

Supplementary Material for:  
Phosphorelay in non-orthodox two component  
systems in vivo functions through an bi-molecular  
mechanism: the case of ArcB

February 27, 2015

## Comparison of Chromosomal and Plasmid Expression Levels

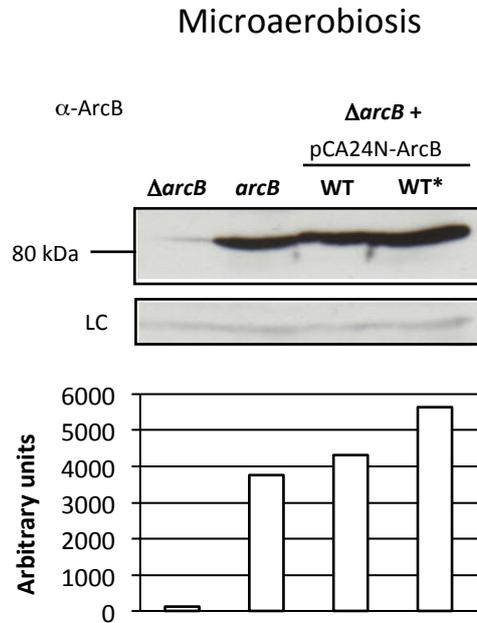


Figure S 1: The expression levels of chromosomal and plasmid-based ArcB. The level of expression of MG1655 chromosomal ArcB and plasmid (pCA24N) borne ArcB (WT) or ArcB\* (WT\*) in a  $\Delta arcB59$  (MVA92) strain grown in microaerobiosis (control was vector pCA24N alone) was assessed using Western blotting and antibodies against ArcB ( $\alpha$ -ArcB). LC, loading control, the protein band from crude cell extract that shows non-specific cross-reaction with the ArcB antibody.

## WT and phosphorelay residue deletion strain activities

### Mathematical Models of Phosphorelay Mechanisms

Here we provide the expressions that are used to analyze the mechanistic models for the ArcB phosphorelay.

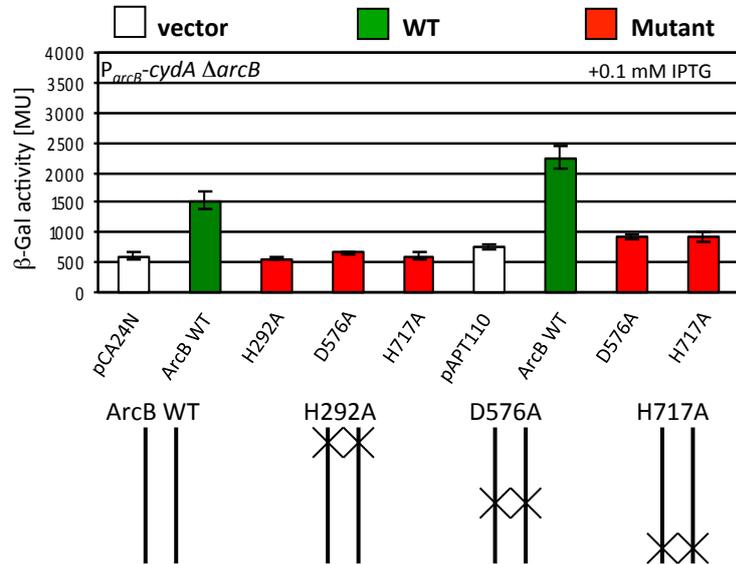
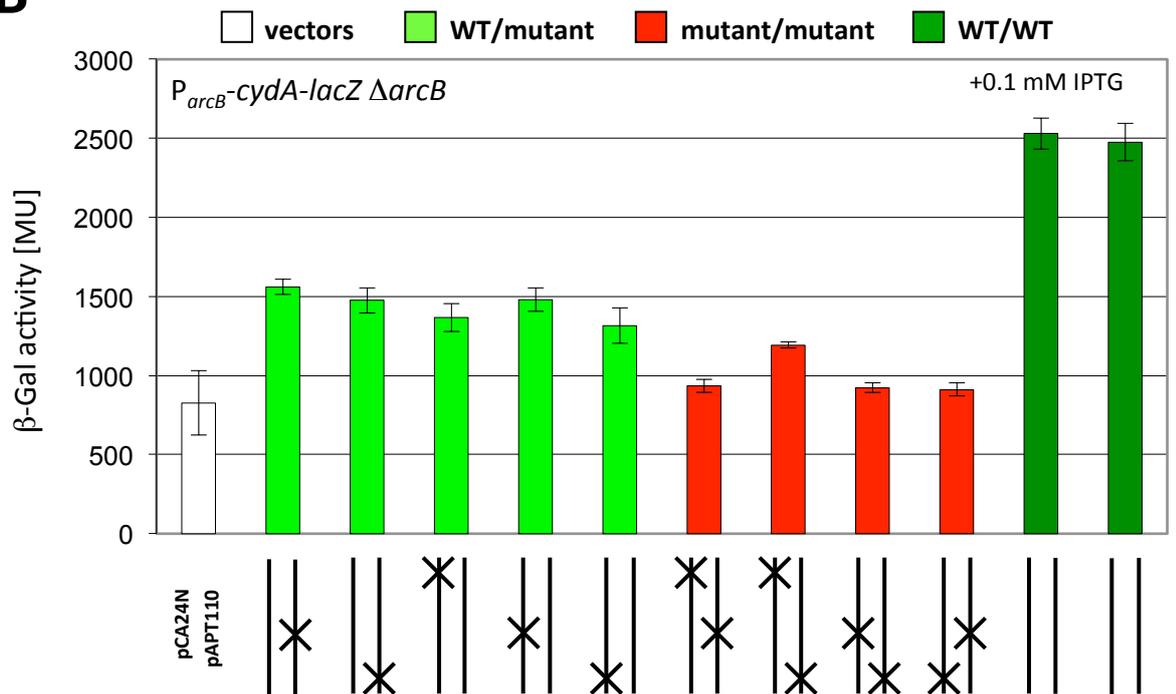
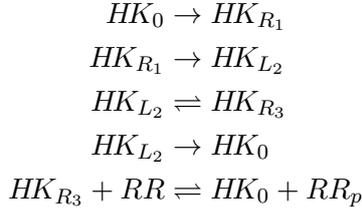
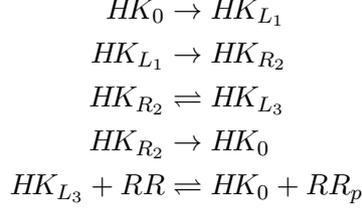
**A****B**

