

# SCIENTIFIC REPORTS



OPEN

## The malate sensing two-component system MaeKR is a non-canonical class of sensory complex for C4-dicarboxylates

L. Miguel-Romero<sup>1</sup>, P. Casino<sup>2,3</sup>, J. M. Landete<sup>4,6</sup>, V. Monedero<sup>4</sup>, M. Zúñiga<sup>4</sup> & A. Marina<sup>1,5</sup>

Microbial colonization of different environments is enabled to a great extent by the plasticity of their sensory mechanisms, among them, the two-component signal transduction systems (TCS). Here, an example of TCS plasticity is presented: the regulation of L-malate catabolism via malic enzyme by MaeKR in *Lactobacillales*. MaeKR belongs to the citrate family of TCS as the *Escherichia coli* DcuSR system. We show that the *Lactobacillus casei* histidine-kinase MaeK is defective in autophosphorylation activity as it lacks a functional catalytic and ATP binding domain. The cognate response regulator MaeR was poorly phosphorylated at its phosphoacceptor Asp *in vitro*. This phosphorylation, however, enhanced MaeR binding *in vitro* to its target sites and it was required for induction of regulated genes *in vivo*. Elucidation of the MaeR structure revealed that response regulator dimerization is accomplished by the swapping of  $\alpha 4$ - $\beta 5$ - $\alpha 5$  elements between two monomers, generating a phosphoacceptor competent conformation. Sequence and phylogenetic analyses showed that the MaeKR peculiarities are not exclusive to *L. casei* as they are shared by the rest of orthologous systems of *Lactobacillales*. Our results reveal MaeKR as a non-canonical TCS displaying distinctive features: a swapped response regulator and a sensor histidine kinase lacking ATP-dependent kinase activity.

To sense and to respond to environmental changes and stress conditions, all living organisms have developed different signal transduction systems. In microorganisms, two-component systems (TCS) represent a major signaling mechanism. A prototypical TCS consists of a sensor histidine kinase (HK) and an effector response regulator (RR). Signal transduction by canonical TCSs is known to occur basically by a mechanism of three reactions where, upon signal detection, the HK is autophosphorylated in a conserved His residue<sup>1,2</sup>. Subsequently, the phosphoryl group is transferred to a conserved Asp residue on a cognate RR. This phosphorylation induces conformational changes in the RR that modulate its activity, generally DNA binding, since most RRs are transcriptional factors. Signal transduction is terminated upon RR dephosphorylation, mediated by endogenous phosphatase activity of its cognate HK, action of auxiliary phosphatases or by RR auto-hydrolysis<sup>3-5</sup>. Since TCS most discriminate and respond to specific stimuli among a vast diversity of signals generating quite diverse responses, the sensor domains in the HKs and the effector domains in the RRs present a high variability in sequence and structure. Contrarily, the catalytic domains involved in the autophosphorylation, phosphotransfer and phosphatase reactions are highly conserved both in HKs and RRs<sup>6</sup>. The catalytic core of HKs consists of two domains. The first one, the dimerization and histidine-phosphotransfer (DHP) domain, holds the phosphorylatable His and mediates HK dimerization. This domain is folded as two long antiparallel helices connected by a variable loop that forms a

<sup>1</sup>Department of Genomic and Proteomic, Instituto de Biomedicina de Valencia (IBV-CSIC), Jaime Roig 11, 46100, Valencia, Spain. <sup>2</sup>Departamento de Bioquímica, Universitat de València, Dr Moliner 50, 46100, Burjassot, Spain. <sup>3</sup>Estructura de Recerca Interdisciplinaria en Biotecnología i Biomedicina (ERI BIOTECMED), Universitat de València, Dr Moliner 50, 46100, Burjassot, Spain. <sup>4</sup>Departamento de Biotecnología de Alimentos, Instituto de Agroquímica y Tecnología de Alimentos (IATA-CSIC), Av. Agustín Escardino 7, 46980, Paterna, Valencia, Spain. <sup>5</sup>Group 739 of the Centro de Investigación Biomédica en Red sobre Enfermedades Raras (CIBERER) del Instituto de Salud Carlos III, -, Spain. <sup>6</sup>Departamento de Tecnología de Alimentos, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Carretera de La Coruña Km 7.5, 28040, Madrid, Spain. Miguel-Romero L and Casino P contributed equally to this work. Correspondence and requests for materials should be addressed to M.Z. (email: [amarina@ibv.csic.es](mailto:amarina@ibv.csic.es)) or A.M. (email: [btzman@iata.csic.es](mailto:btzman@iata.csic.es))

stable four-helix bundle in the HK homodimer<sup>5,6</sup>. The second conserved HK domain, named CA (catalytic and ATP binding), binds the ATP molecule and catalyzes the phosphorylation of the histidine in the DHp domain. CA domains present a Bergerat or GHKL fold shared with other ATP-binding proteins (gyrases, Hsp90, MutL). The structure is an  $\alpha/\beta$  sandwich consisting of a five-stranded mixed  $\beta$  sheet and three  $\alpha$  helices in HKs<sup>5,7</sup>. On the other hand, the RR presents a conserved receiver (REC) domain containing the phosphoacceptor Asp. The REC domain is folded with a  $(\beta\alpha)_5$  topology where a central five-stranded parallel  $\beta$ -sheet is surrounded by three  $\alpha$ -helices in one side and two in the other<sup>8</sup>. REC domains present the molecular machinery to recognize and catalyse the phosphoryl group transfer from the cognate HK. Asp phosphorylation of the RR REC domain triggers a change in the rotameric state of highly conserved Thr/Ser and Tyr/Phe residues in  $\beta 4$  and  $\beta 5$ , respectively, inducing the conformational changes observed between the active and inactive states<sup>8,9</sup>. This activation mechanism, named Y-T switch, seems to be predominant among RRs although alternative activation modes have been described for some RR subfamilies such as NtrC<sup>9,10</sup>.

As our knowledge about TCS increases, variations of this basic signalling paradigm, including new regulatory steps, are described. For example, ancillary proteins can regulate any of the three signalling reactions or work as co-sensors together with the HK, or unphosphorylated RRs can also play a critical role in the regulation of gene expression<sup>11–15</sup>. In this way, the combination of a highly conserved catalytic machinery with variable sensor and effector domains together with the participation of ancillary proteins, provide TCSs of an extremely high signalling plasticity allowing these systems to generate species-specific responses to similar stimulus.

One example would be the regulation of C4-dicarboxylate uptake, which in *Escherichia coli* involves the TCS DcuS/DcuR and the secondary transporters DctA or DcuB<sup>16</sup>. In lactic acid bacteria (LAB), the metabolism of the C4-dicarboxylate L-malate takes place via two distinct pathways. One allows the conversion of L-malate into L-lactate via the function of the malolactic enzyme (MLE) pathway and the other converts L-malate to pyruvate via the malic enzyme (ME). *Lactobacillus casei* is a facultatively heterofermentative LAB frequently used as a starter culture for the production of fermented food products. *L. casei* is remarkable among LAB since it is one of the few species with the ability to utilize L-malic acid via the ME or the MLE pathways<sup>17–19</sup>. Regulation of the transcription of the genes involved in the utilization of L-malate via the ME pathway is under control of a TCS of the citrate family<sup>18</sup>, whereas the MLE pathway is under control of a LysR-type transcriptional regulator<sup>19</sup>. Although both pathways utilize the same substrate, the transcription of the corresponding genes is independently regulated<sup>19</sup>.

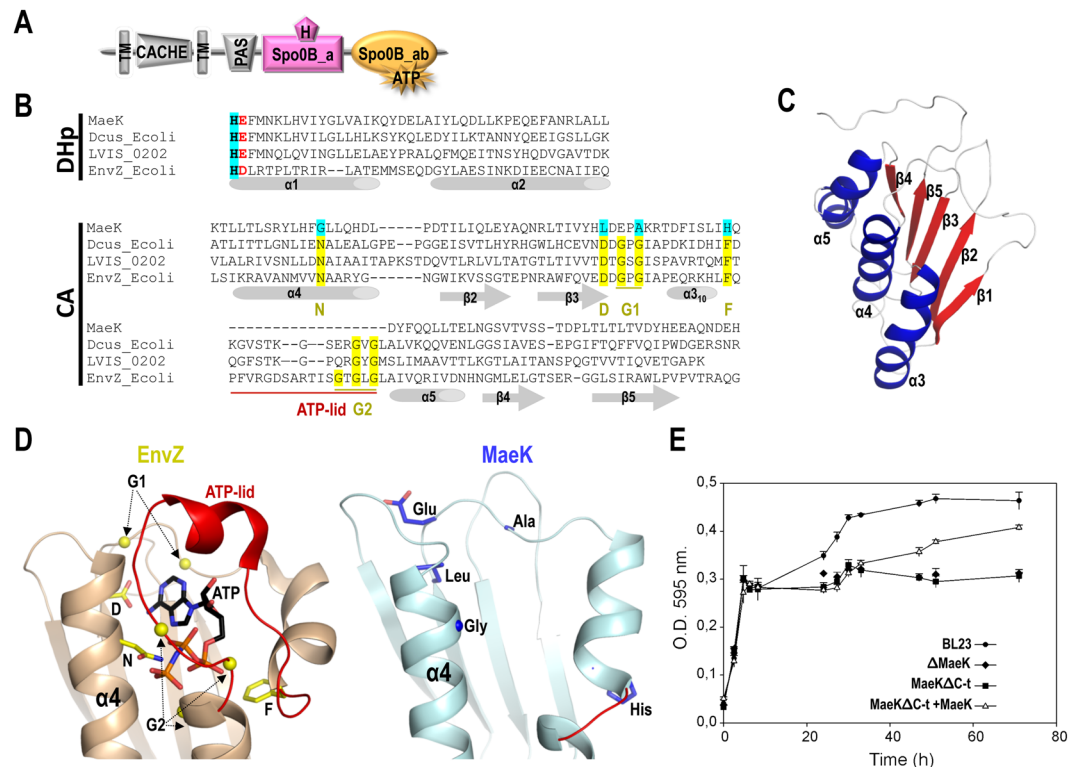
Genes involved in the ME pathway are arranged in two diverging operons, *maePE* and *maeKR*, encoding a putative L-malate transporter (*maeP*), an ME (*maeE*), and a TCS (*maeK* and *maeR*). This cluster is only found in a few species of the *Lactobacillales* order<sup>18,20</sup>. It has been shown that ME is required for growth with L-malate and that both *maeK* and *maeR* are essential for expression of *maePE*<sup>18</sup>. Requirement of *maeR* for *maePE* expression has been also demonstrated in *Enterococcus faecalis* JH2-2<sup>21</sup> and requirement of *maeK* has been demonstrated in *Streptococcus pyogenes*<sup>22</sup>. Transcriptional analyses showed that the expression of the *maePE* operon was induced by L-malate and repressed by glucose in *L. casei*<sup>18,19</sup> and *E. faecalis*<sup>21</sup>. In *S. pyogenes* and *L. casei*, in addition to L-malate, low pH also induced the expression of *maePE*<sup>22</sup>. Interestingly, it has been shown that induction by low pH is independent of L-malate but requires a functional MaeK in *S. pyogenes*<sup>22</sup>.

As noted above, induction by L-malate was dependent on a functional MaeKR TCS. Furthermore, inactivation of the genes encoding L-malate transporters in *L. casei* (*maeP* and *mleT*; the latter associated to the malolactic gene cluster) did not affect the transcription of *maeE* in the presence of L-malate<sup>19</sup>. This result, together with the absence of genes encoding putative solute binding proteins, suggests that no co-sensor proteins are participating in the regulation of MaeK activity, in contrast to other TCS of the citrate family (e.g. DcuS)<sup>23</sup>.

In order to understand these differences in the mechanism of regulation by MaeKR of *L. casei*, we sought to characterize the activity of both proteins *in vitro* and *in vivo*. We showed that MaeK forms a particular group within the HKs, characterized by a lack of a functional ATP-binding domain. We also solved the crystal structure of the REC domain of MaeR in its activated state as well as the structure of a mutant version defective in the phosphoacceptor Asp. Finally, we characterized the interaction of MaeR with the *mae* promoter region. Altogether, our analysis has revealed that MaeKR is a non-canonical TCS conserved in *Lactobacillales* and that the structure of MaeR acknowledges the  $\alpha 4\beta 5\alpha 5$ -swapped dimers as a new class of dimerization for RRs.

