

DECKBLATT

An analysis of two-component regulatory systems in *Myxococcus xanthus*

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Xingqi Shi
aus Anhui, China

Marburg / Lahn, Oktober 2008

Die Untersuchungen zur vorliegenden Arbeit wurden von Oktober 2005 bis Oktober 2008 am Max-Planck-Institut für terrestrische Mikrobiologie unter der Leitung von Prof. Dr. MD Lotte Sogaard-Andersen durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg als Dissertation angenommen am:

Erstgutachter: Prof. Dr. MD Lotte Sogaard-Andersen

Zweitgutachter: Prof. Dr. Erhard Bremer

Tag der mündlichen Prüfung am:

The first part of results in this thesis has been published:

Shi, X., S. Wegener-Feldbrugge, S. Huntley, N. Hamann, R. Hedderich & L. Sogaard-Andersen, (2008) Bioinformatics and experimental analysis of proteins of two-component systems in *Myxococcus xanthus*. *J Bacteriol* 190: 613-624.

List of contents	
Abbreviations	1
1 Abstract	2
Zusammenfassung	4
2 Introduction	7
2.1 Two-component regulatory system	7
2.1.1 General introduction	7
2.1.2 Domains of HPK.....	8
2.1.3 Domains of RR.....	12
2.1.4 Architecture of TCS pathways	14
2.1.5 Recognition specificity of TCS proteins	15
2.2 Development of <i>Myxococcus xanthus</i>	19
2.2.1 Life cycles of <i>M. xanthus</i>	19
2.2.2 Intercellular signal transduction pathways during development	20
2.2.3 Motility of <i>M. xanthus</i>	25
2.3 TCS in <i>M. xanthus</i>	27
3 Results	29
3.1 An analysis of two-component regulatory systems in <i>M. xanthus</i>	29
3.1.1 Identification of TCS genes in <i>M. xanthus</i>	29
3.1.2 Genetic organization of TCS genes in <i>M. xanthus</i>	30
3.1.3 Histidine protein kinases	31
3.1.4 Histidine protein kinase-like proteins	32
3.1.5 Response regulators and output domains	33
3.1.6 Transcriptional regulation of TCS genes during development	34
3.1.7 Genetic analysis of transcriptionally up-regulated, orphan <i>hpk</i> genes	38
3.2 In search of the FruA kinase using a candidate approach	43
3.2.1 A candidate approach to identify for the FruA kinase	43
3.2.2 Y2H analyses to test the interactions between FruA and 25 developmentally up-regulated orphan HPKs.....	43
3.2.3 Analyses of the FruA kinase candidates.....	47
3.2.4 Epistasis analyses of <i>fruA</i> and <i>fruA</i> kinase candidates.....	49
3.2.5 FrzCD methylation and FruA accumulation in the mutants of FruA kinase candidates	51
3.2.6 Developmental gene expression in the mutants of FruA kinase candidates	53
3.2.7 Purification of FruA and the FruA kinase candidates	57
3.2.8 Autophosphorylation assay of Hpk8	59
3.2.9 Chemical stability of phosphorylated Hpk8	60
3.2.10 Preliminary data of phosphotransfer from Hpk8 to FruA.....	64
3.2.11 Accumulation of Hpk8 during development	65

3.2.12	Preliminary data of autophosphorylation of SdeK and Hpk12.....	66
3.2.13	Preliminary analyses of the redundant FruA kinase candidates	67
3.3	Characterization of Hpk37	70
3.3.1	Analysis of Hpk37.....	70
3.3.2	Motility assay of $\Delta hpk37$	73
3.3.3	FrzCD methylation and FruA accumulation in $\Delta hpk37$	74
3.3.4	Developmental gene expression in $\Delta hpk37$	75
3.3.5	Purification of Hpk37 ₁₁₄₈₋₁₉₆₇	76
3.4	In search of the FruA kinase with bioinformatics method	78
3.5	Miscellaneous <i>hpk</i> mutants	79
4	Discussion	81
4.1	Two-component regulatory systems in <i>M. xanthus</i>	81
4.2	In search of the FruA kinase	85
4.3	Hpk37.....	92
5	Material and methods	94
5.1	Reagents, Enzymes and Kits	94
5.2	Microbiological methods	94
5.2.1	<i>Escherichia coli</i> strains	94
5.2.2	<i>Saccharomyces cerevisiae</i> strain	95
5.2.3	<i>Myxococcus xanthus</i> stains	95
5.2.4	Cultivation of <i>E. coli</i>	97
5.2.5	Cultivation of <i>S. cerevisiae</i>	97
5.2.6	Cultivation of <i>M. xanthus</i>	97
5.2.7	Development assay and spore assay of <i>M. xanthus</i>	98
5.2.8	Motility assay of <i>M. xanthus</i>	98
5.3	Molecular biological methods.....	99
5.3.1	Oligonucleotide and plasmids.....	99
5.3.2	Preparation of chromosomal DNA from <i>M. xanthus</i>	111
5.3.3	PCR reaction , digestion and ligation.....	111
5.3.4	Transformation of <i>E. coli</i>	112
5.3.5	Electroporation of <i>M. xanthus</i>	112
5.3.6	Co-transformation of <i>S. cerevisiae</i>	113
5.3.7	RNA preparation from <i>M. xanthus</i>	113
5.3.8	RNA clean up, cDNA synthesis and qRT-PCR.....	114
5.3.9	Construction of in-frame deletion strain in <i>M. xanthus</i>	114
5.4	Biochemical methods	116
5.4.1	Immunoblot analysis.....	116
5.4.2	Over-expression of proteins in <i>E. coli</i>	117
5.4.3	The solubility test of over-expressed proteins.....	118

5.4.4	Purification of His-tag protein under native conditions	118
5.4.5	Gel filtration chromatography	119
5.4.6	Purification of His-tag protein under denature conditions	119
5.4.7	Refolding of protein	120
5.4.8	Purification of FruA and FruA _{D59N} with strep-tag	120
5.4.9	Autophosphorylation of Hpk8	121
5.4.10	Chemical stability of phosphoryl group	121
5.4.11	Phosphotransfer reaction from phosphorylated Hpk8 to FruA	122
5.5	Bioinformatics methods	122
5.5.1	Sequence retrieval and domain analysis of protein structure	122
5.5.2	Identification and phylogenetic analysis of TCS proteins.....	122
5.5.3	Classification of TCS proteins in <i>M. xanthus</i> based on genetic organization	123
6	Supplementary data.....	125
6.1	TCS proteins in <i>M. xanthus</i>	125
7	References	140
	Acknowledgments	150
	Curriculum Vitae	152
	Erklärung	153

Abbreviations

bp	Base pair
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetic acid
hr(s)	Hour(s)
IPTG	Isopropyl-1-thio-D-galactopyranoside
kD	Kilodalton
LB medium	Luria-Bertani medium
Min	Minute
MOPS	3-(N-morpholino)propanesulfonic acid
Ni-NTA	Ni-nitrilotriacetic
OD	Optical density
PMSF	Phenylmethylsulphonyl fluoride
rpm	Rounds per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
X-gal	5-bromo-4-chloro-3-hydroxyindole

1 Abstract

Proteins of two-component regulatory systems (TCS) have essential functions in the sensing of external and self-generated signals in bacteria as well as in the generation of appropriate output responses. Accordingly, in *Myxococcus xanthus* TCS are important for fruiting body formation and sporulation as well as normal motility. In this study, I analyzed the *M. xanthus* genome for the presence and genetic organization of genes encoding TCS. 272 genes that encode TCS proteins were identified including 21 genes in eight loci, which encode TCS proteins that are part of chemotaxis-like systems. Subsequent analyses focused on 251 TCS proteins (non chemotaxis-like) consisting of 118 histidine protein kinases (HPKs), 119 response regulators (RRs) and 14 HPK-like genes. 71% of the TCS genes are organized in unusual manners as orphan genes or in complex gene clusters whereas the remaining 29% display the standard paired gene organization. Bioinformatics analyses suggest that TCS proteins encoded by orphan genes and complex gene clusters are functionally distinct from TCS proteins encoded by paired genes. Experimentally, microarray data and quantitative real-time PCR suggest that orphan TCS genes are overrepresented among TCS genes that display altered transcription during fruiting body formation. The genetic analysis of 25 orphan HPKs, which are transcriptionally up-regulated during development, led to the identification of two HPKs that are likely essential for viability and seven HPKs including four novel HPKs that have important function in fruiting body formation or spore germination.

As an attempt to identify functional partners of orphan TCS proteins in *M. xanthus*, I focused on the RR FruA, which has a key role in the C-signal transduction pathway. To identify the FruA kinase, two candidate approaches were used. The first candidate approach is based on the hypothesis that a FruA kinase gene shares characteristics with the *fruA* gene, i.e. it is orphan, developmentally up-regulated at the transcriptional level and a null mutant is deficient in development. Yeast two-hybrid analysis was used to investigate potential interactions between FruA and developmentally regulated orphan HPKs. Three best FruA kinase candidates (SdeK, Hpk8 and Hpk12) and four potentially redundant candidates (Hpk9, Hpk11, Hpk13 and Hpk29) were

identified. *In vivo* analyses of the three best FruA kinase candidates support a model in which SdeK is the main FruA kinase, Hpk12 is a minor FruA kinase and Hpk8 is a phosphatase of FruA~P. Furthermore, SdeK may have other downstream targets in addition to FruA and there may be other HPKs that phosphorylate or cross talk to FruA. To obtain direct evidence for an interaction between FruA and the FruA kinase candidates *in vitro*, the relevant proteins have been purified. To date, the Hpk8 and Hpk12 proteins have been shown to autophosphorylate *in vitro*. Intriguingly, Hpk8 does not appear to be phosphorylated on the conserved His residue but is likely phosphorylated on a Tyr residue. Preliminary phosphotransfer assay suggests that Hpk8 engages in phosphotransfer to or phosphorylation of FruA. A possible interaction *in vitro* between SdeK and Hpk12 with FruA still remains to be shown.

Hpk37 belongs to the group of orphan HPKs that are transcriptionally up-regulated during development and essential for development. However, the yeast two-hybrid analyses to determine a possible direct interaction with FruA were inconclusive. *In vivo* analyses demonstrated that Hpk37 is likely involved in the production or response to (p)ppGpp or the A-signal suggesting that Hpk37 is not a FruA kinase. Domain analyses of Hpk37 and analyses of the genetic organization of the *hpk37* locus suggest that regulation of Hpk37 activity could involve a unique methylation/demethylation mechanism similar to that resulting in adaptation in chemosensory pathways.

In a second candidate approach to identify a FruA kinase, candidates were predicted using an *in silico* method (White *et al.*, 2007). *In vivo* analyses of mutants carrying mutations in the genes encoding the six best candidates strongly suggest that these HPK are not FruA kinases.

Zusammenfassung

Proteine in Zweikomponentensystemen (Two-component systems, TCS) haben essentielle Funktionen bei der Detektion externer und interner Signale in Bakterien sowie bei der Erzeugung geeigneter Reiz-Antworten. Dementsprechend erfüllen Zweikomponentensysteme auch bei der Fruchtkörperbildung, Sporulation und Motilität von *Myxococcus xanthus* wichtige Aufgaben. In der vorliegenden Arbeit wurde das Genom von *M. xanthus* auf das Vorhandensein und die genetische Organisation von Zweikomponentensystemen untersucht. 272 Gene für Proteine aus Zweikomponentensystemen wurden identifiziert, darunter 21 Gene in acht verschiedenen Loci für Che-ähnliche Systeme. Die weitere Analyse wurde auf die 251 nicht Che-ähnlichen Proteine konzentriert, wovon 118 Histidinkinasen (Histidine Protein Kinases, HPKs), 119 Regulatoren (Response Regulators, RRs) und 14 HPK-ähnliche Gene sind. 71% der Gene für Zweikomponentensysteme sind ungewöhnlich als verwaiste Gene (orphan genes) organisiert oder in komplexen Gen-Clustern angeordnet, während die verbleibenden 29% der Gene eine gewöhnliche paarweise Organisation aufweisen. Bioinformatische Analysen legen nahe, dass sich TCS-Proteine, die von verwaisten Genen kodiert werden oder aus komplexen Gen-Clustern stammen, funktionell von TCS-Proteinen unterscheiden, die von Genpaaren kodiert werden. Microarray-Daten und qRT-PCR-Experimente weisen darauf hin, dass verwaiste TCS-Gene unter den TCS-Genen überrepräsentiert sind, deren Transkription während der Fruchtkörperbildung verändert ist. Die genetische Analyse von 25 verwaisten HPKs, die während der Entwicklung verstärkt gebildet werden, führte zur Identifikation zweier HPKs, die wahrscheinlich für die Lebensfähigkeit von *M. xanthus* essentiell sind sowie sieben HPKs, darunter vier neue, die eine wichtige Funktion bei der Fruchtkörperbildung oder Sporenkeimung haben.

Im Bestreben, die verwaisten TCS-Proteine von *M. xanthus* funktionell miteinander zu verbinden, habe ich mich in dieser Arbeit auf den Response-Regulator FruA konzentriert, der eine Schlüsselrolle bei der Übertragung des C-Signals spielt. Zwei verschiedene Ansätze wurden verfolgt, um die FruA-Kinase zu identifizieren. Der erste Ansatz beruht auf der Annahme, dass ein

FruA-Kinase-Gen Eigenschaften vom *fruA*-Gen aufweist, das heißt, dass es verwaist ist, seine Transkription während der Entwicklung verstärkt ist und eine Nullmutante in der Entwicklung gehemmt ist. Zusätzlich wurde das Hefe-Zwei-Hybrid-System verwendet, um mögliche Interaktionen zwischen FruA und den entwicklungsbedingt regulierten, verwaisten HPKs zu untersuchen. Drei hervorragende FruA-Kinase-Kandidaten (SdeK, Hpk8 und Hpk12) und vier möglicherweise redundante Kandidaten (Hpk9, Hpk11, Hpk13 und Hpk29) wurden identifiziert. *In vivo*-Analysen der besten drei FruA-Kinase-Kandidaten führten zu einem Modell, in dem SdeK die Haupt-FruA-Kinase ist, Hpk12 eine untergeordnete FruA-Kinase und Hpk8 eine Phosphatase von FruA~P ist. Darüber hinaus könnte SdeK weitere Zielproteine haben, die im Entwicklungszyklus FruA nachfolgen, und es könnte zudem weitere HPKs geben, die FruA phosphorylieren oder mit FruA interagieren. Um die Interaktion zwischen FruA und den FruA-Kinase-Kandidaten *in vitro* direkt nachzuweisen, wurden die entsprechenden Proteine aufgereinigt. Es wurde gezeigt, dass Hpk8 und Hpk12 *in vitro* autophosphorylieren. Bemerkenswerterweise wird Hpk8 offensichtlich nicht am konservierten Histidin-Rest, sondern wahrscheinlich an einem Tyrosin-Rest phosphoryliert. Vorläufige Ergebnisse von einem Phosphotransfer-Essay legen nahe, dass Hpk8 entweder am Phosphoryltransfer zu FruA beteiligt ist oder FruA direkt phosphoryliert. Eine Interaktion von SdeK und Hpk12 mit FruA *in vitro* muss noch nachgewiesen werden.

Hpk37 gehört zur Gruppe der verwaisten HPKs, deren Transkription während der Entwicklung verstärkt wird und die für die Entwicklung essentiell sind. Hefe-Zwei-Hybrid-System-Analysen zum Nachweis einer direkten Interaktion mit FruA waren jedoch nicht aussagekräftig. *In vivo*-Analysen zeigten, dass Hpk37 sowohl an der Produktion von (p)ppGpp als auch an der Antwort auf (p)ppGpp oder am A-Signaltransduktionsweg beteiligt sein könnte, sodass es wahrscheinlich keine FruA-Kinase ist. Untersuchungen der Protein-Domänen von Hpk37 und der genetischen Organisation des *hpk37*-Locus weisen übereinstimmend darauf hin, dass die Regulation der Aktivität von Hpk37 Methylierungs- bzw. Demethylierungsreaktionen einschließt, ähnlich den Adaptierungsreaktionen bei der Chemotaxis.

In einem zweiten Ansatz wurde eine *in silico*-Methode (White *et al.*, 2007) zur Identifikation von FruA-Kinasen verwendet. *In vivo*-Analysen von Mutanten der sechs besten FruA-Kinase-Kandidaten weisen stark darauf hin, dass diese HPKs keine FruA-Kinasen sind.

2 Introduction

2.1 Two-component regulatory system

2.1.1 General introduction

Bacteria possess elegant signal transduction pathways to sense and respond to external environmental stimuli as well as internal physiological changes. Few mechanisms are used in bacterial signal transduction pathways: one-component systems, methyl-accepting chemotaxis protein, regulation of the level of secondary messengers and protein phosphorylation. Two-component regulatory systems (TCS) are the predominant mechanism in bacterial signal transduction pathways depending on phosphorylation. TCS has been identified in nearly all prokaryotes, which are involved in regulating different processes such as sporulation, motility, cell division, virulence, antibiotic resistance, metabolism and the response to different environmental stresses (Falke *et al.*, 1997). TCS have been found in some eukaryotes and play a role in yeast osmoregulation, fungi hyphal development, plant ethylene response and Dictyostelium development (Dutta *et al.*, 1999). To date, few mammalian proteins have been identified to be structurally similar to HPKs, but TCS has not been conclusively confirmed in mammals (Besant & Attwood, 2005).

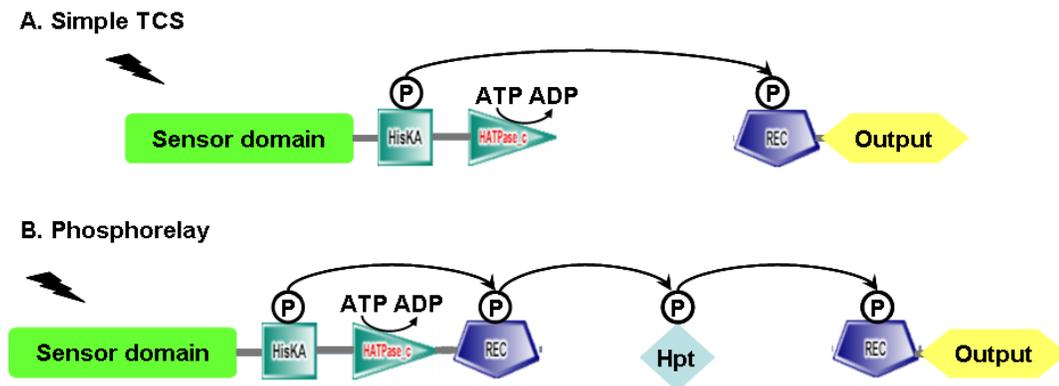


Figure 1. Architecture of the two-component regulatory systems.

A. Simple TCS. B. Linear phosphorelay. Detailed descriptions are in text.

The simple TCS consists of a histidine protein kinase (HPK) and a response regulator (RR) (Figure 1A). The sensor domain of the HPK senses an intercellular or intracellular cue, subsequently the kinase domain of the HPK catalyzes autophosphorylation on a conserved histidine residue, and finally the

phosphoryl group is transferred to a conserved aspartate residue in the receiver domain of the cognate RR. Phosphorylation of the RR activates its activity leading to an output response, which typically involves regulation of gene expression, changes in enzymatic activity or in protein-protein interactions. Upon removal of the stimulating signal, the phosphoryl group on the RR is hydrolyzed via a phosphatase activity in the receiver domain, a phosphatase activity in the cognate HPK or by a separate phosphatase. In addition to the simple TCS, some TCS are organized in so-called phosphorelays. Typically, a phosphorelay is composed of a hybrid HPK, which contains a receiver domain (Figure 1B). In response to stimulation, a hybrid HPK catalyzes the autophosphorylation on the conserved histidine, the phosphoryl group is then transferred to the receiver domain, then to a histidine phosphotransfer (Hpt) protein or domain, and eventually to the final RR, which effects the response to change cellular physiology or behavior (Stock *et al.*, 2000).

2.1.2 Domains of HPK

The majority of HPKs are inner membrane proteins (West & Stock, 2001), which are composed of a non-conserved N-terminal sensor domain and a conserved kinase core containing a HisKA domain and a HATPase domain. Hybrid HPKs contain a receiver domain and some also contain an Hpt domain. The characteristics of main domains in HPK are as follows.

The sensor part of HPK is highly diverse. The typical cytoplasmic domains are GAF, PAS (PAC) and HAMP domains whereas most extracytoplasmic domains still remain unclassified due to the low sequence conservation (Szurmant *et al.*, 2007). The PAS domain (period circadian protein, Ah receptor nuclear translocator protein and single-minded protein) is found in more than 33% of all HPKs in the SMART database and it is also present in other signal transduction proteins in archaea, bacteria and eukaryotes (Szurmant *et al.*, 2007). The PAS domain structure of more than 15 characterized proteins reveals a similar $\alpha\beta$ fold with a five-stranded anti-parallel β -sheet core in which the anti-parallel β -sheets form a groove for binding a cofactor (Taylor, 2007). Different cofactors of PAS domains have been identified: heme in FixL (Gong *et al.*, 1998), FAD in Aer and NifL (Bibikov *et al.*, 1997, Repik *et al.*, 2000), 4-hydroxycinnamic acid in photoactive yellow protein (Pellequer *et al.*, 1998) and flavin mononucleotide in

the plant NPH1 photoreceptor (Christie *et al.*, 1999), whereas the HERG cardiac K⁺ channel lacks a cofactor (Morais Cabral *et al.*, 1998). In some HPKs, the PAS domain is immediately followed by a PAC domain. It is proposed that PAC domain contributes to the PAS domain fold.

GAF domains (cGMP phosphodiesterases, adenylyl cyclases and FhIA (formate hydrogen lyase transcriptional activator)) are present in about 10% of all HPKs in the SMART database (Szurmant *et al.*, 2007). This domain is not only found in TCS proteins but is ubiquitous in many other signaling proteins in both eukaryotes and prokaryotes (Aravind & Ponting, 1997). GAF domains are diverse at the primary sequence level, but they share a specific secondary structure with α helices and β sheets (Aravind & Ponting, 1997). The crystal structure of the GAF domain in YKG9 of *Saccharomyces cerevisiae* displays a fold similar to PAS domains (Ho *et al.*, 2000). It is proposed that in general GAF domains could bind diverse ligands (nucleotides or small molecules) and regulate the activities of catalytic domains including the ATPase activity of HPK (Aravind & Ponting, 1997). Different binding ligands of GAF domains have been identified: cGMP in mammalian PDE5 (Rybalkin *et al.*, 2003), haem in the sensor kinase DosS of *Mycobacterium tuberculosis* (Sardiwal *et al.*, 2005), tetrapyrrole chromophore in SyPixJ1 of *Synechocystis sp.* (Yoshihara *et al.*, 2004) and in TePixJ of *Thermosynechococcus elongates* (Ishizuka *et al.*, 2006), nitric oxide in NorR of *Escherichia coli* (D'Autreaux *et al.*, 2005), 2-oxoglutarate in NifA of *Azotobacter vinelandii* (Little & Dixon, 2003) and cAMP in cyanobacterium *Anabaena* PCC7120 *cyaB2* adenylyl cyclase (Bruder *et al.*, 2005).

HAMP (histidine kinases, adenylyl cyclases, methyl binding proteins, phosphatases) domains are recognized in about 30% of all HPKs in the SMART database (Szurmant *et al.*, 2007). The crystal structure of the HAMP domain from *Archaeoglobus fulgidus* Af1503 reveals that the HAMP domain forms a dimer with a parallel four-helix bundle fold (Figure 2A) (Hulko *et al.*, 2006). The knobs-to-knobs packing of the hydrophobic core of the interface in this bundle can be converted into a knobs-to-holes packing by 26° rotations of each helix in the four-helix bundle (Figure 2B). Single amino acid substitutions have revealed that the proposed rotation of HAMP could activate or inactivate the signal

transduction protein (Hulko *et al.*, 2006). It has been suggested that the HAMP domain converts conformational changes in the sensing part of an HPK to the stimulation of kinase activity (Hulko *et al.*, 2006).

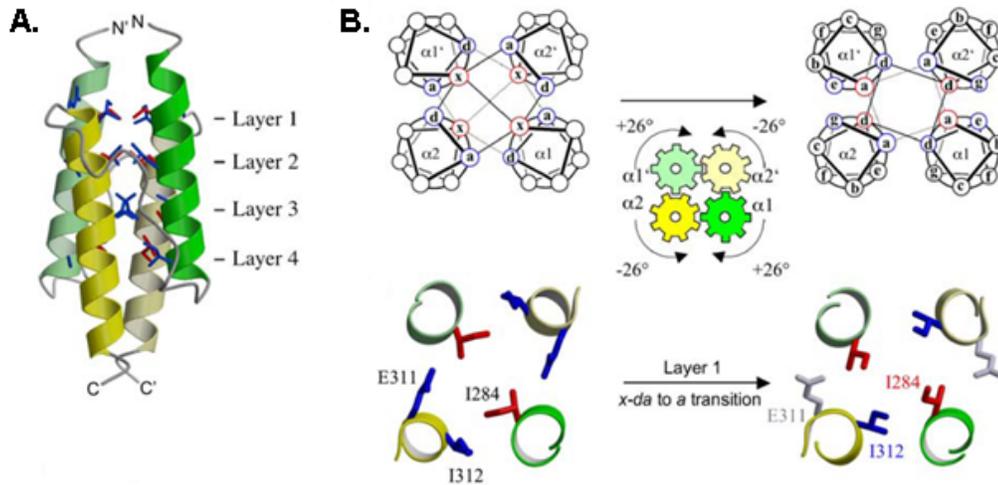


Figure 2. Structure of the Af1503 HAMP domain.

A. Structure of the dimer of Af1503 HAMP domain (monomers in yellow and green). The residues at the interface in x-layer geometry (see B) are in red and residues of interface in d-layer geometry (see B) are in blue. Four layers of interface are packed. B. The knobs-to-knobs converted to knobs-to-holes packing by a 26° rotation of all helices. The first layer of interfacing residue is shown. Modified from (Hulko *et al.*, 2006).

HisKA and HATPase domains are the core part of HPK and they are highly conserved at the sequence level. The HisKA domain usually consists of about 60-80 amino acids and represents the dimerization and phosphoacceptor domain of HPK (also named DHp domain). It contains the conserved H box with the autophosphorylation site (His), from which the phosphoryl group is transferred to the cognate RR. The crystal structure of the cytoplasmic portion of *Thermotoga maritima* TM0853 (Marina *et al.*, 2005) containing the HisKA domain and HATPase_c as well as the nuclear magnetic resonance structure of the HisKA domain of *E. coli* EnvZ (Tomomori *et al.*, 1999) revealed that the dimers of these HisKA domains form similar antiparallel four-helix bundles (Figure 3A), whereas the twist angles of the four-helix bundles and the connections between the two hairpin helices are different in these two structures (Marina *et al.*, 2005). The HATPase_c domain is not only found in HPK, but also in DNA gyrase B, topoisomerases, heat shock protein HSP90, phytochrome-like ATPases and DNA mismatching repair proteins. The

HATPase_c domain is the conserved catalytic domain of HPK, which consists of four conserved motifs: N, D, F and G boxes. The structures of the HATPase_c domains of *T. maritima* TM0853 (Marina *et al.*, 2005), *E. coli* EnvZ (Tomomori *et al.*, 1999), *E. coli* CheA (Bilwes *et al.*, 1999) and *E. coli* PhoQ (Marina *et al.*, 2001) reveal a similar α/β -sandwich composed of five to seven antiparallel β sheets and three α helices. The conserved sequence boxes of N, D, F and G form the ATP binding pocket (Figure 3B). The HisKA domain forms a dimer. The HATPase_c domain catalyzes the trans-phosphorylation of the His in the HisKA domain of the other subunit.

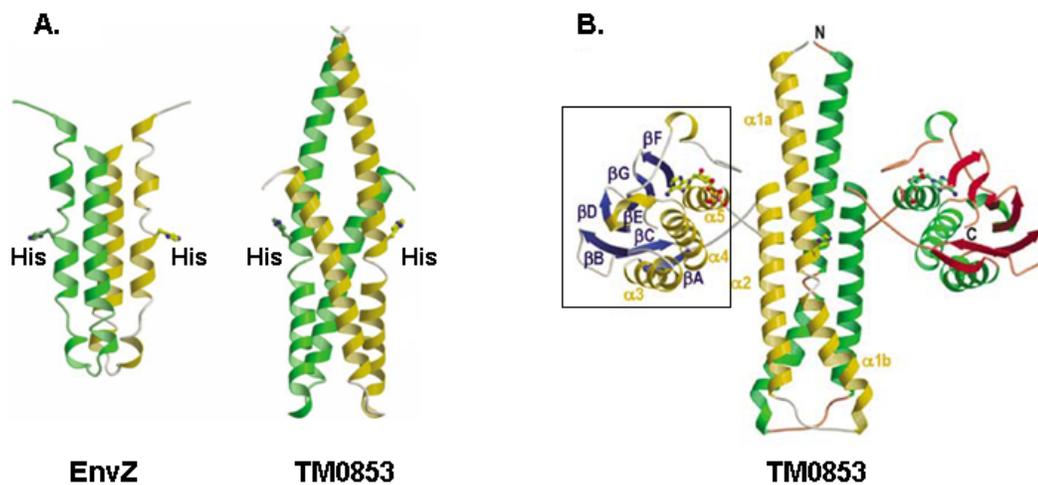


Figure 3. The ribbon diagram of HisKA domain and HATPase_c domain.

A. Dimer of HisKA domains of *E. coli* EnvZ and *T. maritima* TM0853. Green and yellow indicate separate monomers. The site of autophosphorylation (His) is marked in the figure. B. Dimer of the cytoplasmic part of TM0853 containing HisKA domain and HATPase_c domain. Green and yellow indicate separate monomers. The HATPase_c domain of one monomer is highlighted with black outline. Reprinted from (Marina *et al.*, 2005).

Hpt domains exist as independent proteins, attached to hybrid HPKs or in CheA. The Hpt domain contains the conserved His for the phosphotransfer reaction and does not display ATP-dependent catalytic activity. Hpt domains share only a low similarity in the overall sequence, whereas a few residues in the region flanking the conserved His (H box) are conserved in different Hpt domains. The consensus sequences of the H box of Hpt domains are different from those of the H box in HisKA domains (West & Stock, 2001). However, Hpt domains have a similar four-helix bundle structure to that of the HisKA domain as shown for the monomer of *S. cerevisiae* YPD1, the *E. coli* ArcB Hpt domain,

the *E. coli* CheA Hpt domain and the dimer of *Bacillus subtilis* Spo0B (Figure 5A) (Stock *et al.*, 2000).

Previously, different classifications were performed to group HPKs. Based on the order of the conserved H, N, D, F, G boxes, all of the HPKs can be divided into two classes (Dutta *et al.*, 1999, Bilwes *et al.*, 1999). In Class I kinases, such as TM0853 and EnvZ, the H box is immediately adjacent to the HATPase_c domain and the dimerization region contains the conserved His autophosphorylation site. In class II kinases, exclusively represented by CheA orthologs, the H box in the Hpt domain is separated from the HATPase_c domain by distinct domain insertions. The dimerization domain of class II HPK is located between the Hpt domain and the HATPase_c domain. CheA orthologs are part of chemosensory systems. The difference in the domain organization of the two classes of HPKs suggest that they have distinct functions and regulatory mechanisms (Dutta *et al.*, 1999). Another generally used classification is based on the sequence similarity of the HisKA and HATPase_c domains and the consensus sequence of the conserved functional boxes. By multisequence alignment, the HPKs were grouped into 11 different subfamilies HPK1-HPK11 on the basis of the different consensus sequences of the conserved boxes (Grebe & Stock, 1999).

2.1.3 Domains of RR

RR typically either consists of an N-terminal highly conserved receiver domain and a C-terminal variable output domain or only consists of the receiver domain. The receiver domain contains a conserved Asp residue, which is the phosphorylation site receiving the phosphoryl group from a HisKA domain, an Hpt domain or from small molecules (acetyl phosphate and phosphoamidate). Phosphorylation of RR typically mediates dimerization of the proteins. The crystal structures of several receiver domains reveal a similar $(\beta\alpha)_5$ fold with a central five parallel β -sheet flanked by helices (Figure 4A).

In FixJ of *Sinorhizobium meliloti* (Figure 4A) (Birck *et al.*, 1999, Gouet *et al.*, 1999), three residues (Asp10, Asp11 and Asp54) coordinate an Mg^{2+} ion are required for phosphorylation of the conserved Asp54. Three residues, Lys104, Thr82 (could be Ser/Thr in β_4 in other RRs), and Phe101 (could be Phe/Tyr in

$\beta 5$ in other RRs), are involved in the conformational changes during phosphorylation. In the phosphorylated structure, Lys104 forms an ion pair with the phosphate on Asp54, three hydrogen bonds are formed with three oxygen of phosphate (one by Thr82 and two by active-site carboxylate side chain) and Phe101 fills in the space occupied by Thr82 in the unphosphorylated structure.

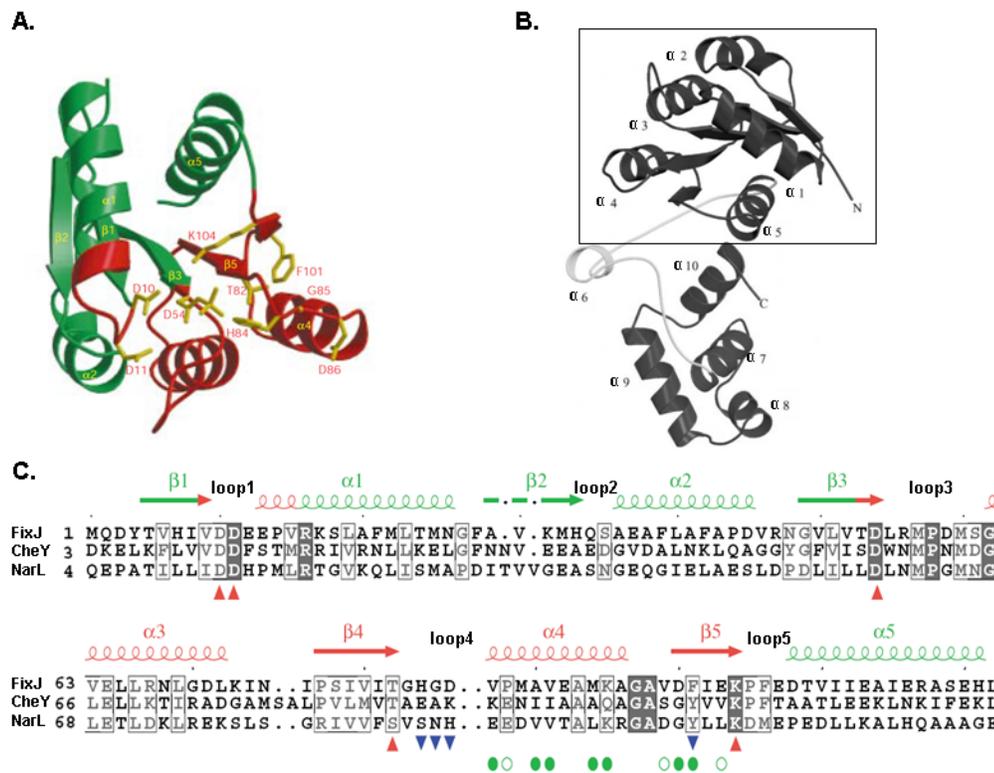


Figure 4. Structure of FixJ from *S. meliloti*.

A. Diagram of the phosphorylated receiver domain of FixJ. The regions that are significantly different in unphosphorylated and phosphorylated receiver domains are highlighted in red and the remaining part is highlighted in green. Reprint from (Birck *et al.*, 1999). B. Ribbon diagram of the FixJ model. The receiver domain is highlighted with black outline, the linker region is in gray and the remaining part is the DNA-binding domain. The helix-turn-helix motif is in the region of $\alpha 8$ - $\alpha 9$. Reprint from (Birck *et al.*, 2002). C. Sequence alignment of receiver domains of FixJ with *E. coli* CheY and *E. coli* NarL. The secondary structure of FixJ is marked in the figure, which is in the same color as (A). The blue triangles indicate the residues that move more than 4 Å upon phosphorylation. Residues that are involved in dimer formation of the phosphorylated FixJ are indicated by solid green circles (large contribution) and open green circles (small contribution). Reprint from (Birck *et al.*, 1999).

The conformational changes between unphosphorylated and phosphorylated FixJ (Figure 4A) involves a large change in the surface region of $\alpha 3$, $\beta 4$, $\alpha 4$, $\beta 5$, $\alpha 5$ as well as the loop regions between these structure elements with the major structural changes in the loop4. The surface of $\alpha 4$ - $\beta 5$ mediates dimerization of

phosphorylated FixJ. The surface of $\alpha 4$ - $\beta 5$ - $\alpha 5$ has been identified to be involved in the phosphorylation-regulated interactions, which control the activity of the C-terminal DNA-binding domain. In the unphosphorylated state of FixJ, the N-terminal receiver domain inhibits the C-terminal DNA-binding activity, whereas in the phosphorylated form, the N-terminal receiver domain releases the inhibition for the DNA-binding domain (Birck *et al.*, 1999, Birck *et al.*, 2002).

The majority of output domains in RRs have DNA-binding activity including OmpR-like, NarL-like, LytR-like, PrrA-like and YesN-like. Some output domains have RNA-binding activity such as NasR-like. Different output domains with enzymatic function have been identified such as CheB (methyltransferase), GGDEF domain (diguanylate cyclase), EAL domain (c-di-GMP phosphodiesterase), HD-GYP domain (c-di-GMP phosphodiesterase) and PP2C domain (protein phosphatase). Some RRs have CheW domains as output domain involved in protein-protein interaction. Some RRs like CheY do not have an output domain and they generate an output response through protein-protein interaction (Galperin, 2006).

In Grebe and Stock's study, RRs were divided into different families (A- H) based on multiple sequence alignments of the receiver domains (Grebe & Stock, 1999). RRs also could be divided into different families based on the different function of the output domains (Galperin, 2006).