Figure S 2: Phosphorelay of wild type ArcB variants lacking some of the key phosphorelay residues. A) Expression of a *cydA-lacZ* chromosomal transcription fusion was measured using a  $\beta$ -Gal assay in  $\Delta$ *arcB* (MVA104) cells grown in microaerobiosis in the presence of 0.1 mM IPTG co-expressing wild type ArcB or its variants from either pCA24N- (left hand part) or pAPT110-based (right hand part) plasmids (see schematic presentation below graph). B) As in A) except cells expressed different combinations of the ArcB and/or its variants from co-transformed pCA24N- and pAPT110-based plasmids (see schematic presentation below graph; the last two columns represent ArcB WT replicate experiment). For all  $\beta$ -Gal assays shown mean values of six independent assays taken from technical duplicates of three independently grown cultures of each strain were used to calculate activity. The data are shown as a mean values with SD error bars.

## ***Trans*–Phosphorelay Model**

For the *trans*–phosphorelay we have the following reactions

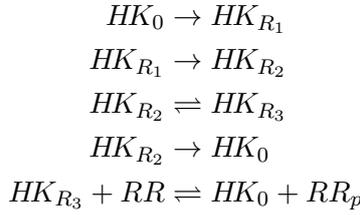
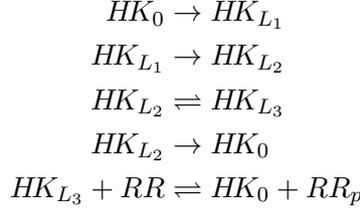


The corresponding ODEs are thus,

$$\begin{aligned}
 d(HK_0)/dt &= -k_{0l1}HK_0 - k_{0r1}HK_0 + k_{l20}HK_{L_2} + k_{r20}HK_{R_2} \\
 &\quad + k_{l30}HK_{L_3}RR + k_{r30}HK_{R_3}RR - k_{0l3}HK_0RR_p \\
 &\quad - k_{0r3}HK_0RR_p + k_{l10}HK_{L_1} + k_{r10}HK_{R_1} \\
 d(HK_{L_1})/dt &= k_{0l1}HK_0 - k_{l10}HK_{L_1} - k_{l1r2}HK_{L_1} \\
 d(HK_{R_1})/dt &= k_{0r1}HK_0 - k_{r10}HK_{R_1} - k_{r1l2}HK_{R_1} \\
 d(HK_{L_2})/dt &= k_{r1l2}HK_{R_1} - k_{l2r3}HK_{L_2} + k_{r3l2}HK_{R_3} - k_{l20}HK_{L_2} \\
 d(HK_{R_2})/dt &= k_{l1r2}HK_{L_1} - k_{r2l3}HK_{R_2} + k_{l3r2}HK_{L_3} - k_{r20}HK_{R_2} \\
 d(HK_{L_3})/dt &= k_{r2l3}HK_{R_2} - k_{l3r2}HK_{L_3} - k_{l30}HK_{L_3}RR + k_{0l3}HK_0RR_p \\
 d(HK_{R_3})/dt &= k_{l2r3}HK_{L_2} - k_{r3l2}HK_{R_3} - k_{r30}HK_{R_3}RR + k_{0r3}HK_0RR_p \\
 d(RR)/dt &= -k_{l30}HK_{L_3}RR - k_{r30}HK_{R_3}RR + k_{0l3}HK_0RR_p + k_{0r3}HK_0RR_p \\
 d(RR_p)/dt &= k_{l30}HK_{L_3}RR + k_{r30}HK_{R_3}RR - k_{0l3}HK_0RR_p - k_{0r3}HK_0RR_p
 \end{aligned}$$

## ***Cis*–Phosphorelay Model**

For the *cis*–phosphorelay model we have:

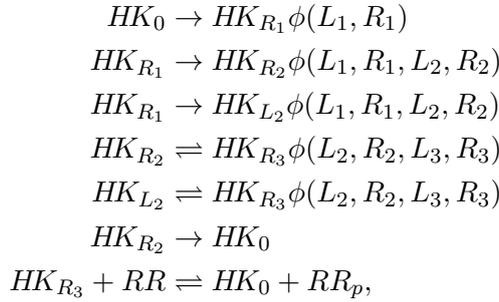
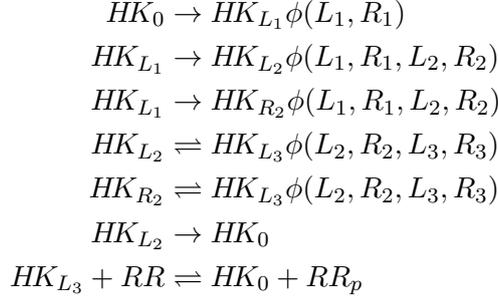