## Results

***L. casei* MaeK lacks the essential conserved motifs for ATP binding.** In order to dissect the functional mechanism of MaeKR, we analysed the domain architecture of both proteins using Pfam and SMART servers. The search revealed that the sensor MaeK presents two transmembrane helices connected through an extracellular sensor domain consisting of a single Cache domain type 3 (PF17203), followed by a cytoplasmic PAS sensor domain (CL0183; residues 212–285), an alpha-helical domain SPOB\_a (PF14689; residues 308–375) and a phospho-transferase B, C-terminal SPOB\_ab domain (PF14682; residues 383–514) (Fig. 1A). However, and surprisingly for a putative sensor HK, the analysis did not reveal the presence of DHp phosphoacceptor (PFAM: HisKA or HisKA\_3) or CA ATP-binding (PFAM: HATPase\_c) domains, the canonical HK catalytic domains. Further comparative sequence analysis with canonical HKs allowed us to identify a putative phosphorylatable histidine (His334) in the SPOB\_a helical domain, which corresponds to the phosphoacceptor His in Spo0B-type phosphotransferases (Fig. 1B). In addition, this His is followed by a catalytic acidic residue (Glu335) as is observed in canonical HKs<sup>2</sup> (Fig. 1B), further supporting the hypothesis that this domain (residues 308–375) is a functional DHp domain. Contrarily, the sequence analysis of the C-terminal region of MaeK revealed the absence of the conserved motifs or “boxes” characteristics of HK CA domains necessary for ATP binding (Fig. 1B). These boxes provide key residues for the ATP-binding pocket and give specificity for this nucleotide. The N box is named after a conserved Asn that coordinates the magnesium ion for catalysis, D box after an Asp that gives specificity for adenine, the G boxes after two Gly rich regions that confer flexibility to the characteristic loop



**Figure 1.** Domain composition and structural description of MaeK. **(A)** Graphical representation of the domains contained in MaeK, obtained with the SMART database and Pfam. TM correspond to transmembrane helix. **(B)** Sequence alignment of the Spo0B\_a domain of MaeK with DHP domain of DcuS and EnvZ of *E. coli* containing the phosphorylatable His (in bold with blue pattern) and the subsequent catalytic acidic residue (colored in red). Below is represented the sequence alignment of the Spo0B\_ab domain (degenerated CA domain) of MaeK with the CA domain of CitB from *Lactobacillus brevis* (LVIS\_0202), and DcuS (DcuS\_Ecoli) and EnvZ (EnvZ\_Ecoli) from *E. coli* highlighting the conserved boxes for a canonical CA domain with their conserved specific residues in yellow pattern. Absence of conserved specific residues in MaeK contain blue pattern. **(C)** 3D structural model for the degenerated CA domain of MaeK generated by the program I-Tasser (helices in blue and beta strands in red). **(D)** Superposed 3D CA domain structures for EnvZ and MaeK to highlight the absence of conserved boxes and ATP-lid in MaeK with respect to EnvZ (colored as in B). **(E)** Growth of wild-type *L. casei* (BL23) and MaeK derivative strains. Wild-type (BL23), MaeK gene defective ( $\Delta$ MaeK), MaeK lacking CA domain (MaeK $\Delta$ C-t) and MaeK $\Delta$ C-t complemented with a plasmid expressing MaeK (MaeK $\Delta$ C-t + MaeK) strains were grown in MEI medium supplemented with 5 g liter<sup>-1</sup> of L-malic acid. Values represent the means of three independent experiments; error bars represent standard deviations.

covering the ATP pocket (ATP-lid) and the F box after a Phe positioned at the beginning of the ATP-lid<sup>7</sup> (Fig. 1B). To further confirm the absence of the canonical ATP-binding pocket in this domain and in order to avoid a possible misalignment of the sequence, we generated a structural model of the C-terminal domain of MaeK, using the I-TASSER server<sup>24</sup>. The model presents the characteristic architecture of HK CA domains with five  $\beta$  strands and three  $\alpha$  helices forming an  $\alpha\beta$  sandwich (Fig. 1C). Superposition with the CA domain of the prototypical HK EnvZ (PDB:4KP4), confirmed the absence in MaeK of the conserved boxes essential for ATP binding, as well as the characteristic ATP-lid (Fig. 1D). Furthermore, docking an ATP molecule in the modelled putative CA domain of MaeK, according to the structure of EnvZ bound to AMPPNP (non-hydrolysable ATP analogue; PDB 4KP4), showed that the nucleotide accommodation would be sterically precluded by the presence of residues that occupies the binding pocket (Supplementary Fig. 1).

In order to check, *in vitro*, the autophosphorylation activity of MaeK we produced the recombinant complete cytoplasmic portion of MaeK (MaeK<sub>C</sub>). The protein eluted in gel filtration chromatography as a single peak corresponding to the expected size of the dimer species (Supplementary Fig. 1), a quaternary structure observed in functional HKs<sup>25</sup>. Furthermore, analysis of the circular dichroism spectra for MaeK<sub>C</sub> estimated a secondary structure composition of 32%  $\alpha$ -helix and 17.8%  $\beta$ -strand which correlated with the secondary structure composition predicted from the *in silico* model (37%  $\alpha$ -helix and 22%  $\beta$ -strand) (Supplementary Fig. 1), supporting that MaeK<sub>C</sub> is well-folded. We checked *in vitro* the kinase capacity of MaeK<sub>C</sub> using [ $\gamma$ -<sup>32</sup>P]ATP as substrate, and no autophosphorylation was observed, even at high concentrations of radiolabeled ATP, long incubation times or in the presence of different divalent ions, conditions where the corresponding portions of canonical HKs showed autophosphorylation (Supplementary Fig. 2). We also measured by isothermal calorimetry (ITC) the binding affinity of MaeK<sub>C</sub> for AMPPNP in the presence of 10 mM MgCl<sub>2</sub>, a divalent cation required for catalysis

(Supplementary Fig. 2). Kd constant values higher than 1 mM were obtained, which reflects non-binding in the ITC system used, either for AMPNP-MgCl<sub>2</sub> at high concentrations or MgCl<sub>2</sub> alone (Supplementary Fig. 2). Parallel assays with HK853, a canonical HK with autophosphorylation capacity, were carried out to validate the assay (Supplementary Fig. 2). This result confirms that MaeK<sub>C</sub> is unable to bind nucleotide or metal ion even at physiological concentrations. Altogether, the *in silico* and *in vitro* analysis of MaeK strongly suggest that, even though it conserves a degenerated CA domain, it has lost the capacity to bind ATP and consequently, its putative autophosphorylation activity. However, MaeK seems to maintain an active phosphorylatable His, suggesting that it might conserve phosphotransferase and/or phosphatase activities. In fact, MaeK contains the E/DxxT/N motif, after the catalytic His, which has been demonstrated to be critical for the phosphatase activity in sensors that dephosphorylate receiver domains<sup>26</sup> (Fig. 1B).

**Deletion of the degenerated CA domain of *L. casei* MaeK prevents growth in malic acid.** We had previously observed that a complete deletion of *maeK* ( $\Delta$ *maeK*) in *L. casei* led to an absence of growth in medium containing L-malate as a carbon source, and this deficiency was totally overcome by expression of MaeK from a plasmid<sup>18</sup>. Here we sought to analyse whether the C-terminal domain of MaeK, which is devoid of all the characteristics of an ATP-binding domain and, consequently, kinase activity, was necessary for activation of the *mae* promoter. To this end, a *L. casei* strain (MaeK $\Delta$ C-t) expressing a MaeK variant lacking the C-terminal domain (residues 378–514) was obtained. As previously observed with the  $\Delta$ *maeK* strain, MaeK $\Delta$ C-t strain could not grow on L-malate (Fig. 1E). This observation suggests that the deleted domain is necessary for MaeR activation, although an incorrect membrane insertion/folding of MaeK $\Delta$ C-t cannot be excluded. In contrast to the results obtained with the  $\Delta$ *maeK* strain, complementation of MaeK $\Delta$ C-t with a plasmid expressing MaeK only showed partial restoration of growth on L-malate (Fig. 1E). The partial restoration of growth observed in this strain could be explained by the formation of MaeK/MaeK $\Delta$ C-t heterodimers with reduced activity compared with wild-type homodimers.

**The *L. casei* *mae* homologous genes share a common evolutionary history in Lactobacillales.** Previous phylogenetic analysis of *maeK* and *maeR* genes showed that the *maeK* and *maeR* homologs, that are present exclusively in some species of *Lactobacillales*, shared a common origin and are related to *maeK* and *maeR* genes present in some species of bacilli<sup>18</sup>. In addition, the same gene organization is found in all gene clusters of *Lactobacillales* belonging to this group which consist of *maeK* and *maeR* and, in the opposite orientation, *maeP* and *maeE* (Fig. 2A). As shown above, the analysis of the structure of MaeK revealed a distinctive feature that is shared by all MaeK belonging to this group: a degenerated CA domain. For clarity of exposition, we will name the group of gene clusters encompassing *maeK* homologs with a degenerated CA domain as the Mae1 group. In *Lactobacillales* another group of homologous *maeK* and *maeR* genes can be found (Mae2). These genes are arranged in a cluster together with a gene encoding a putative D-lactate dehydrogenase and a putative L-malate transporter (Fig. 2B) homologous to the *mleT* gene found in the malolactic cluster of *L. casei*<sup>19</sup>. Contrarily to Mae1 group, HKs belonging to Mae2 contain a canonical CA ATP-binding domain (Fig. 1B).

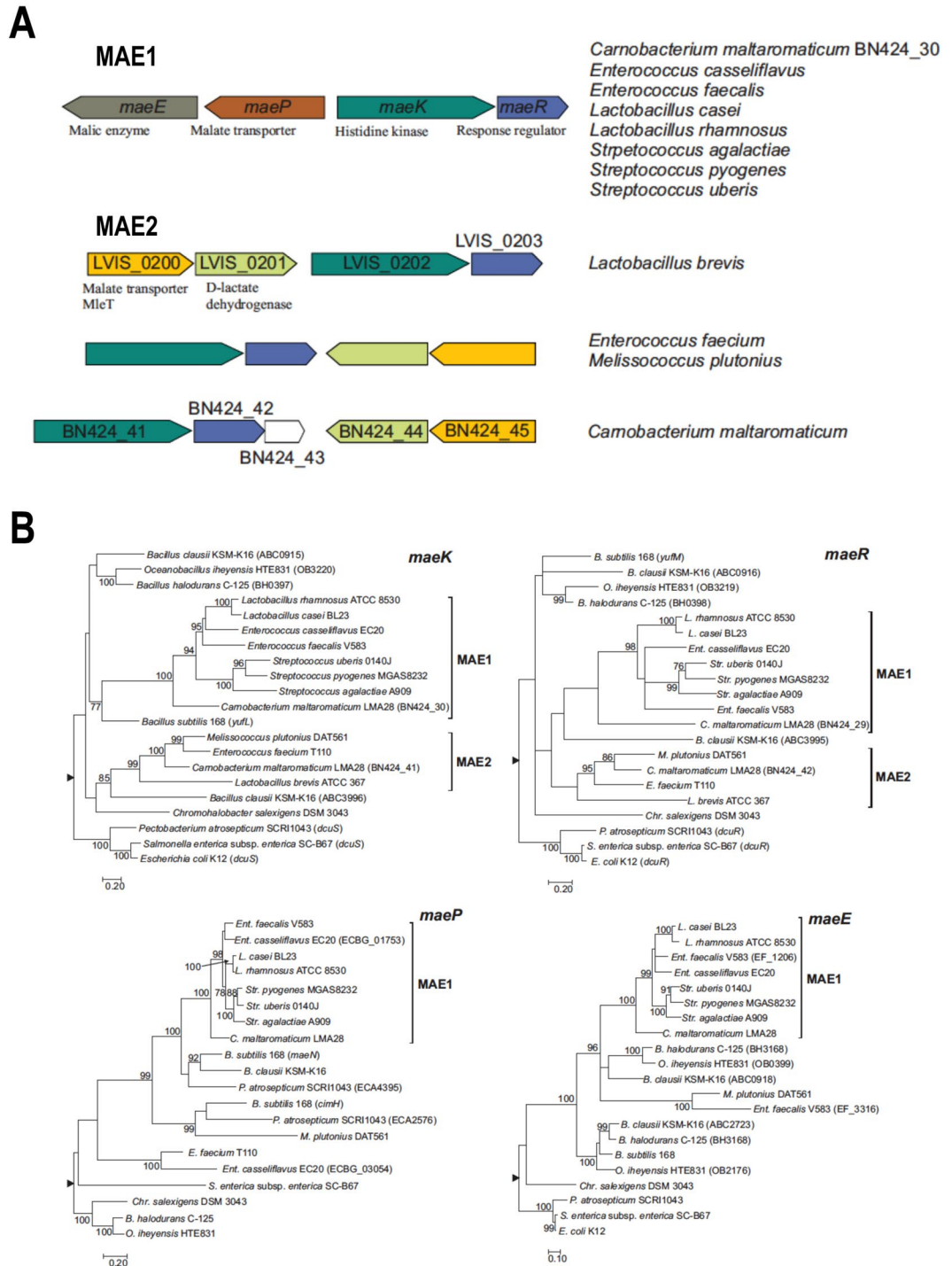
In order to gain insight into the evolution of the *mae* gene cluster we analysed the phylogenetic relationships of *maeK*, *maeR*, *maeP* and *maeE* in a group of species selected on the basis of our previous analysis<sup>18</sup>. The phylogenetic reconstructions showed that genes belonging to the Mae1 group constitute well-supported groups (Fig. 2B) excepting the *maeR* tree where the group can still be identified but with low support. Nevertheless, it must be taken into account that the information content of the *maeR* dataset estimated in number of fully resolved quartets by likelihood mapping was lower than those of the other three genes analysed (79% fully resolved quartets for *maeR*, 87.25% for *maeK*, 91.99% for *maeP* and 92.16% for *maeE*). Furthermore, we tested the congruency of the topologies of the *maeR* and *maeK* phylogenetic trees attending to the tests implemented in Treepuzzle<sup>27</sup>. This analysis showed that for the *maeR* dataset, the topology of the *maeR* tree was not significantly better than the topology of the *maeK* tree (Supplementary Table 1). In contrast, for the *maeK* dataset, the topology of the *maeK* tree was significantly better than the topology of the *maeR* tree. This result indicates that the topology of the *maeR* tree is possibly identical to the topology of the *maeK* tree but due to the low information content of the *maeR* dataset, it could not be retrieved by the phylogenetic reconstruction procedure. In summary, the phylogenetic analyses indicate that *maeR* and *maeK* homologous genes included in this analysis have possibly evolved as a single unit.