2.1.4 Architecture of TCS pathways

In addition to the mentioned simple TCS and linear phosphorelay (Figure 1), some other TCS pathways are branched. The TCS pathways with a branched structure either have a one-to-many architecture or many-to-one architecture. One-to-many is illustrated by the chemotaxis system of *E. coli*, in which CheA phosphorylates two RRs, CheB (control the methyltransferase activity) and CheY (control the motility through binding to the flagellar motor) (Li *et al.*, 1995). Another example illustrating this architecture is HPK SLN1 in *S.cerevisiae*, which phosphorylates the Hpt protein YPD1. Subsequently the phosphoryl group is transferred to two different RRs, SSK1 and SKN7 (Li *et al.*, 1998).

One example of a many-to-one architecture is the sporulation pathway of *B. subtilis*. Five sensor kinases (KinA, KinB, KinC, KinD and KinE) converge on the

same RR, Spo0F, from which the phosphoryl group is subsequently transferred to the Hpt protein Spo0B and then to the final RR Spo0A, which controls the entry into sporulation (Jiang *et al.*, 2000). Another example of this architecture is the quorum sensing systems of *Vibrio harveyi*. At low cell density, three hybrid HPKs (LuxN, LuxQ and CqsS) sense three different autoinducers, autophosphorylate, transfer the phosphoryl group to their own receiver domains, subsequently they transfer the phosphoryl group to the same Hpt domain protein LuxU. Finally, the phosphoryl group is passed to the RR LuxO to inhibit the quorum-sensing regulon. In contrast, at high cell density, LuxN, LuxQ and CqsS switch their activity from kinase to phosphatase of LuxO (Henke & Bassler, 2004). Branched TCS pathways allow integration of signals or generation of different outputs to one signal.

2.1.5 Recognition specificity of TCS proteins

The number of TCS proteins is correlated with the genome size in most of prokaryotes (Galperin, 2005). Typically, bacteria possess many TCS proteins, for instance, *E. coli* contains 64 TCS proteins (Yamamoto *et al.*, 2005), *B. subtilis* contains 63 TCS proteins (Fabret *et al.*, 1999), *Caulobacter crescentus* contains 106 TCS proteins (Skerker *et al.*, 2005), *M. xanthus* contains 272 TCS proteins (Shi *et al.*, 2008, Whitworth & Cock, 2008). Given the large number of TCS proteins and the high sequence as well as structural similarity of TCS proteins, a challenging issue is how TCS pathways are connected, and how TCS proteins determine the recognition specificity and avoid unwanted cross-talk.

Connectivity of TCS proteins

To determine the connection between TCS proteins, genetic, biochemical and bioinformatics methods have been used. Many cognate pairs of TCS are encoded in the same operon. However, many orphan HPKs and RRs (not flanked by a cognate TCS partner in the genome, detailed description in Results) have also been identified in various genomes (Rodrigue *et al.*, 2000, Skerker *et al.*, 2005). In these cases, the genetic context does not provide information about a possible cognate partner. Many studies have used biochemical approaches *in vitro* to identify or confirm cognate TCS partners. Systematic *in vitro* phosphotransfer profiling of the interaction of TCS proteins in

C. crescentus (Skerker *et al.*, 2005) and *E. coli* (Yamamoto *et al.*, 2005) showed that HPK displays a kinetic preference *in vitro* for their cognate RR *in vivo* compared to other RRs.

Bioinformatics' studies have used *in silico* methods to determine interactions between HPKs and RRs. Grebe and Stock (Grebe & Stock, 1999) divided HPKs into 11 subfamilies HPK1-Hpk11 and RRs into 8 subfamilies A-H using multiple sequence alignment and phylogenetic analyses of kinase domains (HisKA/Hpt and HATPase_c domains) of HPK as well as receiver domains of RR. Certain subfamilies of HPK tend to interact with certain subfamilies of RR. But the specific interaction of individual HPKs and RRs is still difficult to determine by this method.

Recently, an *in silico* study tried to determine cognate TCS partners by identifying specific residues that covary between HPKs and RRs via analysis of 2500 paired TCS proteins from approximately 200 genomes (White *et al.*, 2007). A log-likelihood scoring procedure was applied to these residues and then a predictive tool was built for assigning signaling mate (White *et al.*, 2007). However, this method might be limited by lacking structure correction of HPK-RR complex.

Recognition specificity of TCS proteins

To identify the recognition specificity of TCS proteins, a few studies have used structural analyses. So far, no crystal structure of a HPK-RR complex has been solved except the low resolution X-ray crystallography structure of *Thermotoga maritima* ($\Delta 408\text{ThkA}$)₂/ 2TrrA (Yamada *et al.*, 2006). However, the resolution of this complex is too low to determine the interaction surface. The interaction between a HPK-RR complex can be exemplified by the structure of the Spo0B-Spo0F complex (Zapf *et al.*, 2000). Spo0B contains an N-terminal Hpt domain with an α -helical hairpin and a C-terminal domain with an $\alpha\beta$ fold. The Hpt domain of Spo0B forms a dimer with a four-helix bundle with two His residues ($\alpha 1$, $\alpha 2$, $\alpha 1'$ and $\alpha 2'$) (Figure 5A), which is similar to the HisKA domain of TM0853 (Figure 2A). Spo0F is a single receiver domain protein and it has a $(\beta\alpha)_5$ fold with 5 β - α loops like other RR (Figure 5B). Both of the active sites in Spo0B dimer bind one Spo0F in the crystal complex (Figure 5C).

Utsumi *et al.*, 1989). In a recent study (Skerker *et al.*, 2008), a covariation approach was used to identify the residues in HisKA domains likely to interact specifically with receiver domains. Then the identified covarying residues were corrected by using the structure of Spo0F-Spo0B complex. Predictions were tested experimentally *in vitro* as well as *in vivo*. This study indicates that the recognition specificity of HPKs lie in the C-terminal end of α 1 helix and the loop region between two α helices in the HisKA domain.

Avoiding unwanted cross-talk of TCS

Another challenge for bacteria with many TCS is to avoid unwanted cross-talk between different TCS pathways. Generally, the kinetic preference of HPK for its cognate RR is a fundamental mechanism to maintain the specificity of TCS pathways. Many studies reveal that HPK exhibits a kinetic preference *in vitro* for its cognate RR. One example is VanS, which exhibits a 10^4 fold higher kinetic preference for its cognate RR VanR than that for PhoB (Fisher *et al.*, 1996). Another example is that KinA has a more than 50,000-fold preference for its cognate RR Spo0F than for Spo0A (Grimshaw *et al.*, 1998).

In addition to the kinetic preference, a large number of HPKs display a bifunctional activity on its cognate RR. These bifunctional HPKs can act either as a kinase or as a phosphatase on the cognate RR to prevent unwanted cross-talk (Laub & Goulian, 2007, Alves & Savageau, 2003). One of the examples is the PhoR-PhoB and the VanS-VanR TCS in *E. coli*. In the wild type, no cross-talk of these two systems has been observed. However, PhoR could cross talk with VanR in the absence of VanS and VanS could cross talk with PhoB in the absence of PhoR (Haldimann *et al.*, 1996, Fisher *et al.*, 1995). The phosphatase activity of unstimulated VanS and unstimulated PhoR on their cognate RR diminishes the cross-talk between these two systems (Haldimann *et al.*, 1997).

Other mechanisms have been identified to reduce the cross-talk involve the differential spatial localization of TCS proteins. Also, different times of expression of different TCS proteins may contribute to reduce cross-talk (Laub & Goulian, 2007).

2.2 Development of *Myxococcus xanthus*

2.2.1 Life cycles of *M. xanthus*

Myxococcus xanthus, a gram-negative δ -proteobacterium, is characterized by social behaviour and two complicated life cycles, growth and development (Shimkets, 1990, Downard *et al.*, 1993, Dworkin, 1996). In the presence of nutrients the motile, rod-shaped vegetative cells grow and divide. When nutrients are limiting, growth of the cells ceases and the developmental program is triggered. The cells change their movement pattern from expansive swarming to aggregation. After 24 hrs of starvation, approximately 10^5 cells have aggregated to form a multicellular structure called a fruiting body (Figure 6). During the process of aggregation, 65% to 90% of the cells undergo autolysis, where the level of autolysis depends on the starvation conditions (Dworkin, 1996). If starvation continues, the rod-shaped cells inside the fruiting bodies differentiate into spherical, non-motile, and dormant spores, which are resistant to various physical or chemical stresses. While the aggregation process takes about 24 hrs to complete, the spore-maturation is concluded 72 hrs after initiation of starvation (Figure 6). When nutrients become available, the spores germinate and a vegetative life cycle resumes. The developmental life cycle is controlled by a number of intercellular signal transduction pathways and also depends on two motility systems, the A-motility and S-motility systems.

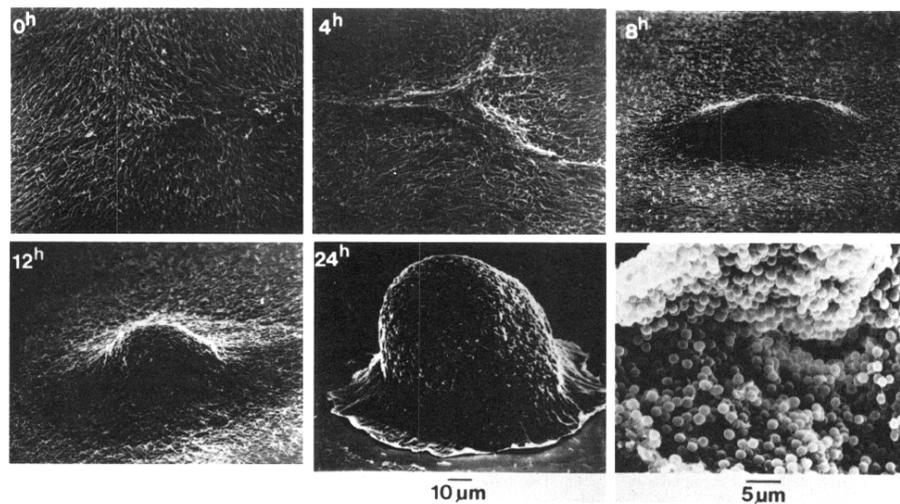


Figure 6. Scanning electron micrographs of *M. xanthus* cells during development.

The first five photos display the formation of a fruiting body. The last photo exhibits the spores inside of a fruiting body after 72hrs of development. Reprint from (Shimkets, 1990).

2.2.2 Intercellular signal transduction pathways during development

At the onset of development, starvation leads to a transient increase in the intracellular guanosine-5'-(tri)di-3'-diphosphate ((p)ppGpp) level in a RelA-dependent manner (Harris *et al.*, 1998). The accumulation of (p)ppGpp initiates the developmental program (Manoil & Kaiser, 1980, Harris *et al.*, 1998). At least five intercellular signals, known as the A-signal, B-signal, C-signal, D-signal and E-signal, are required for the developmental process (Downard *et al.*, 1993, Hagen *et al.*, 1978). Studies on developmental gene expression in the signaling mutants showed that each signaling system begin to function at a well defined developmental time point. The A-signal and B-signal are required for normal gene expression beginning 1 to 2 hrs into development (Gill & Cull, 1986, Kroos & Kaiser, 1987, Kuspa *et al.*, 1986). The D-signal and E-signal are required at 3 to 5 hrs into development (Cheng & Kaiser, 1989, Downard *et al.*, 1993), and the C-signal is needed after 6 hrs of development (Kroos & Kaiser, 1987). Among these signals, the A-signal and the C-signal are the best studied.

The A-signal is a mixture of amino acids and peptides, which is generated by protease cleavage during the early stages of development. A-signal is produced in proportion to cell density and it serves as a quorum sensing system to ensure a sufficient cell density before fruiting body development is initiated (Kuspa *et al.*, 1992). If the cells are at a density greater than about 10^8 cells/ml, the concentration of A-signal exceeds the critical threshold concentration and subsequently the A-signal-dependent genes are expressed (Kuspa *et al.*, 1992).

The C-signal is the latest-acting signal of the five intercellular signals. It is required for aggregation, sporulation and full expression of genes induced after 6 hrs of development (Jelsbak & Sogaard-Andersen, 2000). Moreover, C-signal coordinates these events spatially and temporally (Jelsbak & Sogaard-Andersen, 2000). The C-signal is a cell surface-associated protein encoded by the *csgA* gene (Kim & Kaiser, 1990). The active C-signal is a 17-kD protein (Lobedanz & Sogaard-Andersen, 2003) which is synthesized by N-terminal proteolytic processing of the full-length 25-kD CsgA protein by a serine protease, PopC (Rolbetzki *et al.*, 2008). *csgA* is developmentally regulated by the *act* operon encoding ActA-D (Gronewold & Kaiser, 2001). Intermediate and

increased FrzCD methylation results in inhibition of the FrzE kinase activity (McCleary & Zusman, 1990). The reduction in the reversal frequency allows cells to aggregate into fruiting bodies. In the other branch, activated FruA triggers the transcription of downstream target genes, which are required for the developmental process.

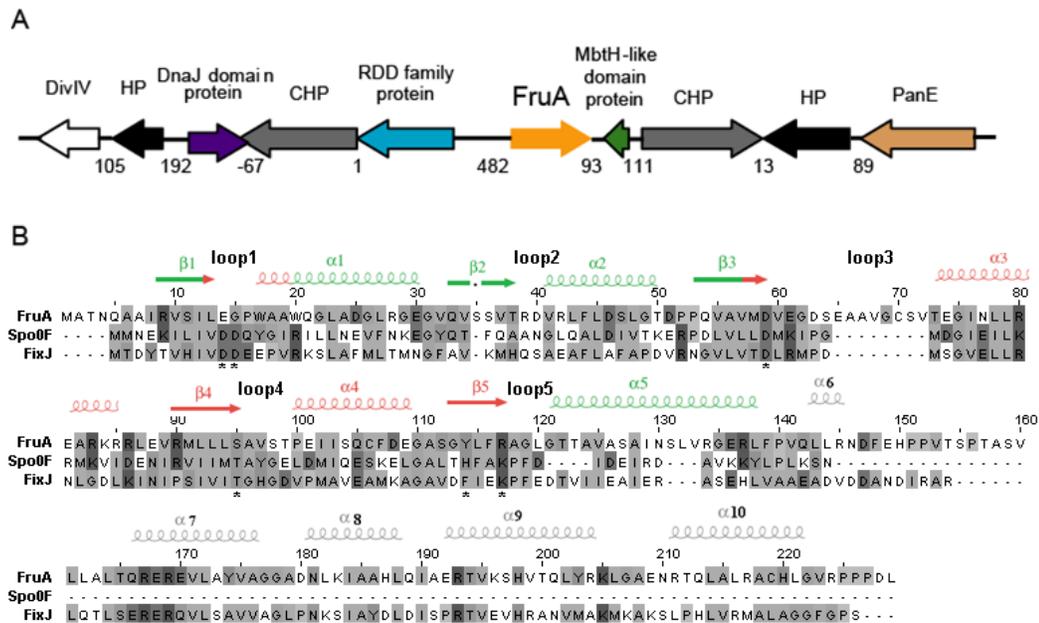


Figure 8. Flanking region of *fruA* (A) and sequence alignment of FruA (B).

In A, arrows indicate open reading frames and the direction indicates the transcriptional direction. The proteins encoded by each gene are marked on the ORFs. The numbers between ORFs indicate the bp between two genes. Negative number means an overlap between two genes. CHP: conserved hypothetical protein. HP: hypothetical protein. In B, FruA is aligned with *S. meliloti* FixJ, and *B. subtilis* Spo0F. Asterisks indicate the highly conserved residues involved in the phosphorylation or the conformational changes during phosphorylation of RR. The secondary structure of FixJ is highlighted and is in the same color as in Figure 4C, in which red indicates the regions that are significantly different in unphosphorylated and phosphorylated receiver domains, green indicate the remaining part of receiver domain and gray indicate DNA-binding domain and the link region between the receiver domain and the DNA-binding domain. The helix-turn-helix motif of the DNA-binding domain of FixJ is in the region of $\alpha 8$ - $\alpha 9$.

FruA is an orphan RR and there is no gene encoding a HPK in the region flanking *fruA* (Figure 8A). It is a homologue of *S. meliloti* FixJ containing a helix-turn-helix DNA-binding domain. The sequence characteristics of the receiver domain of FruA are unique (Figure 8B) (Ellehaug *et al.*, 1998). Of six conserved residues found in most receiver domains, FruA contains the conserved phosphorylatable Asp59 and two conserved Ser95 and Tyr114 residues. In the remaining three positions, FruA contains E14 (D in most of

RRs), G15 (D in most of RRs) and R117 (K in most of RRs). In addition, FruA contains eight additional amino acids inserted in loop 3 of the receiver domain. These eight amino acids are required for the activity of FruA (Rasmussen, 1998). A similar insertion is found in *Nostoc punctiforme* (Hagen & Meeks, 1999) and *Anabaena sp.* DevR and DevR is phosphorylated by HepK *in vitro* (Zhou & Wolk, 2003). Despite these abnormalities of FruA, the *in vivo* genetic evidences suggest that FruA is a functional RR (Ellehaug *et al.*, 1998).

Four direct DNA-binding targets of FruA, *dev* (Viswanathan *et al.*, 2007b), $\Omega 4400$ (Yoder-Himes & Kroos, 2006), *fdgA* (Ueki & Inouye, 2005b) and *dofA* (Ueki & Inouye, 2005a), have been confirmed by *in vitro* or *in vivo* experiments. The *dev* operon consists of eight genes and three of them (*devR*, *devS*, and *devT*) were characterized as important for development (Viswanathan *et al.*, 2007a, Boysen *et al.*, 2002). A Tn5 insertion mutation in *devR* reduces spore formation (Thony-Meyer & Kaiser, 1993). DevS negatively autoregulates *dev* operon expression (Viswanathan *et al.*, 2007a) and DevT stimulates transcription of *fruA* (Boysen *et al.*, 2002). However, the detailed regulation mechanism of the *dev* operon during development is still unknown. FruA and LadA, a transcriptional factor of the LysR family, act together to regulate the proper transcription of the *dev* operon (Viswanathan *et al.*, 2007b). Genetic evidence suggests that the increase in *dev* expression during development depends on phosphorylation of FruA (Ellehaug, 1999) and C-signal (Kruse *et al.*, 2001). $\Omega 4400$ expression depends on FruA and C-signal but the function of this gene remains to be identified (Yoder-Himes & Kroos, 2006). FdgA is a homologue of the outer membrane auxiliary family protein involved in polysaccharide export system (Ueki & Inouye, 2005b). Deletion of *fdgA* leads to aggregation and sporulation defects. DofA does not exhibit a significant similarity to any known protein. It is neither important for development nor for growth (Horiuchi *et al.*, 2002a). From the sequences of the promoter region of these four genes, the consensus of FruA binding region was proposed as GGGC/TA/G(N₄₋₆)C/TGGG. Many sequences match this consensus in the *M. xanthus* genome, which suggests that FruA may directly regulate many genes (Viswanathan *et al.*, 2007b).

Previous studies proposed that the phosphorylated FruA is responsible for all C-signal dependent responses (Sogaard-Andersen *et al.*, 1996, Ellehauge *et al.*, 1998). The graded increase in C-signaling levels during development has been hypothesized to induce increased levels of FruA phosphorylation, which regulate different responses. By 2D-PAGE analysis, 54 up-regulated proteins have been shown to be dependent on FruA as well as on the C-signal (Horiuchi *et al.*, 2002b). On the other hand, few proteins have been identified that are differentially synthesized in *csgA* and *fruA* mutants. These observations suggest that unphosphorylated FruA or low levels of phosphorylated FruA have a function in the regulation of FruA dependent and C-signal independent genes.

The transcription of *fruA* is up-regulated during development and is dependent on the A-signal and E-signal but independent of C-signal (Ellehauge *et al.*, 1998). Further studies (Ueki & Inouye, 2003, Nariya & Inouye, 2006) demonstrated that the *fruA* promoter is a direct DNA-binding target of the transcriptional regulator MrpC. MrpC belongs to the catabolite gene activator protein family of transcriptional regulators containing a cyclic nucleotide binding domain and a DNA-binding domain (Sun & Shi, 2001b). MrpC also activates the *mrpC* expression. MrpC exists in two forms, MrpC and MrpC2. MrpC2 lacks the N-terminal 25 residues of MrpC, which contains 2 potential phosphorylation sites. MrpC2 exhibits higher activity towards the *fruA* promoter and the *mrpC* promoter (Nariya & Inouye, 2006). It was hypothesized that MrpC undergoes LonD dependent proteolysis cleavage to generate MrpC2 (Nariya & Inouye, 2006).

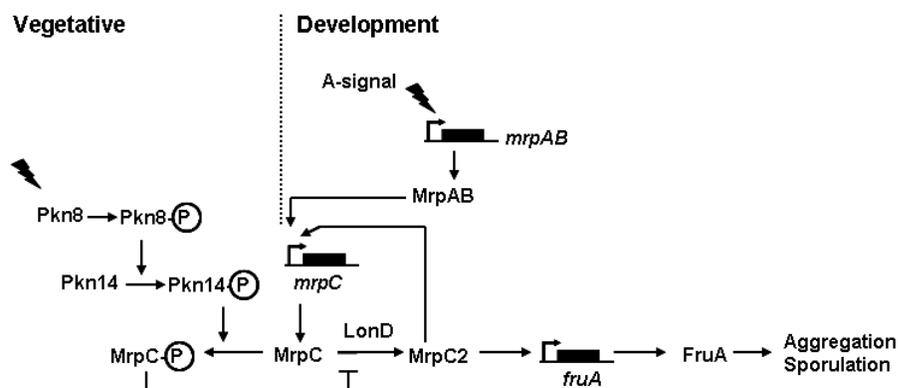


Figure 9. Regulation of *fruA* expression in *M. xanthus*.

Detailed description in the text. Modified from (Ueki & Inouye, 2006).

The proposed model for the regulation of *fruA* expression (Figure 9) is as follows (Ueki & Inouye, 2006). During vegetative growth, *mrpC* gene is transcribed at low levels. MrpC is produced and subsequently is phosphorylated by a Ser/Thr protein kinase Pkn14, which is activated by phosphorylation from Pkn8. Phosphorylation of MrpC decreases its DNA binding activity. This suggests that Pkn8-Pkn14 cascade negatively regulates *mrpC* expression to prevent initiation of development. In early development, *mrpAB* expression is activated potentially by A-signal (Sun & Shi, 2001b). MrpA and MrpB are produced and MrpB is phosphorylated by MrpA or other other phosphodonor (Sun & Shi, 2001b). Phosphorylated MrpB activates the transcription of *mrpC* and MrpC stimulates *mrpC* expression. MrpC is most likely not phosphorylated by Pkn14 during development since Pkn14 is expressed mainly during vegetative growth (Nariya & Inouye, 2005). MrpC is likely cleaved into MrpC2 by LonD. MrpC2 activates the expression of *mrpC* and *fruA*, which induce the aggregation and sporulation.

2.2.3 Motility of *M. xanthus*

During development, cell aggregation requires the coordination of motility behaviours. *M. xanthus* moves by gliding, which is powered by two motility systems, the social motility system (S-motility) and the adventurous motility system (A-motility) (Hodgkin & Kaiser, 1979). S-motility is mediated by extension and retraction of type IV pili which are localized to one pole of cells. The mechanism of A-motility remains to be elucidated. One model is that slime released from nozzle-like structures generates the force for A-motility (Wolgemuth *et al.*, 2002). Another model is that multiple force-generating adhesion complexes move on cytoskeleton filaments and function as the engine of A-motility (Mignot *et al.*, 2007). *M. xanthus* mutants lacking type IV pili still move by means of A-motility and *vice versa*.

M. xanthus cells occasionally undergo reversals to change their direction of movement. The reversal frequency is controlled by the Frz pathway, which is also required for development. The Frz system is similar to the *E. coli* chemotaxis system and it consists of FrzCD (a methyl-accepting chemotaxis protein homologue), FrzF (a CheR methyltransferase homologue), FrzG (CheB methylesterase homologue), FrzE (a hybrid CheA and CheY homologue) as well

as FrzA and FrzB (2 CheW-like proteins), which link FrzCD and FrzE (Ward & Zusman, 1999, Inclan *et al.*, 2007, Inclan *et al.*, 2008). FrzCD methylation influences the phosphorylation state of FrzE (Li *et al.*, 2005). Kinase activity of FrzE stimulates cellular reversals and inactive FrzE inhibits cellular reversals (Zusman *et al.*, 2007). C-signal induces FrzCD methylation via FruA during development (Sogaard-Andersen & Kaiser, 1996). Moreover, C-signal inhibits cell reversals suggesting that C-signal inhibits the activity of the Frz system (Jelsbak & Sogaard-Andersen, 1999).

2.3 TCS in *M. xanthus*

Previous studies have shown that 35 TCS proteins are required for motility or development of *M. xanthus* (Table S1, S2 and S3). Most of these characterized TCS proteins in *M. xanthus* were identified by random mutagenesis. Some of the proteins are encoded by adjacent genes and likely constitute typical TCS systems, which are composed of one HPK and one RR with an output domain, such as MrpA-MrpB (Sun & Shi, 2001b), PhoR1-PhoP1 (Carrero-Lerida *et al.*, 2005), PhoR2-PhoP2 and PhoR3-PhoP3 (Moraleta-Munoz *et al.*, 2003).

However, a large proportion of the characterized TCS proteins are orphan, i.e. these proteins are encoded by genes that are not flanked by a TCS gene. The interaction partners of these TCS proteins still remain to be identified. Also, some orphan HPKs are important for development, such as SdeK (Garza *et al.*, 1998), EspA (Cho & Zusman, 1999b) and AsgD (Cho & Zusman, 1999a). Some orphan RR are involved in development, such as FruA (Ogawa *et al.*, 1996, Ellehaug *et al.*, 1998) and Nla26 (Caberoy *et al.*, 2003, Kirby & Zusman, 2003), or in motility, for instance AglZ (Yang *et al.*, 2004), FrzS (Ward *et al.*, 2000), RomR (Leonardy *et al.*, 2007) and DigR (Overgaard *et al.*, 2006, Youderian & Hartzell, 2006). These studies indicate that TCS proteins in *M. xanthus* may be involved in complicate signaling pathways. However, none of the past studies have systematically analyzed TCS proteins in *M. xanthus*. To obtain an overview of TCS proteins in *M. xanthus*, a general analysis of TCS proteins was performed in this study.

By now, none of the interaction partners of orphan TCS proteins in *M. xanthus* have been identified. Here, I focused on FruA, which is an essential RR in the C-signal transduction pathway. FruA is an orphan RR which controls aggregation and sporulation during development. However, no cognate HPK of FruA has been identified. To further understand the C-signal transduction pathway, I used a candidate approach to identify FruA kinase candidates based on the hypothesis that a FruA kinase gene shares key characteristics with the *fruA* gene, i.e. it is orphan, developmentally up-regulated at the transcriptional level and deficient in development in a null mutant. Moreover, yeast two-hybrid (Y2H) analysis was performed to investigate potential interactions between FruA and these developmentally regulated orphan HPKs. Identified FruA kinase

candidates were further analyzed *in vivo* and *in vitro*. In addition, the FruA kinase candidates were predicted by our collaborators using an *in silico* method (White *et al.*, 2007). However, *in vivo* experiments strongly suggest that they are not FruA kinase(s).

3 Results

3.1 An analysis of two-component regulatory systems in *M. xanthus*

3.1.1 Identification of TCS genes in *M. xanthus*

To analyze TCS in *M. xanthus*, TCS proteins were retrieved from the *M. xanthus* proteome (Materials and Methods). 272 genes that encoding TCS proteins were identified including 21 genes in eight loci, which code for TCS proteins in Che-like systems, i.e. these clusters encode homologues of Che proteins and the HPK has a domain structure similar to that found in CheA (See Table S1 and S3 for a list of these genes). In this work, I specifically focused on the 251 TCS genes encoding proteins that are not part of Che-like systems. These 251 genes code for 118 HPK, 119 RR and 14 HPK-like genes (Table 1).

Table 1. Summary of two-component system genes and proteins in *M. xanthus*

	Total	Paired genes	Orphan genes	Complex gene clusters
HPK genes/proteins (hybrid HPK)	118 (31)	32 (1 hybrid)	68 (22 hybrid)	18 (8 hybrid)
RR genes/proteins	119	37	64	18
HPK-like genes/proteins	14	5	6	3
Total TCS genes/proteins	251	74	138	39

HPKs contain a conserved HATPase_c domain and a HisKA domain with the phosphorylatable His residue and HPK genes are proposed to encode *bona fide* kinases. RR either consists of a single receiver domain or is multidomain proteins containing an output domain in addition to the receiver domain and RR genes are predicted to encode *bona fide* RRs. HPKs that contain one or more receiver domains were classified as hybrid HPK. HPK-like genes encode proteins that either contain a HisKA domain with the phosphorylatable His residue and lack the HATPase_c domain or *vice versa*. The HATPase_c domain is also found in other ATPases such as DNA gyrase B and DNA repair protein MutL (Galperin, 2005). To avoid annotating such proteins as HPK-like proteins, proteins were only classified as HPK-like proteins if they are encoded by genes located next to a TCS gene (e.g. *MXAN0461* and *MXAN4203*) or if they contain one or more receiver domains (e.g. *MXAN0230* and *MXAN4432*) (See Table S2 for the domain structure of these proteins).

Given the large number of TCS genes in *M. xanthus*, it was determined whether the kinase and receiver domains belong to the established kinase and receiver families as defined by Grebe and Stock (Grebe & Stock, 1999). Therefore, all kinase domains (encompassing the HisKA and HATPase_c domains) and receiver domains (in RR and in hybrid HPKs and in HPK-like proteins) were aligned with those used by Grebe and Stock (Grebe & Stock, 1999), and phylogenetic trees were generated (done by Stuart Huntley, data not shown). Analysis of the alignments and trees showed that all the kinase and receiver domains in the *M. xanthus* TCS proteins belong to the established Grebe and Stock families (See Table S1 and Table S3 for family assignment). Thus, despite the large number of TCS proteins, there is no evidence that *M. xanthus* has evolved new families of kinase or receiver domains. Rather *M. xanthus* appears to use domains also found in other species.

3.1.2 Genetic organization of TCS genes in *M. xanthus*

To analyze the genetic organization of TCS genes, a set of criteria was developed according to which TCS genes were divided into three categories: paired genes, complex gene clusters and orphan genes (Figure 10). Two adjacent genes encoding an HPK and a RR, or an HPK-like protein and a RR and transcribed in the same direction were grouped into paired genes. TCS genes in gene clusters containing two or more RR genes, two or more HPK or HPK-like genes, or three or more TCS genes independent of their transcriptional direction, were grouped into complex genes. All other gene organizations of TCS genes were grouped into orphan genes. The genetic organization of TCS genes in *M. xanthus* diverges significantly from the standard description of TCS genes (Table 1, Figure 11, Figure 12, and Figure 13): 55% (138 genes out of 251 total) of TCS genes are orphans, 16% (39 out of 251) of TCS genes are located in complex gene clusters, and only 29% (74 out of 251) are found as paired genes. A comparison with other bacterial genomes illustrates that the percentage of orphan TCS genes shows large inter-species variations. For example, in *C. crescentus* 57% of the 106 TCS genes are orphans (Skerker *et al.*, 2005) whereas 14% of the 36 HPK in *B. subtilis* genes are orphans (Fabret *et al.*, 1999). Thus, in terms of the percentage of orphan TCS genes, *M. xanthus* is not exceptional.

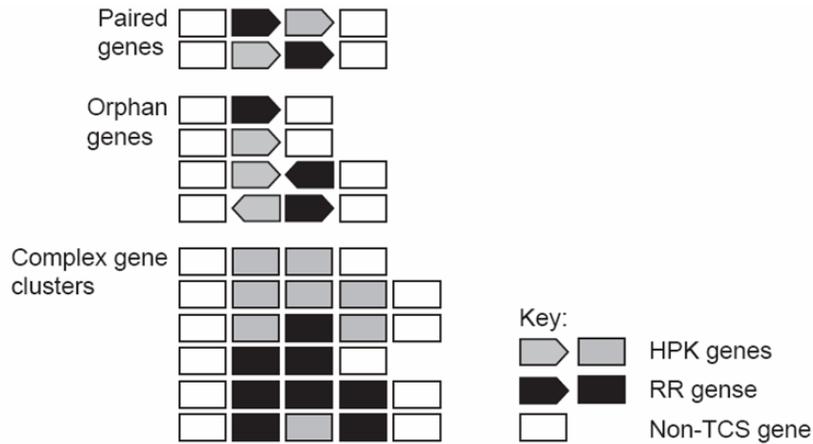


Figure 10. Classification scheme for two-component regulatory system genes.

Schematic diagram of classification schemes for TCS genes. The transcription direction is indicated by the arrow symbols. The definitions of different groups are as follows. Two adjacent genes encoding an HPK, or an HPK-like protein and a RR and transcribed in the same direction were grouped into paired genes. Complex TCS gene clusters include clusters containing two or more RR genes, two or more HPK or HPK-like genes, and clusters with three or more TCS genes irrespectively of transcription direction as indicated by the box symbols. All other gene organizations of TCS genes were grouped into orphan genes. For complex gene clusters only the most common gene organizations are shown (See Table S1, S2 and S3 for all gene organizations found in these clusters).

3.1.3 Histidine protein kinases

As a first step to understand how *M. xanthus* TCS proteins are connected, the 118 HPKs were divided into HPKs and hybrid HPKs, which also contain one or more receiver domains in addition to the HisKA and HATPase_c domains. 31 hybrid kinases (26% of the total number of HPKs) were identified, which contain between one to three receiver domains (Table 1, Figure 11, see Table S1 for detailed domain organization). The distribution of the 31 hybrid HPKs is highly biased. Only one hybrid HPK is encoded by a paired gene (corresponding to 3% of all HPKs encoded by these genes), 22 of the hybrid HPKs are encoded by orphan genes (corresponding to 32% of all HPKs encoded by these genes) and 8 of the hybrid HPKs are encoded by complex gene clusters (corresponding to 44% of all HPKs encoded by these genes). Among the 31 hybrid HPKs only one, MXAN2317, is predicted to contain an Hpt domain.

HPKs are generally described as being integral membrane proteins (Stock *et al.*, 2000). To determine whether HPKs in *M. xanthus* follow this general description, all of HPKs were divided into two groups based on the presence or

absence of trans-membrane helices. The presence of trans-membrane helices indicates that the HPK is an integral inner membrane protein. Among the 118 HPKs, 45 (38% of the total number of HPKs) are likely to be integral membrane proteins and the remaining 73 are likely to be cytoplasmic (Table 2, Figure 11, see Table S1 for detailed results). Also in this analysis a biased distribution of the two types of HPKs was found. Thus, only 8 of the predicted cytoplasmic HPKs are encoded by paired genes (corresponding to 25% of all HPKs encoded by these genes), 54 of the predicted cytoplasmic HPKs are encoded by orphan genes (corresponding to 79% of all HPKs encoded by these genes), and 11 of the predicted cytoplasmic HPKs are encoded by complex gene clusters (corresponding to 61% of all HPKs encoded by these genes). The large fraction of HPKs encoded by paired HPKs which are integral membrane proteins suggests that these systems are primarily involved in monitoring external stimuli. Conversely, a high percentage of orphan HPKs are cytoplasmic suggests these systems are mainly involved in sensing internal stimuli.

Table 2. Summary of potential localization of HPKs encoded in the *M. xanthus* genome

HPKs	Total	HPKs encoded by paired genes	HPKs encoded by orphan genes	HPKs encoded by complex gene clusters
Inner membrane	45	24	14	7
Cytoplasmic	73	8	54	11

3.1.4 Histidine protein kinase-like proteins

In addition to 118 HPKs, 14 HPK-like proteins, which only contain a HisKA domain, a HATPase_c domain fused with a receiver domain or a HATPase_c domain chromosomally adjacent to other TCS proteins (See Table S2, Figure 12), were identified in *M. xanthus* genome. Five of them (4 are cytoplasmic and 1 are membrane-integral) are encoded by paired genes, 6 of them (5 are cytoplasmic and 1 are membrane-integral) are encoded orphan genes and 3 of them (all cytoplasmic) are encoded by genes in complex gene clusters. Three of these HPK-like proteins, RedE (Higgs *et al.*, 2005), AsgA (Plamann *et al.*, 1995) and MrpA (Sun & Shi, 2001a, Sun & Shi, 2001b), are required for development, which indicates that HPK-like proteins may function in *M. xanthus*.

3.1.5 Response regulators and output domains

As a second step to understand the connectivity of TCS in *M. xanthus*, the RRs were divided into two groups: single domain RRs, which only consist of the receiver domain, and multidomain RRs, which in addition to the receiver domain contain an output domain. The 119 RRs can be divided into 38 without (32% of the total number of RRs) and 81 with output domains (Table 3, Figure 13, see Table S3 for the detailed domain organization). All *che* gene clusters either contain a CheA hybrid kinase or CheY homologs (See Table S1 and S3 for the detailed description of proteins encoded by *che* gene clusters). Thus, the remaining single domain RRs are not likely to be CheY paralogs.

Table 3. Summary of the output domains in RRs encoded in the *M. xanthus* genome

Output domains	Total	RRs encoded by paired genes	RRs encoded by orphan genes	RRs encoded by complex gene clusters
No output domain	38	3	29	6
DNA-binding	50	32	13	5
GGDEF	10	0	5	5
PilZ	2	0	2	0
Other domains ^a	5	0	5	0
DUF ^b	14	2	10	2

Other domain^a: defined domains with known function. ^bDUF: Domains of unknown function.

The distribution of RRs without output domains is highly biased. For the 37 RRs encoded by paired genes, 34 contain an output domain and only three do not have an output domain (corresponding to 9%). The 34 output domains comprise 32 DNA-binding domains and 2 DUF domains. The 64 orphan RR genes include 35 encoding RR with and 29 without output domains (corresponding to 45%). The four largest categories of output domains for the orphan RRs comprise 13 DNA-binding domains, 5 GGDEF domains, 2 PilZ domains that bind c-di-GMP (Romling & Amikam, 2006, Ryjenkov *et al.*, 2006), and 10 DUF domains. Finally, the complex gene clusters include 18 RR genes encoding 12 RRs with and 6 RRs without output domains (corresponding to 33%). The output domains comprise 5 DNA-binding domains, five GGDEF domains, which

are involved in c-di-GMP synthesis and regulation (Romling & Amikam, 2006), and 2 domains of unknown function (DUF). The distribution of the output domains suggests that the two main outputs from TCS in *M. xanthus* are gene regulation (50 domains) and c-di-GMP regulation (12 domains). RRs with DUF domains are overrepresented among RRs encoded by complex gene clusters and orphan genes. The high fraction of DNA-binding domains in paired RRs indicated that the paired TCS are mainly involved in regulation of gene expression. While 45% of all orphan RRs without any output domains indicate that their functions are unknown and the rest orphan RRs with output domains function as regulation of gene expression, regulation of cyclic-di-GMP metabolism, and unknown functions.