while the corresponding ODEs are

$$\begin{aligned}
 d(HK_0)/dt &= -k_{0l1}HK_0 - k_{0r1}HK_0 + k_{l20}HK_{L_2} + k_{r20}HK_{R_2} \\
 &\quad + k_{l30}HK_{L_3}RR + k_{r30}HK_{R_3}RR - k_{0l3}HK_0RR_p \\
 &\quad - k_{0r3}HK_0RR_p + k_{l10}HK_{L_1} + k_{r10}HK_{R_1} \\
 d(HK_{L_1})/dt &= k_{0l1}HK_0 - k_{l10}HK_{L_1} - k_{l1l2}HK_{L_1} \\
 d(HK_{R_1})/dt &= k_{0r1}HK_0 - k_{r10}HK_{R_1} - k_{r1r2}HK_{R_1} \\
 d(HK_{L_2})/dt &= k_{l1l2}HK_{L_1} - k_{l2l3}HK_{L_2} + k_{l3l2}HK_{L_3} - k_{l20}HK_{L_2} \\
 d(HK_{R_2})/dt &= k_{r1r2}HK_{R_1} - k_{r2r3}HK_{R_2} + k_{r3r2}HK_{R_3} - k_{r20}HK_{R_2} \\
 d(HK_{L_3})/dt &= k_{l2l3}HK_{L_2} - k_{l3l2}HK_{L_3} - k_{l30}HK_{L_3}RR + k_{0l3}HK_0RR_p \\
 d(HK_{R_3})/dt &= k_{r2r3}HK_{R_2} - k_{r3r2}HK_{R_3} - k_{r30}HK_{R_3}RR + k_{0r3}HK_0RR_p \\
 d(RR)/dt &= -k_{l30}HK_{L_3}RR - k_{r30}HK_{R_3}RR + k_{0l3}HK_0RR_p + k_{0r3}HK_0RR_p \\
 d(RR_p)/dt &= k_{l30}HK_{L_3}RR + k_{r30}HK_{R_3}RR - k_{0l3}HK_0RR_p - k_{0r3}HK_0RR_p
 \end{aligned}$$

## **Allosteric–Phosphorelay Model**

Because of the inability of the previous models to explain the data we then went on to develop a further model, which exhibits the generic hallmarks of allosteric behaviour [1] (see Manuscript); in particular interactions between (and the presence of) both phosphorylation sites at each level are required for functional phosphorelay. This is described by the reactions



where  $\phi(x, y, \dots) = 1$  if all arguments are true (i.e. if all domains and phosphorylation sites are present and functional) and zero otherwise. For simplicity we drop these indicator functions in the following, whence the corresponding ODEs are given by,

$$\begin{aligned}
d(HK_0)/dt &= -k_{0l1}HK_0 - k_{0r1}HK_0 + k_{l20}HK_{L_2} + k_{r20}HK_{R_2} \\
&\quad + k_{l30}HK_{L_3}RR + k_{r30}HK_{R_3}RR - k_{0l3}HK_0RR_p \\
&\quad - k_{0r3}HK_0RR_p + k_{l10}HK_{L_1} + k_{r10}HK_{R_1} \\
d(HK_{L_1})/dt &= k_{0l1}HK_0 - k_{l10}HK_{L_1} - k_{l1r2}HK_{L_1} - k_{l1l2}HK_{L_1} \\
d(HK_{R_1})/dt &= k_{0r1}HK_0 - k_{r10}HK_{R_1} - k_{r1l2}HK_{R_1} - k_{r1r2}HK_{R_1} \\
d(HK_{L_2})/dt &= k_{r1l2}HK_{R_1} - k_{l2r3}HK_{L_2} + k_{r3l2}HK_{R_3} - k_{l20}HK_{L_2} + k_{l1l2}HK_{L_1} \\
&\quad - k_{l2l3}HK_{L_2} + k_{l3l2}HK_{L_3} \\
d(HK_{R_2})/dt &= k_{l1r2}HK_{L_1} - k_{r2l3}HK_{R_2} + k_{l3r2}HK_{L_3} - k_{r20}HK_{R_2} + k_{r1r2}HK_{R_1} \\
&\quad - k_{r2r3}HK_{R_2} + k_{r3r2}HK_{R_3} \\
d(HK_{L_3})/dt &= k_{r2l3}HK_{R_2} - k_{l3r2}HK_{L_3} - k_{l30}HK_{L_3}RR + k_{0l3}HK_0RR_p \\
&\quad + k_{l2l3}HK_{L_2} - k_{l3l2}HK_{L_3} \\
d(HK_{R_3})/dt &= k_{l2r3}HK_{L_2} - k_{r3l2}HK_{R_3} - k_{r30}HK_{R_3}RR + k_{0r3}HK_0RR_p \\
&\quad + k_{r2r3}HK_{R_2} - k_{r3r2}HK_{R_3} \\
d(RR)/dt &= -k_{l30}HK_{L_3}RR - k_{r30}HK_{R_3}RR + k_{0l3}HK_0RR_p + k_{0r3}HK_0RR_p \\
d(RR_p)/dt &= k_{l30}HK_{L_3}RR + k_{r30}HK_{R_3}RR - k_{0l3}HK_0RR_p - k_{0r3}HK_0RR_p
\end{aligned}$$

In this model kinase activity depends on both phosphorylation sites to be present, while the phosphatase process (dephosphorylation always occurs from the D1 domain) is more flexible.