The analyses also indicate that Mae2 group genes evolved independently of Mae1 group. First, the gene content differs between the two groups (Fig. 2A). Second, the phylogenetic analyses show that *maeK* and *maeR* genes of Mae1 and Mae2 groups are distantly related (Fig. 2B). Third, Mae2 group MaeK homologs conserve the CA domain. In this sense, it is worth noting that *Carnobacterium maltaromaticum* LMA28 harbors both gene clusters.

The close relationship between *mae* genes found in some *Bacilli* and those of the Mae1 group of *Lactobacillales* may indicate a common origin of the gene clusters found in both groups. Our results do not allow ascertaining this point. Notwithstanding, our results strongly suggest that the *Lactobacillales* *mae* gene clusters of the Mae1 group originated from a common ancestor cluster that was rearranged in *Lactobacillales* and has been disseminated as a single unit. This notion is further substantiated by the presence of *maeP* or *maeE* paralogs in some strains of the Mae1 group (Fig. 2B). In all cases, these paralogs are distantly related to their cognate Mae1 group genes.

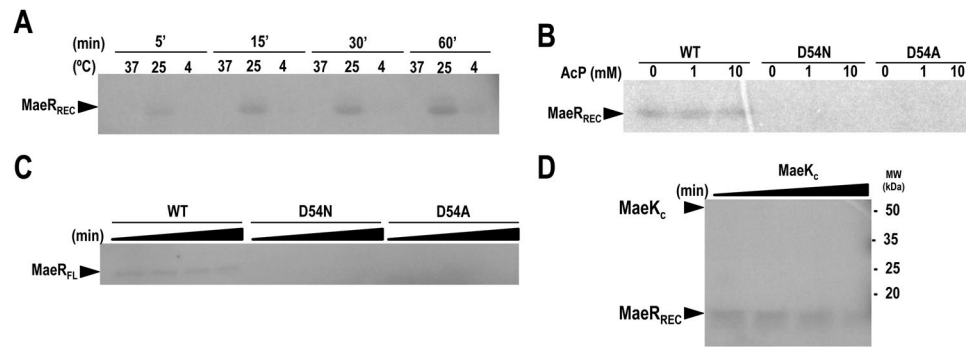
***L. casei* MaeR shows low in vitro phosphorylation levels in the presence of acetyl phosphate.** Since MaeK is a non-canonical HK lacking a functional CA domain, we checked the architecture and functionality of MaeR. *In silico* modeling with I-TASSER server<sup>24</sup> showed that MaeR presents the typical domain architecture of a RR, consisting of a N-terminal phosphoacceptor REC domain (residues 1–122) and a C-terminal DNA-binding domain (type winged HTH; residues 123–229), with high structural homology to the RR PrrA from *Mycobacterium tuberculosis* that belongs to the OmpR/PhoB family (Supplementary Fig. 3). Analysis of this structural model showed that the REC domain of MaeR contains the conserved catalytic residues of RRs, including





**Figure 2.** Clusters of *mae* genes and phylogenetic analysis. **(A)** Arrangements of gene clusters present in the specified species. Colours indicate homologous genes. Arrows indicate the direction of transcription. **(B)** Maximum-likelihood phylogenetic trees of homologous cluster genes. The trees are arbitrarily rooted using the mid-point rooting method implemented in Mega software<sup>64</sup> (black triangle; see Methods). Only bootstrap support values higher than 75% are shown. Genes belonging to Mae1 and Mae2 groups are indicated. Gene names or gene-tags are indicated in brackets when paralogous genes were included in the analysis.

the phosphoacceptor Asp (D54) in a correct spatial disposition (Supplementary Fig. 3), suggesting that MaeR is a functional RR. Frequently, RRs can be phosphorylated *in vivo* and *in vitro* by small phosphodonors, such as acetyl phosphate (acetylP)<sup>28</sup>. Given the particularities of MaeK, we wondered if MaeR could be phosphorylated by acetylP. Therefore, phosphorylation assays of MaeR containing only the REC domain (residues 1–121; MaeR<sub>REC</sub>) were performed in the presence of radioactive <sup>32</sup>P-acetylP at different temperatures (4, 25 and 37 °C) and times (5–60 min). Low phosphorylation levels were observed in MaeR<sub>REC</sub> after incubation at 25 °C and at different time points



**Figure 3.** MaeR phosphorylation by acetylP and dephosphorylation by MaeK. **(A)** Phosphorylation of the REC domain of MaeR in the presence of the radiolabeled phosphodonor acetylP and different temperatures (5, 25 and 37°C) during different incubation times (5, 15, 30 and 60 min). **(B)** Phosphorylation of the MaeR REC domain for wild-type and mutant forms D54A and D54N with radiolabeled acetylP at room temperature for 1 h in the presence of 1 and 10 mM of cold acetylP. **(C)** Phosphorylation of MaeR full-length protein wild-type and mutant forms with radiolabeled acetylP at different incubation times (5, 10, 20 and 40 min). **(D)** Incubation of phosphorylated REC domain of MaeR in the presence of radiolabeled acetylP for 1 h and subsequent incubation with MaeK<sub>C</sub> during 1, 2.5, 5 and 15 min. Full-length autoradiographies are shown in Supplementary Fig. 11

(Fig. 3A). Unexpectedly, lower phosphorylation levels were observed at 37°C, probably indicating a reduced stability of MaeR<sub>REC</sub>~P at high temperatures. To assess that MaeR was phosphorylated at D54, two mutant MaeR<sub>REC</sub> variants at the phosphorylatable Asp (D54A and D54N) were obtained. Assays with radiolabeled acetylP using increasing concentrations of cold acetylP (1 and 10 mM) showed that the mutants were not phosphorylated in any condition assayed (Fig. 3B), confirming that the low, but consistent, level of phosphorylation of the wild-type MaeR<sub>REC</sub> occurred at the phosphoacceptor Asp residue. Further phosphorylation assays using full-length MaeR (MaeR<sub>FL</sub>), as well as mutants D54A and D54N of this full-length protein, confirmed that the low level of phosphorylation was independent of the absence or presence of the DNA-binding domain (Fig. 3C).

Usually, phosphorylation of a RR drives its oligomerization, mainly dimerization, yielding the active conformation of the protein<sup>29</sup>. Even though the level of phosphorylation of MaeR was low; we sought to observe the oligomeric (dimeric) phosphorylated form by size-exclusion chromatography using MaeR<sub>FL</sub> and MaeR<sub>REC</sub>. Size-exclusion chromatography with wild-type and mutants D54A and D54N showed that MaeR eluted from the column with a retention volume corresponding to the monomeric species in the absence or presence of acetylP (Supplementary Fig. 4). These results opened the possibility that phosphorylation did not induce oligomerization on MaeR, as it has been described for some RRs lacking DNA-binding domain, or that the low level of the phosphorylation, and consequently of the dimer, precluded the detection of this oligomeric form.

As MaeK was defective in *in vitro* autophosphorylation activity but contained the E/DxxT/N motif present in HKs with the capacity to dephosphorylate RR~P, MaeK phosphatase activity towards phosphorylated MaeR was assayed. Phosphatase assays were performed by incubating wild-type MaeR<sub>REC</sub>~P with MaeK<sub>C</sub>. In this experiment dephosphorylation of MaeR<sub>REC</sub>~P was not observed in the course of the reaction (Fig. 3D), supporting that MaeK<sub>C</sub> does not show phosphatase activity towards MaeR~P under our assay conditions. In addition, no new band corresponding to MaeK<sub>C</sub>~P (Fig. 3D) was observed, suggesting that this protein cannot be phosphorylated by acetylP, which was present in the reaction. Similar results for MaeK<sub>C</sub> phosphorylation were obtained when phosphoenolpyruvate was used as small phosphodonor (data not shown). Further native-PAGE experiments indicated that the MaeK<sub>C</sub>-MaeR<sub>REC</sub> complex showed low stability or that the complex was transient (Supplementary Fig. 4), suggesting that the full-length membrane-embedded MaeK or additional partners might be needed for a productive interaction. Therefore, lack of phosphatase activity could be intrinsic to MaeK or be related to the low stability showed by the MaeK-MaeR complex.