3.1.6 Transcriptional regulation of TCS genes during development

35 TCS genes have been identified as important for fruiting body formation and sporulation in previous studies (See Table S1, Table S2 and Table S3 for the specific genes). Several of these genes, such as *sdeK*, *fruA* and *espA*, are developmentally regulated at the transcriptional level during development. Thus, it was reasoned that one approach to identify TCS proteins that may function in fruiting body formation would be to analyze the expression profiles of TCS genes during development. The *M. xanthus* DNA microarray data (<http://www.ncbi.nlm.nih.gov/geo/>), which was carried out using total RNA isolated at 9 time points (0, 2, 4, 6, 9, 12, 15, 18 and 24 hrs) of development in wild type DK1622, were used for further analyses. Genes called to be significantly regulated during development were selected by a delta value of the SAM analysis where the false discovery rate became 0% in combination with a 1.5-fold cutoff and data points for all time points. Among the 200 TCS genes present on the array, 50 displayed altered expression during development (Figure 11, Figure 12 and Figure 13; Table S1, S2, and S3 for details on the expression of individual genes). Developmentally regulated genes exhibited expression ratios in the range of 10.9 fold up-regulation to 6.6 fold down-regulation.

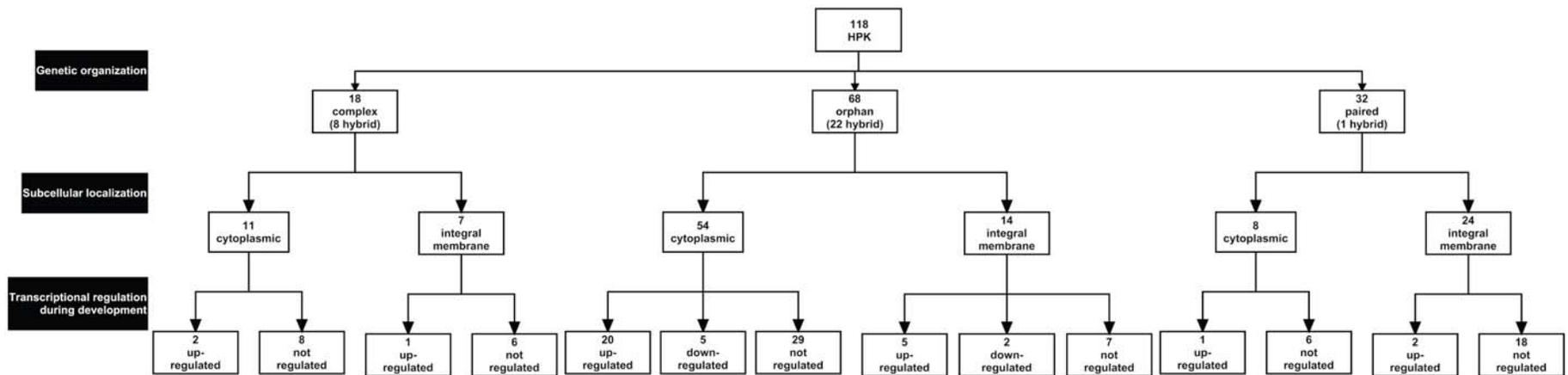


Figure 11. Organization diagram of histidine protein kinase genes and proteins in the *M. xanthus* genome.

HPK genes were divided into three categories based on genetic organization. The number of hybrid HPKs is indicated in brackets for each category. In the second layer, HPKs are divided according to their likely subcellular localization based on the presence (integral membrane) or absence (cytoplasmic) of trans-membrane helices. In the third layer, HPK genes are divided according to transcriptional regulation during development. Note that in the third layer, not all HPK genes are included because they are not all present on the *M. xanthus* DNA microarray or they were not tested by qRT-PCR.

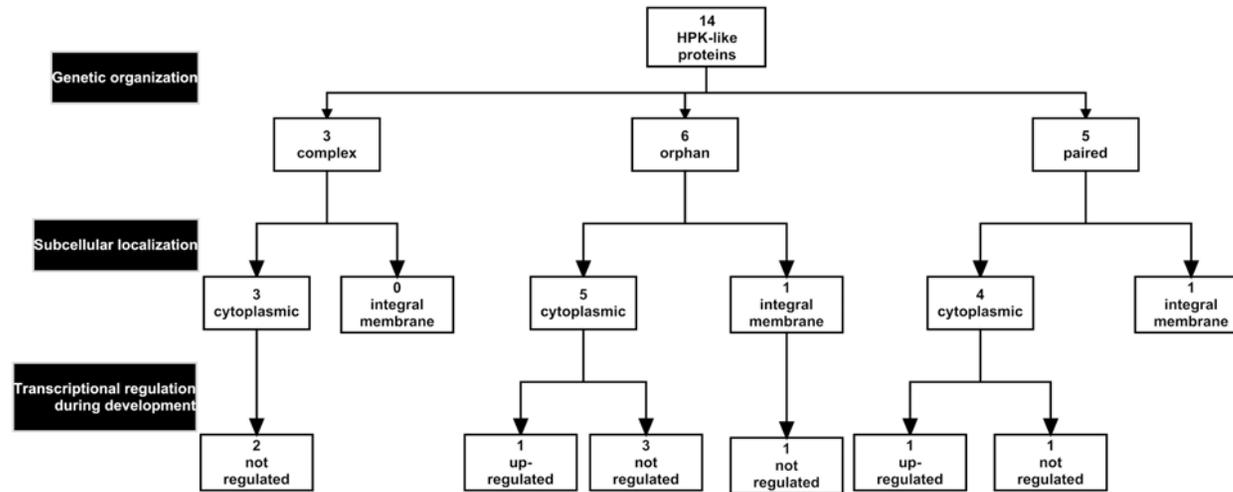


Figure 12. Organization diagram of histidine protein kinase-like genes and proteins in the *M. xanthus* genome.

Genes encoding HPK-like proteins were divided into three categories based on genetic organization. In the second layer, HPK-like proteins are divided according to their likely subcellular localization based on the presence (integral membrane) or absence (cytoplasmic) of trans-membrane helices. Note that in the third layer, not all HPK-like genes are included because they are not all present on the *M. xanthus* DNA microarray or they were not tested by qRT-PCR.

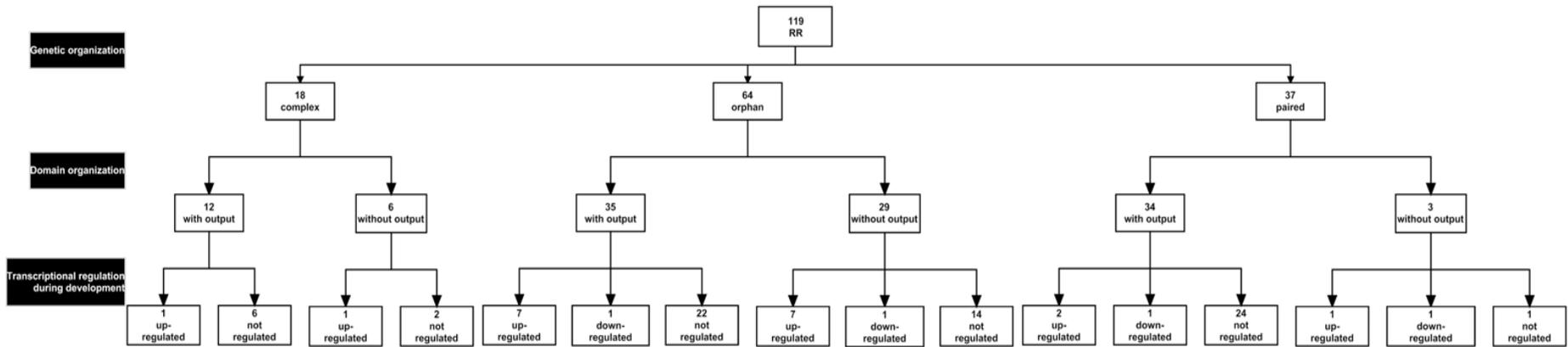


Figure 13. Organization diagram of response regulator genes and proteins in the *M. xanthus* genome.

RR genes were divided into three categories based on genetic organization. In the second layer, RRs are divided according to the presence or absence of output domains. In the third layer, RRs are divided according to transcriptional regulation during development. Note that in the third layer, not all RR genes are included because they are not all present on the *M. xanthus* DNA microarray or they were not tested by qRT-PCR.

The changes in gene expression during development were asymmetric: 46 genes were up-regulated and four genes were down-regulated. To validate the significance of the expression data obtained from the DNA microarrays, qRT-PCR was applied to 11 genes (10 genes up-regulated during development and one gene not showing regulation). The transcriptional differences determined in the microarray experiments were confirmed by the qRT-PCR analysis (See Table S1, Table S2 and Table S3 for details of qRT-PCR data).

To generate a complete data set on the expression profile for an entire category of TCS genes, I tested the expression by qRT-PCR during development of the 13 orphan HPK genes for which no expression data were available from the DNA microarray experiments. Among these genes, six were up-regulated (Table 4) and seven were down-regulated during development (see Table S1 for details of the qRT-PCR data).

In total, 63 TCS genes out of 213 TCS genes tested in either microarray analyses or by qRT-PCR were transcriptionally regulated during development (Figure 11, Figure 12, and Figure 13; Table S1, S2, and S3 for the expression of individual genes). 52 genes were up-regulated and 11 genes down-regulated. 36% and 13% of the tested orphan and complex genes respectively were transcriptionally regulated during development whereas only 12% of the tested paired genes were transcriptionally regulated during development. The higher percentage of developmentally regulated orphan TCS genes compared to that of complex or paired TCS genes indicates that the orphan TCS genes could be more important for development than the complex or paired TCS genes.

3.1.7 Genetic analysis of transcriptionally up-regulated, orphan *hpk* genes

The gene expression profiling experiments indicate that orphan TCS may have important functions in fruiting body formation. To test this hypothesis and to potentially identify novel TCS genes required for development, I focused on the 25 orphan HPK genes that are transcriptionally up-regulated during development (see Table 4 for a list of all 25 genes incl. expression data). These 25 genes include *espA* (Cho & Zusman, 1999b), *sdeK* (Garza *et al.*, 1998), *espC* (Lee *et al.*, 2005), *asgD* (Cho & Zusman, 1999a) and *mokA* (Kimura *et al.*, 2001), which have previously been shown to be important for development.

Table 4. 25 developmentally up-regulated orphan *hpk* genes

TIGR_MXAN	Gene symbol	Max. expression ratio (microarray) /Induction time (hrs) ^{a, b}	Max. expression ratio (qRT-PCR) /Induction time (hrs) ^{a, b}	Dev. defects in mutant
<i>MXAN0060</i>	<i>hpk5</i>	NOA	19.3xup/0-6	No
<i>MXAN0176</i>	<i>hpk36</i>	NOA	14.8xup/6-12	No
<i>MXAN0340</i>	<i>hpk11</i>	2.7xup/4-6	ND	No
<i>MXAN0571</i>	<i>hpk25</i>	1.9xup/2-4	ND	No
<i>MXAN0706</i>	<i>hpk39</i>	NOA	3.4xup/0-6	No
<i>MXAN0712</i>	<i>hpk37</i>	1.6xup/12-15	ND	Yes
<i>MXAN0736</i>	<i>hpk12</i>	2.2xup/0-2	128xup/0-6	Yes
<i>MXAN0928</i>	<i>hpk14</i>	10.9xup/2-4	29.7xup/0-6	No
<i>MXAN0931</i>	<i>espA</i>	4.4xup/0-2	ND	Yes
<i>MXAN1014</i>	<i>sdeK</i>	2.6xup/0-2	ND	Yes
<i>MXAN2763</i>	<i>hpk18</i>	2.4xup/0-2	ND	No
<i>MXAN3036</i>	<i>hpk23</i>	2.1xup/4-6	29.8xup/0-6	essential
<i>MXAN3098</i>	<i>hpk24</i>	1.9xup/0-2	4.1xup/0-6	No
<i>MXAN3290</i>	<i>hpk8</i>	2.8xup/0-2	4.4xup/0-6	Yes
<i>MXAN4465</i>	<i>hpk30</i>	2.1xup/4-6	ND	Yes
<i>MXAN4988</i>	<i>hpk27</i>	1.6xup/0-2	2.8xup/0-6	essential
<i>MXAN5483</i>	<i>hpk20</i>	2.2xup/4-6	ND	No
<i>MXAN6315</i>	<i>hpk9</i>	4.2xup/0-2	36.5xup/0-6	No
<i>MXAN6855</i>	<i>espC</i>	1.7xup/12-15	ND	No
<i>MXAN6941</i>	<i>hpk34</i>	NOA	7.5xup/6-12	No
<i>MXAN6994</i>	<i>hpk15</i>	4.4xup/2-4	76.4xup/0-6	No
<i>MXAN6996</i>	<i>asgD</i>	1.8xup/12-15	ND	Yes
<i>MXAN7123</i>	<i>hpk13</i>	4.0xup/2-4	143xup/0-6	No
<i>MXAN7206</i>	<i>mokA</i>	NOA	9.2xup/0-6	No
<i>MXAN7398</i>	<i>hpk29</i>	NOA	3.8xup/0-6	No

^a Expression ratios were calculated as the expression in developing cells over the expression in vegetative cells. Maximum expression ratios indicate the maximum expression ratio for a particular gene during all time points tested using DNA microarrays (0, 2, 4, 6, 9, 12, 15, 18 and 24 hrs) and qRT-PCR (0, 6, 12, 18, 24 hrs). Induction time indicates the time interval in which the expression ratio began to change relative to that observed in vegetative cells; ^b NOA: Gene not represented or not detected on DNA microarray; ND, not determined; NA, not applicable.

To analyze the importance of the remaining 20 genes for fruiting body formation, I sought to generate in-frame deletions in 19 of these genes using a two-step recombination procedure in the wild type DK1622 (Material and Methods); for *MXAN0060* an insertion mutant was created. In-frame deletions were preferred over insertion mutants to avoid polar effects on downstream genes. In addition, I generated or obtained in-frame deletions of the five previously identified orphan HPKs important for development in order to systematically compare developmental defects. For the genes *MXAN3036* and *MXAN4988* galactose resistant clones (Materials and Methods) consistently contained the intact HPK gene. These observations suggest that *MXAN3036* and *MXAN4988* are essential genes for viability. For the remaining 22 genes, stable in-frame deletion mutants were obtained.

Next, the phenotypes of the wild-type strain (DK1622) and the 23 HPK mutants were examined. To test the mutants for developmental defects, cells were exposed to starvation under three different conditions, i.e. CF starvation agar, TPM starvation agar and in submerged culture, and the sporulation frequency determined after 72 hrs and 120 hrs of starvation. Moreover, levels of germinating spores were determined after 72 hrs and 120 hrs of starvation. Individual mutants displayed similar phenotypes under all three developmental conditions (data from development on CF agar and in submerged culture are shown in Figure 14, Table 5). DK1622 wild-type cells formed fruiting bodies after 24 hrs and at 48 hrs the fruiting bodies had darkened (Figure 14). 16 mutants displayed normal development, sporulation and spore germination (data not shown). As previously reported (Garza *et al.*, 1998, Pollack & Singer, 2001), MS1512 carrying $\Delta sdek$ was unable to aggregate and was strongly decreased in sporulation. Also as previously reported, a mutant carrying $\Delta espA$ (SA2139) displayed early aggregation with the formation of many, small and irregularly shaped fruiting bodies; also, SA2139 displayed early sporulation with many spores localized outside the fruiting bodies (Figure 14B). SA2144 carrying $\Delta asgD$ displayed delayed aggregation but normal levels of sporulation. This is in disagreement with a previous report in which an insertion in *asgD* was reported to cause aggregation as well as sporulation defects (Cho & Zusman, 1999a). These differences may be attributed to differences in strain

backgrounds and mutations being analyzed in the previous report and the data presented here.

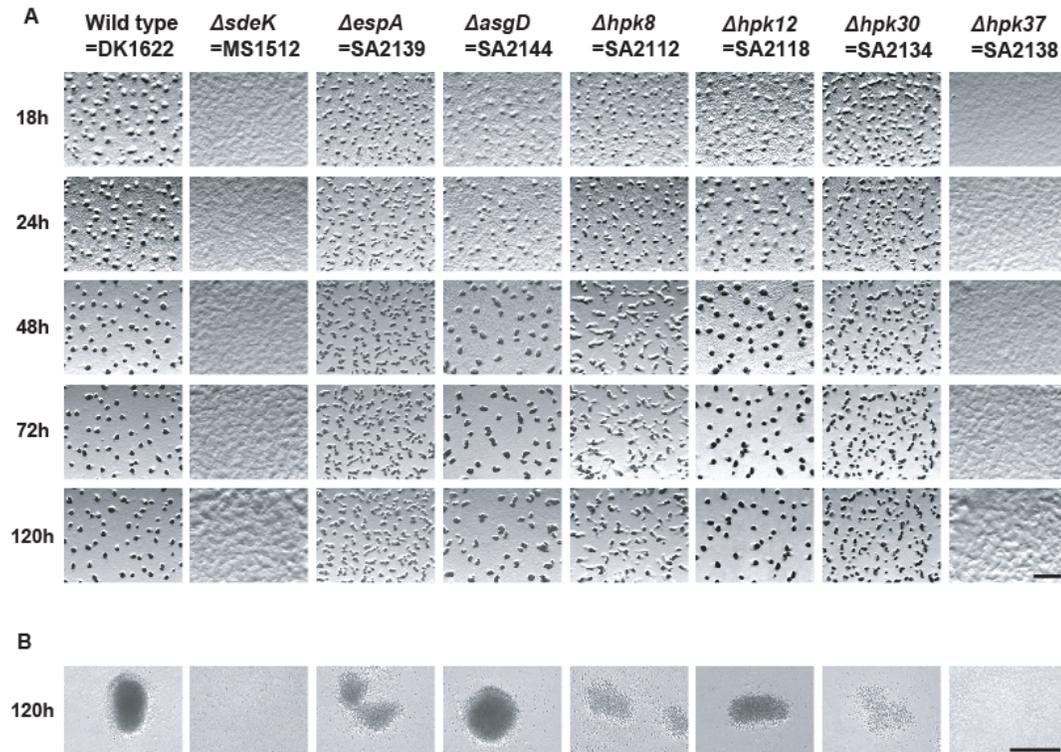


Figure 14. Developmental phenotypes of *hpk* in-frame deletion mutants with developmental defects.

A. Developmental phenotypes on CF agar. The strains were starved on CF agar for the indicated periods of time. Scale bar: 1.0 mm. B. Developmental phenotypes in submerged culture. The same strains as in (A) were exposed to starvation in submerged culture for 120 hrs. Scale bar: 100 μ m.

Table 5. Sporulation frequencies of *M. xanthus* wild type and the HPK mutants

Strain	Sporulation frequency ^a	
	72 hrs	120 hrs
DK1622(wild type)	66 \pm 4%	100 \pm 6% ^b
Δ sdek	<0.001%	<0.001%
Δ espA	94 \pm 9%	151 \pm 15%
Δ asgD	56 \pm 8%	96 \pm 2%
Δ hpk8	47 \pm 9%	80 \pm 16%
Δ hpk12	49 \pm 9%	71 \pm 8% ^c
Δ hpk30	56 \pm 17%	78 \pm 19%
Δ hpk37	<0.001%	<0.001%

^a Sporulation frequencies are presented relative to the sporulation level in the wild-type strain DK1622 after 120 hrs of starvation as mean values and standard deviations from three experiments. ^b The absolute sporulation level of DK1622 was 14.7 \pm 1.0% at 120 hrs. ^c Spores of Δ hpk12 germinated at a 3-fold lower frequency than wild-type.

SA2112 carrying *Δhpk8* (=ΔMXAN3290) displayed delayed aggregation and formation of abnormally shaped fruiting bodies. SA2112 sporulated at a wild-type level, however, many of the spores were localized outside fruiting bodies. SA2118 carrying *Δhpk12* (=ΔMXAN0736) displayed normal aggregation and fruiting body formation and sporulated at wild-type levels. However, these spores germinated at a level 3-fold lower than those observed in wild-type. SA2134 carrying *Δhpk30* (=ΔMXAN4465) displayed normal timing of aggregation and sporulation. However, this mutant formed 1.7 ± 0.4 -fold more fruiting bodies than wild-type and individual fruiting bodies were smaller than those formed by wild-type covering only an area of $60 \pm 7\%$ of that of a wild-type fruiting body. Moreover, the fruiting bodies formed by SA2134 in submerged culture were less condensed than those formed by wild-type (Figure 14B). Finally, SA2138 carrying *Δhpk37* (=ΔMXAN0712) was unable to aggregate to form fruiting bodies and was strongly decreased in sporulation. It should be noted that the two mutants carrying *ΔespC* (PH1017) and *ΔmokaA* (SA2143) developed in a manner indistinguishable from that of the DK1622 wild-type under all conditions tested and sporulated at wild-type levels with the formation of germination proficient spores (data not shown). Inactivation of these two genes has previously been reported to cause developmental defects (Kimura *et al.*, 2001, Lee *et al.*, 2005), the difference between published results and my results could be attributed to differences in strain backgrounds used as well as different mutations analyzed. To exclude the possibility that the developmental defects were due to motility defects in the mutants, motility assays were performed for the strains with developmental defects. None of these mutants displayed motility defects on the two types of surfaces (Material and Methods, data not shown).

Taken together, the systematic analysis of the 25 orphan genes encoding HPKs that are transcriptionally up-regulated during development showed that 2 such genes are likely essential for viability and 7 HPKs were identified, including 4 not previously characterized that have important functions in fruiting body formation or spore germination. The data suggest that not all of the transcriptional up-regulated of TCS genes are necessary for development (at least not under the three conditions tested here).

3.2 In search of the FruA kinase using a candidate approach

3.2.1 A candidate approach to identify for the FruA kinase

FruA is an orphan RR with an N-terminal receiver domain and a C-terminal DNA-binding domain. FruA is essential for development and provides a branch point in our model of the C-signal transduction pathway (Figure 7). In one branch, FruA activates aggregation through the Frz system (Sogaard-Andersen & Kaiser, 1996) and in the other branch, FruA stimulates sporulation by direct activating gene expression, e.g. *fdgA* (Ueki & Inouye, 2005b), *Ω4400* (Yoder-Himes & Kroos, 2006) and *dev* (Ellehaug *et al.*, 1998, Viswanathan *et al.*, 2007b). Previous analysis (Ellehaug *et al.*, 1998) showed that the transcription of *fruA* is up-regulated during development and genetic evidence suggested that the activity of FruA is controlled by phosphorylation. However, a cognate HPK of FruA has not yet been identified.

To identify FruA kinase candidates, I used a candidate approach, which is based on the hypothesis that the FruA kinase shares key characteristics with the *fruA* gene: orphan, developmentally up-regulated at the transcriptional level, and deficient in development in a null mutant. Moreover, yeast two-hybrid (Y2H) analyses were performed to investigate potential interactions between FruA and the developmental up-regulated orphan HPKs. From the first part of my studies (Results 3.1), 25 developmentally up-regulated orphan HPKs were identified in the *M. xanthus* genome and 7 of them were shown to be important for development. For the next step, the interaction of FruA and these 25 developmentally up-regulated orphan HPKs was investigated.

3.2.2 Y2H analyses to test the interactions between FruA and 25 developmentally up-regulated orphan HPKs

To identify FruA kinase candidates, the interaction between FruA and the developmentally up-regulated orphan HPKs was examined using the Y2H system. To ensure the specificity of the Y2H analysis, all 25 HPKs were included to obtain a measure of the specificity of the interaction observed in the Y2H system. The Y2H test was performed as follows. For example, to test the interaction between FruA and Hpk8 (Figure 15), the full length of *fruA* and *hpk8* were cloned into the vectors pGBKT7 and pGADT7, respectively. Three reporter genes in yeast strain AH109 (*HIS3* with GAL1 UCS origin, *ADE2* with

GAL2 UCS origin and *lacZ* with *MEL1* UCS origin) are regulated by GAL4. For testing interactions, transformants were screened for growth on an agar lacking histidine and adenine and for β -galactosidase activity by adding X-gal to the agar. To analyze whether constructs could self-activate in the Y2H, the constructs were cotransformed with the corresponding empty vectors into the yeast report strain. Negative results (transformants did not grow in the absence of histidine and adenine) in these tests showed that these constructs are self-activate in the Y2H system. For the FruA-Hpk8 example interactions were observed in four combinations: FruA and FruA, Hpk8 and Hpk8, FruA and Hpk8 in both directions (Figure 15). These data show that both the FruA and Hpk8 constructs self-interact and FruA interacts with Hpk8 in Y2H. Typically, HPKs form homodimers and the phosphorylated RRs function as dimers *in vivo* (Stock *et al.*, 2000). Hence, the interactions of FruA with itself and Hpk8 with itself suggest that both proteins can form dimers and that the constructs are functional in Y2H.

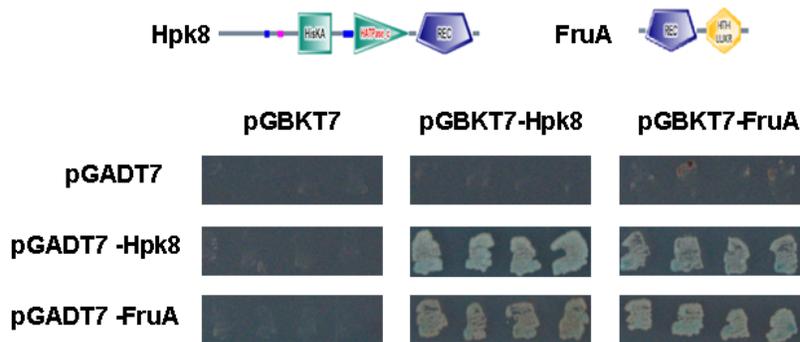


Figure 15. Y2H test of the interaction between FruA and Hpk8.

All of the tests were screened for growth on an agar lacking histidine and adenine and for β -galactosidase activity by adding X-gal to the agar. For each combination at least four colonies were restreaked on the screen plates. Growth of the colonies shows the interaction between two tested proteins: Hpk8 against Hpk8, Hpk8 against FruA, FruA against Hpk8, FruA against FruA. Absence of growth means that interaction was not observed between the tested constructs in Y2H.

The same tests were conducted for the remaining 24 HPKs. For the cytoplasmic HPKs (SdeK, Hpk9, Hpk11, Hpk13, Hpk14, Hpk15, Hpk20, Hpk24, Hpk25, Hpk29, Hpk30, Hpk36 and EspA), full length genes were cloned into the Y2H vectors for these analyses. For the inner membrane HPKs (Hpk5, Hpk12, Hpk18, Hpk23, Hpk27, Hpk34, Moka and EspC), only the cytoplasmic part of these HPKs were cloned in the Y2H vectors.

Table 6. Summary of the Y2H tests and phenotypes in the in-frame deletion mutants of 25 developmentally up-regulated orphan HPKs

TIGR_MXAN	Gene sym	Constructs size ^a	Self-activation ^b	Self-interaction ^c	Interaction with FruA ^d	Dev. defects in mutant	FruA kinase candidate ^e
MXAN0060	<i>hpk5</i>	538/758	-/-	+	-/±	No	-
MXAN0176	<i>hpk36</i>	642/642	-/-	+	-/-	No	-
MXAN0340	<i>hpk11</i>	562/562	-/-	+	±/+	No	Potential redundant
MXAN0571	<i>hpk25</i>	222/755	+/-	+	+/-	No	-
MXAN0706	<i>hpk39</i>	ND	ND	ND	ND	No	?
MXAN0712	<i>hpk37</i>	276/1967	-/-	-	-/-	Yes	?
MXAN0736	<i>hpk12</i>	531/614	-/-	+	±/+	Yes	Top candidate
MXAN0928	<i>hpk14</i>	334/334	-/-	+	-/-	No	-
MXAN0931	<i>espA</i>	232/777	+/-	+	+/-	Yes	-
MXAN1014	<i>sdeK</i>	513/513	-/-	+	±/+	Yes	Top candidate
MXAN2763	<i>hpk18</i>	475/521	-/-	+	-/-	No	-
MXAN3036	<i>hpk23</i>	221/514	-/-	-	-/-	essential	?
MXAN3098	<i>hpk24</i>	390/390	-/-	+	-/-	No	-
MXAN3290	<i>hpk8</i>	530/530	-/-	+	+/+	Yes	Top candidate
MXAN4465	<i>hpk30</i>	517/517	-/-	+	-/-	Yes	-
MXAN4988	<i>hpk27</i>	220/535	-/-	-	-/-	essential	?
MXAN5483	<i>hpk20</i>	578/578	-/-	+	-/-	No	-
MXAN6315	<i>hpk9</i>	632/632	-/-	+	+/+	No	Potential redundant
MXAN6855	<i>espC</i>	225/842	-/-	+	-/-	No	-
MXAN6941	<i>hpk34</i>	226/574	-/-	-	-/-	No	?
MXAN6994	<i>hpk15</i>	419/419	-/-	+	-/-	No	-
MXAN6996	<i>asgD</i>	234/773	-/-	+	-/-	Yes	-
MXAN7123	<i>hpk13</i>	533/533	-/-	+	+/+	No	Potential redundant
MXAN7206	<i>mokA</i>	234/1379	-/-	+	-/-	No	-
MXAN7398	<i>hpk29</i>	999/999	-/-	+	-/+	No	Potential redundant

^a: a size of HPK in the final tested constructs over the size of full length of HPKs. ^b: result before slash is the combination of a HPK construct in pGBKT7 with the empty plasmid of pGADT7. Result after slash is the combination of a HPK construct in pGADT7 and the empty plasmid of pGBKT7. ^c: result from the combination of an HPK construct in pGBKT7 and the same HPK construct in pGADT7. ^d: result before slash is the combination of an HPK construct in pGBKT7 and pGADT7-FruA. Result after slash is the combination of an HPK construct in pGADT7 and pGBKT7-FruA. In ^b, ^c and ^d, + means positive interaction, ± means the insignificant interaction and - means no interaction. ^e: - means not a FruA kinase candidate, ?: inconclusive from Y2H analysis but phenotypes in the mutant indicating that they are not top candidates for being a FruA kinase.

Some of these constructs (Hpk25, Hpk23, Hpk27, Hpk34, EspA, EspC and Moka) did not self-interact in Y2H. For these HPKs shorter constructs only containing the kinase domains were generated. For the HPKs Hpk37 and AsgD, the full length constructs were not obtained. Also, for these 2 HPKs, short constructs only containing the kinase domains were generated. Only Y2H constructs for Hpk39 were not obtained. Subsequently the interactions between HPKs and FruA were examined in Y2H.

Seven HPKs (Hpk11, Hpk12, SdeK, Hpk8, Hpk9, Hpk13 and Hpk29) displayed positive interaction with FruA in Y2H system (Table 6). The Hpk25 and EspA constructs in the plasmid of pGBKT7 could self-activate in the Y2H system but no interaction was observed in the combination of Hpk25 or EspA in the plasmid of pGADT7 with pGBKT7-FruA, which indicated that Hpk25 and EspA did not interact with FruA in Y2H. The remaining 15 HPKs showed no interaction with FruA in Y2H system including four HPKs (Hpk23, Hpk37, Hpk27 and Hpk34), which did not self-interact indicating that the Y2H constructs might not be functional.

To summarize the results of the Y2H test and the results of the phenotypic analyses of the *hpk* in-frame deletion mutants (Table 6), SdeK, Hpk8 and Hpk12, which displayed developmental defects in null mutants and interactions with FruA in Y2H system, were selected as the top candidates of the FruA kinase for the further study. Hpk9, Hpk11, Hpk13 and Hpk29 interacted with FruA in Y2H system while they were not important for development. These observations suggest that these HPKs are either not the FruA kinase or they could be the FruA kinases but their function is redundant.

The remaining 17 HPKs exhibited no interaction with FruA in Y2H. Four HPKs (Hpk23, Hpk27, Hpk37 and Hpk34), however, the Y2H constructs did not self-interact suggesting that the Y2H constructs could be non-functional. Of these HPKs, Hpk23 and Hpk27 are essential for cell viability. Therefore, the phenotypes of the two mutants indicate that these two HPKs may be involved in different pathways and they are unlikely to be the FruA kinase. A deletion of Hpk37 displayed strong developmental defects. Therefore, Hpk37 cannot be ruled out from the list of FruA kinase candidates. Hpk34 exhibit no

developmental defect in the null mutant and no interaction with FruA in Y2H. It is not considered as the FruA kinase candidate.

Hpk30 had developmental defects in the null mutants and did not interact with FruA in the Y2H system. It is not selected as FruA kinase candidate for the further analyses. The remaining 13 HPKs had normal developmental phenotypes in the null mutants and did not interact with FruA in the Y2H system, which indicate that these kinases are not candidates for the FruA kinase. Taken together, SdeK, Hpk8 and Hpk12 were selected as the top candidates of the FruA kinase for further studies while Hpk9, Hpk11, Hpk13 and Hpk29 were considered as potentially functionally redundant FruA kinase candidates and, finally, Hpk37 could not be eliminated from the list of FruA kinase candidates.

3.2.3 Analyses of the FruA kinase candidates

To determine whether the FruA kinase candidates share high homology, the sequences of the kinase domains of these eight HPKs (Figure 16A) were analyzed by multiple sequence alignments. Following the classification of Grebe and Stock's TCS protein families, the three FruA top candidates (SdeK, Hpk8 and Hpk12) as well as the four potential redundant candidates (Hpk9, Hpk11, Hpk13 and Hpk29) belong to the same sub-family HisKA 3f (this family contains 39 out of all 68 orphan HPKs in *M. xanthus*, data not shown) whereas Hpk37 belongs to the different sub-family HisKA 1b. This indicates that the function of Hpk37 may be distinct from the remaining 7 FruA kinase candidates.

Previous studies in which the specificity of EnvZ was rewired indicate that the recognition specificity of partner HPK and RR lie in the C-terminal end of $\alpha 1$ helix and the loop region between the two α -helices in the HisKA domain (Skerker *et al.*, 2008). To determine whether the FruA kinase candidates share similarity in this region, the kinase domains of the three FruA kinase top candidates, the four potential redundant FruA kinase candidates as well as Hpk37 were aligned with the kinase domains of *E. coli* EnvZ and KinA, KinB, KinC, KinD as well as KinE from *B. subtilis* (these five HPKs phosphorylate the same RR Spo0F) (Figure 16B). The FruA kinase candidates share certain similarities in the end of $\alpha 1$ helix and the loop region between the two α -helices in the HisKA domain but this similarity is lower than that between KinA, KinB, KinC, KinD and KinE. Since the exact residues that determine the recognition

The two-helices in the HisKA domain of EnvZ are marked on the figure. Asterisks indicate the residues potentially determine the recognition specificity of EnvZ. The conserved boxes (N, D, F and G) of HATPase_c domain are marked on the alignment.

3.2.4 Epistasis analyses of *fruA* and *fruA* kinase candidates

To determine whether FruA and the FruA kinase candidates are in the same genetic pathway, double mutants of *fruA* and *sdeK*, *fruA* and *hpk8* as well as *fruA* and *hpk12*, double mutants and triple mutant of *sdeK*, *hpk8* and *hpk12* were constructed and their phenotypes determined in developmental assays under three different conditions: CF agar, submerged culture and TPM agar (Figure 17A, B and C). The sporulation efficiency on CF agar is shown in the Figure 17A. Wild type DK1622 formed fruiting bodies after 24 hrs of starvation and fruiting bodies became dark after 120 hrs of development under all three conditions. The *fruA* insertion mutant (DK11063) displayed no fruiting body formation and a low sporulation efficiency. The $\Delta sdeK$ strain exhibited no aggregation and a strong reduction of the sporulation efficiency on all tested conditions. The double mutant of *fruA* $\Delta sdeK$ displayed no formation of fruiting bodies and low sporulation under all tested conditions. The aggregation phenotype on all three media seemed more similar to that of the $\Delta sdeK$ mutant. However, due to the similarity of the phenotypes of the *fruA* and *sdeK* mutants, it was difficult to judge whether the double mutant was more similar to one of the two single mutant phenotypes or a mixture of these two phenotypes. Thus, the phenotype of the double mutant did not contradict that SdeK could be a FruA kinase.

$\Delta hpk8$ displayed slightly delayed aggregation and formation of abnormally shaped fruiting bodies whereas the double mutant *fruA* $\Delta hpk8$ exhibited no fruiting body formation and a low sporulation phenotype, which is similar to the phenotype of the *fruA* mutant. This suggests that *fruA* is epistatic to *hpk8* and that they lie in the same genetic pathway. The $\Delta hpk12$ mutant could aggregate and sporulate at wild type levels but the spores had a germination defect while the double mutant *fruA* $\Delta hpk12$ exhibited no formation of fruiting bodies and almost no sporulation similarly to the *fruA* mutant under all tested conditions. This indicates that *fruA* is epistatic to *hpk12* and that they are in the same genetic pathway.