## Non-Identifiability of Models from Wildtype Data

The three models differ in the way that the phosphate groups move along the (*cis*) or between (*trans*) the two monomers making up the functional HK homo-dimers. As is apparent from the reaction schemes in the previous sections we cannot distinguish between the three different models. Their main difference is in the precise way in which the phosphate group moves along (*cis*) or between (*trans*) the two monomers making up the functional HK homo-dimers.

For two state variables  $x$  and  $y$  with associated ODEs  $\frac{dx}{dt} = f(x, y; \theta)$  and  $\frac{dy}{dt} = g(x, y; \theta)$  we have for their sum  $z = x + y$

$$\frac{dz}{dt} = \frac{dx}{dt} + \frac{dy}{dt}$$

because of the linearity of differentiation. Therefore we can consider the states of the domains we have for example,

$$\frac{dHK_i}{dt} = \frac{dHK_{L_i}}{dt} + \frac{dHK_{R_i}}{dt}$$

For the *cis* and *trans* models a simple relabelling of the domains suffices to show that their wild-type dynamics are, as far as measurable activity levels are concerned) can be written e.g. as

$$\begin{aligned} d(HK_0)/dt &= -k_{01}HK_0 + k_{20}HK_2 + k_{30}HK_3RR - k_{03}HK_0RR_p + k_{10}HK_1 \\ d(HK_1)/dt &= k_{01}HK_0 - k_{10}HK_1 - k_{12}HK_1 \\ d(HK_2)/dt &= k_{12}HK_1 - k_{23}HK_2 + k_{32}HK_3 - k_{20}HK_2 \\ d(HK_3)/dt &= k_{23}HK_2 - k_{32}HK_3 - k_{30}HK_3RR + k_{03}HK_0RR_p \\ d(RR)/dt &= -k_{30}HK_LRR + k_{03}HK_0RR_p \\ d(RR_p)/dt &= k_{30}HK_3RR - k_{03}HK_0RR_p \end{aligned} \quad (1)$$

The measured read-out is the activity of  $RR_p$  and therefore these two models are non-identifiable. To show that the same model also describes the effective, measurable wild-type activities of the allosteric model we need to exploit the symmetries in the parameter set, where e.g.  $k_{lij} = k_{rij}$  for all  $0 \leq i, j \leq 3$ . The three models are thus mathematically indistinguishable given data collected from wild-type ArcB/ArcA. This then determined the development of mutants that in combination allow us to distinguish between the different models.

Even if we could read out the activity of ArcB directly *in vivo* we cannot track reactions along single molecules and would therefore not be able to

distinguish between the different mechanisms in wildtype ArcB. The models do, however, differ in the extent to which mutants where the phosphorylation sites have been ablated can activate ArcA; this insight has therefore guided the experimental design.

## Model Parameters

We start our analysis from the parameter values in Table S1 (taken from [2]); these were, however, only starting points and we used an extensive maximum-likelihood analysis (as described in the Manuscript) to calibrate our models and against the data. Additionally we have used Latin-Hypercube sampling [3] to ensure that the inference was not stuck in local extrema of the likelihood surface.

Parameter	Value	Description
$k_{0y1r}, k_{x1y2r}, k_{x2y3}$	$0.00132 \text{ s}^{-1}$	Auto-phosphorylation and forward phosphotransfer in HK
$k_{x20r}, k_{x3y2}$	$0.001 \text{ s}^{-1}$	Dephosphorylation and reverse phosphotransfer in HK
$k_{x30}$	$0.5 \text{ s}^{-1}$	Phosphotransfer rate from HKp to RR
$k_{0x3}$	$0.05 \text{ s}^{-1}$	Phosphotransfer rate from RRp to HK
$k_{x10}$	$0.0001 \text{ s}^{-1}$	Spontaneous dephosphorylation rate of HK1
Initial [HK]	$1\mu\text{M}$	Initial unphosphorylated HK concentration
Initial [phosphorylated HK]	$0\mu\text{M}$	Initial concentration of all phosphorylated HK
Initial [RR]	$8\mu\text{M}$	Initial unphosphorylated RR concentration
Initial [RRp]	$2\mu\text{M}$	Initial concentration phosphorylated HK (this is also the value for negative control)

\*: x, y could be either "L" or "R".

Table S 1: List of parameter values used in the phosphorelay models. Forward phosphorelay includes the reactions the transfer the phosphate group toward response regulator. Reverse phosphorelay includes the reactions the transfer the phosphate group away from response regulator until dephosphorylation. These are only initial starting values used in the parameter search and are based on analyses carried out in [4], starting from the work of Kim & Cho [5].

## More General Models

In the main paper we have focussed on phosphorelay mechanisms that allow for one phosphate group along each dimer. Other models, with more phosphate groups per dimer, however, yield the same type of behaviour and can effectively be subsumed into the same simplified monomeric model given by the Eqns. (1). In Figure 1 in the main paper we summarize, for example, *cis* and *trans* models for the phosphorelay and their respective transitions diagrammatically. In addition to an algebraic analysis which shows that these models and also mathematically indistinguishable, we can also show their equivalence in light of the data. This is done in Figure S 3, where we show the best fits of the *cis* and *trans* models to the mutant data for this model (with up to 2 phosphates per dimer), which recapitulates the results obtained in the main manuscript.