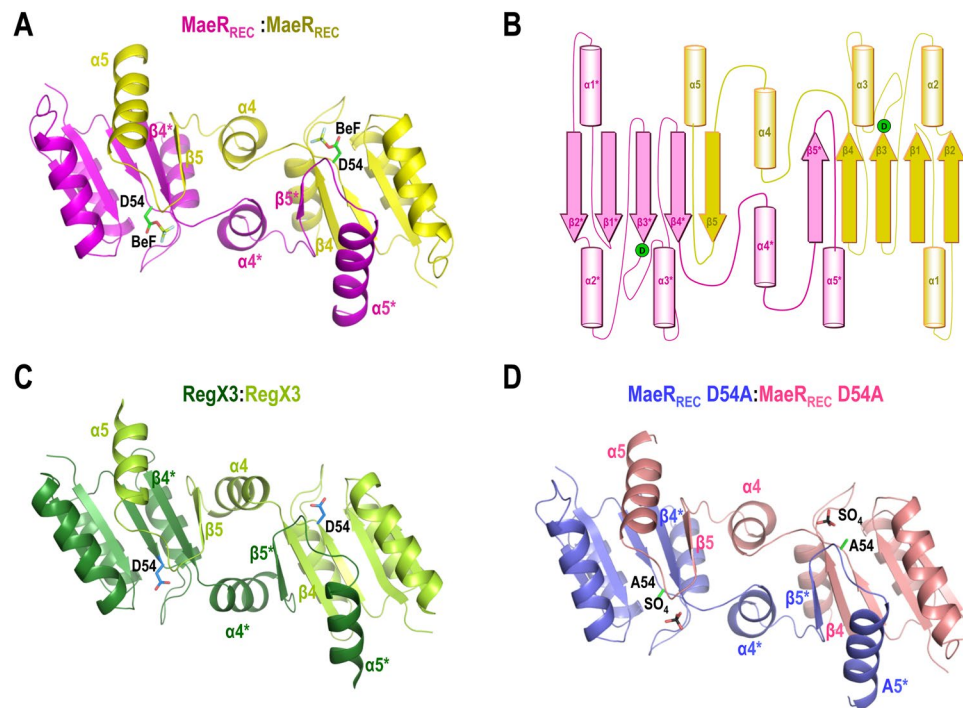
***L. casei* MaeR<sub>REC</sub> dimerizes in a swapped conformation.** To gain insight into the molecular basis of the activation mechanism of MaeR, we performed the structural characterization of MaeR<sub>REC</sub>, either in the phosphorylated or unphosphorylated conformation. To induce the phosphorylated conformation we used the phosphomimetic beryllium trifluoride (BeF<sub>3</sub><sup>-</sup>), that is known to bind the phosphorylatable Asp leading to a structure that mimics the phosphorylated species and that has been broadly used with several RRs to this end<sup>30,31</sup>. Crystals were only obtained for MaeR<sub>REC</sub> in the presence of BeF<sub>3</sub><sup>-</sup>, suggesting that this form (the phosphorylated REC domain) could represent the most compact and stable conformation. Crystals diffracted to 2.1 Å (Table 1) and the asymmetric unit contained two monomers of MaeR<sub>REC</sub>. Remarkably, these two monomers were organized as a dimer but by swapping some structural elements in a conformation just observed, so far, in the activated conformation of the full-length structure of the RR RegX3 from *Mycobacterium tuberculosis*, which contains the REC and DNA binding domains (RegX3<sub>FL</sub>; PDB:2OQR)<sup>32</sup> (Fig. 4). Although the presence of this type of conformation in other RR supports its biological significance, we drove our attention to crystal packing in the MaeR structure. The solvent content of the unit cell was remarkably high (~67%), leaving a considerable free space for the dimeric quaternary structure observed in the crystal. This fact supports that packing forces do not induce the swapped conformation (Supplementary Fig. 5). The dimerization interface involves α4β5α5 structural elements, the main dimerization surface exploited by canonical RRs, such as the OmpR/PhoB family<sup>29</sup>, but in MaeR β5

Processed data	MaeR wild type	MaeR-D54A
ESRF Beamline		I29
DLS beamline	I04	
Wavelength (Å)	0.97624	0.97627
Resolution (Å)	27.91–2.11 (2.22–2.11)	43.85–3.10 (3.27–3.10)
Rmerge (%)	4.4 (67.4)	6.1 (54.0)
Rpim (%)	1.7 (26.3)	4.1 (37.3)
Mean I/σ(I)	22.1 (3.2)	11.2 (2.2)
N° reflections (observed/unique)	179137/23899 (25645/3477)	23722/7772 (3367/1119)
Completeness (%)	99.8 (98.9)	99.2 (99.1)
Redundancy	7.5 (7.4)	3.1 (3.0)
Space group	I4 <sub>1</sub>	I4 <sub>1</sub>
Cell dimensions (Å)	a = b = 75.79 c = 147.55	a = b = 76.64 c = 149.26
Angles (°)	α = β = γ = 90	α = β = γ = 90
<b>Refined data</b>		
R <sub>crys</sub> (%)	20.3	22.6
R <sub>free</sub> (%)	23.1	28.6
<i>N° atoms</i>		
Protein	1984	1964
Ligand/ion	10	10
Water	86	8
<i>B-factors (Å<sup>2</sup>)</i>		
Protein	54.2	116.8
Ligand/ion	37.8	95.6
Water	57.3	76.7
<i>r.m.s deviations</i>		
Bond lengths (Å)	0.007	0.008
Bond angles (°)	1.27	1.29
<i>Ramachandran (%)</i>		
Favoured	97.06	95.42
Allowed	2.94	4.58
Disallowed region	0.0	0.0
PDB code	5LWK	5LWL

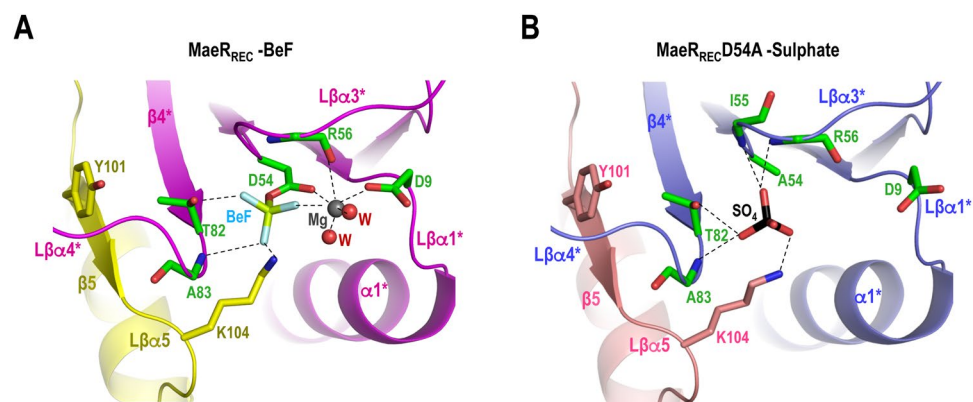
**Table 1.** Crystallographic data and refinement statistics of MaeR.

and α5 are swapped between subunits (Fig. 3A,B and Supplementary Fig. 6). This swapped conformation is due to a motion of the C-terminal β5α5 structural elements that swap from one monomer to the other by a rotation of about 178° (calculated by Dyndom<sup>33</sup>), using the loop β4-α4 and α4 as a hinge (Supplementary Fig. 7). To accomplish this swapping, α4 is reoriented and placed perpendicular to α3 instead of parallel, as is observed in the OmpR/PhoB family of RR (Supplementary Figs 6 and 8). This motion allows the interaction of α4 from one subunit with α4 from the other subunit (α4\*) and promotes new interactions that involve mainly β4 with β5\* (and β4\* with β5) (Fig. 4a and Supplementary Fig. 9). Studies of conformational dynamics in RegX3 point to the α4β5 region as critical to acquire the swapped conformation<sup>34</sup>. Although this swapped conformation exploited the α4β5α5 elements to generate the dimer, the relative disposition of the monomers in the dimer is completely different to that observed in the prototypical OmpR/PhoB family. Contrary to OmpR/PhoB dimers, where both active centres are placed in the same face of the dimer, they are located in opposite faces of the dimer in MaeR (Supplementary Fig. 6). The swapped MaeR<sub>REC</sub> dimer displays an interface area of about 2018 Å<sup>2</sup> and a solvation free energy Δ<sup>i</sup>G of about −34 kcal.mol<sup>−1</sup>, suggesting that it is a tight dimer. The fact that MaeR<sub>REC</sub> presents the same type of dimerization by domain swapping as RegX3<sub>FL</sub> (Fig. 4A and C), supports its biological relevance and reveals that the capacity of swapping seems to be intrinsic to the REC domain, since it is driven independently of the presence of the DNA binding domain. Furthermore, the identical structural organization of MaeR and RegX3, two phylogenetically distant RRs (MaeR belongs to the Cit family and RegX3 to the OmpR family of RRs) from unrelated microorganisms, support that domain swapping may represent a new paradigm of activation for a specific group of RRs.

A close view of the active site of each monomer shows the presence of BeF<sub>3</sub><sup>−</sup> bound to the catalytic Asp54 (Fig. 5A). The binding of BeF<sub>3</sub><sup>−</sup> to MaeR is identical to that observed in complexes of this phosphomimetic with prototypical RRs<sup>35</sup>. In this way, BeF<sub>3</sub><sup>−</sup> contacts with the conserved catalytic Lys104 in the loop connecting β5-α5 (Lβα5), Thr82 at the end of β4, the main-chain amide of Ala83 in loop connecting β4-α4 (Lβα4) and to a magnesium ion. Furthermore, the magnesium ion shows the characteristic interactions with the phosphoacceptor Asp54, the conserved Asp9, the main chain carbonyl oxygen of Arg56 and two water molecules, which are



**Figure 4.** Crystal structure of MaeR<sub>REC</sub> wild-type and D54A mutant. **(A)** Cartoon representation of the dimeric MaeR<sub>REC</sub> wild-type structure highlighting the structural elements involved in the swapped dimerization surface (one monomer in magenta, the other in yellow) as well as the catalytic Asp bound to BeF<sub>3</sub><sup>-</sup> (in green). **(B)** Schematic representation of the MaeR dimer showing the structural elements involved in the swapped dimerization. **(C)** Cartoon representation of RegX3 REC domain (PDB: 2OQR) which presents the same swapped dimerization architecture as MaeR. **(D)** Cartoon representation of the dimeric MaeR<sub>REC</sub>-D54A structure that shows the same swapped dimer organization (one monomer in blue, the other in pink) as MaeR. The Ala residue replacing the phosphoacceptor Asp and the sulfate ion occupying the active site are highlighted in sticks.



**Figure 5.** Active site organization of MaeR<sub>REC</sub> wild-type and D54A mutant. **(A)** Cartoon representation of the active site in the REC domain of wild-type MaeR structure showing as sticks the catalytic residues as well as the BeF<sub>3</sub><sup>-</sup> (carbons in green for the magenta coloured monomer and in yellow for the yellow coloured monomer; nitrogen, oxygen, beryllium and fluor in blue, red, pale-green and cyan, respectively). The magnesium ion (in grey) and the water molecules (in red) that stabilize the active site are also shown. **(B)** Cartoon representation as in A of the active site in the mutant D54A with the presence of a sulphate ion in the active site. Catalytic residues are shown as sticks (carbons in green for the blue coloured monomer and in pink for the pink coloured monomer; sulfur in back).

coordinated by the conserved Asp9 and Lys104, and one fluoride from BeF<sub>3</sub><sup>-</sup> (Fig. 5A). The residues Thr82 and Tyr101, which correspond to the Y-T or aromatic switch, are oriented towards (inward) the active site. The Y-T switch is the general mechanism for intermolecular signalling in RR and the inward conformation is adopted upon Asp phosphorylation, as it is observed in other structures of RRs phosphorylated or bound to BeF<sub>3</sub><sup>-</sup><sup>36</sup>.



Interestingly, Lys104 and Tyr101, which contribute to the active site and to the Y-T switch respectively, are provided from swapped elements of the second monomer (Fig. 5A). In this way, the phosphorylated state cannot be reached unless the REC domain is swapped. Therefore, MaeR should adopt the swapped conformation to be competent for phosphorylation. Similarly, the swapped conformation has been proposed to represent the active (phosphorylated) conformation of RegX<sup>32</sup>.

***L. casei* MaeR<sub>REC</sub>-D54A also dimerizes in a swapped conformation.** In an attempt to determine the MaeR unphosphorylated monomeric conformation, we performed crystallization studies of a MaeR<sub>REC</sub> mutant form where D54 was replaced by an alanine (MaeR<sub>REC</sub>-D54A). MaeR<sub>REC</sub>-D54A crystallized in the same space group and similar size dimension than the wild-type MaeR<sub>REC</sub> (Table 1). The determination of the structure of MaeR<sub>REC</sub>-D54A by molecular replacement revealed the presence of a swapped dimer into the unit cell identical to the one obtained for the wild type MaeR<sub>REC</sub> (RMSD is 0.52 Å for 242 residues) (Fig. 4A and D). It was surprising that this non-phosphorylatable form could adopt the phosphorylated conformation but a close view of the active centre revealed the presence of a sulphate ion that was occupying the same position than the BeF<sub>3</sub><sup>-</sup> ion in the wild-type structure (Fig. 5). The crystals were obtained at high concentrations of sulfate (1.6M) and it is well known that sulphates can mimic phosphate ions<sup>37</sup>. Indeed, this sulphate established interactions with the conserved catalytic residues Lys104, Thr82 and the main chain amide groups of Ala83 and Arg56 as BeF<sub>3</sub><sup>-</sup> did in the wild-type structure (Fig. 5). No interactions with the conserved Asp9 were observed due to the absence of a magnesium ion in the active site. The residues Thr82 and Tyr101 forming the Y-T switch present the inward conformation as in the MaeR<sub>REC</sub>-BeF structure, confirming an activated state for MaeR<sub>REC</sub>-D54A (Fig. 5B). The fact that the non-phosphorylatable mutant also adopts a dimeric swapped conformation indicates that this conformation is not induced upon phosphorylation. On the contrary, it supports that the swapping by itself generates the conformation competent for phosphorylation where the active site displays an organization able to accept the phosphoryl group. This notion is further supported by the coordination of a sulphate ion in the absence of a phosphorylatable Asp in the MaeR<sub>REC</sub>-D54A crystals.