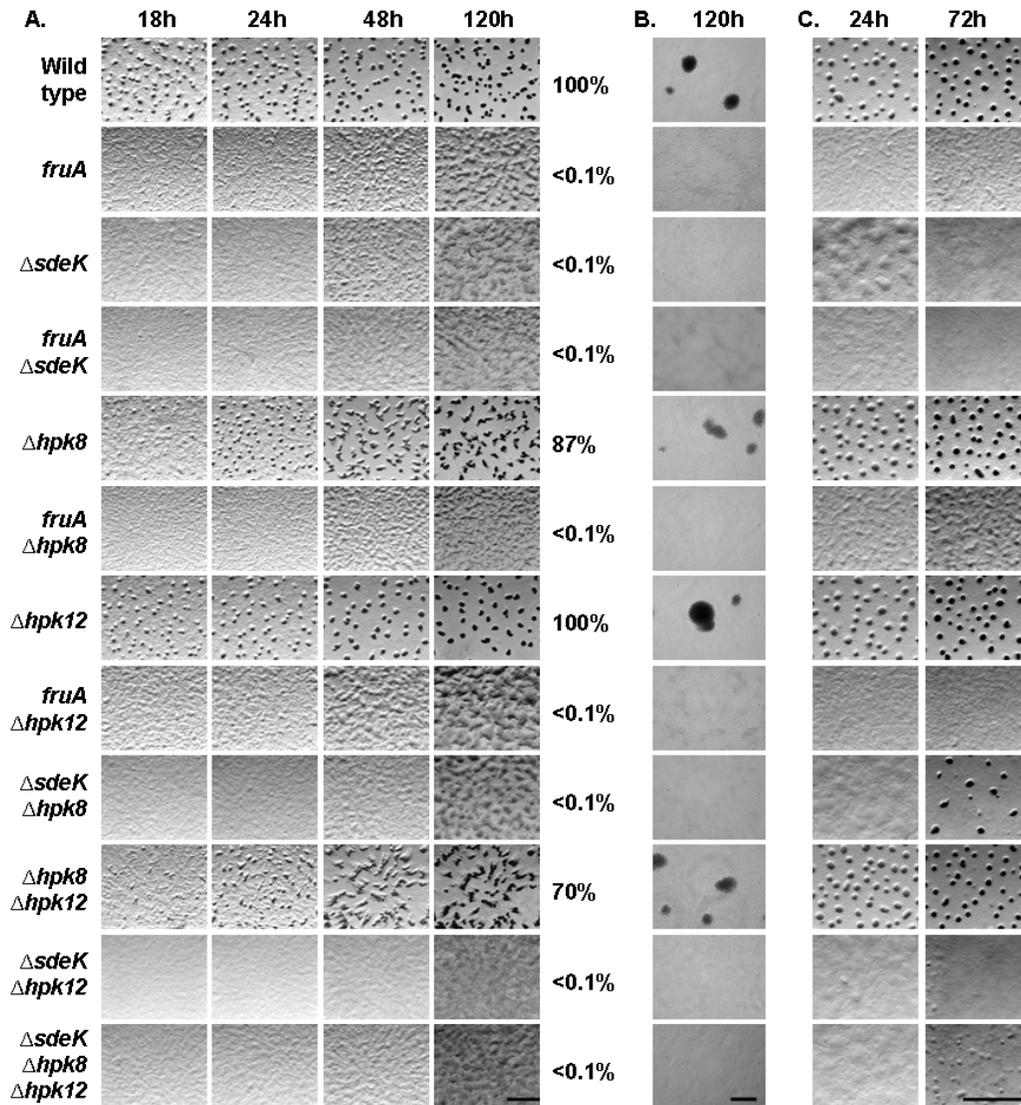


Figure 17. Developmental phenotypes of single, double and triple mutants of *fruA*, *hpk8*, *sdeK*, and *hpk12*.

A. Developmental phenotypes on CF agar. The strains were starved on CF agar for the indicated period of time. Scale bar: 1.0 mm. The sporulation efficiency of each strain after 120 hrs starvation on the CF agar compared with that of wild type (as 100%) are marked in the right side of photos. B. Developmental phenotypes in submerged culture. The same strains as in (A) were exposed to starvation in submerged culture for 120 hrs. Scale bar: 100 μ m. C. Developmental phenotypes on TPM agar. The same strains as in (A) were starved on TPM agar for indicated time. Scale bar: 1.0 mm.

The double mutant $\Delta sdeK \Delta hpk8$ displayed no formation of fruiting bodies and a low sporulation level on CF agar and in the submerged culture similarly to the *sdeK* mutant. On TPM agar, the double mutant displayed no fruiting body formation during the first 48 hrs of development, however, a few fruiting bodies were formed after 72 hrs of development. This demonstrates that the phenotype

of $\Delta sdeK \Delta hpk8$ mutant is a mixed phenotype of the two phenotypes caused by the single mutation. This suggests that *sdeK* and *hpk8* lie in parallel genetic pathways.

Under all three tested conditions, the double mutant $\Delta sdeK \Delta hpk12$ exhibited a phenotype similar to the $\Delta sdeK$ single mutant. The double mutant of *hpk8* and *hpk12* on CF agar exhibited a phenotype similar to that of $\Delta hpk8$ and spores of this strain had a slight defect in germination as $\Delta hpk12$. This implies that *hpk8* and *hpk12* may lie in different pathways. The phenotype of the triple mutant $\Delta sdeK \Delta hpk8 \Delta hpk12$ displayed no fruiting body formation and low sporulation efficiency on CF agar and submerged culture. But on TPM agar this strain could form small fruiting bodies after 72 hrs of development and became dark after 120 hrs of development. Therefore, phenotype of the triple mutant is similar to the phenotype of *fruA* mutant on CF agar and in submerged culture but it is not equal to the phenotype of the *fruA* mutant on the TPM agar.

All together, the data from the above epistasis tests suggest that FruA may lie in the same pathway as Hpk8 and Hpk12, the relation of SdeK and FruA is not conclusive, SdeK and Hpk8 may lie in parallel pathways, *sdeK* is epistatic to *hpk12* or SdeK and Hpk12 are in parallel genetic pathways, and the function of the sum of SdeK, Hpk8 and Hpk12 is not equal to FruA. These results can be understood within the framework of a model in which SdeK is the main FruA kinase, Hpk12 is a minor FruA kinase and Hpk8 is a phosphatase of FruA. Moreover, there may be other HPK that phosphorylate FruA when these three *hpk* genes are deleted.

3.2.5 FrzCD methylation and FruA accumulation in the mutants of FruA kinase candidates

To monitor at the molecular level whether the FruA kinase candidates act in the C-signal transduction pathway, FrzCD methylation profiling and FruA accumulation profiling during development was performed in the mutants of the FruA kinase candidates and compared with wild type and the *fruA* insertion mutant (Figure 18A). FrzCD is a cytoplasmic homologue of methyl-accepting chemotaxis proteins and it is methylated during aggregation in a FruA dependent manner (McCleary *et al.*, 1990, Ellehaug *et al.*, 1998). Moreover, genetic data suggest that phosphorylation of FruA is required for the FrzCD

methylation during development (Ellehaug, 1999). At the initiation of development in wild type (Figure 18A), FrzCD was present as a mixture of methylated and unmethylated forms, the methylated FrzCD shifted to unmethylated FrzCD after 1 hr of starvation, afterwards the level of FrzCD methylation increased and most FrzCD was converted to the methylated form after 9 hrs of development. After 24 hrs of development, both methylated and unmethylated FrzCD were undetectable by immunoblot analysis. In the *fruA* mutant, the shift from unmethylated FrzCD to methylated FrzCD was strongly delayed during development. This suggests that FrzCD methylation in the mutants of the FruA kinases should be delayed (if the kinase is important for the aggregation) or similar to that in the wild type (if the kinase is important for late gene expression or sporulation).

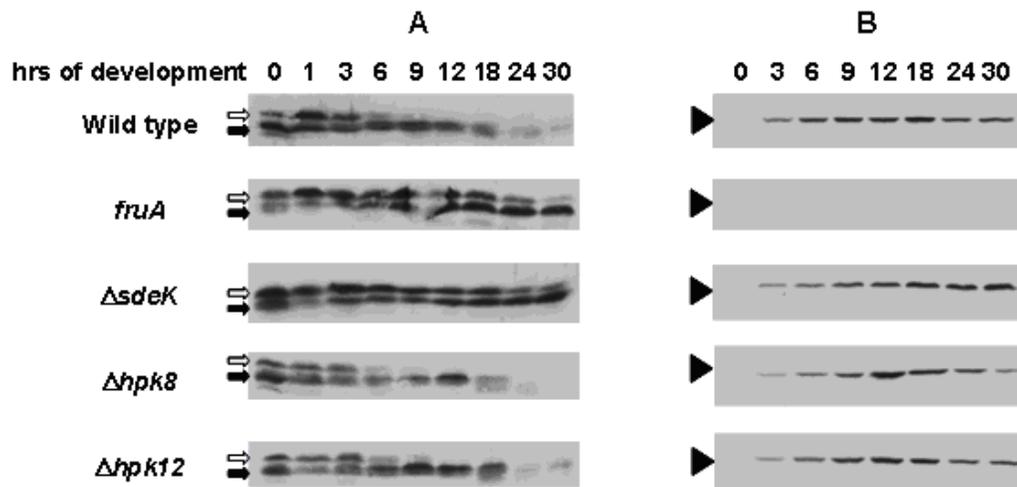


Figure 18. FrzCD methylation and FruA accumulation in wild type, *fruA* mutant and the mutants of FruA kinase candidates.

A. FrzCD methylation in wild type, *fruA* mutant and the mutants of FruA kinase candidates ($\Delta sdeK$, $\Delta hpk8$ or $\Delta hpk12$). The FrzCD methylation profiles were determined by immunoblot analysis with FrzCD antibodies. The blank arrow stands for unmethylated FrzCD. The solid arrow means methylated FrzCD. B. FruA accumulation in the same strains as in (A). The FruA accumulation profiles were determined by immunoblot analysis with FruA antibodies. The solid triangle stands for FruA.

The FrzCD methylation profile of $\Delta sdeK$ was strongly delayed compared to that of wild type (Figure 18A) even after 30 hrs of development not all of the unmethylated FrzCD had shifted to the methylated form. The FrzCD methylation pattern in $\Delta sdeK$ correlated with the phenotype of $\Delta sdeK$ in which cells is unable to aggregate into fruiting bodies. FrzCD methylation in $\Delta hpk8$ and

$\Delta hpk12$ were similar to that of wild type (Figure 18A), which are in agreement with the observation that aggregation of these mutants was not affected during development. These observations are not in disagreement with the notion that SdeK and Hpk12 might be FruA kinases and Hpk8 might be a FruA phosphatase. Accordingly, SdeK may function as the main FruA kinase which is required for early development, while Hpk8 could function as a FruA phosphatase and Hpk12 could be a minor FruA kinase which is required for the later developmental gene expression.

Previous studies (Ellehaug *et al.*, 1998) demonstrated that the FruA accumulation depends on A-signal but is independent of C-signal while phosphorylation of FruA is C-signal dependent. This suggests that the FruA kinases would not be required for FruA accumulation during development.

In wild type (Figure 18B), FruA accumulation was developmentally regulated. FruA protein was detectable after 3hrs of development, reached a peak at about 12 hrs to 18 hrs of development and decreased afterwards. In the $\Delta sdeK$ mutant, the FruA accumulation profile was similar to that in wild type during the first 18 hrs of development and the level of FruA in $\Delta sdeK$ was slightly higher than that in wild type at late developmental time points of 24 hrs and 30 hrs. This implies that SdeK is not important for FruA accumulation during development but it is required for the decreased stability of FruA directly or indirectly after 24 hrs of development. In the null mutants of *hpk8* or *hpk12*, the FruA accumulation profiles were similar to that in wild type. These results are not in disagreement that SdeK and Hpk12 could be FruA kinases and Hpk8 could function as a FruA phosphatase.

3.2.6 Developmental gene expression in the mutants of FruA kinase candidates

To further determine the function of the FruA kinase candidates in development, the expression profiles of five developmental marker genes (*spi*, *sdeK*, *fruA*, *devR* and *exo*), which are turned on at different time points after starvation, were examined in wild type, *fruA* mutant and the mutants of the FruA kinase candidates by qRT-PCR. RNA was prepared from cells harvested at different time points of development in submerged culture, cDNA was synthesized and

expression of the marker genes in the different strains was examined by qRT-PCR.

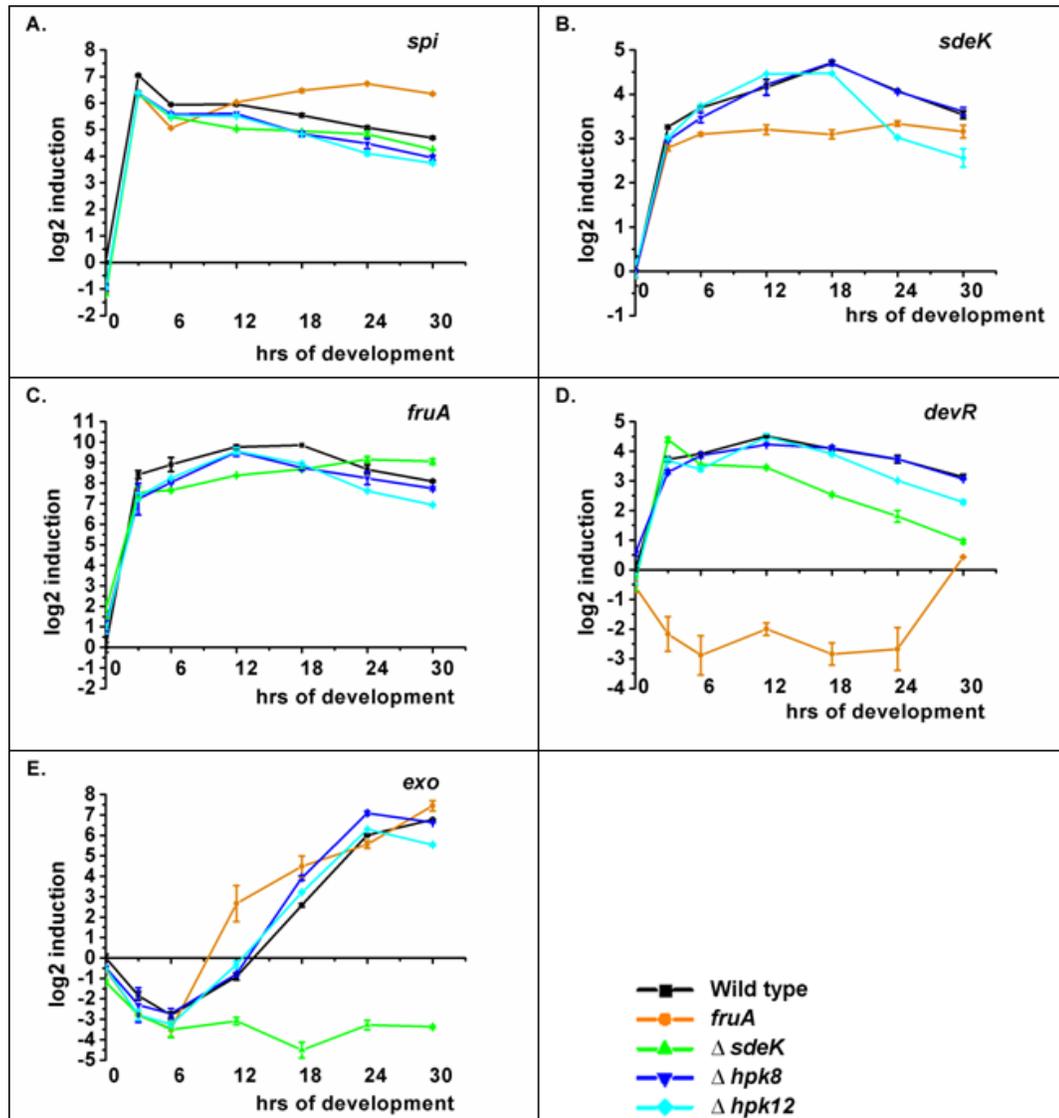


Figure 19. Five marker gene expression profiles in wild type, *fruA* and the mutants of FruA kinase candidates.

x axis represents hrs of development, y axis represents the log₂ ratio of the transcriptional level in comparison to that of wild type at 0 hr. Expression in the mutants was regarded as significantly altered if at least two consecutive time points different by at least 3.5 fold (log₂ ratio = 1.8) from wild type.

The expression levels of the marker genes during development in all strains were compared with that of the wild type at 0 hr of development. Expression in the mutants was regarded as significantly altered if at least two consecutive time points different by at least 3.5 fold (log₂ ratio = 1.8) from wild type.

spi (Ω 4521) is an A-signal dependent gene and it is turned on after 1.5 hrs of development as shown using a *spi* Tn5lac fusion (a transposon coupled with β -galactosidase expression to endogenous promoter) (Kuspa *et al.*, 1986). In this study, the transcription of *spi* in wild type was turned on after 3 hrs of development at the tested time points (1.5 hrs was not tested in this study) and decreased slightly at the later time points (Figure 19A). The *spi* expression profiles in the *fruA* mutant as well as in the $\Delta sdeK$, $\Delta hpk8$ and $\Delta hpk12$ mutants were similar to that of wild type (Figure 19A). This indicates that SdeK, Hpk8 and Hpk12, like FruA, are not involved in the A-signal pathway.

The transcription of *sdeK* (Ω 4408) is independent of A-signal but depends on the (p)ppGpp level during the early development (Garza *et al.*, 1998). In this study, expression of *sdeK* in wild type was turned on after 3 hrs of development at the tested time points, reached a peak after 18 hrs of development and afterwards the level of expression decreased slightly (Figure 19B). With the criteria for significant difference of the profiles, the expression profiles of *sdeK* in the $\Delta hpk8$, $\Delta hpk12$ and *fruA* mutants were similar to that in wild type. This suggests that Hpk8 and Hpk12 as well as FruA are not important for the production of (p)ppGpp.

The transcriptional of *fruA* (Ω 7540) is controlled by A-signal (Ellehaug *et al.*, 1998). The *fruA* expression in wild type was turned on 3 hrs after onset of development, reached a peak around 12 hrs to 18 hrs of development and afterwards the expression level decreased (Figure 19C). The expression profiles of *fruA* in $\Delta sdeK$, $\Delta hpk8$ and $\Delta hpk12$ were similar to that of wild type. This indicates that SdeK, Hpk8 and Hpk12, are not important for the expression of *fruA* and they are not involved in the production of A-signal.

The transcription increase of *devR* (Ω 4414) depends on phosphorylation of FruA (Ellehaug *et al.*, 1998, Ellehaug, 1999). *devR* expression in wild type was turned on after 3 hrs of development, reached a peak around 12 hrs of development and afterward the expression level decreased (Figure 19D), which was slightly different from a previous report (Thony-Meyer & Kaiser, 1993). In that report, the time of turning on expression of *devR* was at 12 hrs of development. The difference between the result of my analysis and the published data could be attributed to the fact that in the present study *devR*

expression was examined in *dev+* background while in the published study it was determined in a *dev* insertion mutant and a different method was used to examine the expression profile.

In the *fruA* mutant, the expression of *devR* decreased after 3 hrs of development and afterwards remained at a low level, which was in agreement with previous data (Ellehaug *et al.*, 1998). The expression of *devR* in Δ *sdeK* was turned on after 3 hrs of development as in wild type, afterwards expression decreased, then became significantly lower compared to wild type after 18 hrs of development (Figure 19D). The profile of *devR* expression in Δ *sdeK* was similar to previous results (Pollack & Singer, 2001). The expression profiles of *devR* in Δ *hpk8* and Δ *hpk12* were similar to that in wild type (Figure 19D).

All together, these data demonstrate that SdeK and FruA are required for transcription of *devR* but FruA is more important than SdeK and that Hpk8 as well as Hpk12 are not essential for the transcription of *devR*. These data support a model in which SdeK could be a FruA kinase but in addition other kinases could phosphorylate FruA in absence of SdeK. Moreover, these data are not in disagreement with that Hpk12 still could be a minor FruA kinase and Hpk8 could be a FruA phosphatase but their function alone are not important in the regulation of *devR* expression.

exo (Ω 7536) is involved in sporulation (Licking *et al.*, 2000). Previously, *exo* expression in wild type was shown to be turned on after 18 hrs of development and the expression level increased at 24 hrs and 30 hrs of development using *exo-lacZ* fusion (Licking *et al.*, 2000). In my study, the same profile of *exo* expression was observed in wild type (Figure 19E). Surprisingly, *exo* expression in the *fruA* mutant was similar to that in wild type. Previous study *exo* expression using *exo-lacZ* fusion is dependent on DevR (Licking *et al.*, 2000) and the transcription of *devR* depends on FruA using a *devR-lacZ* fusion (Ellehaug *et al.*, 1998). Therefore, it was previously proposed that *exo* expression depends on FruA (Licking *et al.*, 2000) but this hypothesis has never been tested experimentally. The expression of *exo* in Δ *sdeK* was down-regulated after 3 hrs of development and remained at low expression levels afterwards. From this study, FruA is important for the *devR* expression but not essential for the *exo* expression, whereas SdeK is partially required for the

expression of *devR* and fully required for the *exo* expression. This suggests that SdeK and FruA might be involved in different signaling pathways or that SdeK has other downstream targets in addition to FruA. The transcription profiles of *exo* in $\Delta hpk8$ and $\Delta hpk12$ were similar to that in wild type (Figure 19E). This demonstrates Hpk8 and Hpk12 are not important for the *exo* expression. These data are not in disagreement with that Hpk12 could still be a minor FruA kinase and Hpk8 could be a FruA phosphatase.

3.2.7 Purification of FruA and the FruA kinase candidates

So far, all the data of the *in vivo* analyses to determine whether SdeK, Hpk8 and Hpk12 are FruA kinases are not in disagreement with a model in which SdeK is a main FruA kinase, Hpk12 is a minor FruA kinase and Hpk8 is a FruA phosphatase. But the data are inconclusive. Therefore, to directly test the interaction between FruA and FruA kinase candidates, I performed *in vitro* phosphorylation assay. For that, first of all, the relevant proteins were purified.

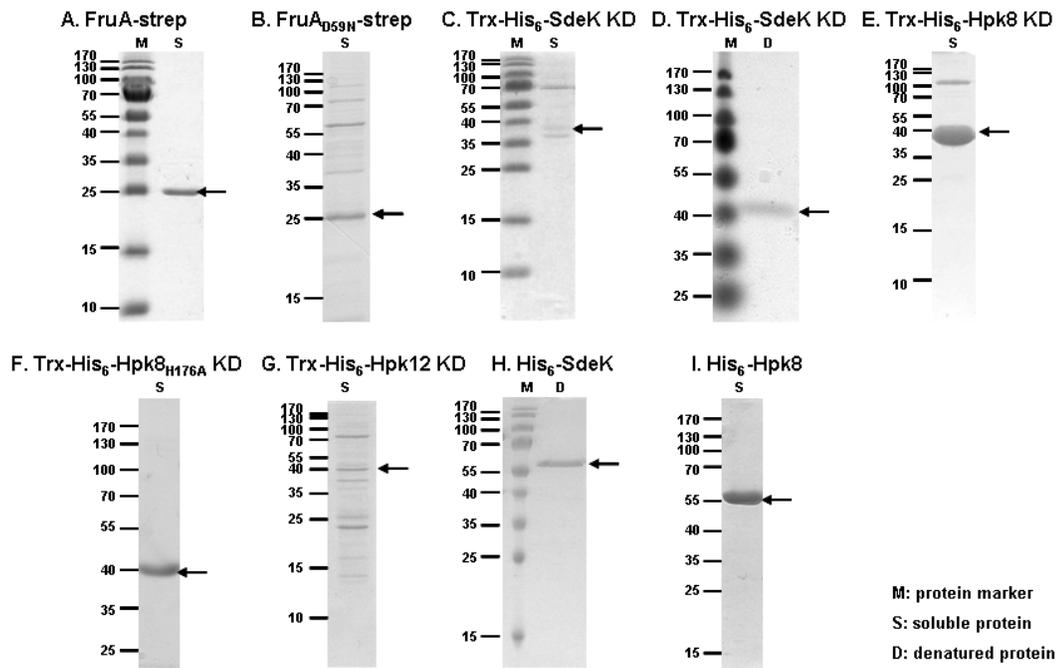


Figure 20. Purification of FruA and the FruA kinase candidates from *E. coli*.

A. Purified native FruA-strep; B. Purified native FruA_{D59N}-strep; C. Purified native Trx-His₆-SdeK KD; D. Purified denatured Trx-His₆-SdeK KD; E. Purified native Trx-His₆-Hpk8 KD; F. Purified native Trx-His₆-Hpk8_{H176A} KD; G. Purified native Trx-His₆-Hpk12 KD; H. Purified denatured His₆-SdeK; and I. Purified native His₆-Hpk8. Lanes marked with M: prestained protein marker; S: purified soluble protein; D: purified denatured protein. The arrows point to the expected size of the purified proteins.

The full length FruA with a C-terminal Strep-tag (FruA-strep) was over-expressed in *E. coli* using the autoinduction procedure (Material and Methods) and purified under native conditions (Figure 20A). The full length FruA protein containing the point mutant (D59N) with a C-terminal Strep-tag (FruA_{D59N}-strep) was mostly insoluble under the same induction conditions. FruA_{D59N}-strep was purified under the same conditions as FruA-strep but only a small amount of protein was obtained with a low level of purity (Figure 20B).

The kinase domain of SdeK with an N-terminal Trx/His₆/thrombin/S/enterokinase-tag (Trx-His₆-SdeK KD) and full length SdeK with an N-terminal His₆-Tag (His₆-SdeK) were over-expressed in *E. coli*. Both were mostly insoluble under all tested induction conditions (Material and Methods). Trx-His₆-SdeK KD was partially purified under native conditions by Ni-NTA affinity-chromatography eluted with a gradient imidazole followed by gel filtration chromatography. However, only a low amount of protein with a low level of purity was obtained (Figure 20C). Insoluble Trx-His₆-SdeK KD was solubilized with guanidine hydrochloride and refolded (Figure 20D) for the *in vitro* phosphorylation assay. His₆-SdeK was over-expressed in *E. coli* and solubilized with urea, purified with Ni-NTA affinity chromatography by elution via a pH gradient procedure and then repurified with Ni-NTA affinity chromatography using a imidazole gradient (Figure 20H). This protein was used as an antigen to produce antibodies against SdeK.

The kinase domain of Hpk8 with an N-terminal Trx/His₆/thrombin/S/enterokinase-tag (Trx-His₆-Hpk8 KD), the kinase domain of Hpk8 containing the point mutation H172A with an N-terminal Trx/His₆/thrombin/S/enterokinase-tag (Trx-His₆-Hpk8_{H172A} KD) and full length Hpk8 with an N-terminal His₆-Tag (His₆-Hpk8) were over-expressed in *E. coli*. All of them were soluble with the autoinduction procedure (Material and Methods) and they were purified under native conditions (Figure 20E, F and I). These proteins were used for *in vitro* assays and the purified His₆-Hpk8 (Figure 20I) was used as an antigen to produce antibodies against Hpk8.

For the kinase domain of Hpk12 with an N-terminal Trx/His₆/thrombin/S/enterokinase-tag (Trx-His₆-Hpk12 KD), only a small portion remained soluble under all tested conditions (Material and Methods). The soluble fraction was

purified under native conditions (Figure 20G, Material and Methods) and only a small amount of protein with a low level of purity was obtained.

3.2.8 Autophosphorylation assay of Hpk8

To test for phosphotransfer between FruA and the FruA kinase candidates, the next step is to autophosphorylate the FruA kinase candidates *in vitro*. The described kinase buffer (Porter *et al.*, 2007, Rasmussen *et al.*, 2006) with either MgCl₂ or MnCl₂ was tested to set up the conditions for autophosphorylation of the Hpk8 proteins. The band with a size of Trx-His₆-Hpk8 KD was detected by phosphoimaging when MnCl₂ was added to the reaction instead of MgCl₂ (Figure 21). Different concentrations of the MnCl₂ (5 mM, 10 mM, 15 mM and 20 mM) were tested in the *in vitro* phosphorylation assay and the signal from the phosphorylated Trx-His₆-Hpk8 KD became more intense when higher concentrations of MnCl₂ was included in the reaction system (Figure 21). With the final concentration 20mM MnCl₂, the phosphorylation of Hpk8 was clearly detected by phosphoimaging (Figure 21). The time course of the autophosphorylation of Trx-His₆-Hpk8 KD is shown in Figure 22A.

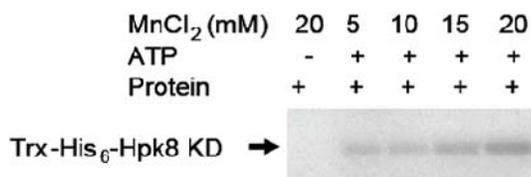


Figure 21. Autophosphorylation of Hpk8 in present of different concentration of MnCl₂.

The phosphorylation reactions in each lane were performed as following: 10μM Trx-His₆-Hpk8 KD was incubated in the reaction buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 50 mM KCl, 0.5 mM ATP and 0.5 mM [³²P] ATP) (one reaction without ATP as control) with the indicated concentration of MnCl₂ at 25 °C for 1h. The reactions were stopped by adding SDS loading buffer, loaded on SDS-PAGE and analyzed with phosphoimaging. The black arrow points to the size of Trx-His₆-Hpk8 KD.

Typically, HPK is phosphorylated at a conserved histidine residue in the HisKA domain (Stock *et al.*, 2000). To determine whether the conserved His172 residue of Hpk8 is the phosphorylation site, Trx-His₆-Hpk8_{H172A} KD was purified (Figure 20F) and the protein was used in the autophosphorylation assay under the same conditions as for autophosphorylation of Trx-His₆-Hpk8 KD. Surprisingly, similar phosphorylation signals were detected with

phosphoimaging for both proteins (Figure 22A and B). This suggested that the Trx-His₆-Hpk8 KD was not phosphorylated on the conserved His residue *in vitro*. To exclude the possibility that the long N-terminal tags of the Trx-His₆-Hpk8 KD and Trx-His₆-Hpk8_{H172A} KD proteins were phosphorylated in the *in vitro* phosphorylation assay, the time course assay of autophosphorylation was performed in parallel with the full length His₆-Hpk8. An autophosphorylation signal from His₆-Hpk8 was detected under the same reaction conditions. This suggested that the autophosphorylation of Trx-His₆-Hpk8 KD and Trx-His₆-Hpk8_{H172A} KD was not due to the long N-terminal tag. Be aware that Trx-His₆-Hpk8 KD, Trx-His₆-Hpk8_{H172A} KD and His₆-Hpk8 still share the same sequence (17 AA) in the His₆-tag region. However, further experiments (see below) suggested that Hpk8 proteins are phosphorylated on a residue in the Hpk8 part.

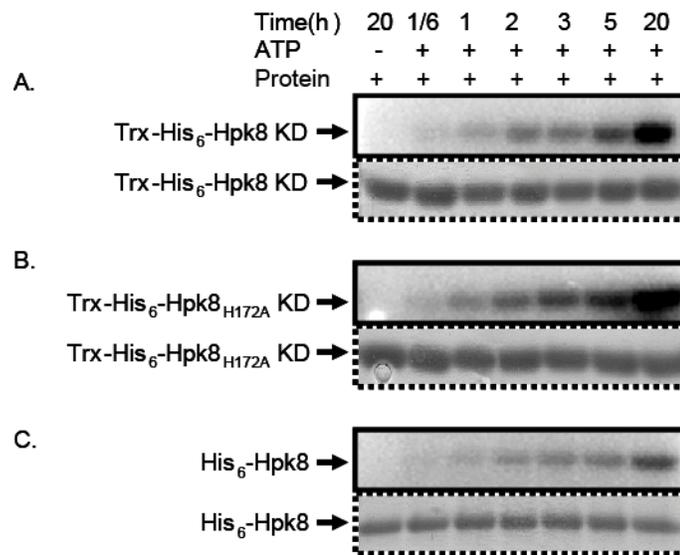


Figure 22. *In vitro* autophosphorylation assay of the Hpk8 proteins.

Time courses of the autophosphorylation of Trx-His₆-Hpk8 KD (A), Trx-His₆-Hpk8_{H172A} KD (B) and His₆-Hpk8 (C). The panels with the solid outlines are the phosphoimaging and the panels with the dashed outline are the SDS-PAGE staining. The arrows point to the size of the Hpk8 proteins. The reactions were performed as follows: 10 μM protein was incubated at 25 °C for the indicated time in the autophosphorylation reaction with [γ-³²P] ATP (one reaction without ATP as control). The reactions were stopped by SDS loading buffer, separated by SDS-PAGE, and analyzed by phosphoimaging and pageblue staining.

3.2.9 Chemical stability of phosphorylated Hpk8

To determine which amino acid was phosphorylated in the Hpk8 fusion proteins, the chemical stability of the phosphoryl group in the phosphorylated Hpk8

proteins was examined. As a control protein, the chemical stability of phosphoryl group in the phosphorylated His₆-RodK_{HNNN} was also examined in parallel. The chemical stability of the different phosphorylated amino acids in the presence of alkali, acid, H₂O, pyridine and hydroxylamine was previously described (Table 6) (Duclos *et al.*, 1991). The control His₆-RodK_{HNNN} is phosphorylated at a His residue (Rasmussen *et al.*, 2006).

Table 7. Chemical stability of phosphorylated amino acid^a. Reprint of (Duclos *et al.*, 1991)

Nature of phosphoamino acid	Stability in			
	alkali	acid	pyridine	hydroxylamine
O-phosphates				
Phosphoserine	-	+	+	+
Phosphothreonine	±	+	+	+
Phosphotyrosine	+	+	+	+
N-phosphates				
Phosphohistidine	+	-	-	-
Phospholysine	+	-	-	-
Phosphoarginine	-	-	-	-
Acyl phosphates				
Phosphoaspartate	-	-	-	-
Phosphoglutamate	-	-	-	-
S-phosphates				
Phosphocysteine	+	+	+	+

^a + indicates that the phosphoamino acid is stable, – means that it is labile, and ± the liability of the phosphoryl groups influenced by the neighboring amino acids. The possible phosphoamino acids of Hpk8 protein are highlighted with grey background.

In this study, three Hpk8 proteins and the control of His₆-RodK_{HNNN} were autophosphorylated *in vitro* with [γ -³²P] ATP and the reactions were quenched with SDS loading buffer. 10 μ l of the quenched proteins was incubated at 42°C for 1 hr with 2 μ l of 1 N NaOH, H₂O, 1 N HCl, 99% pyridine or 1 M hydroxylamine separately. Reactions were loaded on SDS-PAGE and detected by phosphoimaging and pageblue staining. The phosphoryl group on phosphorylated His₆-RodK_{HNNN} was stable under alkaline conditions and labile in the presence of acid, pyridine and hydroxylamine, which is in agreement with the chemical stability of a phosphohistidine (Figure 23A). The phosphorylated proteins of Trx-His₆-Hpk8 KD, Trx-His₆-Hpk8_{H172A} KD and His₆-Hpk8 exhibited the same patterns of stability and were resistant to the five tested conditions as follows (Figure 23A).

reactions were separated with SDS-PAGE, detected by phosphoimaging and pageblue staining. B. The alignment of the sequences of Trx-His₆-Hpk8 KD, Trx-His₆-Hpk8_{H172A} KD, His₆-Hpk8 with the kinase domain of *C. crescentus* DivL and the STK domain of *M. xanthus* MasK. The sequences of Trx-His₆-Hpk8 KD and Trx-His₆-Hpk8_{H172A} KD in the black box are the Trx-His₆-tag region. The sequences in the red box of His₆-Hpk8 are His₆-tag region. The predicted two α -helices on the HisKA domain are highlighted. The conserved His residues in the HisKA domains are highlighted with asterisk. This residue has been substituted to Ala in the Trx-His₆-Hpk8_{H172A} KD protein while the residue at this position in the kinase domain of *C. crescentus* DivL is Tyr. The tyrosine kinase specific active-site signature (VxHRDISxxNL) of MasK is highlighted with black underline.

The phosphoryl signal of Hpk8 proteins was reduced after incubation with 1 N NaOH and this reduction was equal to the decrease in protein levels as judged from SDS-PAGE staining. After the incubation with 1 N HCl, the phosphoryl signal of the Hpk8 proteins was reduced but clearly part of the phosphoryl signal remained. When the protein was incubated with 0.5 N HCl under the same conditions, a smaller decrease in the phosphoryl signal was observed (data not shown here). This indicates that the phosphoryl group in the Hpk8 proteins is resistant to acid conditions. In the presence of hydroxylamine and pyridine, the level of phosphorylation signal was not affected. These data suggest that the Hpk8 proteins are likely phosphorylated on Thr, Tyr or Cys. Phosphocysteine is relatively less abundant in proteins than phosphothreonine and phosphotyrosine (Duclos *et al.*, 1991). There are no Thr, Tyr or Cys residues in the His₆-tag region of His₆-Hpk8, thus these data suggest that the Hpk8 proteins are phosphophorylated on a Thr, Tyr or Cys residue in Hpk8 part but not in the tag region.

In preliminary immunoblot analysis with a phosphotyrosine monoclonal antibody, the phosphorylated Trx-His₆-Hpk8 KD and Trx-His₆-Hpk8_{H172A} KD were detected but not the phosphorylated His₆-RodK_{HNNN} (data not shown) indicating that Hpk8 proteins are possibly phosphorylated on a tyrosine residue. To determine whether Hpk8 contains the conserved motifs of tyrosine protein kinases, the sequences of purified Hpk8 proteins were aligned with the kinase domain of *C. crescentus* DivL and the tyrosine/serine/threonine kinase (STK) domain of *M. xanthus* MasK (Figure 23B). DivL is a HPK homolog and the conserved His residue is substituted with Tyr, which is the phosphorylation site *in vitro* assay (Wu *et al.*, 1999). MasK contains the tyrosine kinase specific active-site signature (VxHRDISxxNL) (Thomasson *et al.*, 2002). The sequence

analysis of Hpk8 proteins suggests that Hpk8 does not contain the tyrosine kinase domain or tyrosine kinase specific active-site signature. Moreover, Hpk8 shares a high level of homology with HPKs, which contain the conserved His residue in the H box. The possible mechanism of phosphorylation of Hpk8 remains to be further investigated and the phosphoamino acid needs to be determined.