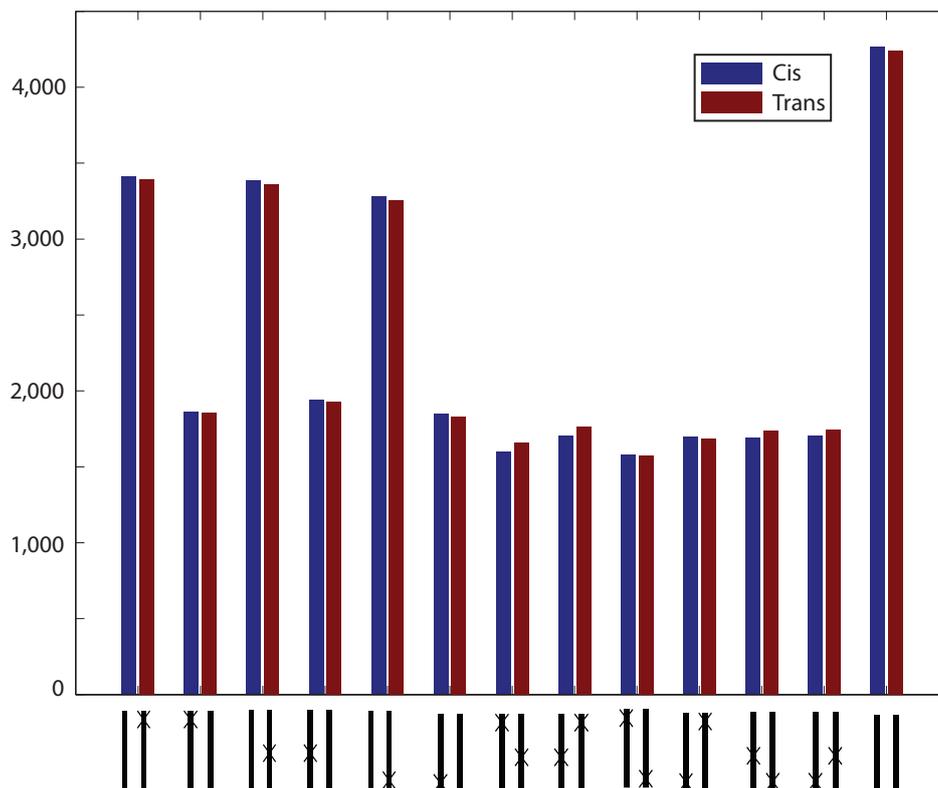


Figure S 3: Steady-state intensities of the two-component system's output (activity of RR) for the *cis* and *trans* models given diagrammatically in Figure 1 in the paper; parameters were again determined in a maximum-likelihood framework.

We also consider a simplified model where the reaction rates are constrained; the results are shown in Figure S 4. In these simulations there are only 4 parameters for each of the three models,  $k_f$ ,  $k_r$ ,  $kt_1$  and  $kt_2$ , as described in Materials and Methods section. The top panel shows the contents of the transfected plasmids in each lane. This model (discussed in the paper) confirms the results seen in the more general model and are discussed in the main paper.

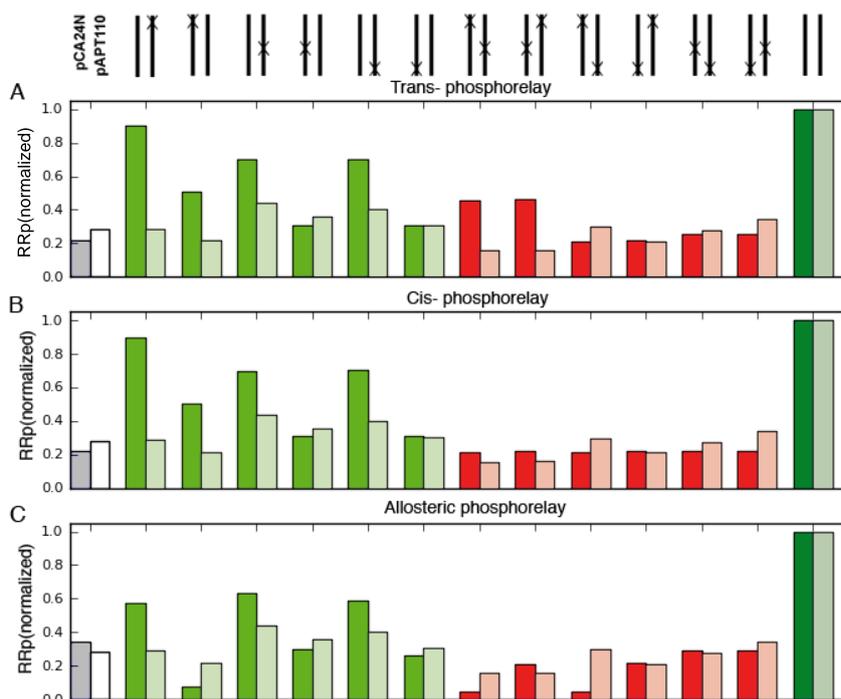


Figure S 4: Simulation of three phosphorelay models with constrained parameter settings.

The extensive and laborious analysis of further models — both more complex as well as simpler models — only served to reinforce the central result of the main paper: pure *cis* and *trans* mechanisms are unable to explain the data obtained from our mutants. Thus model-misspecification is unlikely to be the cause of this shortcoming and a single intra- or inter-molecular phosphorelay mechanism appears extremely unlikely. Instead the allosteric mechanism appears to be required to facilitate the phosphorelay.

## Sensitivity and Robustness Analysis

In order to substantiate our results and safe-guard against model misspecification and effects of parameter uncertainty we also performed a detailed sensitivity analysis (discussed in the main paper) centered around the Fisher information matrix (FIM). In Figure S 5 we first determine the contributions the different parameters make to the eigenvalues of the FIM. The first two eigenvectors (which correspond to the largest eigenvalues of the FIM) identify the “stiff” eigen-parameters [6, 7, 8], which are the directions of greatest sensitivity. These two pairs are very similar for the three models, *i.e.* the projections of the raw parameters onto the eigen-parameters are highly conserved between models. This is, of course, also partly a manifestation of the indistinguishability of the models. The last two eigenvectors correspond the “sloppiest” eigen-parameters. These are the ones least constrained by the data (*i.e.* the directions where the variance of the system output under parameter variation is largest).