**Asp54 participates in MaeR activation and its phosphorylation enhances binding to the *mae* promoter.** Previous genetic evidence indicated that MaeR acts as a transcriptional activator of the expression of *mae* catabolic genes in *L. casei*<sup>18</sup>. In order to test *in vivo* whether MaeR activity was regulated via phosphorylation of Asp54, two mutant *L. casei* strains in which this residue was replaced by Ala (D54A) or Asn (D54N) were constructed (strains MaeRDA and MaeRDN, respectively). These two mutant strains lost the capacity to grow on L-malate, (Fig. 6A) confirming the importance of Asp54, presumably through its phosphorylation, for MaeR activity *in vivo*.

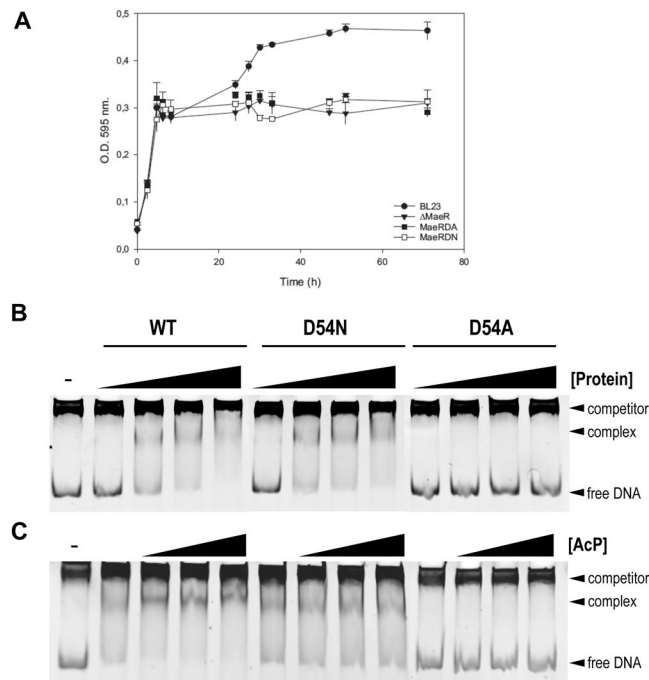
We also studied the *in vitro* binding activity of MaeR<sub>FL</sub> and its derivative D54A and D54N variants to the intergenic region containing the *mae* promoter as it was previously reported<sup>18</sup> (Fig. 6B). Despite the low *in vitro* phosphorylation observed for MaeR with acetylP, the presence of this compound enhanced the binding activity of the wild-type protein to the *mae* promoter region (Fig. 6C). This enhancement was not observed when other possible phosphoryl donors such as ATP or phosphoenolpyruvate were used in binding reactions (data not shown). EMSA assays showed that the D54A mutant had completely lost the capacity to bind the *mae* promoter. On the contrary, binding of the D54N variant was still observed, although, as expected, its binding was not enhanced in the presence of acetylP (Fig. 6C).

**The three direct repeats in the *mae* promoter region are involved in *L. casei* MaeR recognition.** MaeR binding to the *mae* promoter protects a DNA stretch that contains three tandem repeats of the consensus sequence 5'-TTATT(A/T)AA-3'<sup>18</sup>. Similar repeats are present in the *mae* promoters of other LAB carrying homologous *mae* clusters, although differences in their spacing are found<sup>18</sup>. In order to study the unusual binding of a RR to three repeats in a promoter region, we performed mutations of the central TT pair in any of the two first sites, which are separated by 1 bp. These mutations resulted in diminished MaeR binding activity, which was mostly detected in the presence of acetylP (Fig. 7A and B). Specifically, compared to the wild-type sequence, mutation of two (or the three) sites resulted in a complete loss of MaeR binding (Fig. 7A). However, the mutation of just the distal third site did not produce a drastic effect in binding, although the pattern of DNA retardation changed (Fig. 7B and C). These results confirmed that phosphorylation of MaeR increased DNA binding affinity and that the studied DNA repetitions were the target for MaeR recognition. Albeit, they suggested a major contribution of the two first sites in MaeR recognition.

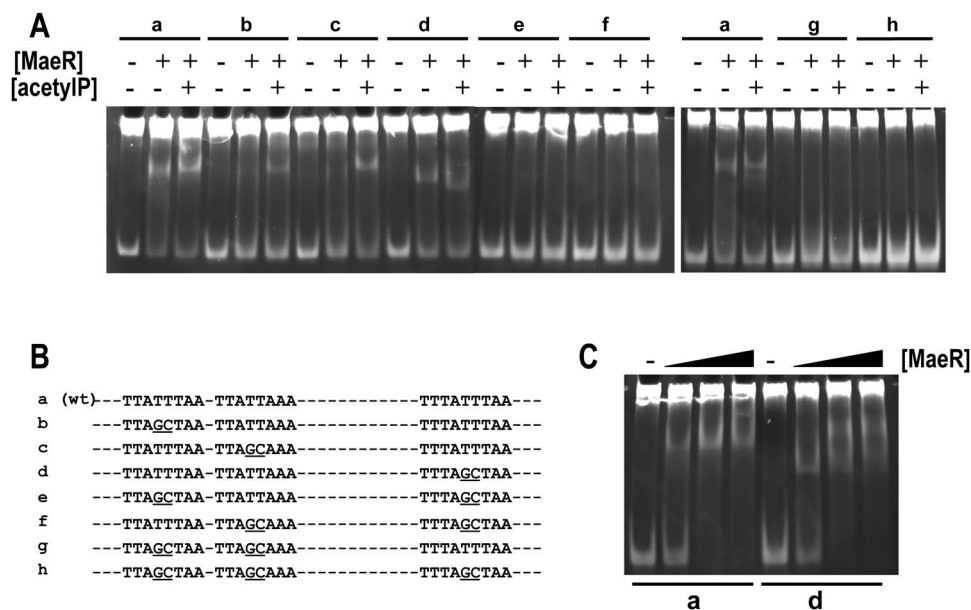
## Discussion

Structural and phylogenetic analyses show that *L. casei* MaeKR proteins belong to the citrate family of TCS, as well as the characterized DcuSR and CitAB TCS. These systems are usually involved in the regulation of the utilization of dicarboxylic and tricarboxylic acids<sup>18, 38–43</sup>. Notwithstanding, there are remarkable differences between the Mae1 group and other previously characterized TCS, most notably, the presence of a degenerated CA domain in the MaeK-like sensor proteins, which precludes MaeK ATP-dependent autophosphorylation. Furthermore, in characterized sensor kinases of this family, a cognate transporter usually functions as a co-sensor<sup>23</sup>, but apparently this is not the case for MaeK since the absence of L-malate transporters did not affect the induction of *mae* genes<sup>19</sup>.

The phylogenetic analyses carried out here and in a previous study<sup>18</sup> have shown that Mae1 group homologs are most closely related to cognate TCS of *Bacillaceae* involved in the regulation of L-malate utilization. The analyses also indicate that the *mae* gene clusters present in *Lactobacillales* proceed from an ancestral cluster that was assembled or rearranged in *Lactobacillales*. The phylogenetic relationships between the genes of this group and the gene arrangement are conserved thus indicating that the Mae1 gene cluster has evolved as a unit. This observation also suggests that the peculiarities of the Mae1 TCS evolved as an adaptation to the genetic and



**Figure 6.** Effect of MaeR mutants on growth and DNA binding. **(A)** Growth of wild-type *L. casei* and derivative strains expressing D54A (MaeRDA) and D54N (MaeRDN) MaeR mutants in MEI medium supplemented with 5 g liter<sup>-1</sup> of L-malic acid. Values represent the means of three independent experiments; error bars represent standard deviations. **(B)** Binding of MaeR and its derived mutants in Asp54 to the *mae* promoter region. Reactions were carried out in 15  $\mu$ l of buffer containing 100 ng of the target DNA (*mae* promoter) and 62, 125, 250 and 375 ng of protein; (-) indicates no protein addition. **(C)** Effect of acetylP on MaeR binding. The *mae* promoter fragment was incubated with 125 ng of protein and increasing concentrations of acetylP (0, 1, 5 and 10 mM). Full-length gels are shown in Supplementary Fig. 11



**Figure 7.** The effect of mutations in the direct repeats of the *mae* promoter on MaeR binding. **(A)** Different *mae* promoter fragments (80 bp; a, wild-type promoter; b to h, mutant forms shown in **(B)**) were incubated with MaeR (125 ng) in the presence (+) or absence (-) of 10 mM acetylP. **(C)** Binding of MaeR to a fragment of the *mae* promoter: wild-type (fragment a) and a mutant in the third repeated sequence (fragment d). 125, 250 and 375 ng of MaeR were added and the binding was performed in the presence of 10 mM acetylP. Full-length gels are shown in Supplementary Fig. 11

physiological environment of an ancestor of *Lactobacillales* and has been conserved during the evolutionary history of this gene cluster. Therefore, the mechanism of signal transduction is possibly identical in all MaeI TCS.

MaeK and MaeR exert their function by activating the expression of L-malate catabolic genes (*maeP* and *maeE*) and the expression of their own genes upon the presence of L-malate in the medium<sup>18,19</sup>. Replacement of the phosphorylatable Asp residue in MaeR resulted in the loss of the ability to grow on L-malate of the corresponding *L. casei* mutant strains. Furthermore, phosphorylation of MaeR with acetylP, albeit low, enhanced MaeR DNA-binding activity *in vitro*, which suggests that MaeR~P has higher affinity for its cognate DNA binding sites than unphosphorylated MaeR. EMSA experiments also showed that the D54N MaeR mutant was not sensitive to acetylP (Fig. 6C). Taken together, these results strongly suggest that phosphorylation of MaeR at Asp 54 is required to function *in vivo*.

The structure of the MaeR REC domain has provided some clues on the phosphorylation pathway of this RR. The structure revealed that dimerization of MaeR involves the swapping of structural elements between the REC domains of the MaeR monomers. The analysis of the structural constraints indicated that the swapped form of MaeR is competent for accepting a phosphate at Asp 54. Since MaeR was predominantly present as a monomer *in vitro*, the low number of MaeR dimers would explain the low level of MaeR~P observed during phosphorylation assays. The possibility to catch in the crystal structure this reduced pool of swapped MaeR molecules is explained by the dynamics of the crystallization process. Although the equilibrium between the monomer and the swapped dimer is displaced toward the monomer in solution, the crystallization conditions, including the presence of BeF<sub>3</sub><sup>-</sup>, can stabilize the swapped form, reaching enough concentration to nucleate. Once this solid nucleus has been formed, it would work as a protein sink shifting the equilibrium in the solution towards the swapped form that is added to the solid crystal. The fact that identical  $\alpha 4\beta 5\alpha 5$ -swapping is present in the activated forms of RegX<sub>3FL</sub> and MaeR<sub>REC</sub> supports that this type of dimerization represents the paradigm of an alternative mechanism of RR activation. We anticipate that this activation mechanism by domain swapping of  $\alpha 4\beta 5\alpha 5$  elements will be found in more RR. Indeed, the domain swapping has been observed in the other RRs as in the bacteriophytochrome RRs RtBRR (*Ramlibacter tataouinensis*; PDB 5IC5) and AtBRR (*Agrobacterium tumefaciens*; 5BRJ)<sup>44</sup> or the master regulator for sporulation Spo0A (*Bacillus stearothermophilus*; PDB 1DZ3)<sup>45</sup>, although in these cases different structural elements were swapped, suggesting that variations of this mechanism could be common among RRs.

