly) and position 17 of the RR (Leu) (the combination found in PhoPR), are shown as ball and stick models in red. It is apparent that the amino acids at these positions are juxtaposed in the kinase–regulator complex. Thus, the size and shape of the complementary side chains (Ile : Ser of WalRK, and Leu : Gly of PhoPR) at this location within the complex are of critical importance to correct docking of the two proteins. This view gives some insight into our finding that changing the identity of the amino acid at this position alone changed discrimination to some extent,
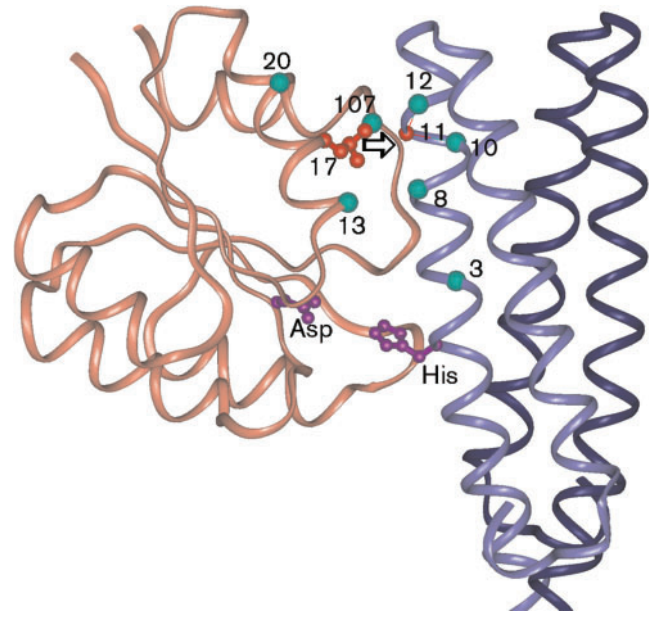


**Fig. 6.** The crucial anchoring site for histidine kinase–RR association. The mode and the geometric locations of sites mutated within WalK and PhoP are shown here on the premise that the Spo0F : Spo0B binding is a prototype for histidine kinase–RR interactions (Skerker *et al.*, 2008). A total of five amino acid residues of WalK (purple) were mutated in order to alter specificity. These positions are represented as spheres on the four-helix bundle. Position 11 was found to be the most crucial position, and it is shown in red. Four positions of PhoP (pink) were mutated, and residue 17, which had the most conspicuous effect, is shown in red. The other three positions are shown as blue–green spheres. The side chain of the residue at position 17 (RR) stacks against position 11 of the histidine kinase, and is indicated by an arrow. The sites of phosphorylation, Asp and His, are shown in dark purple.

although it was not optimal. It also addresses why the contribution of amino acids at the other positions is dependent on the correct amino acid pair at positions 11 and 17. The nuanced nature of the structure at this location within each kinase–regulator complex is illustrated by the disparity in the effect that these amino acid changes might be expected to have on protein structure. The Ser→Gly difference in the WalK and PhoR kinases is a major structural change: Gly has no side chain, and so generates a 'hole' in the protein interacting surface, contrasting with Ser, which has a single carbon side chain of a hydrophilic nature by virtue of the hydroxyl group. Thus, the Ser11Gly change, in terms of the size, shape and hydrophilic/hydrophobic nature of the amino acid side chains, is a major difference within the interacting surfaces of the WalK and PhoR kinases. In contrast, the Leu17Ile change would be viewed as a minimal structural change: both amino acids have a $C_4$ carbon side chain that is hydrophobic and of a similar nature. The amino acids differ only in the positioning of a methyl group, which is located on the $\beta$-carbon in Ile, but on the $\gamma$-carbon in Leu.

Nevertheless, the small structural difference at this position of the RRs is sufficient to be a major determinant of discrimination, requiring a particular amino acid to be located on the corresponding face of the cognate histidine kinases. When the amino acids at both positions are correctly paired, they make a crucial contribution to the high level of phosphotransfer discrimination observed in the WalRK and PhoPR TCSs.