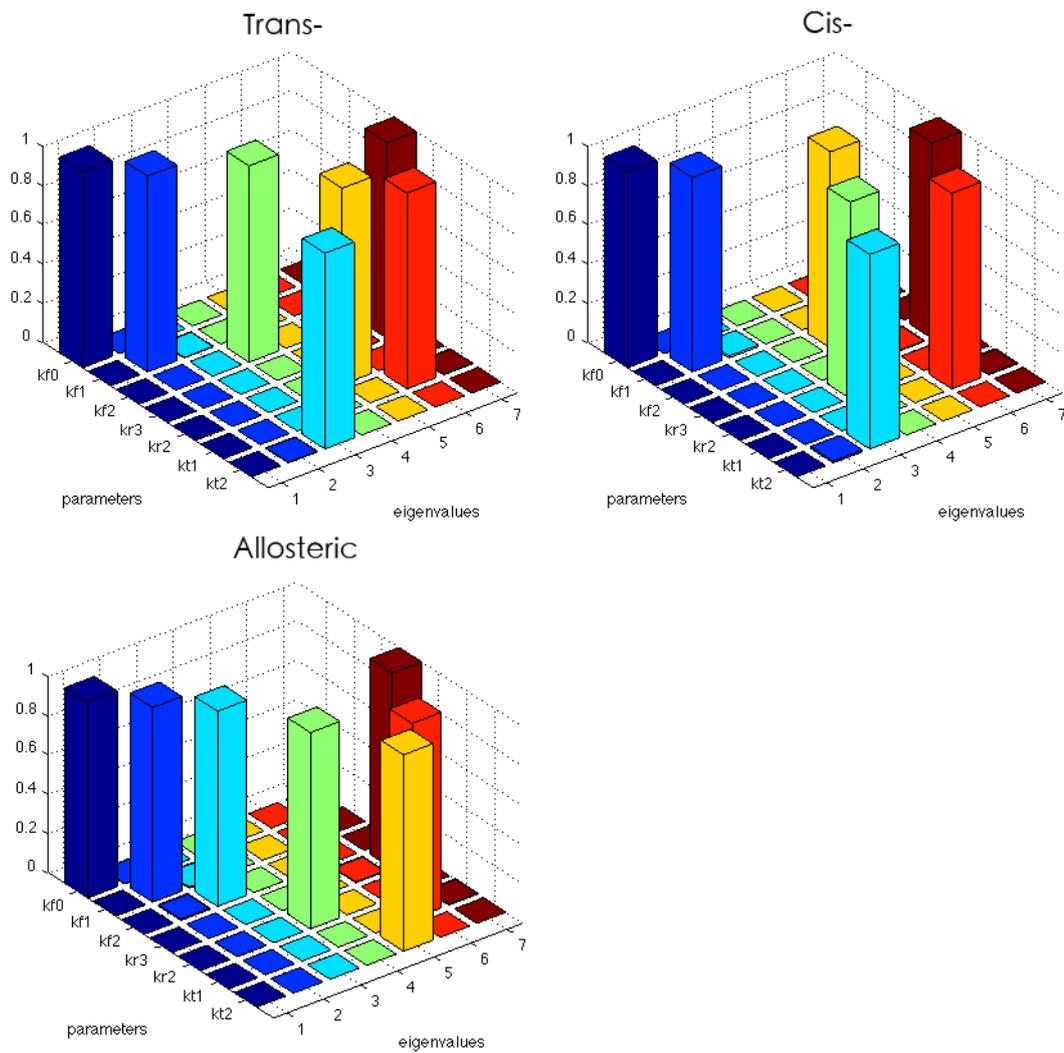


Figure S 5: Contributions of individual parameters into each eigenvalue for all three models.

This finding is further substantiated in figures S 6-S 8, where we show the FIM-derived sensitivity [9, 10] plots as contour plots which indicate the curvature of the likelihood surface around the maximum-likelihood estimate of the model parameters. These results are again in very good agreement, but some parameter combinations appear more constrained by the data in the different models.