The low proportion of MaeR~P obtained via acetylP possibly accounts for the enhanced DNA-binding activity of MaeR observed in EMSA assays, but, how is MaeR phosphorylated *in vivo* and which is the role of MaeK in this process? The fact that *maeK* mutants in different species do not activate *mae* genes in response to the presence of L-malate<sup>18,22</sup> clearly indicate that MaeK is a functional L-malate sensor. However, it is intriguing how a HK devoid of its ATP-dependent kinase activity can operate in signalling. One possibility is that MaeK could promote MaeR swapping and dimerization in response to L-malate, fostering the phosphoacceptor competent conformation of this RR. A comparable situation has been proposed for DesKR TCS, which regulates genes involved in membrane fluidity in *Bacillus subtilis*<sup>46</sup>. In this system, the  $\alpha 1\alpha 5$  surface in the RR DesR is allosterically coupled to homodimerization, favouring the RR autophosphorylation. This surface is also mediating the association of the HK DesK to DesR and it participates in a process of RR activation in the absence of phosphorylation<sup>46</sup>. The possibility of the involvement of MaeK in allosteric regulation of MaeR is also suggested by studies on RegX3 (the other RR with an identical swapped conformation), which requires the presence of the HK SenX3 in order to be phosphorylated and to bind to its target promoters<sup>47</sup>. However, these examples differ from MaeK in the fact that DesK and SenX3 are functional kinases.

If MaeK was participating in a process promoting the phosphoacceptor-competent swapped conformation of MaeR, phosphoryl donors could reach and react effectively with MaeR. However, the nature of these phosphoryl donors and whether MaeK induces the swapped conformation is uncertain. In *S. pyogenes* (MaeI group), it has been described that *ptsI* and *ptsH* mutants, defective in the Enzyme-I and HPr proteins of the phosphoenolpyruvate: sugar phosphotransferase system (PTS), are unable to induce *maePE* genes and have a consequent growth defect on L-malate<sup>22</sup>. The PTS is responsible for taking up carbohydrates with their concomitant phosphorylation by using phosphoenolpyruvate, a glycolytic intermediate, as a phosphoryl donor. It is conceivable that the physiological changes imposed by the lack of a functional PTS will produce changes in the concentration of intracellular metabolites, some of which may be potential phosphoryl donors (e.g. acetylP), that would result in an altered MaeR phosphorylation status. PTS components might also directly phosphorylate or affect the activity of other factor(s) (including MaeK) involved in MaeR phosphorylation. In fact, the Enzyme-I and HPr PTS components have been demonstrated to be involved in many protein regulatory processes<sup>48</sup>. The elucidation of these two extremes deserves further investigation in the *mae* system of *L. casei*.

Our results show that MaeKR form a TCS with a non-canonical signalling mechanism, expanding the variety of mechanisms by which HK transfer the sensed signals to RR. How a HK devoid of an ATP-binding domain performs a role in the phosphorylation of its cognate RR constitutes a new scenario in the field of TCS function. In addition, the domain swapping seems to represent a new paradigm for RR activation that waits to be extended to other TCS. Finally, why the particular characteristics of MaeKR have been developed for L-malate sensing and utilization in *Lactobacillales* and at which point and how MaeKR diverged from its relatives in the TCS citrate family are stimulating questions.

## Methods

**Bacterial strains and growth conditions.** Strains used in this work are listed in Table 2. Lactobacilli were grown in MRS medium (Difco) at 37 °C under static conditions. Growth assays in the presence of L-malic acid were carried out at 30 °C in malic enzyme induction medium (MEI) as previously described<sup>18</sup>. Inoculation was performed with cells grown in MRS for 16 h and washed twice with 1 volume of sterile distilled water. Growth

was monitored by measuring O.D. at 595 nm. At least three independent replicates of each growth curve were obtained. Results were expressed as averages and plus and minus standard deviations.

*E. coli* strains were grown in LB medium at 37 °C under agitation (200 rpm). When required antibiotic concentration used were 5 µg/ml erythromycin for *Lactobacillus*, 100 µg/ml ampicillin, 50 µg/ml kanamycin and 20 µg/ml chloramphenicol for *E. coli*.

**Deletion of MaeK C-terminal domain.** In order to obtain a BL23 derivative strain harbouring a deletion of the C-t domain of *maeK*, flanking fragments of the region to be deleted were amplified using the primer sets MaeK-Ct-F1/MaeK-Ct-R1 and MaeK-Ct-F2/MaeK-Ct-R2, respectively (Supplementary Table 2). As *maeK* 3'-end overlaps the *maeR* 5'-end, a stop codon was introduced after the amino-acid 423, and a region upstream *maeR* encompassing the RBS was conserved (Supplementary Fig. 10). The corresponding pairs of PCR fragments were combined into one fragment by PCR, digested with SpeI/XhoI, and cloned in pRV300 to generate the plasmid pRVDmaeK-Ct. This plasmid was transformed by electroporation in *L. casei* BL23 as described previously<sup>18</sup> and one erythromycin-resistant clone carrying the plasmid integrated by a single crossover was grown in MRS without erythromycin for ~200 generations, antibiotic-sensitive clones were isolated and, among them, one was selected in which a second recombination event led to the deletion of the C-terminal domain of *maeK*, as subsequently confirmed by sequencing of a PCR-amplified fragment spanning the deleted region. The construction resulted in a derivative strain (MCt) encoding a truncated MaeK protein (1–423 aa; Supplementary Fig. 10). Complementation of the deletion of the C-terminal domain of *maeK* was achieved by transforming the plasmid pT1maeK which harbours a *maeK* gene under control of its own promoter<sup>18</sup> into strain MCt.

**Replacement of the MaeR phosphorylatable Asp54 residue by Ala or Asn.** Two BL23 derivative strains where the phosphorylatable amino acid of MaeR (D54) was replaced by Ala or Asn were obtained. To this end, a 1.2 kb region encompassing the entire *maeR* gene was amplified by PCR with primers CitbUp and CitbDown. The PCR product was digested with XhoI and SacI and cloned into pRV300 digested with the same enzymes. This plasmid was used as a template for inverse PCR reactions with oligonucleotide pairs Mut-citb1/Mut-citb2 and Mut-citb3/Mut-citb4, respectively. The amplified products were treated with DpnI and used to transform *E. coli* DH5α. The presence of the required mutations was confirmed by sequencing and selected plasmids (pRVmaeRDA and pRVmaeRDN) were transformed into *L. casei* BL23 by electroporation. Strains with chromosomally integrated plasmids were selected by erythromycin resistance and the second recombination event leading to the replacement of wild type *maeR* by the mutated alleles (strains MDA and MDN) was achieved as described for the MaeKΔC-t strain.

**Expression plasmids.** A *maeR* fragment encoding the receiver (REC) domain (1 to 122) was amplified by PCR with oligonucleotides maeRR-F/maeRR-R. The PCR product was digested with BamHI/SacI and cloned in pQE80 digested with the same enzymes, giving pQEmaeRR. Expression plasmids for purification of mutated variants of MaeR in D54 were constructed by inverse PCR with Mut-citb1/Mut-citb2 and Mut-citb3/Mut-citb4 oligonucleotide pairs as described for the integrative plasmids using pQEmaeR (expressing full-length MaeR<sup>18</sup>), and pQEmaeRR as templates. The cytoplasmic part of MaeK protein, MaeK<sub>C</sub>, were cloned in pNIC28-Bsa4 vector by ligase-independent cloning (LIC) using DNA fragments amplified with the oligonucleotide pairs MaeK<sub>C</sub>-pNIC28/MaeK<sub>C</sub>-pNIC28.

**Protein expression and purification.** All proteins with the exception of MaeR full length protein and MaeR<sub>REC</sub> were expressed in the BL21-codonplus(DE3)-RIL strain using 1 mM IPTG at 20 °C overnight. Proteins were purified by affinity chromatography in a 5-ml HisTrap FF column (GE Healthcare) equilibrated with buffer A (20 mM Tris-HCl pH 7.5, 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 500 mM NaCl), washed with buffer A plus 30 mM of imidazol and eluted with buffer A plus 300 mM of imidazol. MaeR full length protein and the MaeR<sub>REC</sub> domain protein were expressed in DH5α strain using 1 mM IPTG at 20 °C overnight. The purification of proteins was performed using 1 ml HisTrap FF columns (GE Healthcare) equilibrated with buffer C (20 mM Bis-Tris pH 6, 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 500 mM NaCl, 10% glycerol), washed with buffer C plus 30 mM of imidazol and eluted with buffer C plus 300 mM of imidazol. The MaeR mutants used (D54A and D54N) were expressed and purified as wild type MaeR proteins. All the proteins used for *in vitro* assays were loaded in a gel filtration column, Superdex75 16/60 (GE Healthcare) for MaeR proteins and Superdex200 16/60 (GE Healthcare) for MaeK<sub>C</sub> protein. The proteins were concentrated around 10 mgml<sup>-1</sup> with a 10 K Amicon Ultra Centrifuge Filters (Millipore) and stored at –80 °C.

**Protein crystallization, data acquisition and processing.** Crystals of wild-type MaeR<sub>REC</sub> and the D54A variant were obtained using the vapor diffusion method and using the sitting drop technique in the Crystallogensis facility of IBV. Drops of wild-type MaeR<sub>REC</sub> and the D54A variant were set up by mixing 0.4 µl of protein solution at 4.5 mgml<sup>-1</sup> and 15 mgml<sup>-1</sup>, respectively, in buffer 50 mM Bis-Tris pH 6.0 and 200 mM ammonium sulfate with 0.4 µl of different reservoir solutions. Wild-type MaeR<sub>REC</sub> crystals were obtained in 1.2 M magnesium sulfate and 0.1 M MES pH 6.5 while D54A crystals were obtained in 1 M lithium sulfate, 0.5 M ammonium sulfate and 0.1 M sodium citrate. For cryoprotection, wild type crystals were added to a solution containing the reservoir solution at 0.8 M of magnesium sulfate plus 15% ethylenglycol and 25% sucrose while D54A crystals were soaked in the same reservoir solution plus 13% ethylenglycol, 7.3% sucrose and 3% PEG 6000. Diffraction quality was tested from single crystals at 100 K on European Synchrotron Facility (ESRF, Grenoble, France), Diamond Light Source Synchrotron (DLS, Didcot, UK) and Spanish Synchrotron Radiation Facility ALBA (ALBA, Cerdanyola del Vallès, Spain). The complete diffraction data set was collected at the ESRF in a nitrogen-gas stream at ID29 at 2.8 Å and 3.1 Å for wild type and D54A mutant, respectively. Higher resolution at 2.1 Å was obtained for wild type crystals diffracted at I04 at DLS one year later after first crystals appeared in the plate. Crystallographic data was processed and scaled with the programs MOSFLM and SCALA from



Strain or plasmid	Characteristics or relevant genotype	Source or reference
<i>Escherichia coli</i>		
DH5 $\alpha$	F <sup>-</sup> <i>endA1 hsdR17 gyrA96 thi-1 recA1 relA1 supE44 <math>\Delta</math>lacU169 (F80 lacZ <math>\Delta</math>M15)</i>	Stratagene
BL21-(DE3)-RIL	F <sup>-</sup> <i>ompT hsdS(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) dcm<sup>+</sup> Tet<sup>r</sup> gal endA <math>\lambda</math>(DE3) Hte [argU ileY leuW Cam<sup>r</sup>]</i>	Agilent
<i>Lactobacillus casei</i>		
BL23	wild type strain	65
BL315	BL23 $\Delta$ <i>maeR</i>	18
BL322	BL23 $\Delta$ <i>maeK</i>	18
MaeK $\Delta$ C-t	BL23 harboring a deletion of the C-terminal domain of <i>maeK</i>	This study
MaeK $\Delta$ C-t + MaeK	MaeK $\Delta$ C-t harbouring plasmid pT1maeK	This study
MaeRDA	BL23 <i>maeRD54A</i>	This study
MaeRDN	BL23 <i>maeRD45N</i>	This study
plasmids		
pQE80	<i>cis</i> -repressed, IPTG inducible, N-terminal His <sub>6</sub> -tagged recombinant protein expression vector; Amp <sup>r</sup>	Qiagen
pQEmaeR	pQE80 with cloned <i>maeR</i> gene	18
pQEmaeRDA	pQE80 with cloned <i>maeRD54A</i> gene	This study
pQEmaeRDN	pQE80 with cloned <i>maeRD54N</i> gene	This study
pQEmaeRR	pQE80 with cloned receiver (REC) domain of <i>maeR</i> (amino acids 1 to 121)	This study
pQEmaeRRDA	pQEmaeRR with cloned receiver (REC) domain of <i>maeRD54A</i> (amino acids 1 to 121)	This study
pQEmaeRRDN	pQEmaeRR with cloned receiver (REC) domain of <i>maeRD54N</i> (amino acids 1 to 121)	This study
pNIC28-Bsa4	LIC vector. IPTG inducible, N-terminal His <sub>6</sub> -tagged recombinant protein expression; Km <sup>r</sup>	66
pNIC-MaeK <sub>C</sub>	pNIC28-Bsa4 with cloned cytoplasmic domain of MaeK	This study
pRV300	Insertional vector for <i>Lactobacillus</i> , Amp <sup>r</sup> , Ery <sup>r</sup>	67
pRVDmaeK-Ct	pRV300 containing a 2 kb with a deletion of the C-terminal domain of <i>maeK</i> cloned at SpeI and XhoI sites	This study
pRVmaeRDA	pRV300 with a 1.2 kb fragment cloned at XhoI and SacI containing <i>maeR</i> with a mutated codon 54 (GAT to GCT)	This study
pRVmaeRDN	pRV300 with a 1.2 kb fragment cloned at XhoI and SacI containing <i>maeR</i> with a mutated codon 54 (GAT to AAT)	This study
pT1maeK	pT1NX with cloned <i>maeK</i> expressed from its own promoter. Ery <sup>r</sup>	18