3.2.10 Preliminary data of phosphotransfer from Hpk8 to FruA

Trx-His ₆ -Hpk8KD(μ M)	10	10	10	10	10	10	10	10	10	10	10	10
FruA-strep(μ M)	-	-	-	10	10	10	-	-	-	20	20	20
FruA D59N-strep(μ M)	20	20	20	-	-	-	26	26	26	-	-	-
Time (h)	0,5	1	3	3	1	0,5	0,5	1	3	3	1	0,5

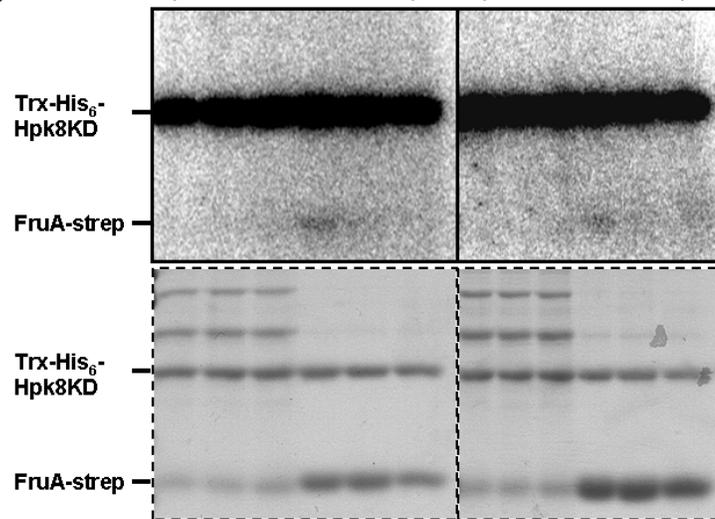


Figure 24. Phosphotransfer assay from Trx-His₆-Hpk8 KD to FruA.

The panels with solid outline are the phosphoimaging and the panels with dashed outline are the SDS-PAGE staining. The equal volume of the indicated concentrations of FruA-strep or FruA_{D59N}-strep was incubated with the autophosphorylated Trx-His₆-Hpk8 KD (10 μ M) by [γ -³²P] ATP for indicated time. The reactions were stopped, separated by SDS-PAGE and detected by phosphoimaging and pageblue staining. The size of Trx-His₆-Hpk8 KD and FruA-strep (same as FruA_{D59N}-strep) are marked on the figure.

To examine whether Hpk8 protein could transfer the phosphoryl group to FruA, a phosphotransfer assay was performed *in vitro*. Trx-His₆-Hpk8 KD (10 μ M) was autophosphorylated with [γ -³²P] ATP for 18 hrs. 5 μ l of the autophosphorylated Trx-His₆-Hpk8 KD (10 μ M) was mixed with an equal volume of FruA-strep (10 μ M or 20 μ M) or FruA_{D59N}-strep (calculated concentration of 20 μ M or 26 μ M) for each phosphotransfer reaction and incubated at 25°C for a defined time. The reactions were stopped, separated by SDS-PAGE and detected by

phosphoimaging and pageblue staining (Figure 24). The phosphorylated band at the expected size of FruA-strep was visible in the reactions from 3 hrs incubation of phosphorylated Hpk8 and 10 μ M or 20 μ M of FruA-strep while this band was not observed in the reactions of incubation of phosphorylated Hpk8 and the calculated concentration of 20 μ M or 26 μ M FruA_{D59N}-strep (whereas less actual amount of the FruA_{D59N}-strep was observed). In the same period, no obvious reduction in the phosphorylation of Hpk8 was observed. A control experiment in which equal amounts of FruA_{D59N}-strep and FruA-strep are added to the phosphotransfer reactions should be performed in future analysis. This preliminary result implies that Hpk8 could phosphotransfer to or phosphorylate FruA with low efficiency or that the observed phosphorylated band of FruA is due to unspecific phosphorylation *in vitro*.

3.2.11 Accumulation of Hpk8 during development

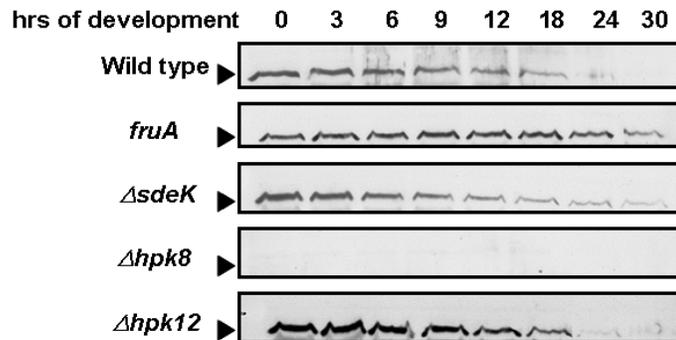


Figure 25. Accumulation of Hpk8 during development.

The Hpk8 accumulation was checked by immunoblot analysis with Hpk8 antibodies. The tested samples were harvested during the development time course of wild type, *fruA* mutant and the mutants of FruA kinase candidate ($\Delta sdeK$, $\Delta hpk8$ or $\Delta hpk12$) starved in the submerged culture. The solid triangle points to the size of Hpk8. In each lane, the lysate of 1×10^7 cells was loaded.

To test the accumulation of Hpk8 during development, cells of wild type, *fruA* mutant and the mutants of the FruA kinase candidates were harvested during development and the samples were examined by immunoblot analysis with the Hpk8 antibodies produced against His₆-Hpk8. The antibodies detected a 57 kD protein corresponding to the expected size of Hpk8 in wild type, which is not present in $\Delta hpk8$ (Figure 25). In wild type, Hpk8 is present at the start point of development and the level of Hpk8 remained at the same level at 3 hrs and 6 hrs after initiation of development, decreased afterwards and was not

detectable after 24 hrs of development. In a *fruA* mutant, the level of Hpk8 was similar to that in wild type at 0 hr of development, remained at the same level until 24 hrs of development and decreased at 30 hrs of the development. In a Δ *sdeK* mutant, the level of the Hpk8 protein in the first 18 hrs of development was similar as that in wild type but Hpk8 was still detectable at a low level at the time points of 24 hrs and 30 hrs. In Δ *hpk12* mutant, the expression profile of Hpk8 was similar to that in wild type at all tested time points. This result indicates that FruA is not essential for accumulation of Hpk8 in the first 6 hrs of development but affects the decrease of HPK8 directly or indirectly at the late time points. Likewise, SdeK is not essential for the accumulation of Hpk8 in the first 18 hrs of development but slightly affect the decrease of HPK8 at late time points. Finally, Hpk12 is not important for the accumulation of Hpk8 during development.

3.2.12 Preliminary data of autophosphorylation of SdeK and Hpk12

Since Trx-His₆-SdeK KD was insoluble in *E. coli*, only a low amount of the protein with a low level of purity was purified under native condition. Although further gel filtration chromatography was performed, still there were some contaminating proteins from *E. coli* in the fraction containing Trx-His₆-SdeK KD protein (Figure 20C). With this purified native protein, the same reaction was conducted as described by Pollack and Singer (Pollack & Singer, 2001). However, no band corresponding to phosphorylated SdeK was observed at the expected size of Trx-His₆-SdeK KD (data not shown). Next, Trx-His₆-SdeK KD was denatured, purified (Figure 20D) and autophosphorylated following the protocol of Pollack and Singer 2001. However, still no phosphorylated band was observed at the expected size of Trx-His₆-SdeK KD (data not shown). By now, the recombinant protein of His₆-SdeK was purified under denaturing conditions (Figure 20H). The protein was refolded (Material and Methods) and was used in autophosphorylation assay. Nevertheless, still no phosphorylated band was observed at the expected size of His₆-SdeK (data not shown).

For the partially purified native Trx-His₆-Hpk12 KD, the described kinase buffer (Porter *et al.*, 2007, Rasmussen *et al.*, 2006) with either MgCl₂ or MnCl₂ was tested for the *in vitro* phosphorylation assay. The autophosphorylation of Trx-His₆-Hpk12 KD was detected in the reaction with MnCl₂ (Figure 26) but not with

that containing $MgCl_2$. The time course of the phosphorylation of Trx-His₆-Hpk12 KD was examined (Figure 26). Since Trx-His₆-Hpk12 KD was not pure, an intense band was also observed by phosphoimaging below the expected size of Trx-His₆-Hpk12 KD. To get a clear *in vitro* phosphorylation assay, the purification of Trx-His₆-Hpk12 KD should be modified and the point mutant of the conserved His residue to Ala of Trx-His₆-Hpk12 KD protein should be constructed for the future assays.

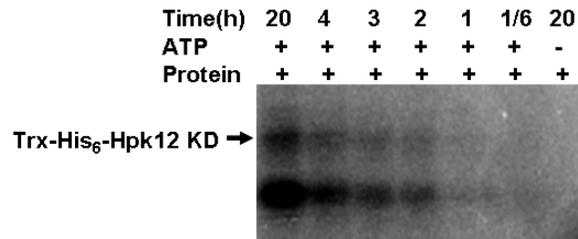


Figure 26. *In vitro* phosphorylation assay of Hpk12.

The reactions in each lane were performed as following: 10 μ M of Trx-His₆-Hpk12 KD was incubated in the phosphorylation reaction (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 50 mM KCl, 20 mM $MnCl_2$, 0.5 mM ATP and 0.5 mM [γ -³²P] ATP) (one reaction without ATP as control) at 25 °C for indicated period of time. The reactions were stopped by adding SDS loading buffer, separated by SDS-PAGE and analyzed with phosphoimaging and pageblue staining. The black arrow points to the expected size of Trx-His₆-Hpk12 KD.

3.2.13 Preliminary analyses of the redundant FruA kinase candidates

Four HPKs, Hpk9, Hpk11, Hpk13 and Hpk29, were identified as potentially functionally redundant FruA kinases from the phenotype of the null mutants and the Y2H analysis. The kinase domains of Hpk11 and Hpk13 are closely related from phylogenetic analysis and share 49% identity, so the double deletion mutant of *hpk11* and *hpk13* was constructed to examine whether the lack of development defects in the single mutants was due to masking by the other kinase still present. The development assay was performed under three different starvation conditions: CF agar, TPM agar and submerged culture (Figure 27).

The double mutant $\Delta hpk11\Delta hpk13$ exhibited the phenotype of wild type on TPM agar and in the submerged culture, while $\Delta hpk11\Delta hpk13$ displayed a slightly delayed aggregation phenotype on CF agar after 18hrs of development. However, this mutant sporulates at a level similar to that of wild type after 120

hrs of development. This suggests that Hpk11 and Hpk13 only have a minor function during development.

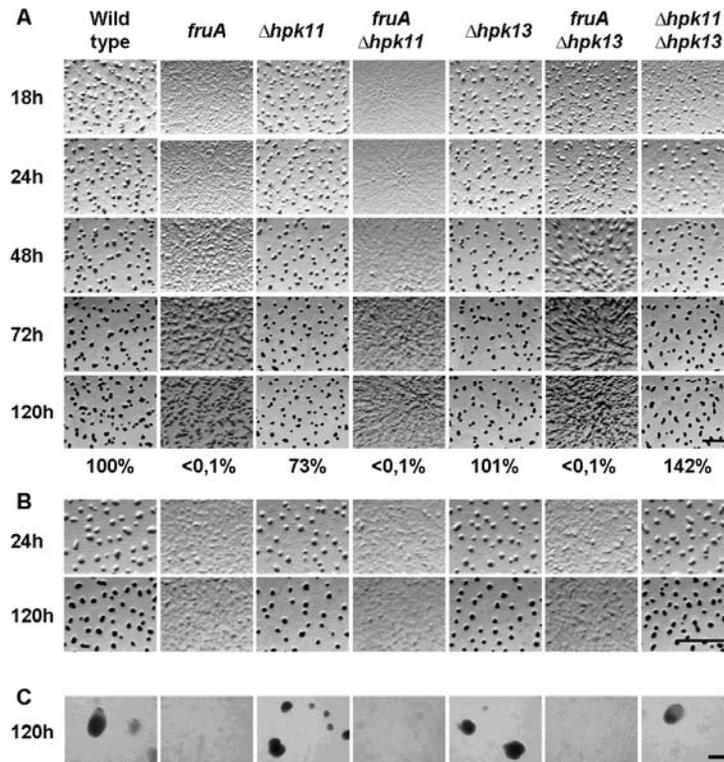


Figure 27. Development assay of relevant double mutants of *fruA*, *hpk11* and *hpk13*.

A. Developmental phenotypes on CF agar. The strains were starved on CF agar for the indicated period of time. Scale bar: 1.0 mm. The numbers in the right side of photos are the percent of sporulation efficiency of each strain after 120 hrs starvation on the CF agar compared with that of wild type. B. Developmental phenotypes on TPM agar. The same strains as in (A) were exposed to starvation on TPM agar for indicated time. Scale bar: 1 mm. C. Developmental phenotypes in submerged culture. The same strains as in (A) were exposed to starvation in submerged culture for 120 hrs. Scale bar: 100 μ m.

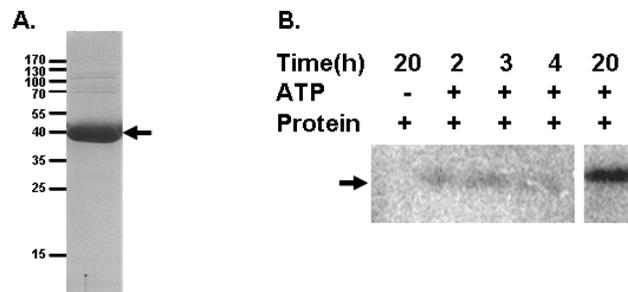


Figure 28. Purification and autophosphorylation assay of Trx-His₆-Hpk13 KD.

A. SDS-PAGE analysis of purified Trx-His₆-Hpk13 KD. The black arrow points to the size of Trx-His₆-Hpk13 KD. B. *In vitro* autophosphorylation assay of Trx-His₆-Hpk13 KD. The reaction was performed as following: 10 μ M of Trx-His₆-Hpk13 KD was incubated in the phosphorylation reaction (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 50 mM KCl, 20 mM MnCl₂, 0.5 mM ATP and 0.5 mM [γ -³²P] ATP) (one reaction without ATP as control) at 25 °C for indicated

period of time. The reactions were stopped by adding SDS loading buffer, separated by SDS-PAGE and analyzed by phosphoimaging. The black arrow points to the expected size of Trx-His₆-Hpk13 KD.

To examine whether the potentially functionally redundant FruA kinase candidates interact with FruA *in vitro*, the kinase domains of Hpk9, Hpk11 and Hpk13 were over-expressed as recombinant proteins in *E. coli*. The recombinant proteins of Trx-His₆-Hpk9 KD and Trx-His₆-Hpk11 KD were insoluble under all tested induction conditions (Material and Methods, data not shown). Trx-His₆-Hpk13 KD was soluble using a salt induction procedure and this protein was purified (Figure 28A) with Ni-NTA affinity chromatography (Material and Methods). *In vitro* phosphorylation assay was performed with the purified Trx-His₆-Hpk13 KD using the same conditions as for autophosphorylation of Trx-His₆-Hpk8 KD, and phosphorylated protein with the expected size of Trx-His₆-Hpk13 KD was observed by phosphoimaging (Figure 28B). The phosphotransfer assay from Hpk13 to FruA will be performed in the future.

3.3 Characterization of Hpk37

3.3.1 Analysis of Hpk37

From the phenotypic analysis of 25 developmentally up-regulated orphan HPKs and Y2H tests of the interaction between Hpk37 and FruA, it was unclear whether Hpk37 is a relevant FruA kinase candidate. Therefore, a further analysis was performed. The genetic organization of *hpk37* in the genome is shown in Figure 29A. A gene encoding a methyltransferase and a gene encoding a methylesterase are located immediately downstream of *hpk37* (Figure 29A). These three genes together with the downstream RR (MXAN0715) are likely to be in the same operon.

The structure of Hpk37 is highly interesting with the following domain organization from the N-terminus: ten tandem HAMP domains, a GAF domain, a coil-coiled region, a kinase domain and three receiver domains (Figure 29B). The sequence of the coiled-coil region of Hpk37 between the GAF domain and the kinase domain is rich in Q and E residues (37% of the all residues in this region is Q or E). Interestingly, methyl accepting chemotaxis protein homolog is methylated on Q or E residue in a part of the protein that forms a long coiled-coil region. The observation that genes encoding a methyltransferase and a methylesterase are located downstream of Hpk37 in the genome, suggests that Hpk37 could be methylated.

To determine whether the coiled-coil region of Hpk37 contains the identified methylation sites, the sequence of this region was aligned with the methylation site sequences of FrzCD from *M. xanthus* and Tar from *E. coli* (Figure 29C). In *E. coli*, four methylation sites of Tar (*E. coli*) are as following Q295, E302, Q309 and E491 (the first three sites have been marked on Figure 29C). The Q residue at the position of 295 and 309 could be rapidly deamidated *in vivo* by the chemotaxis methylesterase (Terwilliger & Koshland, 1984). The consensus sequences of the methylation sites were identified as E-E-x-x-A-T/S (x stands for the any amino acids and the second E is the methylation site) (Terwilliger & Koshland, 1984).

Hpk37 contains a sequence of the methylation site of EQQAT at the position 1201-1206, which is identical to the consensus of the third methylation site EQxxAT of Tar that has been determined as the most rapid methylation and

al., 1983). The amino acids circled with black box are the confirmed methylation sites. The potential methylation site sequence of Hpk37 is circled with yellow box and the potential methylation site is pointed with black arrow. D. Alignment of 10 HAMP domains of Hpk37, the HAMP domain of Tar from *E. coli* and Af1503 from *A. fulgidus*. Residues of Af1503 involved in contacts between the helices and the connector are marked with asterisks. Positions of Af1503 form the *a* and *d* core layers of the alternate structure. Positions forming the *x* and *da* core layers of structure were indicated on the figure. E. The alignment of three receiver domains of Hpk37, the receiver domain of downstream response regulator (MXAN0175) and CheY from *E. coli*. The conserved residues, which are important for the phosphorylation of CheY, are marked with black triangles.

The different methylation sites (E168, E175 and E182) of the *M. xanthus* FrzCD have been identified (Scott *et al.*, 2008). One methylation site E182 of FrzCD shares the same paradigm as the methylation site Q295 of *E. coli* Tar, while two unusual methylation sites at E168 is the methylated at the first E residue of the consensus sequence of the methylation site and E175 is flanked by H and T instead of flanking by E or Q (Scott *et al.*, 2008). In the coiled-coil region of Hpk37, no sequences are identical to the methylation site sequences of FrzCD. However, more methylation sites with unidentified consensus may remain this coiled-coil region (91aa total). In conclusion, Hpk37 contains at least one potential methylation site of EQQAT at the position 1201-1206 from the sequence analysis. This suggests that Hpk37 might function as a HPK, which could adapt by altering methylation state through the downstream methyltransferase and methylesterase to control kinase activity.

Hpk37 contains ten HAMP domains in the N-terminal region while no transmembrane helix was detected (Table S1). Those 10 HAMP domains of Hpk37 were aligned in comparison with HAMP domains of Tar from *E. coli* and Af1503 from *A. fulgidus* (Figure 29D). The sequence of HAMP4 is exactly identical to HAMP7. The identities between all of the possible combination of HAMP2, HAMP3, HAMP4, HAMP5, HAMP6, HAMP7 and HAMP8 are above 88%. HAMP1, HAMP9 and HAMP10 are not as closely related as the rest seven HAMP domains of Hpk37. Previous studies of the structure of HAMP domain suggested that a HAMP domain could act as a converter of conformational changes between two parts of a protein by a rotational mechanism (Hulko *et al.*, 2006). The ten HAMP domains of Hpk37 may provide multiple switches of conformational changes in response to a certain stimuli.

The alignment of the three receiver domains of Hpk37, the receiver domain of downstream RR (MXAN0715) and the receiver domain of CheY from *E. coli* is shown in Figure 29E. The phylogenetic analysis of RRs in *M. xanthus* demonstrated that the three receiver domains of Hpk37 and the receiver domain of MXAN0715 belong to different subfamilies (A2, B1, B3 and A4) following the Grebe and Stock's classification, which suggest that they might be functionally distinct.

3.3.2 Motility assay of $\Delta hpk37$

As described in the text 3.1.7, in-frame deletion mutant of *hpk37* has normal motility on 0.5% CTT medium containing 0.5% agar (Figure 30A), which favours motility by means of the S-motility system, and normal motility (Figure 30B) on 1.5% agar, which favours motility by means of the A-motility system. This suggested that Hpk37 is not required for motility and that the strong developmental defect is caused by the in-frame deletion of *hpk37*, which is not due to impaired motility.

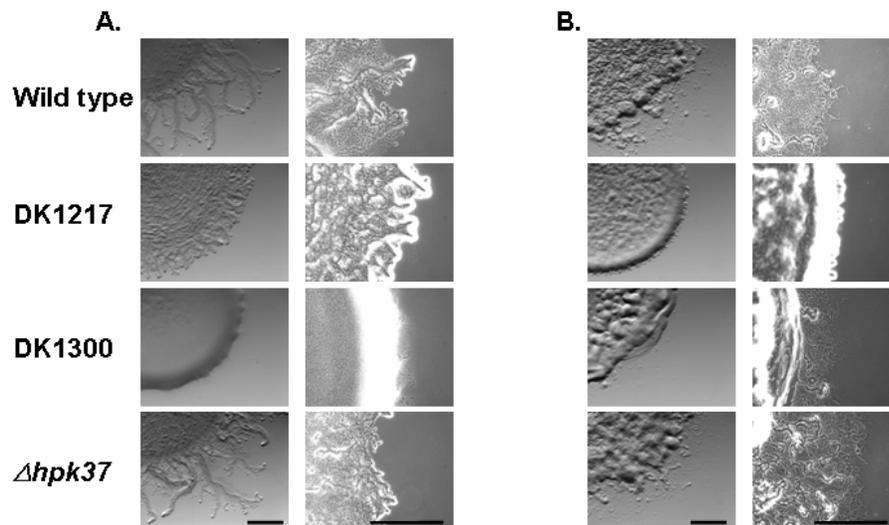


Figure 30. Motility assay of $\Delta hpk37$.

A. Motility phenotypes of $\Delta hpk37$ comparing with wild type, DK1217 (A-S+) and DK1300 (A+S) on 0.5% CTT medium containing 0.5% agar. The cells were in favor of S motility on this surface. The scale bar on the left side photos is 1 mm and the scale bar on the right side of photos is 0.1 mm. B. Motility phenotypes of $\Delta hpk37$ comparing with wild type, DK1217 (A-S+) and DK1300 (A+S-) on 0.5% CTT medium containing 1.5 % agar. The cells were in favor of A motility on this surface. The scale bar on the left side photos is 1 mm and the scale bar on the right side of photos is 0.1 mm.

3.3.3 FrzCD methylation and FruA accumulation in $\Delta hpk37$

To determine the function of Hpk37 at the molecular level during development, the FrzCD methylation profile and FruA accumulation profile during development were examined in $\Delta hpk37$ compared to wild type. As described in 3.2.5, FrzCD in wild type is present as a mixture of a methylated and an unmethylated form at the initiation of development, then FrzCD shifts to the unmethylated form after 1 hr, afterwards the unmethylated FrzCD is converted to the methylated form (Figure 31A). In the $\Delta hpk37$ mutant, the FrzCD methylation pattern was similar to that of wild type during the first 6 hrs of development, afterwards the conversion from unmethylated FrzCD to methylated FrzCD was 6 hrs delayed and the methylated FrzCD was still present at 24 hrs and 30 hrs (Figure 31A). This demonstrates that Hpk37 act upstream of FrzCD methylation during development.

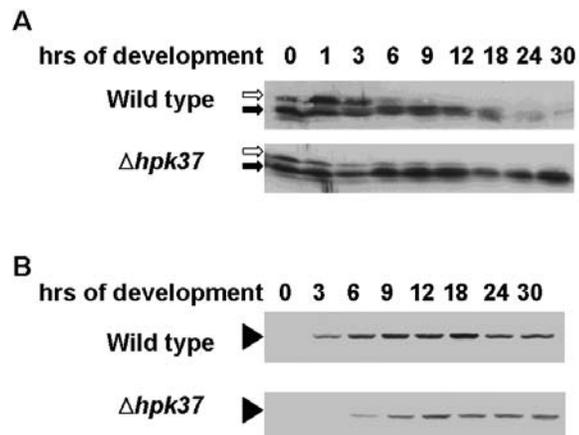


Figure 31. FrzCD methylation and FruA accumulation during development in $\Delta hpk37$.

A. FrzCD methylation during development in $\Delta hpk37$ compared with wild type by immunoblot analysis. The blank arrow points to the unmethylated FrzCD. The solid arrow indicates the methylated FrzCD. B. FruA accumulation during development $\Delta hpk37$ comparing that in wild type by immunoblot analysis. The solid triangle points to FruA.

The accumulation of FruA in wild type is developmentally regulated (Figure 31A) as previously described (Results 3.2.5). FruA was first detectable in $\Delta hpk37$ by immunoblot analysis after 6 hrs of development, which was three hours later compared with that in wild type. Afterwards the level of FruA in $\Delta hpk37$ was lower than that in wild type at all tested time points. This indicates that Hpk37

act upstream of FruA accumulation and that Hpk37 is involved in FruA accumulation.

3.3.4 Developmental gene expression in $\Delta hpk37$

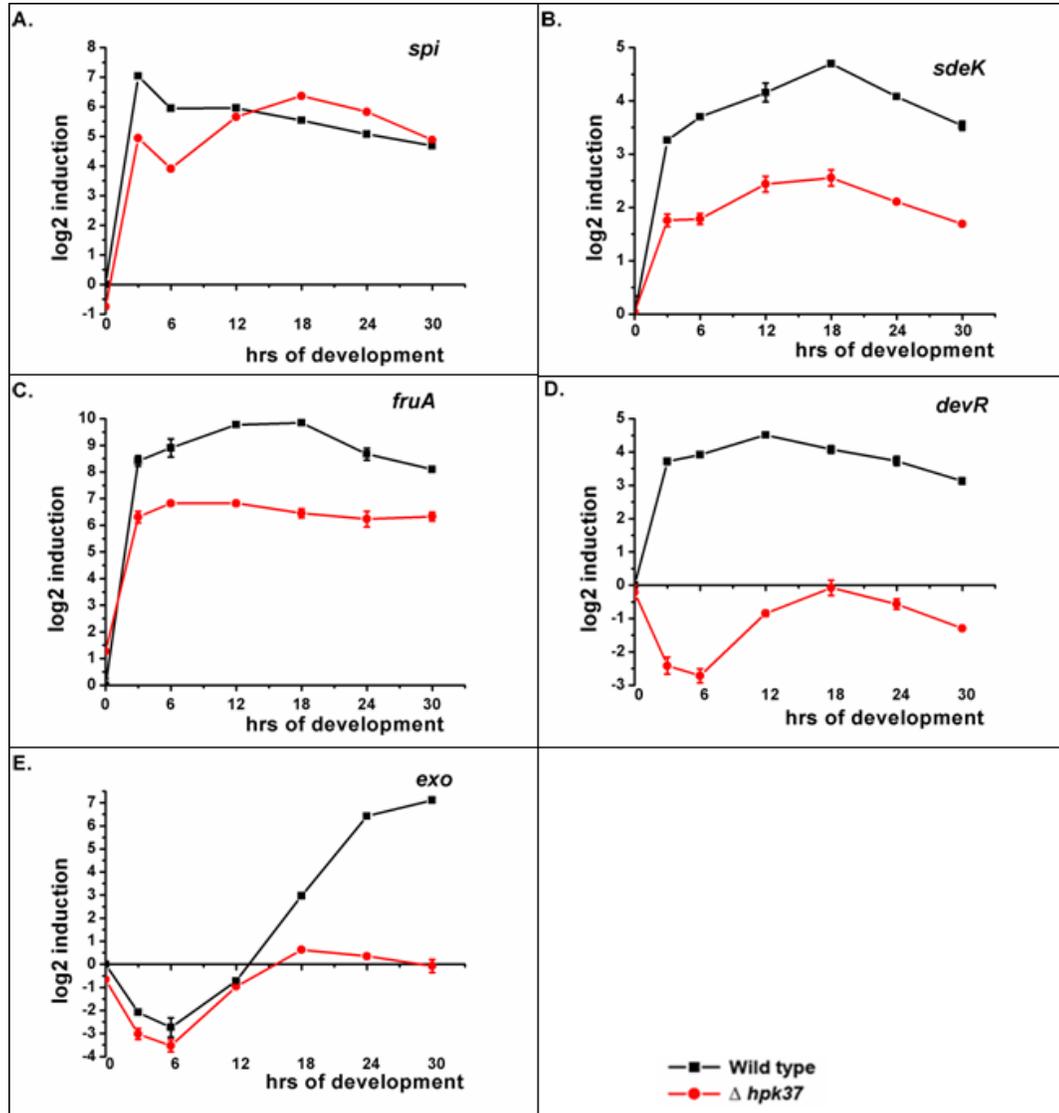


Figure 32. qRT-PCR analysis of five marker genes expression in wild type (black square) and $\Delta hpk37$ (red circle).

x axis represents the time after initiation of development, y axis represents the log₂ ratio of the transcriptional level in comparison to that of wild type at 0 hr. Expression in the mutant was regarded as significantly altered if at least two consecutive time points different by at least 3.5 fold (log₂ ratio = 1.8) from wild type.

To determine the function of Hpk37 during development, the transcription profiles of five marker genes (*spi*, *sdeK*, *fruA*, *devR* and *exo*) were examined in the in-frame deletion mutant of *hpk37* by qRT-PCR (Figure 32) with the same

procedure described previously (3.2.6). Expression in the mutant was regarded as significantly altered if at least two consecutive time points different by at least 3.5 fold (\log_2 ratio = 1.8) from wild type. The *spi* gene (the A-signal dependent gene) expression level in $\Delta hpk37$ at 3 hrs and 6 hrs of development was around 4 fold lower than that in wild type, while afterwards the expression profile of *spi* in $\Delta hpk37$ was similar to that in wild type. *sdeK* (a (p)ppGpp dependent gene) and *fruA* (a A-signal dependent gene) expression in $\Delta hpk37$ was developmentally up-regulated but the expression level of *sdeK* and *fruA* in $\Delta hpk37$ was lower than that in wild type at all tested time points. The *devR* and *exo* (the C-signal dependent genes) expression was significantly low in $\Delta hpk37$ compared with wild type. Taken together, these results indicated that Hpk37 act upstream of the C-signal pathway and that Hpk37 is required for production or receiving of (p)ppGpp level and/or A-signal. These results seem to exclude that Hpk37 could be a FruA kinase.

3.3.5 Purification of Hpk37₁₁₄₈₋₁₉₆₇

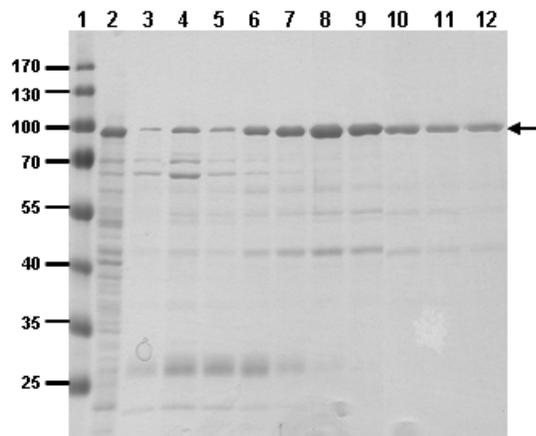


Figure 33. Purification of His₆-Hpk37₁₁₄₈₋₁₉₆₇

The samples in each lane were from different steps of the purification of His₆-Hpk37₁₁₄₈₋₁₉₆₇ as follows. 1. Prestained protein ladder; 2. The lysate of overexpressed *E. coli* cells; 3. Wash; 4. Elution 1 with 50 mM imidazole; 5. Elution 2 with 100 mM imidazole; 6. Elution 3 with 150 mM imidazole; 7. Elution 4 with 150 mM imidazole; 8. Elution 5 with 200 mM imidazole; 9. Elution 6 with 250 mM imidazole; 10. Elution 7 with 250 mM imidazole; 11. Elution 8 with 250 mM imidazole; 12. Elution 9 with 250 mM imidazole. The black arrow points to the calculated size of His₆-Hpk37₁₁₄₈₋₁₉₆₇.

To investigate whether Hpk37 is methylated during development *in vivo*, first, Hpk37 antibodies are required. For production of Hpk37 antibodies, the last 820aa of Hpk37 with N-terminal His₆-tag (His₆-Hpk37₁₁₄₈₋₁₉₆₇) was purified

under native conditions. This protein was purified with Ni-NTA affinity chromatography and eluted with a gradient imidazole from 50mM to 250mM. The last four fractions of the purification (Figure 33) were combined. The MS Maldi-Tof analysis showed that all of the bands below the expected size of His₆-Hpk37₁₁₄₈₋₁₉₆₇ were derived from Hpk37. This protein was used as an antigen to generate antibodies. The attained antibodies detect the band corresponding to the calculated size of Hpk37 protein in wild-type cell lysis, which is not presented in the cell lysis of $\Delta hpk37$ mutant (tested by Susanne Kneip).

3.4 In search of the FruA kinase with bioinformatics method

To identify the FruA kinase, a bioinformatics approach was also used in parallel with the candidate approach described above. The bioinformatics prediction of FruA kinase candidates was done by our collaborator, Robert White, with the following approach. Briefly, the covarying residues involved in the interaction between HPK and RR were identified by multi-alignment of the 2,700 paired TCS proteins from approximately 200 genomes and determination of the substitution constraints between TCS pairs. A log-likelihood scoring procedure applied to these residues and then a predictive tool was built for assigning signaling mate (White *et al.*, 2007). This method was applied for the prediction of the interaction of TCS proteins in *M. xanthus*. For the prediction of the FruA kinase, the log-likelihood scores of six HPKs were above 0 (Table 8). None of these six HPKs overlap with the FruA kinase candidates identified in the previous candidate approach (Results 3.2).

Table 8. Summary the analyses of the FruA kinase candidates predicted by *in silico* method

Log-likelihood Score	TIGR_MXAN	Gene sym	Dev. defects in mutants ^a	Motility defects in mutants ^a	Interaction with FruA in Y2H ^b
6,498	MXAN4640	<i>hpk40/sgmT</i>	+	+	ND
3,775	MXAN4465	<i>hpk30</i>	+	-	-
1,759	MXAN0229	<i>hpk41</i>	-	-	ND
1,206	MXAN6734	<i>hpk38</i>	-	-	ND
1,103	MXAN4988	<i>hpk27</i>	essential?	essential?	-*
0,644	MXAN6015	<i>hpk42</i>	-	-	ND

^a: +: defects, -: no defects; ^b: - means no interaction with FruA in Y2H; *mean the HPK constructs did not self-interact.

To test whether they are the FruA kinases, in-frame deletion mutants of three *hpk* genes were constructed and the stable insertion mutant of *hpk40* was obtained (*hpk40*, Δ *hpk41* and Δ *hpk42* were constructed by Tobias Petters). The insertion mutant of *hpk40* exhibited strong motility defects and developmental defects suggesting that the developmental defects of the *hpk40* mutant are due to impaired motility. Hpk40 is important for the motility of *M. xanthus* whereas FruA is important for development, which suggested that Hpk40 is not likely to

be the FruA kinase. $\Delta hpk38$, $\Delta hpk41$ and $\Delta hpk42$ displayed no developmental defects as well as no motility defects. Hpk27 is potentially essential for cell viability as previously described in this study, which indicates that Hpk27 may be involved in different pathways as FruA and Hpk27 is unlikely to be the FruA kinase. Hpk30 had developmental defects in the null mutants and did not interact with FruA in the Y2H system. It is not selected as FruA kinase candidate. From all above analyses, none of these six HPKs are likely to be FruA kinases.

3.5 Miscellaneous *hpk* mutants

In order to examine the potential function of HPKs during development, the miscellaneous mutants of 12 HPKs and 3 HPK-like proteins were constructed (listed in Table 9). None of them displayed growth, developmental or motility defects. Of them, SA2111 carries the deletion of *MXAN2317* (= *hpk7*), encoding the only hybrid HPK in *M. xanthus* containing an Hpt domain.

Table 9. The summary of the phenotypes of miscellaneous *hpk* mutants

Strains	Genotype	Genetic organization	Dev. defects in mutants ^a	Motility defects in mutants ^a
SA2101	<i>MXAN0347</i> (= <i>hpk1 short</i>)	orphan	-	-
SA2102	<i>MXAN0347</i> (= <i>hpk1 long</i>)	orphan	-	-
SA2103	<i>MXAN5184</i> (= <i>hpk2 short</i>)	orphan	-	-
SA2104	<i>MXAN5184</i> (= <i>hpk2 long</i>)	orphan	-	-
SA2105	<i>MXAN3343</i> (= <i>hpk3 short</i>)	orphan	-	-
SA2106	<i>MXAN3343</i> (= <i>hpk3 long</i>)	orphan	-	-
SA2109	<i>MXAN6053</i> (= <i>hpk6 short</i>)	orphan	-	-
SA2110	<i>MXAN6053</i> (= <i>hpk6 long</i>)	orphan	-	-
SA2111	Δ <i>MXAN2317</i> (= Δ <i>hpk7</i>)	orphan	-	-
SA2113	Δ <i>MXAN3343</i> (= Δ <i>hpk3</i>)	orphan	-	-
SA2116	Δ <i>MXAN4432</i> (= Δ <i>hpk10</i>)	orphan	-	-
SA2122	Δ <i>MXAN1553</i> (= Δ <i>hpk16</i>)	paired	-	-
SA2123	Δ <i>MXAN5715</i> (= Δ <i>hpk17</i>)	orphan	-	-
SA2125	Δ <i>MXAN7444</i> (= Δ <i>hpk19</i>)	orphan	-	-
SA2127	Δ <i>MXAN7002</i> (= Δ <i>hpk21</i>)	complex	-	-

SA2128	Δ MXAN7003 (=Δhpk22)	complex	-	-
SA2129	Δ MXAN0230 (=Δhpk31)	complex	-	-
SA2132	Δ MXAN0720 (=Δhpk26)	orphan	-	-
SA2135	Δ MXAN6865 (=Δhpk33)	complex	-	-
SA2140	Δ MXAN5704 (=Δhpk32)	orphan	-	-

^a -stands for no defect.

4 Discussion

4.1 Two-component regulatory systems in *M. xanthus*

In this study, 272 genes that encode TCS proteins were identified including 21 genes in eight loci, encoding TCS proteins, which are part of Che-like systems. The further analyses were focused on 251 TCS proteins (non Che-like) containing 118 HPK, 119 RR and 14 HPK-like genes. TCS genes make up 3.4% of the 7380 ORFs (Goldman *et al.*, 2006) in the *M. xanthus* genome. The number of TCS genes in *M. xanthus* complies with the general rule that the number of TCS proteins per genome increases with the square of the genome size (Galperin, 2005, Galperin, 2006).