The structures of histidine kinase–RR complexes of TCSs from *Thermotoga maritima* have been reported (Casino *et al.*, 2009; Yamada *et al.*, 2009). These studies have shown the importance of interactions between the α1 helices of the histidine kinase and RRs in complex formation, and in achieving specificity. Mutational analysis and domain swapping within the α1 helices of TCSs from *E. coli* and *T. maritima* have identified the amino acid positions that are crucial for the specificity of phosphoryl transfer between cognate proteins (Skerker *et al.*, 2008; Casino *et al.*, 2009; Schug *et al.*, 2009). Casino *et al.* (2009) showed that changing the amino acids at four positions [two within the α1 helix of HK853 (Thr267Asp and Tyr272Asp) and two within the α1 helix of RR468 (Ile17Asn and Phe20Asn)] attenuates or abolishes the capability of these proteins to form complexes and phosphoryl transfer. Also, Skerker *et al.* (2008) have shown that changing three amino acids of EnvZ to those found at the corresponding positions of RstB (Thr250Val Leu254Tyr Ala255Arg) confers on it the ability to phosphorylate RstA at high efficiency. However, in other cases, the specificity set of amino acids (i.e. those required to effect specific and efficient phosphoryl transfer) appears to be more extensive, and requires amino acid changes both within the α1 helix and in the adjacent loop region between the α1 and α2 helices (Skerker *et al.*, 2008). Our results are in agreement with those studies, demonstrating how specificity can be achieved in two closely related TCSs. There is amino acid variance at only some of the positions within the α1 helices of the WalRK and PhoPR TCSs. For example, of seven covariant positions within the C-terminal regions of histidine kinase α1 helices identified by Skerker *et al.* (2008), only five vary between WalK and PhoR, and three were included in this study (one of these being position 11 at which the Ser→Gly difference occurs). Interestingly, the WalK112 (Ser11Gly, Tyr12Phe) kinase constructed in this study had amino acids changed at two of the three positions that were changed in RstB (Thr250Val, Leu254Tyr, Ala255Arg) in the study by Skerker *et al.* (2008), while the amino acid at the third position (250) is the same in WalK and PhoR. Changing these two amino acids of WalK allows phosphorylation of PhoP both *in vivo* and upon extended incubation *in vitro*, consistent with the results obtained by Skerker *et al.* (2008). However, in contrast with the Skerker *et al.* (2008) study, a complete change of specificity was not observed in the present study, and the rate of phosphotransfer remained low. However, the clear conclusion from both studies is that the complete specificity set of each TCS will be nuanced and complex.

The juxtaposition of amino acids at position 17 of the α1 helix of RRs, and position 11 of the α1 helix of histidine kinases, and its suggested importance as an anchoring point within the co-protein complex (Fig. 6), prompted us to examine the amino acid combinations at these positions in the complement of 34 *B. subtilis* TCSs. Gly[HK] occurs at this position in 10 histidine kinases, and is always paired with a hydrophobic amino acid as follows: Leu[RR] (5), Ile (1), Val (1), Ala (1) and Met (2). The reverse situation of having Gly[RR] at position 17 in RRs occurs in eight TCSs, and it is paired with hydrophobic (Ile, Leu, Met or Trp) or small hydrophilic (Ser) amino acids. Thus, pairing Gly with a hydrophobic and/or small hydrophilic amino acid at this critical anchor position occurs in more than half of all *B. subtilis* TCSs, and this supports its importance as a feature in facilitating initial binding of cognate proteins. Ongoing experiments are focused on establishing the hierarchical nature of specificity determination by amino acid combinations at particular sites of the histidine kinase and RR interacting surfaces.

## ACKNOWLEDGEMENTS

## REFERENCES

**Bisicchia, P., Noone, D., Lioliou, E., Howell, A., Quigley, S., Jensen, T., Jarmer, H. & Devine, K. M. (2007).** The essential YycFG two-component system controls cell wall metabolism in *Bacillus subtilis*. *Mol Microbiol* **65**, 180–200.

**Bisicchia, P., Lioliou, E., Noone, D., Salzberg, L. I., Botella, E., Hubner, S. & Devine, K. M. (2010).** Peptidoglycan metabolism is controlled by the WalRK (YycFG) and PhoPR two-component systems in phosphate-limited *Bacillus subtilis* cells. *Mol Microbiol* **75**, 972–989.

**Casino, P., Rubio, V. & Marina, A. (2009).** Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. *Cell* **139**, 325–336.

**Derre, I., Rapoport, G. & Msadek, T. (2000).** The CtsR regulator of stress response is active as a dimer and specifically degraded *in vivo* at 37 °C. *Mol Microbiol* **38**, 335–347.

**Fabret, C., Fehrer, V. A. & Hoch, J. A. (1999).** Two-component signal transduction in *Bacillus subtilis*; how one organism sees the world. *J Bacteriol* **181**, 1975–1983.

**Gao, R. & Stock, A. M. (2009).** Biological insights from structures of two-component proteins. *Annu Rev Microbiol* **63**, 133–154.

**Hoch, J. A. (2000).** Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* **3**, 165–170.

**Hoch, J. A. & Silhavy, T. (1995).** *Two-Component Signal Transduction*. Washington, DC: American Society for Microbiology.