**Table 2.** Strains and plasmids used in this study.

the Collaborative Computational Project, Number 4, CCP4 suite<sup>49</sup> or the XDS Program Package<sup>50</sup>. Details of the data-collection are shown in Table 1. The structure of wild-type MaeR<sub>REC</sub> was determined by molecular replacement using the program MrBUMP<sup>51</sup> and the structure of the D54A mutant was obtained by molecular replacement with the program Phaser<sup>52</sup> using the refined structure of the wild-type protein. The final models were obtained using a combination of refinement cycles with the program Refmac5<sup>53</sup> and manual building with the program Coot<sup>54</sup>. Details of the refinement are shown in Table 1. Structural analysis of the structures was performed using other programs such as Ncont and Superpose contained in CCP4 suite<sup>49</sup>. Figures were produced using PyMOL (<http://www.pymol.org>).

**Autophosphorylation assays.** The MaeK<sub>C</sub> autophosphorylation assays were performed with [ $\gamma$ -<sup>32</sup>P] ATP (3000 Ci/mmol Perkin Elmer) at concentrations of 0.5 mM of cold ATP and 0.1 and 0.5  $\mu$ Ci/ $\mu$ l of radioactive ATP during 10 and 60 min. Samples were subjected to SDS-PAGE on 15% gel and phosphorylated proteins were visualized by phosphorimaging using a Fluoro Image Analyzer FLA-5000 (Fuji) and evaluated with the MultiGauge software (Fuji).

**Isothermal titration calorimetry.** The binding affinity measurements for MgCl<sub>2</sub> and AMPPNP and the indicated protein were taken with a Nano-ITC (TA Instruments). Experiments were carried out at 25 °C with 0.03 mM of protein in the cell and mixed with 0.7 or 5 mM nucleotide in the syringe by 25 injections of 2  $\mu$ l each at intervals of 180s under continuous stirring. Data integration, correction and analysis were carried out using NanoAnalyze program (TA Instruments) with a single-site binding model.

**Circular Dichroism.** Circular Dichroism spectroscopy (CD) experiments were performed using a JASCO J-810 spectrometer that was coupled to a Peltier temperature controller (PTC-423S). Protein solutions in 50 mM phosphate buffer, pH 7.5, were adjusted to a protein concentration of 10  $\mu$ M and then placed in a round cuvette Helma with a 0.1 cm path length. Data were collected over a wavelength range of 190–250 nm in milliabsorbance units at 25 °C, averaged and background corrected using the Spectra Manager software provided with the instrument. These data were transformed to ellipticity  $\theta$  using the relationship [ $\theta$  = milligrades  $\times$  10lCp] where l is the path length, C is the concentration of the protein (M) and p is the number of peptidic bonds. The analysis was done using DICHROWEB<sup>55</sup> and VADAR<sup>56</sup>.

**Synthesis of radioactive acetylP.** The radioactive [ $^{32}\text{P}$ ]acetylP used in the MaeR phosphorylation was obtained incubating for 2 h at room temperature 1.5U of acetate kinase with 100  $\mu\text{Ci}/\mu\text{l}$  of [ $\gamma\text{-}^{32}\text{P}$ ] ATP (3000 Ci/mmol Perkin Elmer) in 2.5 mM Tris pH 8, 6 mM potassium acetate, 1 mM  $\text{Cl}_2\text{Mg}$  buffer. The [ $^{32}\text{P}$ ]-acetylP was freed from acetate kinase by filtering using Microcon-10 kDa Centrifugal Filter Unit (GE Healthcare). The pure [ $^{32}\text{P}$ ]-acetylP was stored at  $-20^\circ\text{C}$ .

**MaeR phosphorylation assays.** MaeR was phosphorylated incubating 1  $\text{mg ml}^{-1}$  of the protein at room temperature for the indicated incubation times with 12.5 mM [ $^{32}\text{P}$ ]-acetylP in a solution containing 50 mM Tris-HCl pH 8.0, 500 mM KCl, 5 mM  $\text{MgCl}_2$ , 20 mM DTT and 1 mM EDTA. The samples were subjected to SDS-PAGE on 15% gel and phosphorylated proteins were visualized by phosphorimaging using a Fluoro Image Analyzer FLA-5000 (Fuji) and evaluated with the MultiGauge software (Fuji).

**Phosphatase assays.** The MaeR<sub>REC</sub> protein phosphorylated for 1 h as described in the MaeR phosphorylation assays above, was incubated with equimolar concentrations of MaeK<sub>C</sub> at room temperature for the indicated incubation times and the samples were subjected to SDS-PAGE on 15% gel. Phosphorylated proteins were visualized by phosphorimaging using a Fluoro Image Analyzer FLA-5000 (Fuji) and evaluated with the MultiGauge software (Fuji).

**Electrophoretic mobility assays.** A 219 bp fragment spanning the *maeP-maeK* intergenic region was amplified by PCR with the oligonucleotide pair FP1/FP2. Different combinations of mutations in the three putative MaeR binding sites (5'-TTATT(A/T)AA-3' direct repeats) in which the central TT pair was replaced by GC were obtained by PCR using oligonucleotides carrying the desired mutations that amplified a 80 bp fragment. The combinations of the different oligonucleotides (see Supplementary Table S2) were: FP5/FP6 (wild type sequence), FP8mut2/FP6 (mutant at site 1), FP9mut3/FP6 (mutant at site 2), FP5/FP7mut1 (mutant at site 3), FP8mut2/FP7mut1 (mutant at sites 1 and 3), FP9mut3/FP7mut1 (mutant at sites 2 and 3), FP10mut4/FP6 (mutant at sites 1 and 2), FP10mut4/FP7mut1 (mutant at the three sites). EMSA assays using these fragments were carried out in 15  $\mu\text{l}$  of 2.5 mM Tris-HCl pH 7.5, 100 mM NaCl, 25 mM  $\text{MgCl}_2$ , 0.25 mM EDTA, 0.25 mM dithiothreitol, 10% glycerol containing 1.5  $\mu\text{g}$  of salmon sperm DNA, 100 ng of each fragment and varying amounts of MaeR or its variants MaeR-D54A and MaeR-D54N (62 to 375 ng) with or without acetylP (1 to 10 mM). These mixes were incubated at  $37^\circ\text{C}$  for 30 min before being resolved in 6% polyacrylamide (19:1) gels with 1X TAE (40 mM Tris-acetate pH 7.6, 1 mM EDTA) as running buffer at 100 V. The gels were stained with ethidium bromide.

**Phylogenetic analysis.** Homologous sequences of *mae* genes (*maeK*, *maeR*, *maeP* and *maeE*) present in species belonging to *Lactobacillales* were identified and retrieved from complete genomes deposited at MBGD<sup>57</sup>. Redundant sequences were excluded from the analysis. Outgroup sequences were selected on the basis of our previous phylogenetic analysis<sup>18</sup>. Sequences were aligned at the T-Coffee server (<http://tcoffee.crg.cat>) with the M-Coffee tool and default options<sup>58</sup>. Gaps and low homology regions were removed using Gblocks<sup>59</sup>. Details on sequences used and alignments are available from the authors upon request.

The best-fit models of amino acid substitution were selected using the ProtTest 2.4 server ([http://darwin.uvigo.es/software/prottest2\\_server.html](http://darwin.uvigo.es/software/prottest2_server.html))<sup>60</sup>. The Akaike information criterion (AIC) was adopted to select the best model. The models selected were used to estimate the corresponding phylogenies by maximum likelihood using the PhyML tool<sup>61</sup> implemented in the T-Rex server (<http://www.trex.uqam.ca>)<sup>62</sup>. Bootstrap support values were obtained from 1000 pseudoreplicates. Phylogenetic information content in the sequence alignments was analyzed by likelihood mapping as implemented in TreePuzzle 5.3rc<sup>63</sup>. Congruence among topologies for *mae* gene trees was evaluated using the tests implemented in TreePuzzle 5.3rc<sup>27</sup>. Phylogenetic trees were visualized and arbitrarily rooted using the mid-point rooting method implemented in Mega software<sup>64</sup>.

## References

- Gao, R. & Stock, A. M. Biological insights from structures of two-component proteins. *Annual review of microbiology* **63**, 133–154 (2009).
- Casino, P., Miguel-Romero, L. & Marina, A. Visualizing autophosphorylation in histidine kinases. *Nat Commun* **5**, 3258, doi:10.1038/ncomms4258 (2014).
- Igo, M. M., Ninfa, A. J., Stock, J. B. & Silhavy, T. J. Phosphorylation and dephosphorylation of a bacterial transcriptional activator by a transmembrane receptor. *Genes Dev* **3**, 1725–1734 (1989).
- Huynh, T. N. & Stewart, V. Negative control in two-component signal transduction by transmitter phosphatase activity. *Mol Microbiol* **82**, 275–286, doi:10.1111/j.1365-2958.2011.07829.x (2011).
- Casino, P., Rubio, V. & Marina, A. The mechanism of signal transduction by two-component systems. *Curr Opin Struct Biol* **20**, 763–771, doi:10.1016/j.sbi.2010.09.010 (2010).
- Zschiedrich, C. P., Keidel, V. & Szurmant, H. Molecular Mechanisms of Two-Component Signal Transduction. *J Mol Biol*, doi:10.1016/j.jmb.2016.08.003 (2016).
- Dutta, R. & Inouye, M. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci* **25**, 24–28 (2000).
- Bourret, R. B. Receiver domain structure and function in response regulator proteins. *Curr Opin Microbiol* **13**, 142–149, doi:10.1016/j.mib.2010.01.015 (2010).
- Cho, H. S. *et al.* NMR structure of activated CheY. *J Mol Biol* **297**, 543–551, doi:10.1006/jmbi.2000.3595 (2000).
- Villali, J., Pontiggia, F., Clarkson, M. W., Hagan, M. F. & Kern, D. Evidence against the “Y-T coupling” mechanism of activation in the response regulator NtrC. *J Mol Biol* **426**, 1554–1567, doi:10.1016/j.jmb.2013.12.027 (2014).
- Dintner, S., Heermann, R., Fang, C., Jung, K. & Gebhard, S. A sensory complex consisting of an ATP-binding cassette transporter and a two-component regulatory system controls bacitracin resistance in *Bacillus subtilis*. *J Biol Chem* **289**, 27899–27910, doi:10.1074/jbc.M114.596221 (2014).
- Tetsch, L. & Jung, K. The regulatory interplay between membrane-integrated sensors and transport proteins in bacteria. *Mol Microbiol* **73**, 982–991, doi:10.1111/j.1365-2958.2009.06847.x (2014).