The *M. xanthus* TCS genes could be divided into three classes based on their genetic organization. 55% and 16% of all the TCS genes are organized as orphan genes or in complex gene clusters, respectively and only 29% are found in the typical organization of TCS genes as paired genes. A comparison with other bacterial genomes illustrates that the percentage of orphan TCS genes shows large inter-species variations. For instance, in *C. crescentus* 57% of the 106 TCS genes are orphans (Skerker *et al.*, 2005) whereas in *Pseudomonas aeruginosa* 16% of the 127 TCS genes are orphans (Rodrigue *et al.*, 2000). Thus, in terms of the percentage of orphan TCS genes, *M. xanthus* is not exceptional.

The genome size of *M. xanthus* is 9.14 Mb and it has been suggested that lineage-specific gene family expansions (LSE) were major contributors to the genomic expansion (Goldman *et al.*, 2006). The preliminary analyses (Huntley and Søgaard-Andersen, In preparation) suggest that a large fraction of TCS genes in *M. xanthus* may have arisen by LSE. It has been reported previously (Alm *et al.*, 2006) that deltaproteobacteria have a propensity for LSE of TCS genes. In agreement with this, preliminary analyses (Huntley and Søgaard-Andersen, In preparation) demonstrate that the complements of TCS genes in the three Myxococcales species, *Sorangium cellulosum*, *Stigmatella aurantiaca* and *Anaeromyxobacter dehalogenans*, also contain a large fraction of TCS genes that likely arose by LSE. Interestingly, the complements of TCS genes that have expanded by LSE in the four Myxococcales species appear to be different suggesting that for each species the particular genes amplified provide

that species with some selective benefits (Huntley and Søggaard-Andersen, In preparation).

Three groups of structurally remarkable TCS proteins were identified in *M. xanthus*. One group consists of 14 HPK-like proteins, which only contain a HisKA domain or a HATPase_c domain (See Table S2). Evidences suggest that these genes are not pseudogenes but code for functional proteins: Firstly, three of these genes (*MXAN0461* (=redE) (Higgs *et al.*, 2005), *MXAN2670* (=asgA) (Plamann *et al.*, 1995) and *MXAN5123* (=mrpA) (Sun & Shi, 2001a, Sun & Shi, 2001b) are required for development. The HPK-like protein LtnC in *Synechococcus elongates* consists of a HisKA domain, a receiver domain and an output domain but it does not contain HATPase_c domain. LtnC could receive the phosphoryl group from a RR LtnA and control the activity of its output domain (Maeda *et al.*, 2006). Secondly, at least nine of the genes were found to be expressed in global transcriptional profiling experiments and two of the genes are transcriptionally up-regulated during development.

A second group of proteins with interesting structural features consists of proteins that have organizations of signal transduction domains, which have not been reported previously and which raise interesting questions in terms of how they function in phosphotransfer reactions. For instance, the orphan HPKs *MXAN2606* and *MXAN2317* are predicted to have the domain structures HisKA-HATPase_c-RR-HisKA-HATPase_c and HisKA-HATPase_c-RR-RR-Hpt respectively and, the RR *MXAN7362*, which is encoded by a complex gene cluster, is predicted to have the domain structure RR-Hpt-RR-RR-GGDEF.

The third group of TCS proteins with interesting structural features are 14 RR with output domains of unknown functions (DUF). These RRs are overrepresented among RRs encoded by complex gene clusters and orphan genes, i.e. 10 and two of these domains are found in RRs encoded by orphan and complex genes, respectively. Interestingly, four of these proteins are orphan RRs involved in regulating gliding motility (*MXAN2991* (=aglZ) (Yang *et al.*, 2004), *MXAN4149* (=frzS) (Ward *et al.*, 2000, Mignot *et al.*, 2005), *MXAN4461* (=romR) (Leonardy *et al.*, 2007) and *MXAN6627* (=sgnC) (Youderian & Hartzell, 2006).

The analysis of *M. xanthus* TCS proteins revealed structural features that have functional implications. First, 73 out of the 118 HPKs are predicted to be cytoplasmic suggesting that the many HPKs in *M. xanthus* may not be involved in monitoring external stimuli or intercellular signals but rather in monitoring cytoplasmic stimuli. Alternatively, they could indirectly be involved in monitoring external stimuli by interacting with membrane proteins. Second, the analysis of output domains in RRs suggests that the output responses from TCS systems in *M. xanthus* center on three types, regulation of gene expression, regulation of di-c-GMP metabolism and unknown functions.

Interestingly, strongly biased distributions of different types of TCS proteins encoded by paired genes, orphan genes and in complex gene clusters were found. For paired TCS genes the main implication is that a large fraction of the corresponding proteins are part of simple 1:1 TCS with an integral membrane HPK and a cognate RR that is involved in regulation of gene expression. The suggested membrane localization of the paired HPKs suggests that they are primarily involved in monitoring external stimuli. Moreover, the under representation of these genes among transcriptionally regulated genes during development indicate that they may be active in vegetative cells. Clearly, this latter implication does not preclude a function during fruiting body formation of these proteins. Consistently, 15 paired TCS genes have been identified, which are important for development (See Table S1, Table S2 and Table S3 for the identity of these genes).

For TCS proteins encoded by orphan genes or in complex gene clusters the biased distribution of protein characteristics and expression profiles suggest that the corresponding HPKs are primarily involved in monitoring cytoplasmic stimuli (due to the overrepresentation of HPKs predicted to be cytoplasmic) and that the main output responses from the corresponding pathways are regulation of gene expression, regulation of di-c-GMP metabolism and unknown functions (as indicated by the overrepresentation of RRs with DUF output domains). Moreover, the overrepresentation of these genes among those that are developmentally regulated during development suggests that many of these genes encode TCS proteins with a function only during development. It should be noticed that the transcriptional regulation during development does not

preclude a function in vegetative cells. Consistently, 16 orphan TCS genes (incl. the four identified in this report) and 6 TCS genes encoded in complex gene clusters have been shown to be important for development or spore germination (See Table S1, Table S2 and Table S3 for the identity of these genes).

A question that remains to be addressed focuses on the connectivity of the TCS proteins in *M. xanthus*. For the paired TCS genes, the almost complete absence of hybrid HPKs and single domain RRs in the corresponding proteins suggests that the paired genes encode proteins that make up simple, linear 1:1 pathways. For TCS proteins encoded by orphan genes and in complex gene clusters, the connectivity has only been analyzed experimentally for the RedCDEF proteins (=MXAN0459-MXAN0462) and the data suggest that these four proteins may constitute a complex phosphorelay (Higgs *et al.*, 2005). As the connectivity of TCS proteins cannot be predicted based on sequence conservation alone (Skerker *et al.*, 2005), this question, therefore, remains open for most of the TCS proteins encoded by orphan genes and in complex gene clusters. The close 1:1 numerical ratio of HPKs and RRs encoded by these genes could lead to the notion that they could also be make up 1:1 pathways. However, two observations argue against this notion. First, hybrid HPKs are overrepresented among these proteins. Second, many of the RRs encoded by these genes are single domain RRs. The overrepresentation of hybrid HPKs and RRs without output domains among the proteins encoded by complex gene clusters and orphan genes strongly suggest that the signal transduction pathways encoded by these genes are structured as phosphorelays and/or are branched. Phosphorelays would likely depend on the presence of Hpt domain containing proteins. In addition to CheA kinases, only two proteins were identified containing Hpt domains, the hybrid, orphan HPK MXAN2317 and the RR MXAN7362, which is encoded in a complex gene cluster. It has been argued that Hpt domains are difficult to identify due to the low level of sequence conservation (Biondi *et al.*, 2007), thus, *M. xanthus* may indeed encode more proteins containing Hpt domains. Clearly, experimental analyzes are needed to address the question of the connectivity of the *M. xanthus* TCS proteins.

This study directly tested the hypothesis that orphan developmentally up-regulated genes could be important for development by focusing on the 25

orphan HPK genes that are up-regulated at the transcriptional level during development. Among these genes, two (*MXAN3036=hpk23* and *MXAN4988=hpk27*) are likely to be essential for viability. Hpk23 and Hpk27 are predicted to contain 3 and 8 transmembrane helices, which suggest that they may be important for the cell envelope integrity or other unknown essential process. Of the remaining 23 HPKs, seven are important for development without having vegetative defects. These seven genes include *MXAN0931* (=espA) (Cho & Zusman, 1999b), *MXAN1014* (=sdeK) (Garza *et al.*, 1998, Pollack & Singer, 2001) and *MXAN6996* (=asgD) (Cho & Zusman, 1999a), which have previously been shown to be important for development. In addition, I identified *MXAN0712* (=hpk37), *MXAN0736* (=hpk12), *MXAN3290* (=hpk8) and *MXAN4465* (=hpk30) as important for development or spore germination. Finally, inactivation of *MXAN6855* (=espC) (Lee *et al.*, 2005) and *MXAN7206* (=mokA) (Kimura *et al.*, 2001), which have previously been reported to be important for development, did not display developmental defects under the conditions used in this study. Clearly, the lack of developmental defects in the remaining 16 mutants could be caused by the corresponding HPKs being functionally redundant. Nevertheless, these data have two implications: First, the transcriptional up-regulation of a TCS gene does not necessarily mean a function during development (at least under the conditions used here). Second, the large number of TCS genes in *M. xanthus* in general and orphan TCS genes in particular have not only evolved to regulate fruiting body formation.

4.2 In search of the FruA kinase

In this study, two candidate approaches were used to identify FruA kinase candidates. One of these approaches was based on the hypothesis that the FruA kinase gene shares key characteristics with the *fruA* gene, i.e. it is orphan, developmentally up-regulated at the transcriptional level, and a null mutant exhibits developmental defects. In the first part of this study, 25 orphan HPKs, which are developmentally up-regulated at the transcriptional level, were identified. Furthermore, Y2H analyses were performed to determine a potential interaction between FruA and these 25 developmental up-regulated orphan HPKs. In total, these analyses resulted in the identification of three groups of interesting HPKs that could be FruA kinase candidates. Three HPKs (SdeK,

Hpk8 and Hpk12) are essential for normal development and interact with FruA in the Y2H. Four HPKs (Hpk9, Hpk11, Hpk13 and Hpk29) are not important for development and interact with FruA in the Y2H. Finally, Hpk37 is essential for development, however, the Y2H analysis was inconclusive.

According to Grebe & Stock's study, members of specific HPK subfamilies often interact with specific subfamilies of RR (Grebe & Stock, 1999). Therefore, I determined which Grebe & Stock groups these eight HPK belong to. SdeK, Hpk8 and Hpk12 as well as Hpk9, Hpk11, Hpk13 and Hpk29 belong to the HisKA 3f sub-family whereas Hpk37 is a member of the HisKA 1b sub-family. FruA belong to U (unclassified subfamily). Members of the RR U group are often hard to determine which sub-family of HPK that they interact with. Previous studies suggest that the recognition specificity between partner HPK and RR lie in the C-terminal end of $\alpha 1$ helix and the loop region between the two helices $\alpha 1$ and $\alpha 2$ in the HisKA domain (Skerker *et al.*, 2008). The eight kinase domains do not share a high level of similarity in this region. If the recognition specificity of HPKs is determined by this region, this may either suggest that not all of these eight kinases are FruA kinases or that the unique structural features of FruA (8 additional amino acids located after the phosphorylation site and lack of conserved amino acids) could allow different regions of the eight HisKA domains to interact with FruA in unique ways. In total, these sequence-based analyses did not provide additional help to eliminate kinases from the FruA kinase candidate list.

Epistasis analyses showed that *fruA* is epistatic to *hpk8* as well as *hpk12*, and thus, suggesting that *fruA* is in the same genetic pathway as both *hpk8* and *hpk12*. The epistasis relationship between *fruA* and *sdeK* could not be determined because the two mutants carrying single mutations and the double mutants have similar phenotypes. Importantly, the phenotype of the *fruA* Δ *sdeK* mutant does not contradict with that SdeK could be a FruA kinase. The epistasis relationship between *sdeK* and *hpk12* indicates that *sdeK* is epistatic to *hpk12* or that SdeK and Hpk12 are in parallel genetic pathways. Epistasis analyses showed that SdeK and Hpk8 are in parallel genetic pathways, *hpk8* and *hpk12* are in parallel genetic pathways, and the phenotype of the triple mutant Δ *sdeK* Δ *hpk8* Δ *hpk12* is different from that of a *fruA* mutant. The results

from the genetic interaction tests are consistent with a model in which SdeK is the main FruA kinase, Hpk12 is a minor FruA kinase and Hpk8 is a FruA~P phosphatase (Figure 34). Furthermore, there may be other HPKs that phosphorylate or cross talk to FruA.

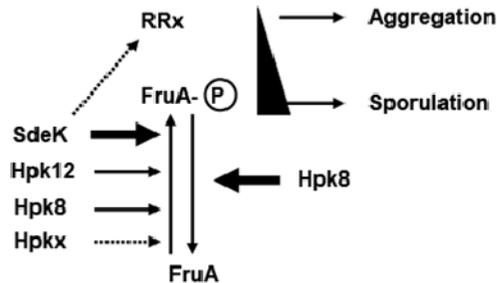


Figure 34. The model of the phosphorylation pathways of FruA.

SdeK is the main FruA kinase, Hpk12 is a minor FruA kinase and Hpk8 is a phosphatase of FruA in total. Moreover, there may be other HPKs that phosphorylate or cross talk to FruA when these three *hpk* genes are deleted. SdeK may have other downstream target except FruA.

To further characterize the potential interactions between FruA and SdeK, Hpk8 and Hpk12, several molecular markers were analyzed in the three kinase mutants. Genetic evidence suggests that FruA phosphorylation is important for the FrzCD methylation during development (Ellehaug, 1999). SdeK is important for FrzCD methylation during development while Hpk8 and Hpk12 are not. The FrzCD methylation pattern is in agreement with the proposed model that SdeK is the main FruA kinase required for FrzCD methylation during aggregation, Hpk12 is a minor FruA kinase, which is not important for aggregation or act after aggregation, and Hpk8 is a minor FruA kinase or FruA~P phosphatase, which is not essential for aggregation or acts after aggregation.

Previous study (Ellehaug *et al.*, 1998) showed that *fruA* transcription and FruA accumulation depend on the A-signal but is independent of the C-signal and FruA suggesting that the FruA kinase(s) would also not be required for FruA accumulation during development. FruA accumulation profiling experiments revealed that neither SdeK, nor Hpk8, nor Hpk12 is required for FruA accumulation during development. These results are in agreement with the proposed model (Figure 34).

To further analyze the function of SdeK, Hpk8 and Hpk12, the expressions of five marker genes (*spi*, *sdeK*, *fruA*, *devR* and *exo*) were examined in the null mutants of the FruA kinase candidates. Of these genes, *spi*, *sdeK*, *fruA* and *exo* are expressed independently of FruA and *devR* is expressed in a FruA dependent manner. These analyses demonstrated that SdeK is partially required for the expression of *devR* and essential for *exo* expression. Hpk8 and Hpk12 were not important for the expression of the marker genes. These data suggest that either SdeK is not a FruA kinase or that SdeK have other downstream targets in addition to FruA, which would be important for *devR* expression (Figure 34). Hpk12 could still be a minor FruA kinase acting late in development and Hpk8 could be a minor FruA kinase or phosphatase acting late in development.

All of the *in vivo* data are in agreement with the proposed model but direct evidence, which SdeK, Hpk8 and Hpk12 interact directly with FruA in phosphotransfer reactions, is still missing. To obtain direct evidences for the interaction between FruA and the FruA kinase candidates, I carried out *in vitro* phosphotransfer experiments with purified proteins. Some data from this part of my work are still preliminary and need further investigations.

Previously, Pollack and Singer (Pollack & Singer, 2001) purified the denatured His₆-SdeK full-length protein, refolded the purified protein and observed autophosphorylated *in vitro*. However, no phosphotransfer was observed between SdeK and FruA. But the negative result in the phosphotransfer reaction may be due to the SdeK protein not being refolded properly. The same authors observed that the *fruA*_{D59E} allele did not lead to bypass of the development defects in the Δ *sdeK* mutant. However, an alternative interpretation of these results is that SdeK has other downstream targets than FruA. I purified a Trx-His₆-SdeK KD protein containing only the SdeK kinase domain. The Trx-His₆-SdeK KD protein was purified under denaturing conditions and refolded, however, no autophosphorylation was observed. Thus, it could not be determined whether SdeK directly phosphorylates FruA. In the future experiments, His₆-SdeK full length protein will be purified, refolded and used in the *in vitro* phosphorylation assay under the same conditions as

described by Pollack and Singer. In total, it remains an open question whether SdeK is a FruA kinase.

Hpk12 may function as a minor FruA kinase. *In vitro* phosphorylation assay was performed with the partially purified Trx-His₆-Hpk12 KD protein containing the Hpk12 kinase domain. This protein was purified under denaturing conditions and refolded. Autophosphorylation of Trx-His₆-Hpk12 KD was detected *in vitro*. However, an additional protein was also observed to be phosphorylated at a high level in these experiments. Therefore, the purification procedure of Trx-His₆-Hpk12 KD should be improved and the point mutant of the conserved His residue to Ala of Trx-His₆-Hpk12 KD protein should be constructed for the future assays.

To analyze phosphotransfer between Hpk8 and FruA, a Trx-His₆-Hpk8 KD protein containing the kinase domain of Hpk8 was purified under native conditions. The protein underwent autophosphorylation *in vitro*. Surprisingly, however, a Trx-His₆-Hpk8_{H172A} KD protein containing a substitution of the potential phosphorylation site H172A was also autophosphorylated in the *in vitro* assay. This suggests that the Trx-His₆-Hpk8 KD is not phosphorylated at the conserved His172 residue *in vitro*. To further explore the site of Trx-His₆-Hpk8 KD autophosphorylation, a His₆-Hpk8 full-length protein was also purified under native conditions. This protein also underwent autophosphorylation *in vitro*. The chemical stability of the phosphoryl group in the three different Hpk8 proteins (Trx-His₆-Hpk8 KD, Trx-His₆-Hpk8_{H172A} KD and His₆-Hpk8) suggests that Hpk8 proteins are likely phosphorylated on either a Thr, a Tyr or a Cys residue in the *in vitro* assay but not on a His or an Asp residue. There is no Thr, Tyr or Cys residue in the His₆-tag region of His₆-Hpk8, suggesting that the phosphorylation of Hpk8 recombinant proteins is on a residue in the Hpk8 moiety.

In preliminary immunoblot analysis with an anti-phosphotyrosine monoclonal antibody, the phosphorylated Trx-His₆-Hpk8 KD and Trx-His₆-Hpk8_{H172A} KD were detected indicating that the Hpk8 proteins are possibly phosphorylated on a Tyr residue. So far, only one bacterial protein kinase, DivL in *C. crescentus* has been shown to be phosphorylated on Tyr residue *in vitro* and this residue is important for the *in vivo* function (Wu *et al.*, 1999). DivL is a homologue of HPK,

however, it contains a Tyr in place of the conserved His residue. DivL has been shown to catalyze the phosphorylation of the RR CtrA *in vitro*. But the molecular mechanism of phosphotransfer from DivL to CtrA is still unknown (Wu *et al.*, 1999). Hpk8 is a typical HPK with a His residue in the H box. Further analysis needs to be performed to investigate the potential novel mechanism of phosphorylation of Hpk8 and to clarify on which residue Hpk8 autophosphorylates.

The genetic analyses indicated that Hpk8 may function as phosphatase on FruA~P. However, based on the observation that some kinases have dual functions acting as both a kinase and a phosphatase on the cognate RR (Igo *et al.*, 1989), I tested the phosphotransfer from Trx-His₆-Hpk8 KD to a FruA-strep protein. A band corresponding to phosphorylation of FruA-strep was observed after 3 hrs incubation with the phosphorylated Trx-His₆-Hpk8 KD. This incubation time is longer than that normally considered to be relevant for specific interaction (Skerker *et al.*, 2005). Importantly, phosphorylation of FruA was not observed with the FruA_{D59N} protein. Obviously, more experiments are needed to determine whether Hpk8 engages in phosphotransfer to FruA or in phosphorylation of FruA.

Hpk8 is a hybrid HPK and contains a C-terminal receiver domain. No reduction of the phosphoryl signal in the chemical stability tests was observed after treatment with pyridine or hydroxylamine, indicating that the *in vitro* phosphorylation signal of His₆-Hpk8 is neither due to nor partially due to the phosphorylation on the Asp residue in the receiver domain of Hpk8. Whether Hpk8 kinase domain can engage in phosphotransfer to its own receiver domain remains to be investigated.

The preliminary result of the *in vitro* phosphotransfer experiments leaves several open ends. First, the unique structure of FruA may cause a slower rate of FruA phosphorylation than that of other RRs. FruA contains substitutions at the place of three conserved position: E14 (D in most RRs), G15 (D in most RRs) and R117 (K in most RRs). Previous studies of mutant CheY proteins with single amino acid substitutions have shown the following: a D15N substitution (same position as G15 in FruA) significantly reduces the rate of phosphotransfer from CheA to CheY (Lukat *et al.*, 1991); a K109R substitution (same position as

R117 in FruA) slightly reduces the rate of phosphotransfer from CheA to CheY and significantly reduces the rate of dephosphorylation of CheY (Lukat *et al.*, 1991). The substitutions in FruA at G15 and R117 may cause a slow rate of the phosphotransfer to FruA. Second, Hpk8 does not appear to be a typical HPK. Therefore, phosphotransfer to or phosphorylation of FruA could be slower than that normally observed for cognate HPK/RR. Third, Hpk8 may primarily function as a phosphatase on FruA~P. Fourth, Hpk8 is a hybrid kinase and contains a C-terminal receiver domain. Therefore, Hpk8 may be part of a phosphorelay and transfer the phosphoryl group to FruA via an Hpt protein. In *B. subtilis*, the HPK KinA phosphorylates RR Spo0F, then the phosphoryl group is shuttled to SpoB (containing the Hpt domain) and finally to RR Spo0A (Burbulys *et al.*, 1991). In *in vitro* assays, KinA can phosphorylate both Spo0A and Spo0F, but KinA has a more than 50,000-fold higher preference for Spo0F than for Spo0A (Grimshaw *et al.*, 1998). At last, it remains a possibility that Hpk8 does not engage in specific phosphotransfer reactions with FruA.

The candidate approach described above is based on the hypothesis that FruA kinase candidates share the same characteristics as FruA. The advantage of this method is that it narrows down the number of candidates step by step rationally. But this method also has potential pitfalls. First, the FruA kinase may not be transcriptionally up-regulated during development. Second, the FruA kinase gene may not be orphan. Third, FruA may be part of a phosphorelay system with FruA receiving the phosphoryl group from an Hpt domain protein. Although the only hybrid HPK (MXAN2317) with an Hpt domain in *M. xanthus* is not important for development, unidentified Hpt domains may still remain to be discovered in the *M. xanthus* genome. Finally, the Y2H analysis may provide false positive or false negative results.

The second candidate approach to identify FruA kinase candidates is based on bioinformatics predictions, which was performed in collaboration with Robert White. By multi-alignment of the 2,700 coupled TCS proteins from around 200 genomes and determination of the substitution constraints between TCS pairs, the covarying residues potentially involved in the interaction between HPK and RR were identified. A log-likelihood scoring procedure was applied to these residues and then a predictive tool was built for assigning signaling mate (White

et al., 2007). For the prediction of the FruA kinase, the log-likelihood scores of six HPKs were above 0. None of these candidates overlapped with those identified in the first candidate approach used in this study. Phenotypic analyses of the six corresponding null mutants suggested that none of these six HPKs are FruA kinases. The negative results in this approach could be explained in at least two scenarios. First, the FruA kinase is in the list of *M. xanthus* TCS proteins but it is missed with this prediction program. This could be caused by the original dataset of TCS sequences not being representative, the identified covariation residues are not sufficient to determine the recognition specificity or the unique sequence features of FruA have a negative effect on the prediction value. Second, FruA may not be phosphorylated directly by an HPK but is part of a phosphorelay and depends on a yet unidentified Hpt domain protein for phosphorylation.

As a general method to investigate the connectivity of orphan TCS proteins, a previous study on TCS proteins of *C. crescentus* used a phosphotransfer profiling method (Skerker *et al.*, 2005). In a recent study (Whitworth *et al.*, 2008), attempts were made to connect 29 HPK kinase domains and 40 receiver domains of *M. xanthus* using the Y2H system. 255 potential interactions between kinase domains and receiver domains were identified but these interactions remain to be confirmed by further studies. To determine the connectivity of TCS proteins in *M. xanthus*, the phosphotransfer profiling method is an obvious choice. However, it will involve a massive effort due to the high number of TCS proteins in *M. xanthus*. Another concern regarding phosphotransfer profiling in *M. xanthus* is the potentially unidentified Hpt domain proteins. But phosphotransfer profiling seems to be a straight forward method compared with the Y2H analyses in general.

4.3 Hpk37

The phenotypic analysis of the Δ *hpk37* mutant and the Y2H analysis does not exclude that Hpk37 could be a FruA kinase. However, FruA accumulation and FrzCD methylation profiling experiments demonstrated that Hpk37 acts upstream of FruA accumulation and FrzCD methylation. Moreover, the expression profiles of marker genes indicate that Hpk37 act upstream of the C-signal transduction pathway and that it is involved in the production or response

to (p)ppGpp or A-signal. These results strongly support that Hpk37 is not a FruA kinase.

However, the structure of Hpk37 is highly interesting. Hpk37 contains ten tandem N-terminal HAMP domains, followed by a GAF domain, a coiled-coil region, a kinase domain and three receiver domains. The sequence of the coiled-coil region is rich in Q and E residues (37% of the all residues in this region are Q or E). By sequence analysis, Hpk37 in this region contains a methylation site, which is identical to the consensus of the third methylation site of *E. coli* Tar (Terwilliger & Koshland, 1984). The third methylation site of *E. coli* Tar has been reported as the fastest methylation site and demethylation site *in vitro* (Terwilliger *et al.*, 1986). In addition, genes encoding a methyltransferase (CheR-like) and a methylesterase (CheB-like) are located downstream of *hpk37* and are likely co-transcribed with *hpk37*, suggesting that Hpk37 could be methylated.

The structure of Hpk37 brings up three main questions for future analyses. First, Hpk37 might function as a HPK, which could adapt to a sensory input by methylation and demethylation. Such a mechanism is at work for MCP proteins, which are part of chemosensory systems (Wadhams & Armitage, 2004). However, regulation of HPK activity by methylation/demethylation has not been reported. Interestingly, bioinformatic analyses have demonstrated that HPKs containing a coiled-coil region rich in E and Q upstream of the HisKA domain is not uncommon (Stuart Huntley and Lotte Sogaard-Andersen, unpublished). This may imply that regulation of HPK activity by methylation/demethylation is a general mechanism. Second, the function of the N-terminal poly-HAMP domains remains enigmatic. They may provide multiple conformational switches to control the activity of Hpk37 in response to certain stimuli. Bioinformatic analyses have demonstrated that HPKs with poly-HAMP domains are not uncommon and all of them are hybrid kinases with two or three C-terminal receiver domains (Stuart Huntley and Lotte Sogaard-Andersen, unpublished). Third, Hpk37 contains three receiver domains. Bioinformatics analyses show that this domain structure is not uncommon among HPKs (Sigrun Wegener-Feldbrügge and Lotte Sogaard-Andersen, submitted). It will be interesting to analyze phosphotransfer within the Hpk37 protein.

5 Material and methods

5.1 Reagents, Enzymes and Kits

The chemicals unless stated in the text were obtained from Roth (Karlsruhe), Sigma (Taufkirchen), Becton Dickerson (Heidelberg), Invitrogen (Karlsruhe) and Merck (Darmstadt). The DNA markers used in the study were purchased from Bioline (Luckenwalde). The prestain protein ladders were from Fermentas (St. Leon-Rot).

Restriction enzymes, the quick ligase and Antarctic phosphatase were obtained from Fermentas (St. Leon-Rot) or New England Biosciences (Frankfurt a. M.). *Pfu* ultra II[®] DNA-Polymerase from Stratagene, Platinum *Pfx* DNA-Polymerase or AccuPrime *Pfx* DNA Polymerase from Invitrogen (Karlsruhe) was used for cloning PCR reaction. Taq DNA polymerase from New England Biosciences (Frankfurt a. M.) or from Fermentas (St. Leon-Rot) was used for the checking PCR reaction in the first period of the PhD work and *Taq*-Polymerase in "Eppendorf[®] MasterMix" from Eppendorf (Hamburg) was used for the checking PCR reaction in the latter period of this PhD work.

QIAquick Gel Extraction Kit from Qiagen (Hilden) and Zymoclean Gel DNA Recovery Kit from Hiss (Freiburg) were used to purify the DNA fragment from the agarose gels. PCR products in the PCR reaction systems were purified with QIAquick PCR Purification Kit from Qiagen (Hilden). QIAprep Spin Miniprep Kit from Qiagen (Hilden) was used for the isolation of plasmids from their *E. coli* host cells. RNeasy Kit from Qiagen (Hilden) was used to purify the RNA from the DNase I digestion reactions. The cDNA Archive kit (ABI) was used to synthesize cDNA from RNA. The PCR product of sequencing reaction was purified with an ethanol precipitation procedure or BigDye XTerminator Purification Kit from Applied Biosystems.

5.2 Microbiological methods

5.2.1 *Escherichia coli* strains

The *E. coli* strains used in this study are listed as follows.

Table 10. *E. coli* strains used in this study

<i>E. coli</i> Strains	Genotype	Reference
TOP10	F ⁻ <i>mcrA</i> Δ(<i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i>) 80 <i>lacZ</i> ΔM15Δ <i>lacX74</i>	Invitrogen (Karlsruhe)

	<i>recA1 deoR araD139 Δ(ara-leu)7697 galU galK, rpsL strR endA1 nupG</i>	
Mach1	<i>ΔrecA1398 endA1 tonA Φ80ΔlacM15 ΔlacX74 hsdR(r_K m_K⁺)</i>	Invitrogen (Karlsruhe)
Rosetta 2(DE3)	F ⁻ <i>ompT hsdS_B(r_B-m_B) gal dcm(DE3) pRARE2 (Cm^R)</i>	Novagen/Merck (Darmstadt)
GJ1158	<i>ompT hsdS gal dcm ΔmalAp510 malP:::(proUp-T7 RNAP) malQ:::lacZhyb11 Δ(zhf-900::Tn10dTet)</i>	(Bhandari & Gowrishankar, 1997)

5.2.2 *Saccharomyces cerevisiae* strain

Table 11. *S. cerevisiae* strain used in this study

Strain	Genotype	Reference
AH109	<i>MA Ta trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80ΔLYS2:::GAL1_{UAS}-GAL1_{TATA}-HIS3 GAL2_{UAS}-GAL2_{TATA}-ADE2 URA3:::MEL1_{UAS}-MEL1_{TATA}-lacZ</i>	Clontech

5.2.3 *Myxococcus xanthus* stains

Table 12. *M. xanthus* strains used in this study

Strains	Genotype	Reference
DK1622	Wild-type	(Kaiser, 1979)
DK1217	<i>aglB1</i>	(Hodgkin & Kaiser, 1979)
DK1300	<i>sglG1</i>	(Hodgkin & Kaiser, 1979)
DK11063	<i>fruA: Ω7540 Tn5 lac</i>	(Sogaard-Andersen <i>et al.</i> , 1996)
MS1512	<i>ΔMXAN1014 (=Δsdek)</i>	(Pollack & Singer, 2001)
PH1017	<i>ΔMXAN6855 (=ΔespC)</i>	B. Lee and P. Higgs
SA2101	<i>MXAN0347:::pXS008 (=hpk1 short)</i>	This study
SA2102	<i>MXAN0347:::pXS001 (=hpk1 long)</i>	This study
SA2103	<i>MXAN5184:::pXS009 (=hpk2 short)</i>	This study
SA2104	<i>MXAN5184:::pXS002 (=hpk2 long)</i>	This study
SA2105	<i>MXAN3343:::pXS010 (=hpk3 short)</i>	This study
SA2106	<i>MXAN3343:::pXS003 (=hpk3 long)</i>	This study
SA2107	<i>MXAN0060:::pXS011(=hpk5 short)</i>	This study
SA2108	<i>MXAN0060:::pXS005(=hpk5 long)</i>	This study
SA2109	<i>MXAN6053:::pXS012 (=hpk6 short)</i>	This study
SA2110	<i>MXAN6053:::pXS006 (=hpk6 long)</i>	This study
SA2111	<i>MXAN2317:::pXS007 (=hpk7)</i>	This study
SA2112	<i>ΔMXAN3290 (=Δhpk8)</i>	This study
SA2113	<i>ΔMXAN3343 (=Δhpk3)</i>	This study
SA2114	<i>ΔMXAN2317 (=Δhpk7)</i>	This study
SA2115	<i>ΔMXAN6315 (=Δhpk9)</i>	This study

SA2116	Δ MXAN4432 (=Δhpk10)	This study
SA2117	Δ MXAN0340 (=Δhpk11)	This study
SA2118	Δ MXAN0736 (=Δhpk12)	This study
SA2119	Δ MXAN7123 (=Δhpk13)	This study
SA2120	Δ MXAN0928 (=Δhpk14)	This study
SA2121	Δ MXAN6994 (=Δhpk15)	This study
SA2122	Δ MXAN1553 (=Δhpk16)	This study
SA2123	Δ MXAN5715 (=Δhpk17)	This study
SA2124	Δ MXAN2763 (=Δhpk18)	This study
SA2125	Δ MXAN7444 (=Δhpk19)	This study
SA2126	Δ MXAN5483 (=Δhpk20)	This study
SA2127	Δ MXAN7002 (=Δhpk21)	This study
SA2128	Δ MXAN7003 (=Δhpk22)	This study
SA2129	Δ MXAN0230 (=Δhpk31)	This study
SA2130	Δ MXAN3098 (=Δhpk24)	This study
SA2131	Δ MXAN0571 (=Δhpk25)	This study
SA2132	Δ MXAN0720 (=Δhpk26)	This study
SA2133	Δ MXAN7398 (=Δhpk29)	This study
SA2134	Δ MXAN4465 (=Δhpk30)	This study
SA2135	Δ MXAN6865 (=Δhpk33)	This study
SA2136	Δ MXAN0176 (=Δhpk36)	This study
SA2137	Δ MXAN6941 (=Δhpk34)	This study
SA2138	Δ MXAN0712 (=Δhpk37)	This study
SA2139	Δ MXAN0931 (=ΔespA)	This study
SA2140	Δ MXAN5704 (=Δhpk32)	This study
SA2141	Δ MXAN6734 (=Δhpk38)	This study
SA2142	Δ MXAN0706 (=Δhpk39)	This study
SA2143	Δ MXAN7206 (=ΔmokA)	This study
SA2144	Δ MXAN6996 (=ΔasgD)	This study
SA2145	Δ MXAN3290 Δ MXAN1014 (=Δhpk8 ΔsdeK)	This study
SA2146	Δ MXAN0736, MXAN0060::pXS011(=Δhpk12 hpk5)	This study
SA2147	Δ MXAN3290, fruA: Ω7540 Tn5 lac (= Δhpk8 fruA)	This study
SA2148	Δ MXAN1014, fruA: Ω7540 Tn5 lac (=ΔsdeK fruA)	This study
SA2149	Δ MXAN0712, fruA: Ω7540 Tn5 lac (= Δhpk37 fruA)	This study
SA2150	Δ MXAN0736, fruA: Ω7540 Tn5 lac (= Δhpk12 fruA)	This study
SA2151	Δ MXAN0340, fruA: Ω7540 Tn5 lac (= Δhpk11 fruA)	This study
SA2152	Δ MXAN7123, fruA: Ω7540 Tn5 lac (= Δhpk13 fruA)	This study
SA2153	Δ MXAN0736 Δ MXAN1014 (= Δhpk12 ΔsdeK)	This study
SA2154	Δ MXAN3290 Δ MXAN0736 (= Δhpk8 Δhpk12)	This study
SA2155	Δ MXAN3290 Δ MXAN1014 Δ MXAN0736 (= Δhpk8 ΔsdeK Δhpk12)	This study
SA2156	Δ MXAN0340 Δ MXAN7123 (=Δhpk11 Δhpk13)	This study

5.2.4 Cultivation of *E. coli*

E. coli cells were grown in LB medium (1% Tryptone, 0.5% yeast extract and 1% NaCl, pH 7.0) or on LB agar plate (1.5% agar in LB medium) following standard protocol (Sambrook & Russell, 2000). Antibiotics were added with following concentrations for selective growth if necessary: ampicillin (100 µg/ml), kanamycin (50 µg/ml), chloramphenicol (50 µg/ml), carbenicillin (100 µg/ml) and X-Gal (20 µg/ml). Liquid culture of *E. coli* was incubated at 37°C with 230 rpm shaking and *E. coli* stains on the solid plates were incubated at 37°C under aerobic conditions. Glycerol stocks were made with overnight culture by adding the glycerol to the final concentration of 10% and stored at -80°C.

5.2.5 Cultivation of *S. cerevisiae*

S. cerevisiae was grown in YPD medium (1% yeast extract, 2 % Difco peptone and 2% glucose, pH 6.5) or YPD agar plate (2% agar in YPD medium) for normal growth and in SD drop out medium (0.67% Yeast nitrogen base without amino acids, 0.154% Sigma Dropout (-Leu-Trp) or 0.06% BD Dropout (-Leu-Trp-His-Ade), pH 5.8) or agar plate (2% agar in corresponding SD drop out medium) for selection growth. Liquid culture of *S. cerevisiae* was incubated at 30°C with 230 rpm shaking and the solid plates of *S. cerevisiae* were incubated at 30°C under aerobic conditions.