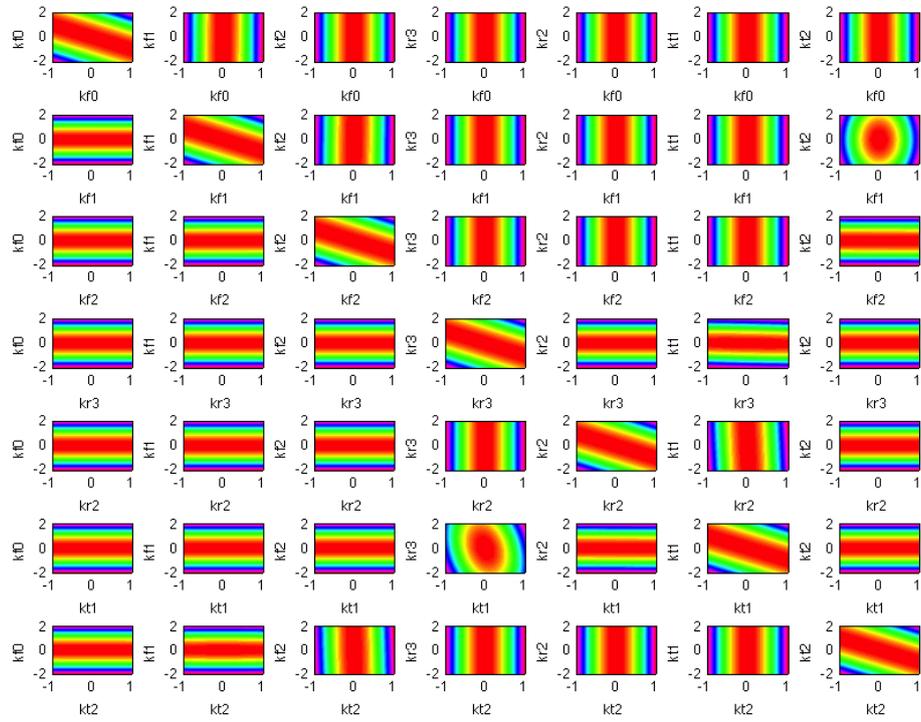


Figure S 6: FIM for all pairs of parameters for the *trans*-phosphorelay model visualised as heat maps.

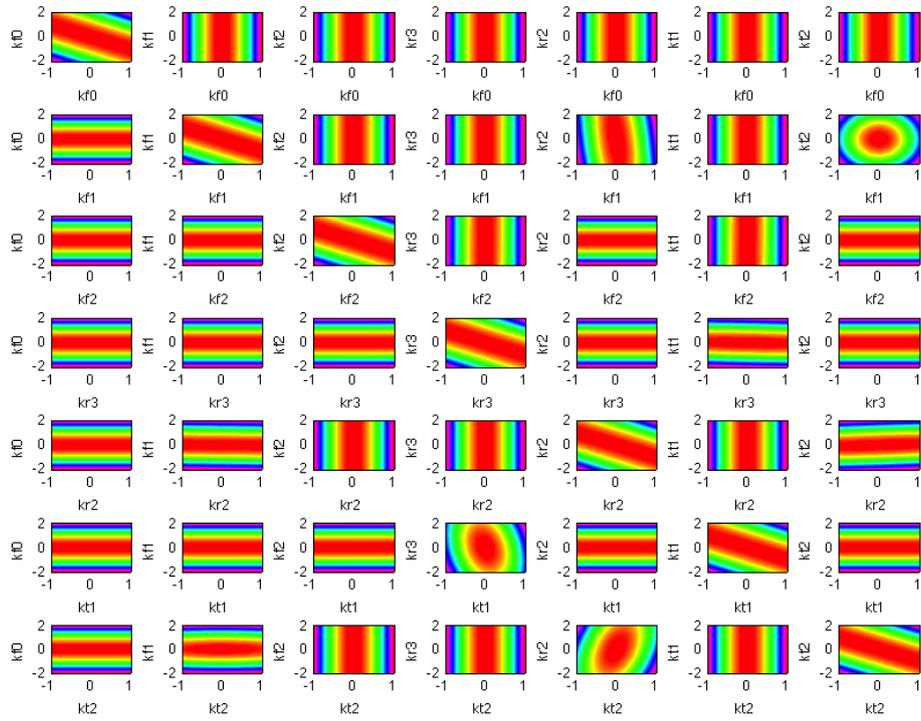


Figure S 7: FIM for all pairs of parameters for the *cis*-phosphorelay model visualised as heat maps.

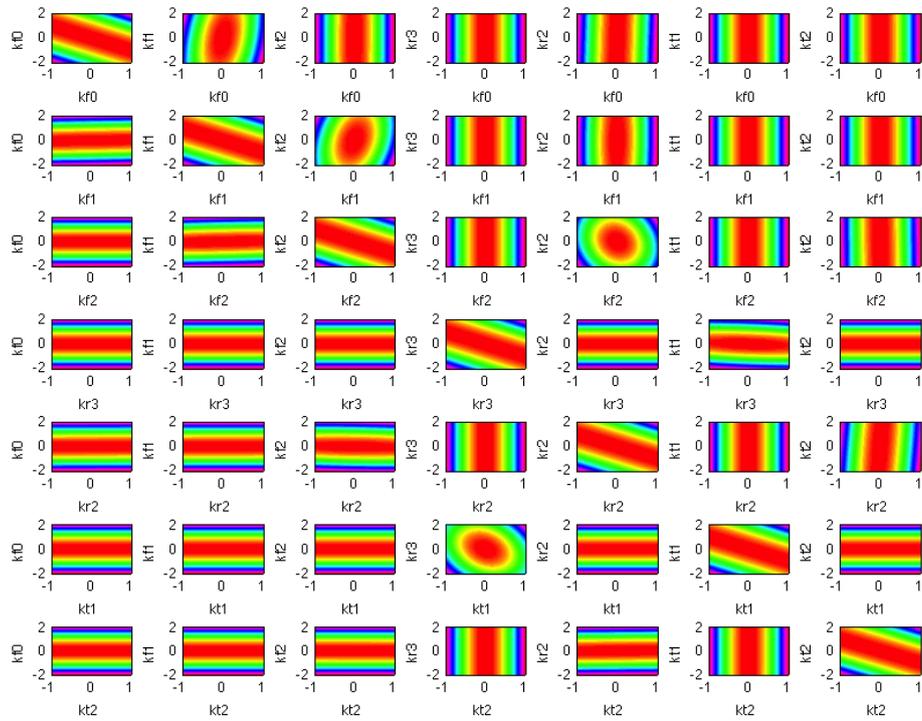


Figure S 8: FIM for all pairs of parameters for the allosteric phosphorelay model visualised as heat maps.