13. Zakikhany, K., Harrington, C. R., Nimtz, M., Hinton, J. C. & Romling, U. Unphosphorylated CsgD controls biofilm formation in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **77**, 771–786, doi:10.1111/j.1365-2958.2010.07247.x (2009).
14. Pioszak, A. A. & Ninfa, A. J. Mutations altering the N-terminal receiver domain of NRI (NtrC) That prevent dephosphorylation by the NRII-PII complex in *Escherichia coli*. *J Bacteriol* **186**, 5730–5740, doi:10.1128/JB.186.17.5730-5740.2004 (2010).
15. Sakayori, T., Shiraiwa, Y. & Suzuki, I. A Synechocystis homolog of SipA protein, Ssl3451, enhances the activity of the histidine kinase Hik33. *Plant Cell Physiol* **50**, 1439–1448, doi:10.1093/pcp/pcp089 (2009).
16. Scheu, P. D. *et al.* Oligomeric sensor kinase DcuS in the membrane of *Escherichia coli* and in proteoliposomes: chemical cross-linking and FRET spectroscopy. *J Bacteriol* **192**, 3474–3483, doi:10.1128/JB.00082-10 (2010).
17. Schutz, M. & Radler, F. [The presence of malic enzyme and malo-lactic enzyme in various lactic acid bacteria (author's transl)]. *Arch Mikrobiol* **96**, 329–339 (1974).
18. Landete, J. M. *et al.* Requirement of the *Lactobacillus casei* MaeKR two-component system for L-malic acid utilization via a malic enzyme pathway. *Appl Environ Microbiol* **76**, 84–95, doi:10.1128/AEM.02145-09 (2010).
19. Landete, J. M., Ferrer, S., Monedero, V. & Zuniga, M. Malic enzyme and malolactate enzyme pathways are functionally linked but independently regulated in *Lactobacillus casei* BL23. *Appl Environ Microbiol* **79**, 5509–5518, doi:10.1128/AEM.01177-13 (2013).
20. Espariz, M. *et al.* Identification of malic and soluble oxaloacetate decarboxylase enzymes in *Enterococcus faecalis*. *FEBS J* **278**, 2140–2151, doi:10.1111/j.1742-4658.2011.08131.x (2011).
21. Mortera, P. *et al.* Fine-tuned transcriptional regulation of malate operons in *Enterococcus faecalis*. *Appl Environ Microbiol* **78**, 1936–1945, doi:10.1128/AEM.07280-11 (2012).
22. Paluscio, E. & Caparon, M. G. *Streptococcus pyogenes* malate degradation pathway links pH regulation and virulence. *Infect Immun* **83**, 1162–1171, doi:10.1128/IAI.02814-14 (2015).
23. Unden, G., Wormer, S. & Monzel, C. Cooperation of Secondary Transporters and Sensor Kinases in Transmembrane Signalling: The DctA/DcuS and DcuB/DcuS Sensor Complexes of *Escherichia coli*. *Adv Microb Physiol* **68**, 139–167, doi:10.1016/bs.ampbs.2016.02.003 (2016).
24. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7–8, doi:10.1038/nmeth.3213 (2015).
25. Marina, A., Waldburger, C. D. & Hendrickson, W. A. Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. *The EMBO journal* **24**, 4247–4259, doi:10.1038/sj.emboj.7600886 (2005).
26. Huynh, T. N., Noriega, C. E. & Stewart, V. Conserved mechanism for sensor phosphatase control of two-component signaling revealed in the nitrate sensor NarX. *Proc Natl Acad Sci USA* **107**, 21140–21145, doi:10.1073/pnas.1013081107 (2010).
27. Schmidt, H. A., Strimmer, K., Vingron, M. & von Haeseler, A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502–504, doi:10.1093/bioinformatics/18.3.502 (2002).
28. Lukat, G. S., McCleary, W. R., Stock, A. M. & Stock, J. B. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc Natl Acad Sci USA* **89**, 718–722 (1992).
29. Gao, R. & Stock, A. M. Molecular strategies for phosphorylation-mediated regulation of response regulator activity. *Curr Opin Microbiol* **13**, 160–167, doi:10.1016/j.mib.2009.12.009 (2010).
30. Yan, D. *et al.* Beryllofluoride mimics phosphorylation of NtrC and other bacterial response regulators. *Proc Natl Acad Sci USA* **96**, 14789–14794 (1999).
31. Wemmer, D. E. & Kern, D. Beryllofluoride binding mimics phosphorylation of aspartate in response regulators. *J Bacteriol* **187**, 8229–8230, doi:10.1128/JB.187.24.8229 (2005).
32. King-Scott, J. *et al.* The structure of a full-length response regulator from *Mycobacterium tuberculosis* in a stabilized three-dimensional domain-swapped, activated state. *J Biol Chem* **282**, 37717–37729, doi:10.1074/jbc.M705081200 (2007).
33. Taylor, D., Cawley, G. & Hayward, S. Quantitative method for the assignment of hinge and shear mechanism in protein domain movements. *Bioinformatics* **30**, 3189–3196, doi:10.1093/bioinformatics/btu506 (2014).
34. Ahmad, A., Cai, Y., Chen, X., Shuai, J. & Han, A. Conformational Dynamics of Response Regulator RegX3 from *Mycobacterium tuberculosis*. *PLoS One* **10**, e0133389, doi:10.1371/journal.pone.0133389 (2015).
35. Toro-Roman, A., Wu, T. & Stock, A. M. A common dimerization interface in bacterial response regulators KdpE and TorR. *Protein Sci* **14**, 3077–3088, doi:10.1110/ps.051722805 (2005).
36. Bachhawat, P. & Stock, A. M. Crystal structures of the receiver domain of the response regulator PhoP from *Escherichia coli* in the absence and presence of the phosphoryl analog beryllofluoride. *J Bacteriol* **189**, 5987–5995, doi:10.1128/JB.00049-07 (2007).
37. Casino, P., Rubio, V. & Marina, A. Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction. *Cell* **139**, 325–336, doi:10.1016/j.cell.2009.08.032 (2009).
38. Bott, M., Meyer, M. & Dimroth, P. Regulation of anaerobic citrate metabolism in *Klebsiella pneumoniae*. *Molecular Microbiology* **18**, 533–546 (1995).
39. Zientz, E., Bongaerts, J. & Unden, G. Fumarate Regulation of Gene Expression in *Escherichia coli* by the DcuSR (dcuSR Genes) Two-Component Regulatory System. *Journal of Bacteriology* **180**, 5421–5425 (1998).
40. Yamamoto, H., Murata, M. & Sekiguchi, J. The CitST two-component system regulates the expression of the Mg-citrate transporter in *Bacillus subtilis*. *Molecular Microbiology* **37**, 898–912, doi:10.1046/j.1365-2958.2000.02055.x (2000).
41. Asai, K., Baik, S. H., Kasahara, Y., Moriya, S. & Ogasawara, N. Regulation of the transport system for C4-dicarboxylic acids in *Bacillus subtilis*. *Microbiology* **146**, 263–271 (2000).
42. Janausch, I. G., Zientz, E., Tran, Q. H., Kröger, A. & Unden, G. C4-dicarboxylate carriers and sensors in bacteria. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1553**, 39–56, doi:10.1016/S0005-2728(01)00233-X (2002).
43. Tanaka, K., Kobayashi, K. & Ogasawara, N. The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YfLS, and is essential for utilization of malate in minimal medium. *Microbiology* **149**, 2317–2329, doi:10.1099/mic.0.26257-0 (2003).
44. Baker, A. W., Satyshur, K. A., Moreno Morales, N. & Forest, K. T. Arm-in-Arm Response Regulator Dimers Promote Intermolecular Signal Transduction. *J Bacteriol* **198**, 1218–1229, doi:10.1128/JB.00872-15 (2016).
45. Lewis, R. J. *et al.* Domain swapping in the sporulation response regulator Spo0A. *J Mol Biol* **297**, 757–770, doi:10.1006/jmbi.2000.3598 (2000).
46. Trajtenberg, F. *et al.* Allosteric activation of bacterial response regulators: the role of the cognate histidine kinase beyond phosphorylation. *mBio* **5**, e02105, doi:10.1128/mBio.02105-14 (2014).
47. Glover, R. T., Kriakov, J., Garforth, S. J., Baughn, A. D. & Jacobs, W. R. Jr. The two-component regulatory system senX3-regX3 regulates phosphate-dependent gene expression in *Mycobacterium smegmatis*. *J Bacteriol* **189**, 5495–5503, doi:10.1128/JB.00190-07 (2007).
48. Deutscher, J. *et al.* The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiology and Molecular Biology Reviews* **78**, 231–256, doi:10.1128/mmb.00001-14 (2014).
49. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242, doi:10.1107/S0907444910045749 (2011).
50. Kabsch, W. X. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132, doi:10.1107/S0907444909047337 (2010).
51. Keegan, R. M. & Winn, M. D. MrBUMP: an automated pipeline for molecular replacement. *Acta Crystallogr D Biol Crystallogr* **64**, 119–124, doi:10.1107/S0907444907037195 (2008).
52. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674, doi:10.1107/S0021889807021206 (2007).

53. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367, doi:10.1107/S0907444911001314 (2011).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501, doi:10.1107/S0907444910007493 (2010).
55. Lobley, A., Whitmore, L. & Wallace, B. A. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18**, 211–212 (2002).
56. Willard, L. *et al.* VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* **31**, 3316–3319 (2003).
57. Uchiyama, I., Mihara, M., Nishide, H. & Chiba, H. MBGD update 2015: microbial genome database for flexible ortholog analysis utilizing a diverse set of genomic data. *Nucleic Acids Research* **43**, D270–D276, doi:10.1093/nar/gku1152 (2015).
58. Di Tommaso, P. *et al.* T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research* **39**, W13–W17, doi:10.1093/nar/gkr245 (2011).
59. Castresana, J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution* **17**, 540–552 (2000).
60. Abascal, F., Zardoya, R. & Posada, D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105, doi:10.1093/bioinformatics/bti263 (2005).
61. Guindon, S. & Gascuel, O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* **52**, 696–704, doi:10.1080/10635150390235520 (2003).
62. Boc, A., Diallo, A. B. & Makarenkov, V. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Research* **40**, W573–W579, doi:10.1093/nar/gks485 (2012).
63. Strimmer, K. & von Haeseler, A. Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences* **94**, 6815–6819 (1997).
64. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870–1874, doi:10.1093/molbev/msw054 (2016).
65. Maze, A. *et al.* Complete genome sequence of the probiotic *Lactobacillus casei* strain BL23. *J Bacteriol* **192**, 2647–2648, doi:10.1128/JB.00076-10 (2010).
66. Savitsky, P. *et al.* High-throughput production of human proteins for crystallization: the SGC experience. *Journal of structural biology* **172**, 3–13, doi:10.1016/j.jsb.2010.06.008 (2010).
67. Leloup, L., Ehrlich, S. D., Zagorec, M. & Morel-Deville, F. Single-crossover integration in the *Lactobacillus sake* chromosome and insertional inactivation of the *ptsI* and *lacL* genes. *Appl Environ Microbiol* **63**, 2117–2123 (1997).

## Acknowledgements

We thank Local Contact and staffs of beamline I04 in Diamond Light Source (Didcot, UK), beamline ID29 in European Synchrotron Radiation Facility (Grenoble, France) and beamline BL13-XALOC in Spanish Synchrotron Radiation Facility ALBA (Cerdanyola del Vallès, Spain) for providing assistance during data collection. This study was financed by grants AGL2007–60975 and AGL2010–15679 from the former Spanish Ministry of Science and by grant ACOMP2012/137 from Generalitat Valenciana to V.M. and M.Z., by grants BIO2013–42619-P and BIO2016–78571-P from Spanish Ministry of Economy and Competitiveness and PrometeoII/2014/029 from the Valencian Government to A.M. and by grant BFU2016–78606-P from the Spanish Ministry of Economy and Competitiveness to P.C. P.C. is the recipient of a Ramon y Cajal contract, and L.M.-R. is the recipient of a FPI fellow contract, both from the Ministry of Economy and Competitiveness. J.M.L. was the recipient of a Juan de la Cierva contract from the former Spanish Ministry of Science and Innovation. X-ray diffraction data collection was supported by Diamond Light Source block allocation group (BAG) Proposal MX14739 and Spanish Synchrotron Radiation Facility ALBA Proposal 2015071314 and 2016071762. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (grant agreement N°283570).

## Author Contributions

Design of the study: P.C., V.M., M.Z. and A.M.; the acquisition, analysis, and interpretation of the data: L.M.-R., P.C., J.M.L., V.M., M.Z. and A.M.; writing of the manuscript: P.C., V.M., M.Z. and A.M.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-02900-z

**Competing Interests:** The authors declare that they have no competing interests.

**Accession codes:** The structures of Maer<sub>Rec</sub> wild-type and D54A mutant have been deposited in the Protein Data Bank under the accession codes 5LWK and 5LWL, respectively

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017