5.2.6 Cultivation of *M. xanthus*

M. xanthus cells were grown in 1% CTT medium (1% casitone, 10 mM Tris-HCl pH 8.0, 1 mM KH₂PO₄ pH 7.6, 8 mM MgSO₄) (Hodgkin & Kaiser, 1977) or on 1% CTT agar plates (1.5% agar in 1% CTT medium) supplied with 1/1000 volume of trace element (1.5% Titriplex III, 1 mM ZnCl₂, 1 mM CuSO₄, 1 mM CoCl₂, 1 mM Na₂MO₄, 1 mM MnSO₄, 10 mM FeSO₄) at 32°C in the dark. Kanamycin at the concentration of 40 µg/ml or oxytetracycline at the concentration of 10 µg/ml was used for selective growth on the CTT agar plates. Liquid culture of *M. xanthus* without antibiotics unless stated was incubated with shaking at 230 rpm in dark. The glycerol stocks were made with the *M. xanthus* culture (OD₅₅₀ between 0.8 and 1.2) by adding the glycerol to 4% and the mixtures were fast frozen in liquid nitrogen and stored at -80°C. *M. xanthus* strains on the plates were stored at 18°C for one month.

5.2.7 Development assay and spore assay of *M. xanthus*

M. xanthus development was examined on the following three different conditions: TPM agar plates (10 mM Tris-HCl, pH 7.6, 1 mM KH_2PO_4 , pH 7.6, 8 mM MgSO_4 , 1.5% agar) (Kuner & Kaiser, 1982), CF agar plates (10 mM Tris-HCl, pH 8.0, 1 mM KH_2PO_4 , pH 7.6, 8 mM MgSO_4 , 0.02% $(\text{NH}_4)_2\text{SO}_4$, 0.1% NaPyruvate, 0.2% NaCitrate, 1.5% agar) (Shimkets & Kaiser, 1982), and submerged in MC7 buffer (10 mM MOPS, pH 7.0, 1 mM CaCl_2). CF and TPM plates were freshly prepared one day before every assay. The stains for the development assay were cultivated in parallel to OD_{550} 0.5~0.9. The cells were harvested and resuspended in MC7 buffer to a calculated density of 5×10^9 cells/ml. 20 μl aliquots were spotted on CF agar and TPM agar. For development in submerged culture, 50 μl of concentrated cells were diluted with 350 μl MC7 and placed in a 15 mm well in a microtiter dish. Aggregation was followed using a Leica MZ8 stereomicroscope and a Leica IMB/E inverted microscope and visualized using a Leica DFC280 CCD camera.

Spore numbers were determined as the number of spores formed after 72 hrs and 120 hrs of starvation by harvesting 5×10^8 cells from each of the three different starvation conditions. Cells were placed for 2 hrs at 50°C and briefly sonicated to disperse fruiting bodies. Spores were counted in a haemocytometer (Depth 0.1 mm, Marienfeld). Three biological experiments were performed for each strain to determine the sporulation efficiency. To determine the number of germinating spores, spore solutions were diluted and plated on 1.0% CTT agar plates with CTT softagar (0.75% agar in 1% CTT medium).

The number of fruiting bodies formed was determined after 120 hrs of starvation by manually counting fruiting bodies formed in three 20 μl cell aliquots on CF agar in three biological experiments. The surface area covered by individual fruiting bodies was calculated by measuring the total area of fruiting bodies formed in three 20 μl cell aliquots on CF agar after 120 hrs of starvation in three biological experiments using the area measurement tool in Metamorph (Molecular Devices) followed by division with the total number of fruiting bodies formed.

5.2.8 Motility assay of *M. xanthus*

Motility was performed on two different agar plates: 0.5% CTT medium (0.5% casitone, 10 mM Tris-HCl, pH 8.0, 1 mM KH₂PO₄, pH 7.6, and 8 mM MgSO₄) with 0.5% agar (Invitrogen, Select agar) plates, on which cells favor S-motility, and 0.5% CTT medium with 1.5% agar (Invitrogen, Select agar), on which cells favor A-motility. The plates were freshly prepared one day before every assay. Cells were grown to OD₅₅₀ 0.5~0.9, harvested, and resuspended in CTT medium to a calculated density of 5.0 ×10⁹ cells/ml. 5 µl aliquots were placed on two different motility plates and incubated at 32°C. After 24 hrs, colony morphology and colony edges were observed visually in a Leica MZ8 stereomicroscope and a Leica IMB/E inverted microscope and visualized using a Leica DFC280 CCD camera.

5.3 Molecular biological methods

5.3.1 Oligonucleotide and plasmids

The oligonucleotide primers for PCR reactions were from Sigma-Aldrich (Steinheim) with reverse phase purification or from Invitrogen (Karlsruhe) with desalt purification. The primers for Quantitative real time PCR were designed with Primer Express 2.0 (ABI) and the remaining primers were designed manually.

Table 13. PCR primers used for the Quantitative real time PCR

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
<i>MXAN0060=hpk5</i>	CCAGGACCCAGACCAGACG	GCCAGAGGTCACCACTGGC
<i>MXAN0168</i>	TTCAGGACAACCTACCGCAGTGT	CGATGATGAAGAGCGCCC
<i>MXAN0176=hpk36</i>	CACGTTCTGACTCGATTGCG	CCCCCAGATGTCGAGGAAG
<i>MXAN0245</i>	GCCACCCGAGTACAGCCTCTT	CAACTTCGCGTAGCCCAACTG
<i>MXAN0706=hpk39</i>	CTGAAGCTGCAACGCAAACC	AACTCCGAGCACGCTTCGT
<i>MXAN0736=hpk12</i>	CACCATCCACGTCCAGAGC	TGCACTTCATTCACTGAACAGGA
<i>MXAN0928=hpk14</i>	GGCGTGGGACACGAAATCA	TGCTCGGACAGGTGCTCCT
<i>MXAN2317=hpk7</i>	AGCGGGCCTTCATTCTCGT	GAGTGCTCGAAAATCGCTGC
<i>MXAN3036=hpk23</i>	GAGAACGTGCGGACGGAG	TCGACCCACGTGTGAGC
<i>MXAN3098=hpk24</i>	CTCCGGCTGTTGATTGAGAGC	CGCTACGTATCCGTCAGGTC
<i>MXAN3290=hpk8</i>	GATCTGCGACCCGAACTCG	CCGAAGGGACGCTGATTGA
<i>MXAN3343=hpk3</i>	TTCGTAGAG GTG CGGGTGTC	CGGTGCCTTTCTGCTTCGTC
<i>MXAN3879</i>	GGGCTCAGCATGGTGTATGG	CGAACGCGGGAGGTAGATC
<i>MXAN4988=hpk27</i>	CGACCAGTCCCAGCTCTCAC	CTCAATGGATGTCCTCAACCG
<i>MXAN5704</i>	AGGCACAGAACGCCGAAC	GACCAGGAACGCGGTGAGCTG
<i>MXAN5715</i>	TACGTGAAGGAACGGGC ATC	CGTCGTGTAGTAGGGCTCGAAGA

<i>MXAN6117</i>	AGACGAACCTCGTCACCGG	GGAATTCTCGTGCCTCCAC
<i>MXAN6315=hpk9</i>	TCAGGCGCAGTTGGGCAT	CGAAAATGGAGCCCATGGAA
<i>MXAN6941=hpk34</i>	CTTCGTCTACCTCCCCCTCG	TCCCGTCGTACACCTCCAG
<i>MXAN6994=hpk15</i>	CAATGTATCGCCGTGGACG	GGAACGGGTCTGA ACAATGTCT
<i>MXAN7027</i>	ACTTCCTGCATGACAAACGGC	GCCGATCATCTCTGAGGGTG
<i>MXAN7123=hpk13</i>	ACTGTCCCCCTCAGTTGATGG	TGAAGGCATTGACCAGGAGGT
<i>MXAN7206=mokA</i>	TGCCTGTGACTCCAGCGTCTT	GTGCTCGACAGACACATCGG
<i>MXAN7398=hpk29</i>	GTCGGTGCAGCACTTCAAGG	ACGATGCGCTTGGTAATCCAC
<i>spi</i>	GGCTGTCTCCCGCTTTCTTC	TGGATGTCGATCTGATGGTTCT
<i>sdeK</i>	GGTCGTCTTGCAGAACGT	TGGCGGAGACACACTCGAA
<i>fruA</i>	ATCATCTCGCAGTGCTTCGA	CCTCGGACCAGGGAGTTGA
<i>devR</i>	AAACATCACCAGCCTCCAGAA	TGCATGGCTCCTGCTCATT
<i>exo</i>	ATGAACCTCTATCCGGACATCGT	AGCTCGAAGGCCGTCTCA

Table 14. PCR primers used for the in-frame deletion constructs

Gene	Primer A (5'-3') ^a	Primer B (5'-3') ^{a,b}	Primer C (5'-3') ^{a,b}	Primer D (5'-3') ^a	Primer E (5'-3')	Primer F (5'-3')	Primer G (5'-3')	Primer H (5'-3')
<i>MXAN01</i> 76= <i>hpk36</i>	<u>TCGGTA</u> CCTTCC CCACGG CAGCAC GTC	<u>ATCGGC</u> TGCAGG AATACC TCCAGC ACCTGC G	<u>ATCGGC</u> TGCAGC GCTTCA CCGTGC CGCTG	<u>ATCGGT</u> CTAGAA CTACGG TCTCCC CGCTC	ACACGA GCACCA GGCGTC C	CCTGTT CGGGAG CGTCAT GG	ACCTCG CGGTAG CTCCAC AC	CGCGCC ACGTTT GACTCG AT
<i>MXAN02</i> 30= <i>hpk31</i>	<u>ATCGGA</u> AGCTTC TGCGCA CCCTGC TGGTGT	<u>TGCGTC</u> <u>CCGGCG</u> <u>GCTGAG</u> GTGCTC GAAGA	<u>CTCAGC</u> <u>CGCCGG</u> <u>GACGCA</u> CTGTCC AAGG	<u>ATCGGT</u> CTAGAG CATGGC CAGGAA CTCGTC C	CCTCTG TCCTGG AAGGCT AC	GCTCCA GGATGC CAATGG CT	TCCACG CCGCGC TGGACA C	GTCCGT CACCCT CACCAC CA
<i>MXAN03</i> 40= <i>hpk11</i>	<u>ATCGGA</u> AGCTTC CATCAT CGCCAG GATGCC	<u>GAAGGT</u> <u>GGAATG</u> <u>GCTGTA</u> ACGGGG CGAGGA	<u>TACAGC</u> <u>CATTCC</u> <u>ACCTTC</u> ACCGTC GTGCTG	<u>ATCGGT</u> CTAGAC GAACGA AAGCAG GACGCC	CGCTGC CGCTGC ATCACC A	CTGGTG TTGGGG GTGGTG CT	CATGCA ACACGA TGGGTG GG	AGGAAC ACGCGG CAGGAA TC
<i>MXAN05</i> 71= <i>hpk25</i>	<u>TCGGTA</u> CCGGAT GCACCG GAGCGT ACG	<u>ATCGGC</u> TGCAGC GGCGTG ACGCTT GGACTC A	<u>ATCGGC</u> TGCAGC AGGAGG GAACCA CGTTCC A	<u>ATCGGT</u> CTAGAT CAGCCC CATGAA CAGGCT C	GAGAGC GGCGAG CGTGT C	ATCCCC TGCGCC TTGAGG C	GTGGGA CGCTCC CGTTTC GA	GATGAC GTCGCT GGTCGC CA
<i>MXAN07</i> 06= <i>hpk39</i>	<u>ATCGGC</u> TGCAGC TTCAAC GATGGC GTGCTC G	<u>GCGAGA</u> <u>CATGGT</u> <u>CGACCA</u> <u>GCGCGT</u> CATGTC	<u>GCTGGT</u> <u>CGACCA</u> <u>TGTCTC</u> <u>GCTGCC</u> CGTGT	<u>ATCGGG</u> <u>ATCCGA</u> CCCGT TACGCG CGGATG	CAACCG GCCAAT GCCACA GG	TGGGAG TGCTGG TGGCCT G	GCTTCA AGGCGT CATGGT GG	CCAGAT CATTGA CCAGCG CG
<i>MXAN07</i> 12= <i>hpk37</i>	<u>ATCGGC</u> TGCAGT ATGCCG TGGACG CGCTCA C	<u>CAGGCT</u> <u>GAGGAG</u> <u>GACCCG</u> CAGCAA CTGG	<u>CGGGTC</u> <u>CTCCTC</u> <u>AGCCTG</u> CTGCCG GTGTG	<u>ATCGGG</u> <u>ATCCTC</u> CAGGTA GCGCTG CGCGTC	ATGGAG ATGCAG GCGCGT G	GAAGGT CAGCAA CGTGTC CC	CATCGC GGAAGT CACCAC GG	TGCGCT CCACCA CCAGCT TG
<i>MXAN07</i> 20= <i>hpk26</i>	<u>TCGGTA</u> CCTGCT CCTCTT CGTGGA CCTG	<u>ATCGGG</u> <u>GATCCG</u> GGCCGC GAGGAT GAAGAA G	<u>ATCGGG</u> <u>GATCCC</u> CCCCGT CCACGG CACAA	<u>ATCGGT</u> CTAGAA CCTCGT GGGCTC ATCCAG	GCTGTG CTTCGT GGGCTT GG	CGGATG TCTCGA GCCGTC G	CCGAAG CGCTGA TGCGCG AG	ACTCGT CGCGCA GACGCA CC
<i>MXAN07</i>	<u>ATCGGA</u>	<u>AGGTCA</u>	<u>CAGCGA</u>	<u>ATCGGT</u>	CTCGTC	GAGGTG	CACCAA	AGTGGG

36=hpk12	<u>AGCTTC</u> ACGTAG CGCAGC ACGTAG C	<u>TACCAC</u> <u>TCGCTG</u> GGCGAA TGTCCA	<u>GTGGTA</u> <u>TGACCT</u> TGTGGG TGGGGC	<u>CTAGAG</u> <u>TTCTTC</u> ACCTCG AGCGCC A	CTCCTC GTCGTT CT	ACGAAG GTGCGG AT	TCCAGC GAGCAG CA	TTGCC AGGTCA TG
MXAN09 28=hpk14	<u>ATCGGA</u> <u>AGCTTG</u> AGGACG AAAGCA CCGTGC	<u>GACGCG</u> <u>GAACGA</u> <u>CATCAG</u> CATCAC CCTCGC	<u>CTGATG</u> <u>TCGTTC</u> <u>CGCGTC</u> ACGTTG CCCGT	<u>ATCGGT</u> <u>CTAGAG</u> CCATTG CCGGAA CAGACA G	CAACCT CTTCGA CGGCAC C	TCACAC TCCGTC GCACAC G	AGGCGG CCAACG TCAAGC TGT	TCGCGA TGCTCC GGGGAA AT
MXAN09 31 =espA	<u>ATCGAA</u> <u>GCTITG</u> CTCCGC CTCGTA CACGC	<u>CAGGAG</u> <u>CTCGAT</u> <u>GAGGAT</u> GAGGCA GGCTCC	<u>ATCCTC</u> <u>ATCGAG</u> <u>CTCGTG</u> GGCCGC TCC	<u>ATCGGT</u> <u>CTAGAG</u> CATCGC GGTAAT CAGCTC C	ATCGAA GCTTGC CCGAAG TCATAC ACGTCC AC	ATCGGT CTAGAT CCGTTT CAGAGA GGCCAC T	CGAGAG CATCTA CGACGG CG	ACGAGC ATGTAG GCCAGC GG
MXAN15 53=hpk16	<u>ATCGGA</u> <u>AGCTTG</u> CCGTCA CCGTCC AAAGCG A	<u>TCCATA</u> <u>GGCGAG</u> <u>CCGTGC</u> TTCGAC GTCTG	<u>GCACGG</u> <u>CTCGCC</u> <u>TATGGA</u> GGCAGG GGTCT	<u>ATCGGT</u> <u>CTAGAG</u> CGGCG GACTTC ATCCAC CA	GCATGT CCAGCA CCACCA CG	CGTGTC GTCCGG GACATG GT	ACGGAA AGGTGG GTCTGC CT	TGACGA CCTCGA AGAGCA GG
MXAN23 17=hpk7	<u>ATCGGA</u> <u>AGCTTG</u> AAGAAG GCGGCT GCGAAG A	<u>GACGCT</u> <u>GTCTGC</u> <u>CACGAG</u> AATGAA GGCCC	<u>CTCGTG</u> <u>GCAGAC</u> <u>AGCGTC</u> CTGGCC CTGT	<u>ATCGGT</u> <u>CTAGAA</u> CCGTGT GCGGGT CCACCA	ACGACG GCGAAG AAGACG AC	CGTGCC AATCAG GATGCG CT	<u>ATCGGA</u> <u>AGCTTC</u> GCATCG TGGAGG TCAACC G	<u>ATCGGT</u> <u>CTAGAA</u> CTGGCG CAGGTC GAAGGG
MXAN27 63=hpk18	<u>ATCGGA</u> <u>AGCTTC</u> GGCACC TGGGAC TGACTT T	<u>CTTCTC</u> <u>CAGGTC</u> <u>CCACAG</u> GACGAT GAGTC	<u>CTGTGG</u> <u>GACCTG</u> <u>GAGAAG</u> CCCTTC GAGCG	<u>ATCGGT</u> <u>CTAGAC</u> AGATTG ACCTTG CCGTCC G	CGTGGA CTTCGG CATCTG GA	AAGTCC CCATGA CCGCCG AA	TGCACG AGCTGT CAGCGG TT	ACGACA ATCTGG GTGAGG CG
MXAN30 36 (long)=hp k23	<u>ATCGGC</u> <u>TGCAGA</u> CCCCCT GACACA TGCCAC G	<u>GAAGAC</u> <u>GCTGCC</u> <u>CGTCAG</u> GTGGAT GAAG	<u>CTGACG</u> <u>GGCAGC</u> <u>GTCTTC</u> ACCCTC TGCCT	<u>ATCGGG</u> <u>ATCCAG</u> GGGTGG CCATGA TTGCC	CCACGG CGTGAA TGGCAT CC	TTGTA GCCACA CGAGGA GG		
MXAN30 36 (short)=h pk23	<u>ATCGGA</u> <u>AGCTTC</u> TGCTCA AGTGCA CCGTGA G	<u>ATCGGC</u> <u>TGCAGA</u> CCCTCC GTCCGC ACGTTT T	<u>ATCGGC</u> <u>TGCAGA</u> GCGTCT TCACCC TCTGCC T	<u>ATCGGT</u> <u>CTAGAG</u> ACGTGG TCCGGA TGGAGG T	CCACGG CGTGAA TGGCAT CC	CCTTTG ACGGAG CTGGCT CG		
MXAN30 98=hpk24	<u>ATCGGA</u> <u>AGCTTT</u> GCGAGA AGGCCG CGGTGA T	<u>ATCGGC</u> <u>TGCAGG</u> GATTCC CCTTGC CCGAA T	<u>ATCGGC</u> <u>TGCAGC</u> CCACGT AGACAC ACGCCA T	<u>ATCGGT</u> <u>CTAGAG</u> CATGGG GATGGA GGCGGA A	TACAGC TCGCCG CCATCC TC	TGAGCA GCTCGG AGGATG CG	GCCATC CTGACG CTCGAC CT	TCGTCC ACCAGC TCCACC AG
MXAN32 90=hpk8	<u>ATCGGA</u> <u>AGCTTA</u> CGGTGG TGCCCT TCCTGG A	<u>CAGCAG</u> <u>CTTCGG</u> <u>CTTGAG</u> GAGTGT GTGGAT	<u>CTCAAG</u> <u>CCGAAG</u> <u>CTGCTG</u> GACACC GTCTCT	<u>ATCGGT</u> <u>CTAGAA</u> GTTGAT GCCTTC CACCCG C	CGGAGA AGAAGG TGTTGT TGC	AACTGT TTGCC CAGTCG CC	TACGTG GC GCGC TACTTC CCA	TTGTGG ATGGCC TCGGCG AT
MXAN33 43 =hpk3	<u>ATCGGA</u> <u>AGCTTT</u> GTCGCC CTCGTC CTTCAG CA	<u>TGGCTT</u> <u>GCCGAT</u> <u>GATGGA</u> TGCCAG CAGCGC	<u>TCCATC</u> <u>ATCGGC</u> <u>AAGCCA</u> TCGCAT GTGGAC	<u>ATCGGT</u> <u>CTAGAG</u> CTCATC GCGGTA CGTCTC CA	GGGATA GTCCGG GACCAG C	TGCTCC AGCCGG AGCTTG C	<u>ATCGGA</u> <u>AGCTTC</u> GGCGAC TACCCG TGGACA TC	<u>TAGCCA</u> <u>AGCTTG</u> CGGCTC CAGCAG GATGAG G
MXAN44 32=hpk10	<u>ATCGGA</u> <u>AGCTTA</u> TGCCGC TGGGGT CGATGA TG	<u>CGCGTG</u> <u>TCACAT</u> <u>GCTTCC</u> ACTCCC TCGCC	<u>GGAAGC</u> <u>ATGTGA</u> <u>CACGCG</u> ATGGCA TCCCTT	<u>ATCGGT</u> <u>CTAGAT</u> ACGAGG GCTTCA CCTGGT C	GATGAC GTGGAT GGCGGA AG	CCACTC CATCTC CTACGG CA	TCAGTC AAGTCG TGGAGG ACC	AACGTG ACGTGT GCGCCA TG
MXAN44 65=hpk30	<u>TCGGTA</u> <u>CCAGGC</u> TGCCCA	<u>ATCGCT</u> <u>GCAGCG</u> GCACGG	<u>ATCGCT</u> <u>GCAGCT</u> GGAGTC	<u>ATCGTC</u> <u>TAGACT</u> TGTCCA	ATGGCG GCGAAG GACCTG	GTTTAC CGTGCC ATCGAG	TGTCCG TGTGCC AGCGCA	GCCAGC TCCAGG CCGGAC

	CGGTGA GG	GTGTTT CCATGT	CCTCCG TGC GG	CCCGCC ACAGC	G	CG	TC	AT
<i>MXAN49</i> 88 (short)=h pk27	<u>TCGGTA</u> CCAGCC TGATTT CGAGGA CCCG	<u>ATCGGC</u> TGCAGC ACGAGG CTGTAG GCGAAC A	<u>ATCGGC</u> TGCAGC CGCAGC CTGCCT CGATTT C	<u>ATCGTC</u> TAGACC GAGATC TGCTCC CAGACC	GGCGTT GAAGTA CCCGGA GC	GCGATG GCCGCC TTGAAG TC		
<i>MXAN49</i> 88 (long)=hp k27	<u>ATCGGC</u> TGCAGC GCTGAT CTGTCA CAGTAC GG	AGGCTG CGCGGA GTAGAC GAGGCT GCGTTC	GTCTAC TCCGCG CAGCCT GCCTCG ATTTC	<u>ATCGGG</u> <u>ATCCTG</u> GGGCTT CCTTCG GGGACG CCCATG	GGCGTT GAAGTA CCCGGA GC	GTGAGC GGCAGC AGAAGC G		
<i>MXAN54</i> 83=hpk20	<u>ATCGGA</u> <u>AGCTTG</u> AGTCCC GCCGCT GGAATG A	AGTGCC GCTCGG TGTGAG AACGGG CCAAG	CTCACA CCGAGC GGCACT ACGGCG GACT	<u>ATCGGT</u> <u>CTAGAT</u> GGGCTT CCACCA CTTCCC C	GCTGCC CGACAA CCTCTA CT	AGCTCA ACATCC GTGACG CC	TGCTGG ACCAGA TGCCGG AA	TCCAGG CGAATG GGATAG CG
<i>MXAN57</i> 04=hpk32	<u>ATCGGA</u> <u>AGCTTG</u> CGTCCT TCTCAA TCTGGA GC	<u>ATCGGC</u> TGCAGC CTGCTT CCTTCG TGGCTT G	<u>ATCGGC</u> <u>TGCAGT</u> CCTTCG TGCTGC GGCTGC	<u>ATCGGT</u> <u>CTAGAG</u> CTCCAA CCTCAA GCTGGC	TCACCG ACAGGG CCTTCT CG	AGGACA TCCTGC GCGACA CC	TGTGCG GACGGA CTCGCC GA	GCAGGC GAGCGT GGGTTG G
<i>MXAN57</i> 15=hpk17	<u>ATCGGA</u> <u>AGCTTC</u> CACCTT CTTCGC CTTCGG T	GTGCAC CGTACC CCTACA TGGTGA CCAGGT	TGTAGG GGTACC GTGCAC AGCCGC GTCAA	<u>ATCGGT</u> <u>CTAGAG</u> CTTCGA CATCGG CGGCAA G	CTGGAC CGTTTC GAGCAG CA	GCGGTC AGGTTT GGTCTG TC	GGCGG GTTATG ACGTGC TGA	ATTGGT GGTGCC CAGCGA CA
<i>MXAN63</i> 15=hpk9	<u>ATCGGA</u> <u>AGCTTA</u> CGGTGA CGATGG TTCCGA TG	GTCCAG CGGCAG CGGCAG GCACGA GAATG	CTGCCG CTGCCG CTGGAC CTGCTC AGCT	<u>ATCGGT</u> <u>CTAGAA</u> GTTGCT CGGCGA GGACGG	CCTGCT GCTTGG ACTTGT CG	CCTGCA CCTCAA CCTCGC C	GTCCCA TACCGG CCACGT CAT	CGCTTG GACTTG AGCACC GT
<i>MXAN67</i> 34=hpk38	<u>ATCGGC</u> TGCAGG GCTCAG CGCAAC GAGAAG G	GTGGAA GGTCTT TTCCTG GTTGCG GTCCG	CAGGAA AGGACC TTCCAC GTCTGG CTGCC	<u>ATCGGG</u> <u>ATCCAT</u> GAGCGC GACCAC GAGGAC	AGCAAC TGCTCA GCGACA CG	CTGGAC CCCGTG CATCAT C	CGGACG GCAGCT TCACCT AC	ACGTAC TGGAGC ACCACC TC
<i>MXAN68</i> 65=hpk33	<u>TCGGTA</u> <u>CCTGGC</u> GAAGCG GGTGAT GGC	<u>ATCGGC</u> TGCAGC AGGGCC TGGAAG AACTCG G	<u>ATCGGC</u> TGCAGA CCACCT TCATCG TGCGCC T	<u>ATCGGT</u> <u>CTAGAA</u> CCTCGC CCTTCA CGGAGA C	GTGCAC CAGCTC TATCCG CC	AGGAGG GTTATG GGGCTT GG	CGCAAG TACCTG GAGGAC C	ATGGCG ATGACG CCCTCC AC
<i>MXAN69</i> 41=hpk34	<u>TCGGTA</u> CCGCTT GACGTA GTCCTC CAGC	<u>ATCGGC</u> TGCAGC TCGTGG ATGCTC TGCTGC A	<u>ATCGGC</u> TGCAGG AGGCTT GACGCT CCCTGC	<u>ATCGGT</u> <u>CTAGAA</u> GCATGG TCCTCC GCATCT C	GTGGTG TCCAGG TGCGTG C	CGGAAA GATGGT CCCCAG CG	GACCTG GAGGTG TACGAC GG	GGCGAA GCTGGC GGTACA GC
<i>MXAN69</i> 94=hpk15	<u>ATCGGA</u> <u>AGCTTT</u> GCAGGA CATCGT GCGCG G	CAAGGG TGCCCA CTTCGA CGACAG ACACAC	TCGAAG TGGGCA CCCTTG GGCGTG AGACAT	<u>ATCGGT</u> <u>CTAGAG</u> CGCTGA ACCACG TGCTGA A	GCTGTC TGCTTT CAACGG CG	GCACCG GACGAT GTATCG C	CACGAT TGGGTC CTGCCT CCA	ATGATG GCCAGG CTGACC GT
<i>MXAN69</i> 96=asgD	<u>ATCGGC</u> TGCAGA CGCGCT TCTACC GCTACG AC	GTAGAG CCCCGG CACCAT CGTCGG AATCC	ATGGTG CCGGG GCTCTA CATCGT GCAGGA C	<u>ATCGGG</u> <u>ATCCTG</u> GAGGCA GGACCC AATCGT C	ACAACT GGTTCG GCTGGC TG	ATGCAC AGCAGC ACGGAC AC	TGCTGG ACGTGC GGATGG TG	CGTGAA GTCCTC GCGTAT CC
<i>MXAN70</i> 02=hpk21	<u>ATCGGA</u> <u>AGCTTG</u> ATGCCT CCGAGC CAGAAG	GATACG TCCGAT GGATTC CCGCTG CTGTAG	GAATCC ATCGGA CGTATC GAGGTG ACCTC	<u>ATCGGT</u> <u>CTAGAG</u> ATCGTG TCGCGG TGGACA	GGCGAT GGCTGA CGATGA TG	TACACA CCCCC AAGACA GG	GGGATA TTCCGG ACCGGC AT	CATGGC CCGGAA GAGGAA CT

	A			T				
<i>MXAN70</i> 03= <i>hpk22</i>	<u>ATCGGA</u> AGCTTC AGCAGC CGGTAC TCCGTC A	GCCGCC ATGGGG AAATCG GTTGAG CGCATG	CGATTT CCCCAT GGCGG CAGCAT CCACCT	<u>ATCGGT</u> <u>CTAGAG</u> CAGCGG TAATCT CGGCC	GTGGAA GACGAG CTCCGC CT	GGACAT GGAAGC GACCTA CG	CATCCC AGGCAT CGACAT CC	GTGACG ATGCTC GCCAGC AA
<i>MXAN71</i> 23= <i>hpk13</i>	<u>ATCGGA</u> AGCTTA GGACGA CTTCCC GAACGG C	GTAGGG CAAGAA GCCGTC ACTGGT GGGCT	GACGGC TTCTTG CCCTAC CGTCCG CAGTAG	<u>ATCGGT</u> <u>CTAGAG</u> GCAGCG TGTTGG TGACGT A	GTGCAA CAGCGT CGCCAG TT	TTCGTG CATCCG GGCTGG TT	CGAAAG GCCCG ACAGAA TA	TGCAGA CATCGA GGAAGG GC
<i>MXAN72</i> 06= <i>mokA</i>	<u>ATCGAA</u> <u>GCTTIG</u> CTCCAG CGCGCG TAACG	CACTTC GGACAT GTCGCC CGCAGT GTCCG	GGCGAC ATGTCC GAAGTG GCGCGT GCTCG	<u>ATCGTC</u> <u>TAGAAG</u> CACGTG CGCTTC TCCTTC	CACCTG TCGCCA GAAGGT CC	CACCGC CTTCCA CGAACT GC	ATCGAA TTCGCG GTGGGA CAGCTC GCGG	ATCGGG ATCCCT ACGCCG CGGCC TCAGCG TCAC
<i>MXAN73</i> 98= <i>hpk29</i>	<u>TCGGTA</u> <u>CCACAT</u> CACCCG CATCCA CGAG	<u>ATCGGC</u> <u>TGCAGC</u> GACATG CGGTCC ACCGCG	<u>ATCGGC</u> <u>TGCAGA</u> CCTTCA CCGTGG AGCTGC C	<u>ATCGGT</u> <u>CTAGAT</u> TCATCC GGTACG CCATCC C	CCAGGA CCTGGT CGTCAT CG	AGGGCA GTCCCT CCTCGT C	TGCGGG ACGTGG TGCCCT C	ATCACC CCCAGG GTGCC AG
<i>MXAN74</i> 44= <i>hpk19</i>	<u>ATCGGA</u> AGCTTA GCCTCC AGGAGC CGTTGG	TTCGAC GATCGA TGCCTC CCTTCG TGCCT	GAGGCA TCGATC GTCGAA CGCCAC GGCG	<u>ATCGGT</u> <u>CTAGAA</u> GCTGAG GGCCTC GTCGAT	TGAGCG AGGACA CGGTGC G	CTCGTC CAGCAC ATAGGC GT	GCTGGG TCAACG GTTGGT GA	GGAATG GTGCGC GTCAGG AA

^a Underlined sequences indicate restriction sites used for cloning. ^b: red sequences in columns B and C indicate the complementary parts of the two flanking PCR products used to generate the full-length in-frame deletion fragment.

Table 15. PCR primers used for insertion mutants

Gene	Forward primer (5'-3') ^a	Reverse primer (5'-3') ^a	Checking Forward primer (5'-3')	Checking Reverse primer (5'-3')
<i>MXAN0347</i> = <i>hpk1 short</i>	<u>ATCGGAAGCTT</u> GGGCCGGA CGTGGAGATG	<u>TAGCCAAGCTT</u> ATGGGAGCTT CGCCTCCAG	TGGGCGCGCT GTACCTCCT	CGCTCACGAA GGTGGCGAAG
<i>MXAN0347</i> = <i>hpk1 long</i>	<u>ATCGGAAGCTT</u> GCCACTCCGTC CGGTTGC	<u>ATCGGTCTAGA</u> GCTTGCCGTC GGTGGTAG	TCCGCCACCTG CTTGACGC	CCCATCGGGG TGGAGGATGA
<i>MXAN5184</i> = <i>hpk2 short</i>	<u>ATCGGAAGCTT</u> CCGCACAGGA GTCCCTGGAG	<u>TAGCCAAGCTT</u> CCAGCAGCAC CGCGAAAGTG	AGCTGGACGC GCGCATGGA	CATCCACCGG GACACCAGCA
<i>MXAN5184</i> = <i>hpk2 long</i>	<u>ATCGGAAGCTT</u> GCTCGCCTTCG CCCTGCT	<u>ATCGGTCTAGA</u> CGCCACTTCT GCCACGC	AGGAGAACGG CATGCGCTTGA	GCGGGTGCTT CGCTTCTTGC
<i>MXAN3343</i> = <i>hpk3 short</i>	<u>ATCGGAAGCTT</u> CGGCGACTAC CCGTGGACATC	<u>TAGCCAAGCTT</u> GCGGCTCCAG CAGGATGAGG	CTTCTTTTACA GCGTCTTCCCG	CCATCTCCTCC ACCTCACCT
<i>MXAN3343</i> = <i>hpk3 long</i>	<u>ATCGGAAGCTT</u> GGGCTGACCG TGGTGTC	<u>ATCGGTCTAGA</u> CCGCCACCTC CAGGATGATG	CTGGGGGTCA GCAAGGGGA	AGACCCGCGG CCATCTCAC
<i>MXAN0060</i> = <i>hpk5 short</i>	<u>ATCGGAAGCTT</u> CAATCACAAAGG CGCGCATCC	<u>TAGCCAAGCTT</u> TCCCGGGTTCC AGCTCTCCAC	TGGTGGCGCT GCTCTCCGT	CCGGTAGCCC ATGATGCGC
<i>MXAN0060</i> = <i>hpk5 long</i>	<u>ATCGGAAGCTT</u> GAGGCCATGC AGGCGCTCA	<u>ATCGGTCTAGA</u> GCCACCTTCT CCATCAACG	TCATCGCCACC TACGAAGCC	TCAGCAGCACA AACACCCCG
<i>MXAN6053</i> = <i>hpk6 short</i>	<u>ATCGGAAGCTT</u> GGCAGACCAC CAGCCTCATCG	<u>TAGCCAAGCTT</u> GCCGACAGCTT CAGTGCGTTG	CTGGATGACCA CCCTGCTGC	TCTTGGCCCGC AGCAGCATC

<i>MXAN6053</i> = <i>hpk6 long</i>	<u>ATCGGAAGCTT</u> CGGTACTCCTT GCGGTTCTG	<u>ATCGGTCTAGA</u> CGGTGATGCTC ACTGGACGG	TGGAGGGCGT GATGCGGTTG	TCGAATCACAC CGCCAGGGC
<i>MXAN2317</i> = <i>hpk7</i>	<u>ATCGGAAGCTT</u> CGCATCGTGG AGGTCAACCG	<u>ATCGGTCTAGA</u> ACTGGCGCAG GTCGAAGGG	TCAAGCGCACA TGCAGCGATT	GAGGTCACAC CCCAGGTCCA
<i>MXAN3036</i> = <i>hpk23</i>	<u>TCGGTACCCAA</u> GTGGCGGTGG TGCTGG	<u>ATCGTCTAGAT</u> CGCCGTCATCC GCCCCAC		

^a Underlined sequences indicate restriction sites used for cloning.