## More General Models

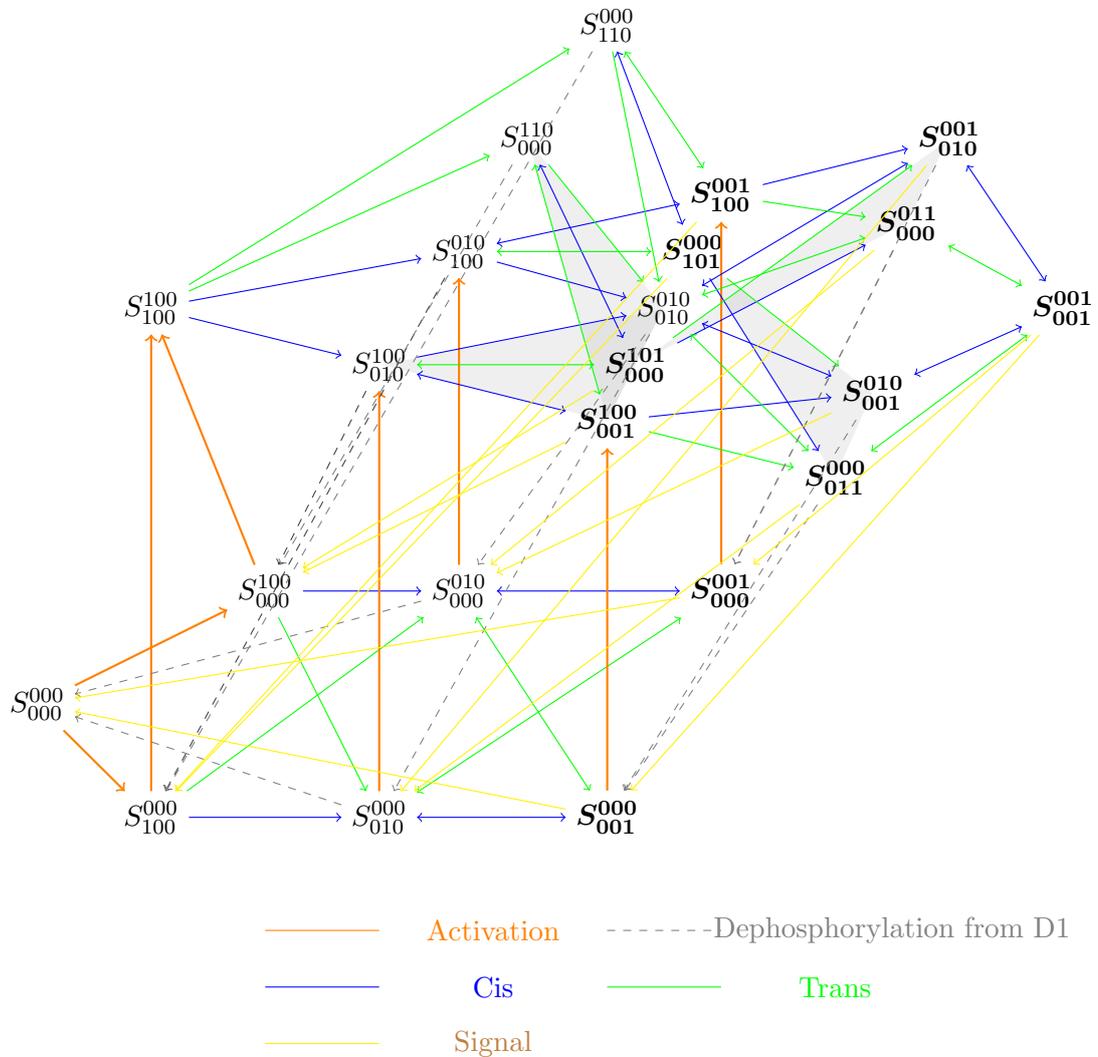


Figure S 9: Graphical representation of all possible transitions in a *cis* and *trans* phosphorelay models that allows for up to two phosphates in each dimer. Upper and lower triplets indicate phosphate occupancy in each protein in the dimer (0= no phosphate, 1=phosphate). Bold ArcB states represent a phosphate group capable of transferring the phosphate to the cognate response regulator, ArcA.

In the main paper we have focussed on phosphorelay mechanisms that allow for one phosphate group along each dimer. Other models, with more phosphate groups per dimer, however, yield the same type of behaviour and can effectively be subsumed into the same simplified monomeric model given by

Eqns. (1). In figure S 9 we summarize, for example, *cis* and *trans* models for the phosphorelay and their respective transitions diagrammatically. In addition to an algebraic analysis which shows that these models and also mathematically indistinguishable, we can also show their equivalence in light of the data.

## References

- [1] J. Monod, J. P. Changeux, F. Jacob, *Journal of molecular biology* **6**, 306 (1963).
- [2] A. M. Kierzek, L. Zhou, B. L. Wanner **6**, 531 (2010).
- [3] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, *Global Sensitivity Analysis: The Primer* (Wiley, 2008).
- [4] X. Sheng, Evolutionary and functional aspects of two-component signalling systems (2013).
- [5] J.-R. Kim, K.-H. Cho, *Computational biology and chemistry* **30**, 438 (2006).
- [6] R. N. Gutenkunst, *et al.*, *PLoS computational biology* **3**, 1871 (2007).
- [7] D. A. Rand, *Journal of the Royal Society Interface* **5**, S59 (2008).
- [8] K. Erguler, M. P. H. Stumpf, *Molecular Biosystems* **7**, 1593 (2011).
- [9] M. Komorowski, M. J. Costa, D. A. Rand, M. P. H. Stumpf, *Proceedings of the National Academy of Sciences of the United States of America* **108**, 8645 (2011).
- [10] M. Komorowski, J. Zurauskiene, M. P. H. Stumpf, *Bioinformatics (Oxford, England)* **28**, 731 (2012).