**Hoch, J. A. & Varughese, K. I. (2001).** Keeping signals straight in phosphorelay signal transduction. *J Bacteriol* **183**, 4941–4949.

**Howell, A., Dubrac, S., Andersen, K. K., Noone, D., Fert, J., Msadek, T. & Devine, K. (2003).** Genes controlled by the essential WalK/YycF two-

component system of *Bacillus subtilis* revealed through a novel hybrid regulator approach. *Mol Microbiol* **49**, 1639–1655.

Howell, A., Dubrac, S., Noone, D., Varughese, K. I. & Devine, K. (2006). Interactions between the YycFG and PhoPR two-component systems in *Bacillus subtilis*: the PhoR kinase phosphorylates the non-cognate YycF response regulator upon phosphate limitation. *Mol Microbiol* **59**, 1199–1215.

Hulett, F. M. (1996). The signal-transduction network for Pho regulation in *Bacillus subtilis. Mol Microbiol* **19**, 933–939.

Jiang, M., Tzeng, Y. L., Feher, V. A., Perego, M. & Hoch, J. A. (1999). Alanine mutants of the Spo0F response regulator modifying specificity for sensor kinases in sporulation initiation. *Mol Microbiol* **33**, 389–395.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessières, P., Bolotin, A. & other authors (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis. Nature* **390**, 249–256.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–688.

Madhusudan, Zapf, J., Whiteley, J. M., Hoch, J. A., Xuong, N. H. & Varughese, K. I. (1996). Crystal structure of a phosphatase-resistant mutant of sporulation response regulator Spo0F from *Bacillus subtilis. Structure* **4**, 679–690.

McCleary, W. R. & Stock, J. B. (1994). Acetyl phosphate and the activation of two-component response regulators. *J Biol Chem* **269**, 31567–31572.

McCleary, W. R., Stock, J. B. & Ninfa, A. J. (1993). Is acetyl phosphate a global signal in *Escherichia coli? J Bacteriol* **175**, 2793–2798.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G. & Dedonder, R. (1990). Signal transduction pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: expression of the regulatory genes and analysis of mutations in *degS* and *degU. J Bacteriol* **172**, 824–834.

Mukhopadhyay, D. & Varughese, K. I. (2005). A computational analysis on the specificity of interactions between histidine kinases and response regulators. *J Biomol Struct Dyn* **22**, 555–562.

Schug, A., Weight, M., Onuchic, J. N., Hwa, T. & Szurmant, H. (2009). High-resolution protein complexes from integrating genomic information with molecular simulation. *Proc Natl Acad Sci U S A* **106**, 22,124–22,129.

Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M. & Laub, M. T. (2008). Rewiring the specificity of two-component signal transduction systems. *Cell* **133**, 1043–1054.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* **69**, 183–215.

Sun, G., Birkey, S. M. & Hulett, F. M. (1996). Three two-component signal-transduction systems interact for Pho regulation in *Bacillus subtilis. Mol Microbiol* **19**, 941–948.

Szurmant, H., Bobay, B. G., White, R. A., Sullivan, D. M., Thompson, R. J., Hwa, T., Hoch, J. A. & Cavanagh, J. (2008). Co-evolving motions at protein–protein interfaces of two-component signaling systems identified by covariance analysis. *Biochemistry* **47**, 7782–7784.

Tzeng, Y. L. & Hoch, J. A. (1997). Molecular recognition in signal transduction: the interaction surfaces of the Spo0F response regulator with its cognate phosphorelay proteins revealed by alanine scanning mutagenesis. *J Mol Biol* **272**, 200–212.

Varughese, K. I. (2002). Molecular recognition of bacterial phosphorelay proteins. *Curr Opin Microbiol* **5**, 142–148.

Varughese, K. I., Madhusudan, Zhou, X. Z., Whiteley, J. M. & Hoch, J. A. (1998). Formation of a novel four-helix bundle and molecular recognition sites by dimerization of a response regulator phosphotransferase. *Mol Cell* **2**, 485–493.

Weigt, M., White, R. A., Szurmant, H., Hoch, J. A. & Hwa, T. (2008). Identification of direct residue contacts in protein–protein interaction by message passing. *Proc Natl Acad Sci U S A* **106**, 67–72.

Yamada, S., Sugimoto, H., Kobayashi, M., Ohno, A., Nakamura, H. & Shiro, Y. (2009). Structure of PAS-linked histidine kinase and the response regulator complex. *Structure* **17**, 1333–1344.

Zapf, J., Sen, U., Madhusudan, Hoch, J. A. & Varughese, K. I. (2000). A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure* **8**, 851–862.

Edited by: J. M. van Dijl