Table 16. PCR primers used for sequencing reaction

Primer name	Primer sequence (5'-3')
BDseq	TTTTCGTTTTAAACCTAAGAGTC
ADseq	AGATGGTGCACGATGCACAG
T7seq	TAATACGACTCACTATAGGGC
006	TGCGACATCATCATCGGAAGAGAGTAG
007	GGGATGTTTAATACCACTACAATGGATGATG
M13-forw	CGCCAGGGTTTTCCAGTCACGAC
M13-rev	TAGCTCACTCATTAGGCACCCAG
hpk37 seq1	GAGGTTGGACGCCATGCTGT
hpk37 seq2	GGGCACCGACGAAAGCTGG
hpk37 seq3	CCTGGGACGTCAGGTTGAC
hpk37 seq4	AGGTGGGGACGGAGGGCAAG
hpk37 seq5	GGAGACAAGGTGCTGCTGGC
hpk37 seq6	CCATCGCGGAGCCAGTCTCC
hpk9-seq1	GCGCGTTCTCCAGCAGGTTT
hpk9-seq2	CTCCAGGTGCGCGAGGAGTT
hpk20-seq1	GAGGTCAGCGCGTGTGAG
hpk20-seq2	CCCATTGCCTGGAGGACGC
espA-seq1	CGAGAGCATCTACGACGGCG
espA-seq2	ACGAGCATGTAGGCCACGGG
espC-seq1	CCATCGGCCAGCTCCAAGT
hpk23-seq1	TGGTCCGCCAGCGTCACGTC
hpk23-seq2	CATCTGGTGGCCGGAGTCCA
hpk25-seq1	GATGACGTCGCTGGTCGCCA
hpk25-seq2	CGGCAGCGTGATGAGTGTGG
hpk25-seq3	GTGGGACGCTCCGCGTTCCA
hpk27-seq1	TGGGCCACTGCGGACGTGAA
hpk27-seq2	TGACGCTGCTGTCGGGCAGC
hpk29-seq1	AGGAGCCGCAGGCGCTGTC
hpk29-seq2	TGCGGGACGTGGTGCCTC
hpk29-seq3	ATCACCCCAAGGGTGCCAG
hpk29-seq4	GACGCATCCATGCGGCGCTG
hpk30-seq1	GCCAGCTCCAGGCCGGACAT
hpk30-seq2	TGTCGGTGTGCCAGCGCATC

hpk34-seq1	GGCGAAGCTGGCGGTACAGC
hpk34-seq2	GACCTGGAGGTGTACGACGG
hpk36-seq1	ACCTCGCGGTAGCTCCACAC
hpk36-seq2	CGCGCCACGTTCTGACTCGAT
hpk5-seq1	GACAACGCCCGGCTCTGGAG

Table 17. PCR primers used for Y2H constructs

Primer name	TIGR_MXAN	Primer sequence(5'-3') ^{a, b}
asgDk-y2h-f	MXAN6996	<u>ATCGGAATTCTCGCAAGAGGAACTGCGCCGG</u>
asgDk-y2h-r	MXAN6996	<u>ATCGGGATCCTCAGCCCGCCGACGCGGCAG</u>
espAk-y2h-f	MXAN0931	<u>ATCGAATTCTCGGTGGGCACGCTGGGC</u>
espAk-y2h-r	MXAN0931	<u>ATCGGGATCCTCAGCGTGGCGGCACAGCACCAC</u>
espA-y2h-forw	MXAN0931	<u>ATCGGCATATGGTGAAACTTCGCACGGAGAG</u>
espA-y2h-rev	MXAN0931	<u>ATCGGGGATCCTCATGCCCCGCTCGGCGGA</u>
espCk-y2h-f	MXAN6855	<u>ATCGCATATGGCGGTGGGGACGCTCGCG</u>
espCk-y2h-r	MXAN6855	<u>ATCGGAATTCAGCGCGCGGGCGGCAACAGCAC</u>
espC-y2h-forw	MXAN6855	<u>ATCGGCATATGATGCCCTTCATGGATACCCG</u>
espC-y2h-forw2	MXAN6855	<u>ATCGGCATATGCGCGCCATCGGCCAGCTCCAA</u>
espC-y2h-rev2	MXAN6855	<u>ATCGGGAATTCAGTGCCGCACGCCGCCAA</u>
fruA-y2h-forw2	MXAN3117	<u>ATCGGGAATTCATGGCAACCAATCAAGCAGCG</u>
fruA-y2h-rev	MXAN3117	<u>ATCGGGGATCCCTAGAGGTCCGGCGGCGGC</u>
hpk11-y2h-forw	MXAN0340	<u>ATCGGCATATGGTGCGTCCGGAACGAACACT</u>
hpk11-y2h-rev	MXAN0340	<u>ATCGGGGATCCTCACCCGCGCGCCCGCGGCAGCA</u>
hpk12-y2h-forw	MXAN0340	<u>ATCGGCATATGATGATGCGCCGTTGGACATTCG</u>
hpk12-y2h-forw2	MXAN0736	<u>ATCGGCATATGCACCTGCAGCTCCTCATGGA</u>
hpk12-y2h-rev	MXAN0736	<u>ATCGGGGATCCTCATACCTGCACCTTCATTAC</u>
hpk13-y2h-forw	MXAN7123	<u>ATCGGCATATGGTGACACTCAGGGACGCGC</u>
hpk13-y2h-rev	MXAN7123	<u>ATCGGGGATCCCTACTGCGGACGGTAGGGCA</u>
hpk14-y2h-forw	MXAN0928	<u>ATCGGCATATGGTGATGACGCTGCGAGCGAG</u>
hpk14-y2h-rev	MXAN0928	<u>ATCGGGGATCCTCAGGCCACCGCTCCGCTT</u>
hpk15-y2h-forw	MXAN6994	<u>ATCGGCATATGATGGGGGTGTGTCTGTCTGTC</u>
hpk15-y2h-rev	MXAN6994	<u>ATCGGGGATCCTCACACGCGATGTCTCACGC</u>
hpk18-y2h-forw	MXAN2763	<u>ATCGGCATATGGGACTCATCGTCTGTGGGA</u>
hpk18-y2h-rev	MXAN2763	<u>ATCGGGGATCCTCATTGGATGAGGTGCTCCAC</u>
hpk20-y2h-forw	MXAN5483	<u>ATCGGCATATGGTGGGCGCGGCGGAAATGTT</u>
hpk20-y2h-rev	MXAN5483	<u>ATCGGGGATCCTCACGCAGGGTGGCGCGGCA</u>
hpk23k-y2h-f	MXAN3036	<u>ATCGAATTCGCGGTGGGGCGGATGACG</u>
hpk23k-y2h-r	MXAN3036	<u>ATCGGGATCCCTAGCAACCCGCGGGAGGCAG</u>
hpk23-y2h-forw	MXAN3036	<u>ATCGGCATATGATGGCCGGGTGAGTCACGCA</u>
hpk23-y2h-forw2	MXAN3036	<u>ATCGGCATATGGGCCGCATCGAAACGCACCTT</u>
hpk23-y2h-rev	MXAN3036	<u>ATCGGGGATCCCTAGAGGGCGACCTCGGCG</u>
hpk24-y2h-forw	MXAN3098	<u>ATCGGCATATGATGGAAGACACCAACGGCTGC</u>

hpk24-y2h-rev	MXAN3098	<u>ATCGGGGATCCCTACGTGGGGCCGAGCGGGA</u>
hpk25k-y2h-f	<i>MXAN0571</i>	<u>ATCGAATTCTTCGAGCAGCTTCTCATCGG</u>
hpk25k-y2h-r	<i>MXAN0571</i>	<u>ATCGGGGATCC</u> TA GGGCTGGCGGGGCAGGCG
hpk25-y2h-forw	MXAN0571	<u>ATCGGCATATGATGAGTCCAAGCGTCAGCC</u>
hpk25-y2h-rev	MXAN0571	<u>ATCGGGGATCCTCACGGAGCGTCGACGTC</u>
hpk27k-y2h-f	<i>MXAN4988</i>	<u>ATCGAATTGAGTTGGGACAATTCACGTCC</u>
hpk27k-y2h-r	<i>MXAN4988</i>	<u>ATCGGGGATCC</u> TA CTGCGCCACGGGTAATTC
hpk27-y2h-forw	MXAN4988	<u>ATCGGCATATGATGGGGCCCTTGTTCGCCTAC</u>
hpk27-y2h-forw2	MXAN4988	<u>ATCGGCATATGTCGGGCAGCGGGCAGCCGTT</u>
hpk27-y2h-rev	MXAN4988	<u>ATCGGGGATCCTCAGGCCAGCGCGGCGAAAT</u>
hpk29-y2h-forw	MXAN7398	<u>ATCGGCATATGATGCGTCCGTGCCGGGGGA</u>
hpk29-y2h-rev	MXAN7398	<u>ATCGGGGATCCTCAAGCCGGCAAGGGCAGCT</u>
hpk30-y2h-forw	MXAN4465	<u>ATCGGCATATGTTGCGTACAAGCGGTTGCCTC</u>
hpk30-y2h-rev	MXAN4465	<u>ATCGGGGATCCTCATGGGGTCTCTCGTTT</u>
hpk34k-y2h-f	MXAN6941	<u>ATCGGAATTCGTGCGCGACGACTTCCTCAG</u>
hpk34k-y2h-r	MXAN6941	<u>ATCGGGGATCC</u> TA GCCGCTCAAGGGCAGCGC
hpk34-y2h-forw	MXAN6941	<u>ATCGGCATATGGTGGTCCTGGTGGGGCTGCT</u>
hpk34-y2h-forw2	MXAN6941	<u>ATCGGCATATGTTCAACGCGCGCCAGGCCTT</u>
hpk34-y2h-rev	MXAN6941	<u>ATCGGGGATCCTCAAGCCTCGGCACGTGGGC</u>
hpk36-y2h-forw	MXAN0176	<u>ATCGGCATATGGTGTCTGGAGGTATTCGTGTT</u>
hpk36-y2h-rev	MXAN0176	<u>ATCGGGGATCCTCACGTGGACGCCAGCGGCA</u>
hpk37k-y2h-f	<i>MXAN0712</i>	<u>ATCGGAATTC</u> AAGCTCGCCCTGGAGGAGAAG
hpk37k-y2h-r	<i>MXAN0712</i>	<u>ATCGGGGATCC</u> TA GGTGGCCAGCGGGTTGTAGTC
hpk39-y2h-f2	<i>MXAN0706</i>	<u>ATCGCATATGGTCTGGAGCTTCGCGACG</u>
hpk39-y2h-r2	<i>MXAN0706</i>	<u>ATCGGGGATCCTCAGGGCCGGGCCTCCCATG</u>
hpk5-y2h-f	<i>MXAN0060</i>	<u>ATCGGAATTCGCCGAGCGCTTTCAGCTCCTC</u>
hpk5-y2h-r	<i>MXAN0060</i>	<u>ATCGGGATCCTCACGAGTGCAGGGGAATATC</u>
Hpk8-y2h-forw	MXAN3290	<u>ATCGGCATATGGTGAAGCGGATGTTGCTCAC</u>
Hpk8-y2h-rev	MXAN3290	<u>ATCGGGGATCCTCACGGCACGGCGGGGGCT</u>
hpk9-y2h-forw	<i>MXAN6315</i>	<u>ATCGGCATATGATGCCGCTTCCCTCCGACGT</u>
hpk9-y2h-rev	<i>MXAN6315</i>	<u>ATCGGGGATCCTCACGTCCGGGAGCTGAGCA</u>
mokAk-y2h-f	<i>MXAN7206</i>	<u>ATCGAATTCGCGGTGGGACAGCTCGCGG</u>
mokAk-y2h-r	<i>MXAN7206</i>	<u>ATCGGGGATCC</u> TA CGCCGCGGCCCTCAGCGTCAC
mokA-y2h-forw3	MXAN7206	<u>ATCGGCATATGCGCGTGGGGCAGTTGAAGGA</u>
SdeK-y2h-forw	<i>MXAN1014</i>	<u>ATCGGCATATGGTGGCTCACAAACGGCAGCAT</u>
SdeK-y2h-rev	<i>MXAN1014</i>	<u>ATCGGGGATCCTCAGGCAGTCTTGCTGGCG</u>

^a Underlined sequences indicate restriction sites used for cloning. ^b: red sequences mean the added stop code.

Table 18. PCR primers used for over-express protein constructs

Name	Primer sequence (5'-3') ^{a, b, c}
FruA-strep-forw	<u>ATCGCAT</u> ATG GCAACCAATCAAGCAGCG

FruA-strep-rev	<u>ATCGCTCGAG</u> ^a <u>CTATTTTCGAACTGCGGGTGGCTCCAAGCGCT</u> ^c GAGGTCCGGCGGCGGCCGGAC
hpk8KD-forw	<u>ATCGGAATTCAACCGCGAGGAGCTGCTGGG</u>
hpk8KDA-forw	<u>ATCGGAATTCAACCGCGAGGAGCTGCTGGGCAACGTCTCG</u> ^b <u>GCCGACCTGAAGAAC</u> CCGCTC
hpk8KD-rev	<u>ATCGCTCGAG</u> ^a <u>T</u> ^b AGGCGAACATGGGGAGCACC
sdeK KD-forw	<u>ATCGGAATTC</u> ^c TTCCAGGAGCGCTTCATCGG
sdeK KD-rev	<u>ATCGCTCGAG</u> ^a <u>T</u> ^b AGGTGACGCGAGGCAGCGTC
hpk9KD-forw	<u>ATCGGAATTC</u> ^c GTGCGCGAGGAGTTCATCTCC
hpk9KD-rev	<u>ATCGCTCGAG</u> ^a <u>T</u> ^b ACGCCATGCGAGGCAGGCTC
hpk11KD-forw	<u>ATCGGGATCC</u> ^c GCGCGGAGGAGCTTTGG
hpk11KD-rev	<u>ATCGAAGCTT</u> ^c <u>T</u> ^b AGCGCGCCCGCGGCAGCACGAC
hpk13KD-forw	<u>ATCGGGATCC</u> ^c GCCGGATGAATTCCTGGC
hpk13KD-rev	<u>ATCGAAGCTT</u> ^c <u>T</u> ^b ACGGACGGTAGGGCAACTCC
hpk12KD-forw	<u>ATCGGGATCC</u> ^c GACGCGCTGGAGCTCATCGG
hpk12KD-rev	<u>ATCGAAGCTT</u> ^c <u>T</u> ^b ATGTCCTGCGGGGAAGTTC
hpk29KD-forw	<u>ATCGGAATTC</u> ^c GCGCGGGACGAGCTGCTGTC
hpk29KD-rev	<u>ATCGCTCGAG</u> ^a <u>T</u> ^b CAAGCCGGCAAGGGCAGC
hpk37bforw	<u>ATCGCATATGGGCTT</u> ^c CTGGAGCAGCTC
hpk37brev	<u>ATCGGGATCC</u> ^c TATCCGGAGACCTCCGTGC

^a Underlined black sequences indicate restriction sites used for cloning. ^b: red sequences mean the added stop code or point mutations. ^c: blue underlined sequences indicate the strep-tag.

Table 19. List of plasmids used in this study

Plasmids	Genotype	Reference
pBGS18	vector for insertion mutants, Km ^R	(Spratt <i>et al.</i> , 1986)
pBJ114	vector for in-frame deletion mutants, <i>galK</i> Km ^R	(Julien <i>et al.</i> , 2000)
pET24b (+)	expression vector, T7 promoter, His ₆ -Tag (C-terminal), Km ^R	Novagen
pET28a	expression vector, T7 promoter, N-terminal His ₆ -Tag/thrombin /T7-tag, C-terminal His ₆ -Tag, Km ^R	Novagen
pET32a	expression vector, T7 promoter, N-terminal Trx-tag/ His ₆ -Tag/ thrombin /S-tag/ enterokinase, C-terminal His ₆ -Tag, Amp ^R	Novagen
pGADT7	LEU2, Amp ^R , vector in Y2H for expression of protein fused to N- terminal GAL4 AD polypeptide with SV40 nuclear localization signal	Clontech (Heidelberg)
pGADT7-RecT	LEU2, Amp ^R , control plasmid Foz Y2H	Clontech (Heidelberg)
pGBKT7	TRP1, Kna ^f , vector in Y2H for expression of protein fused to N- terminal GAL4 DNA-BD	Clontech (Heidelberg)
pGBKT7-53	TRP1, Kna ^f , positive control of Y2H	Clontech (Heidelberg)
pGBKT7-lam	TRP1, Kna ^f , negative control of Y2H	Clontech (Heidelberg)
pXS001	pBGS18-1011 bp fragment of <i>hpk1</i>	This study
pXS002	pBGS18-988 bp fragment of <i>hpk2</i>	This study
pXS003	pBGS18-1012 bp fragment of <i>hpk3</i>	This study
pXS004	pBJ114- <i>hpk3</i> in frame delete fragment	This study
pXS005	pBGS18-1020 bp fragment of <i>hpk5</i>	This study

pXS006	pBGS18-973 bp fragment of hpk6	This study
pXS007	pBGS18-917 bp fragment of hpk7	This study
pXS008	pBGS18-685 bp fragment of hpk1	This study
pXS009	pBGS18-651 bp fragment of hpk2	This study
pXS010	pBGS18-626 bp fragment of hpk3	This study
pXS011	pBGS18- 639 bp fragment of hpk5	This study
pXS012	pBGS18-679 bp fragment of hpk6	This study
pXS013	pBJ114-hpk8 in-frame delete	This study
pXS014	pBJ114-hpk12 in-frame delete	This study
pXS015	pBJ114-hpk13 in-frame delete	This study
pXS016	pBJ114-hpk14 in-frame delete	This study
pXS017	pBJ114-hpk15 in-frame delete	This study
pXS018	pBJ114-hpk16 in-frame delete	This study
PXS020	pBJ114-hpk9 in-frame delete	This study
PXS022	pBJ114-hpk11 in-frame delete	This study
PXS023	pBJ114-hpk7 in-frame delete	This study
pXS024	pBJ114-hpk18 in-frame delete	This study
pXS025	pBJ114-hpk19 in-frame delete	This study
pXS026	pBJ114-hpk20 in-frame delete	This study
pXS027	pBJ114-hpk21 in-frame delete	This study
pXS028	pBJ114-hpk22 in-frame delete	This study
pXS029	pBJ114-hpk24 in-frame delete (with PstI site in the middle)	This study
pXS030	pBJ114-hpk32 in-frame delete (with PstI site in the middle)	This study
pXS031	pBJ114-hpk23 in-frame delete (with PstI site in the middle)	This study
pXS032	pBJ114-hpk29 in-frame delete (with PstI site in the middle)	This study
pXS033	pBJ114-hpk36 in-frame delete (with PstI site in the middle)	This study
pXS034	pBJ114-hpk26 in-frame delete (with BamHI site in the middle)	This study
pXS035	pBJ114-hpk33 in-frame delete (with PstI site in the middle)	This study
pXS036	pBJ114-hpk25 in-frame delete (with PstI site in the middle)	This study
pXS037	pBJ114-hpk30 in-frame delete (with PstI site in the middle)	This study
pXS038	pBJ114-hpk27 in-frame delete (with PstI site in the middle)	This study
pXS039	pBJ114-hpk34 in frame delete (with PstI site in the middle)	This study
pXS040	pBJ114-hpk31 in-frame delete	This study
pXS041	pBJ114-espA in-frame delete	This study
pXS042	pBJ114-hpk37 in-frame delete	This study
pXS043	pBJ114-hpk38 in-frame delete	This study
pXS044	pGBS18-hpk23 insertion fragment	This study

pXS045	pBJ114-hpk27 in-frame delete (long fragment construct)	This study
pXS046	pBJ114-hpk23 in-frame delete (long fragment construct)	This study
pXS047	pBJ114-hpk39 in-frame delete	This study
pXS048	pBJ114-asgD in-frame delete	This study
pXS049	pBJ114-mokA in-frame delete	This study
pXS050	pPGBKT7-hpk8 (start-stop)	This study
pXS051	pGBKT7-hpk9 (start-stop)	This study
pXS053	pGBKT7-hpk13 (start-stop)	This study
pXS054	pGBKT7-hpk14 (start-stop)	This study
pXS055	pGBKT7-hpk15 (start-stop)	This study
pXS057	pGBKT7-hpk20 (start-stop)	This study
pXS058	pGBKT7-sdeK (start-stop)	This study
pXS059	pGBKT7-30 (start-stop)	This study
pXS060	pGBKT7-espA (start-stop)	This study
pXS061	pGBKT7-hpk24 (start-stop)	This study
pXS062	pGBKT7-hpk11 (start-stop)	This study
pXS064	pGBKT7-hpk23 (start-stop)	This study
pXS065	pGBKT7-hpk24 (start-stop)	This study
pXS066	pPGBKT7-hpk25 (start-stop)	This study
pXS068	pGBKT7-hpk27 (start-stop)	This study
pXS069	pGBKT7-hpk29 (start-stop)	This study
pXS072	pGBKT7-hpk34 (start-stop)	This study
pXS073	pGBKT7-hpk36 (start-stop)	This study
pXS074	pGADT7- hpk8(start-stop)	This study
pXS075	pGADT7- hpk9(start-stop)	This study
pXS077	pGADT7- hpk11(start-stop)	This study
pXS078	pGADT7- hpk15(start-stop)	This study
pXS079	pPGADT7- hpk17(start-stop)	This study
pXS080	pGADT7- hpk20(start-stop)	This study
pXS081	pGADT7- sdeK(start-stop)	This study
pXS082	pGADT7- espA(start-stop)	This study
pXS083	pGADT7- hpk30(start-stop)	This study
pXS084	pGADT7- hpk13(start-stop)	This study
pXS085	pGADT7- hpk14(start-stop)	This study
pXS086	pGADT7- hpk24(start-stop)	This study
pXS088	pGADT7-hpk23 (start-stop)	This study
pXS089	pGBAD7-hpk24 (start-stop)	This study
pXS090	pGBAD7-hpk25 (start-stop)	This study
pXS092	pGADT7-hpk27 (start-stop)	This study
pXS093	pGADT7-hpk29 (start-stop)	This study
pXS096	pGADT7-hpk34 (start-stop)	This study
pXS097	pGADT7-hpk36 (start-stop)	This study

pXS098	pGBKT7-hpk12 cytoplasmic part	This study
pXS099	pGBKT7-hpk18 cytoplasmic part	This study
pXS100	pGBKT7-espC cytoplasmic part	This study
pXS101	pGADT7-hpk12 cytoplasmic part	This study
pXS102	pGADT7-hpk18 cytoplasmic part	This study
pXS103	pGADT7-espC cytoplasmic part	This study
pXS105	pGADT7-hpk23 cytoplasmic part	This study
pXS106	pGADT7-hpk27 cytoplasmic part	This study
pXS108	pGADT7- hpk34 cytoplasmic part	This study
pXS109	pGADT7- mokA cytoplasmic part	This study
pXS110	pGBKT7-fruA (start-stop) (in EcoRI and BamHI sites)	This study
pXS111	pGADT7-fruA (start-stop) (in EcoRI and BamHI sites)	This study
pXS112	pGBKT7- mokA (cytoplasmic part)	This study
pXS113	pGADT7-hpk10 kinase domain with stop code (in EcoRI and BamH sites)	This study
pXS114	pGADT7-hpk23kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS115	pGADT7-hpk25kinase domain with stop code (in EcoRI and BamH sites I)	This study
pXS117	pGADT7-hpk27kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS118	pGADT7-hpk34 kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS119	pGADT7-espA kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS120	pGBKT7-hpk10 kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS121	pGBKT7-hpk23 kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS122	pGBKT7-hpk25 kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS124	pGBKT7-hpk27 kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS125	pGBKT7-hpk34 kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS126	pGBKT7-espA kinase domain with stop code (in EcoRI and BamHI sites)	This study
pXS131	pGBKT7-hpk37KD (in BamHI and EcoRI sites)	This study
pXS132	pGADT7- hpk37KD (in BamHI and EcoRI sites)	This study
pXS133	pGBKT7- asgD KD (in BamHI and EcoRI sites)	This study
pXS134	pGADT7- asgD KD (in BamHI and EcoRI sites)	This study
pXS135	pGBKT7- mokA KD	This study
pXS136	pGADT7- mokA KD	This study
pXS137	pGBKT7- espC KD	This study
pXS138	pGADT7- espC KD	This study
pXS139	pGBKT7- hpk5 cytoplasmic part	This study
pXS140	pGADT7- hpk5 cytoplasmic part	This study
pXS141	pET24b-FruA with C terminal strep tag (in NdeI and XhoI sites)	This study

pXS142	pET32a-SdeK kinase domain with stop code (in EcoRI and XhoI sites)	This study
pXS143	pET32a-Hpk8 kinase domain with stop code (in EcoRI and XhoI sites)	This study
pXS144	pET32a-Hpk9 kinase domain with stop code (in EcoRI and XhoI sites)	This study
pXS145	pET32a-Hpk11 kinase domain with stop code (in BamHI and HindIII sites)	This study
pXS146	pET32a-Hpk12 kinase domain with stop code (in BamHI and HindIII sites)	This study
pXS147	pET32a-Hpk13 kinase domain with stop code (in BamHI and HindIII sites)	This study
pXS148	pET28a-Hpk37 last part (in NdeI and XhoI sites)	This study
pXS149	pET28a-SdeK full length	This study
pXS150	pET28a-Hpk8 full length	This study
pXS151	pET24b-FruA D59N	This study
pXS152	pET32a-Hpk8 KD H172A (in EcoRI and XhoI sites)	This study

5.3.2 Preparation of chromosomal DNA from *M. xanthus*

Briefly, chromosomal DNA from *M. xanthus* was prepared as follows. 8 ml of *M. xanthus* culture was grown to OD₅₅₀ above 1.0 and harvested (4,700 rpm, 10 min, 4°C; Heraeus Multifuge 1 S-R). The cell pellet was resuspended in 400 µl TE (10 mM Tris pH 8.0, 1 mM EDTA). Then 80 µl preincubated pronase (2.5 mg/ml) and 80 µl rapid lysis mix (5% SDS, 0.125 M EDTA, 0.5 M Tris-HCl, pH 9.4) were added to the reaction and incubated at 37°C for 1hr. The suspension was aspirated with 1ml syringe (Guage 22) for 10 times and the proteins were sequentially extracted once by equal volume phenol extraction, once by equal volume phenol-chloroform (3:1) extraction and twice by equal volume of chloroform extractions. DNA from the aqueous phase was precipitated by 2 volume of 96% ethanol, washed in 70% ethanol and then dissolved in 200 µl TE at 4°C, overnight. The chromosomal concentration was determined by Nano-Drop ND-10000 Spectrophotometer and the stock was stored at -20°C till use.

5.3.3 PCR reaction , digestion and ligation

The cloning PCR reactions were conducted in one of the following systems: 1). 1x Pfx amplification buffer, 0.4 mM dNTPs (each), 1 mM MgSO₄, 0.4 µM primers (each), genomic DNA 100 ng, Pfx platinum polymerase 2 units and H₂O to the final volume of 50 µl (In case for the PCR reaction which is hard to get the product, up to 1.5x enhancer or 10% DMSO was added to the PCR system). 2) 1x Pfu Ultra II reaction buffer, 0.3 mM dNTPs (each), genomic DNA 100 ng, 0.3

μM primers (each), Pfu Ultra II fusion HS DNA polymerase 1 unit and H_2O to the final volume of 50 μl (In case of the PCR product is difficult to get, 5-10% of DMSO was added to the reaction).

The PCR reaction for checking is performed as follows: 1). 2.5 μl 10x ThermoPol Reaction buffer, 0.4 mM dNTPs (each), 0.4 μM primers (each), 10% DMSO, plasmid or chromosomal DNA 10-100 ng or *M. xanthus* lysate 2 μl , Taq DNA polymerase 2 units and H_2O to the final volume of 25 μl ; 2). 10 μl 2.5x Eppendorf mix, 0.4 μM primers (each), 10% DMSO, plasmid or chromosomal DNA 10-100 ng or *M. xanthus* lysate 2 μl and H_2O to the final volume of 25 μl . The PCR cycles were carried out following the recommended protocols.

The restriction digestion reactions were performed following the catalogs from Fermentas (St. Leon-Rot) or New England Biosciences (Frankfurt a. M.) and the ligation reaction was conducted with recommended protocol of Quick ligase kit.

5.3.4 Transformation of *E. coli*

Appropriate volume of *E. coli* strain (1:100 diluted from overnight culture) was grown up to OD_{450} 0.5~0.7. Cells were harvested (4,700 rpm, 10 min, 4°C, Heraeus Multifuge 1 S-R), washed with half volume of ice cold 50 mM CaCl_2 and resuspended in 1/10 volume of ice cold 50 mM CaCl_2 . 200 μl of competent cells was mixed 100 ng plasmid or 10 μl of ligation product at 4°C for 30 min, and then transferred to 42°C water bath for 2 min, put back on the ice for 5 min and finally spread on corresponding antibiotic resistant plates.

5.3.5 Electroporation of *M. xanthus*

The *M. xanthus* strain for electroporation was grown in the CTT medium to OD_{550} 0.5~0.9. The cells were harvested (4,700 rpm, 10 min, RT, Heraeus Multifuge 1 S-R), washed twice with equal volume of H_2O and once with 1/2 volume of H_2O , and resuspended in 1/100 volume of filtered H_2O . 50 μl competent cells were mixed with 1 μg plasmid (for homologous recombination) or 5 μg chromosomal DNA (for chromosomal transformation) in 0.1 cm ice cold cuvettes. The electroporation was conducted with following pulse parameters: 25 μF of capacitance, 0.65 KV of voltage and 400 Ω of resistance. The time constant should range from 8.5 to 10.5 millisecond. The electroporated cells were grown in CTT medium for 1 to 3 generations and cells were spread on

corresponding selective plates with CTT softagar. The colonies grew up after 4-7 days.

5.3.6 Co-transformation of *S. cerevisiae*

A calculated volume of the yeast culture was diluted from overnight culture and grown up to OD₆₀₀ 0.6~0.8. The cells were harvested (RT, 5 min, 1.000 g), wash with 1/2 volume of H₂O and resuspended in 1/100 volume fresh made 100 mM lithium acetate. 25 µl aliquoted competent yeast cells were mixed with 10 µl herring sperm DNA (10 mg/ml) (boiled at 95°C for 20 min, chill on the ice until use) and 100 ng plasmids (each). Then 300 µl PEG-Solution (40% PEG, 10 mM Tris pH 8.0, 1 mM EDTA, 100 mM lithium acetate) was added to the reaction with vortexing carefully. The reactions were incubated at 30°C for 30 min in waterbath, 1/10 volume DMSO was added to the reaction and the mixture was further incubated in 42°C waterbath for 20-30 min. The cells after transformation were spun down and plated on corresponding Drop-out agar.

5.3.7 RNA preparation from *M. xanthus*

Total RNA was isolated from cell pellets using the hot-phenol method (Overgaard *et al.*, 2006). Briefly, approximately 5×10^9 *M. xanthus* cells (2-4 times more cells from later time points of development) were harvested to a tube containing 1/10 volume of ice-cold ethanol/phenol stop solution (5% saturated acid phenol (pH <6.0) in 96% ethanol) and spun down (4,700 rpm, 10 min, 4°C, Heraeus Multifuge 1 S-R). The pellet was resuspended in 600 µl ice cold solution 1 (0.3 M sucrose, 0.01 M NaAc, pH 4.5) and transferred 300 µl into each of two 1.5 ml tubes containing 300 µl hot (65°C) solution 2 (2% SDS, 0.01 M NaAc, pH 4.5). The cell lysis was conducted twice with equal volume hot phenol (saturated acid phenol (pH <6.0) at 65°C) extraction, once with phenol: chloroform (saturated acid phenol (pH <6.0): chloroform = 5:1) extraction and once with equal volume of chloroform: isoamyl alcohol (24:1) extraction. RNA was precipitated with 1/10 volume of 3 M NaAc pH 4.5 and 2 volume of 96% ethanol for 20 min at -20°C. The RNA pellet was spun down in microcentrifuge (Heraeus Biofuge Fresco) with full speed at 4°C and washed twice with equal volume of ice cold 75% ethanol. The pellet was dried briefly at room temperature and resuspended in about 50 µl RNase-free H₂O. The RNA should be stored at -80°C for longer storage.

5.3.8 RNA clean up, cDNA synthesis and qRT-PCR

The purified 100 µg of total RNA was treated with 20 U RNase-free DNase I (Amersham) for 60 min at 37°C. RNA was purified using the RNeasy Mini Kit (QIAGEN). The absence of DNA was verified by PCR reaction of 32 cycles with Taq polymerase. The presence of PCR products were checked by agarose gel electrophoresis. The above steps should be repeated if there was DNA contamination in the RNA sample. 1.0 µg of DNA-free total RNA was used as the template to synthesize cDNA with the cDNA Archive kit (ABI) following the recommended protocol.

The qRT-PCR reactions were carried out in triplicates in a total volume of 25 µl containing 12.5 µl Sybr green PCR Master Mix (ABI), 1 µl of each primer (10 µM), 0.1 µl cDNA and 11.9 µl H₂O. AB 7300 Real time PCR detection system was used for qRT-PCR reactions with standard conditions.

5.3.9 Construction of in-frame deletion strain in *M. xanthus*

In-frame deletion mutants were constructed by two-step homologous recombination (Figure 35). Briefly, approximately 1100 bp PCR products containing the in-frame deletions were cloned in the plasmid pBJ114 (Julien *et al.*, 2000), which contains the *galK* gene for counter selection. Primers used for the constructions are listed in Table 14. Four primers (A, B, C, D) were designed to amplify the 1100 bp fragment carrying an *hpk* in-frame deletion by PCR with *M. xanthus* chromosomal DNA as template. Shortly, primers A and B were used to amplify the upstream flanking region of the *hpk* gene. Primer A contained a restriction site for cloning in pBJ114 and primer B either contained a restriction site for cloning or a region complementary to the downstream flanking PCR fragment. Primers C and D were used to amplify the downstream flanking fragment of the *hpk* gene. Primer C either contained a restriction site for cloning or a region complementary to the upstream flanking PCR fragment and primer D contained a restriction site for cloning in pBJ114. The fragments AB and CD were used to generate the full-length in-frame deletion fragment either by direct cloning or in a second PCR reaction with primers A and D and the two flanking PCR fragments as templates. The *hpk* in-frame fragments were cloned into the plasmid pBJ114 and then transformed into *E. coli* Top10 and checked by sequencing.

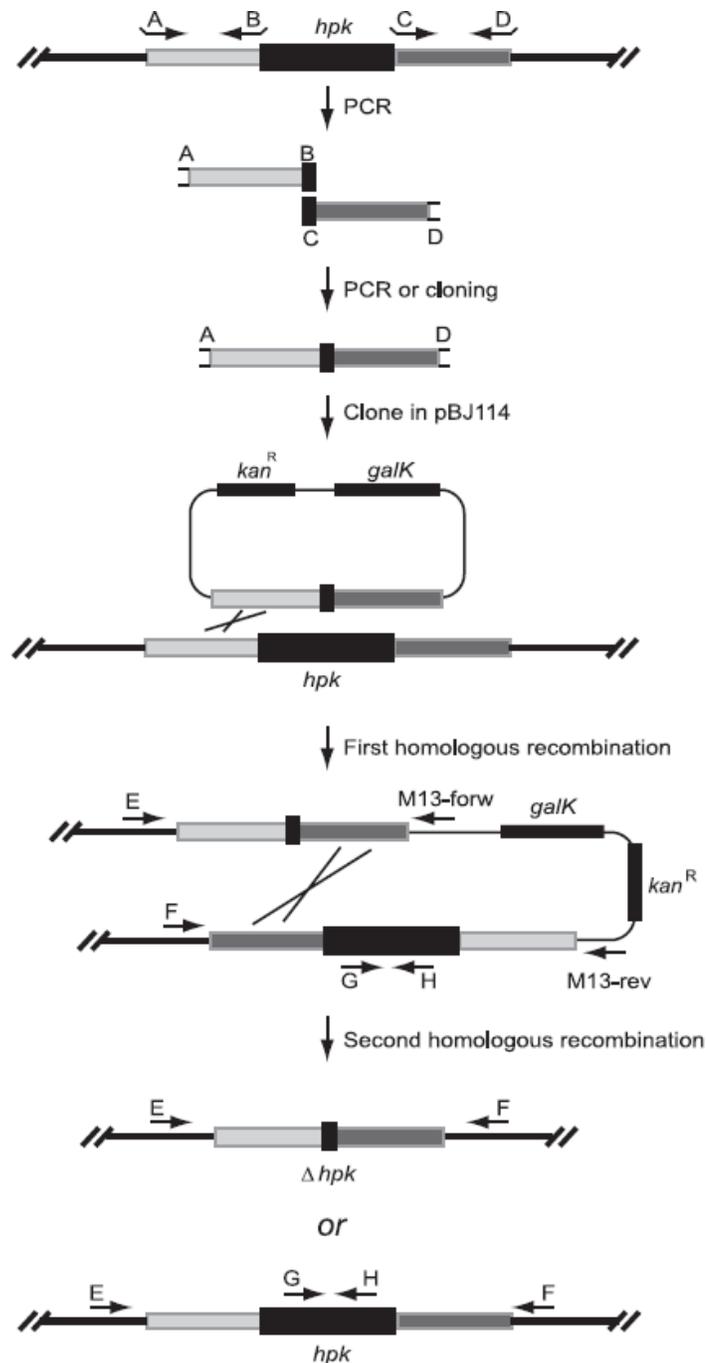


Figure 35. Outline of strategy for generating in-frame deletions in genes encoding HPKs. Detailed description is in the text.

Correct plasmids were introduced into the *M. xanthus* wild type strain DK1622 by electroporation (Kashefi & Hartzell, 1995). The insertion of plasmids after the first homologous recombination was confirmed by three PCR reactions with three primer pair combinations (Primers are listed in Table 14): Primers E (binds

upstream of primer A) and F (binds downstream of primer D), and primers E and M13-forward (hybridizes to pBJ114), and primers F and M13-reverse (hybridizes to pBJ114). For each in-frame construct, at least one clone with the insertion of the plasmid in upstream flanking region of the *hpk* gene and one clone with the insertion in the downstream flanking region of the of *hpk* gene were chosen for the second homologous recombination. To isolate clones containing the in-frame deletion, cells were grown in liquid 1.0% CTT medium with 40 µg/ml kanamycin to mid-log phase. Subsequently, the culture was diluted 100-fold in 1.0% CTT medium without kanamycin and cells grown to mid-log phase, diluted and plated on CTT plate with 1% or 2% galactose (Sigma) for counter-selection. Galactose resistant and kanamycin sensitive colonies were screened out and checked by two PCR reactions with the primers E and F and the primers G and H, which bind to the deleted part of the *hpk* gene) to verify the in-frame deletion. For *MXAN3036* and *MXAN4988* no in-frame deletions were obtained with the 1100 bp in-frame deletion constructs using the primers labelled “short” in Table 14, i.e. all galactose resistant clones contained in the intact genes. Therefore, for both genes in-frame deletion constructs on 1400 bp fragments were generated using the primers labelled “long” in Table 14. Also with these constructs, no in-frame deletions were isolated.

5.4 Biochemical methods

5.4.1 Immunoblot analysis

The immunoblot analysis was performed following standard protocol (Sambrook & Russell, 2000). First, SDS-PAGE was carried out with standard protocol (Sambrook & Russell, 2000) and the SDS-PAGE for FrzCD methylation was performed as described (McCleary *et al.*, 1990). After SDS-PAGE, the separation gel and a piece of slightly bigger size of nitrocellulose membrane (Amersham Biosciences) were soaked with transfer buffer (39 mM Glycin, 48 mM Tris, 0.0375% SDS, 20% methanol) for 5 min. The transfer sandwich was assembled from anode to cathode with 3 pieces of Whatmann 3MM paper soaked with transfer buffer, nitrocellulose membrane, separation gel and 3 pieces of Whatmann 3MM paper soaked with transfer buffer. The transfer were performed in Hoefer TE 77 semi-dry blotting apparatus (Amersham Bioscience,

München) with a constant current of 0.8 mA/cm² of the membrane size for 1-2 hrs. After transfer, the membrane was soaked in TTBS (20 mM Tris-HCl, 137 mM NaCl, 0.05 % Tween 20, pH 7.6) with 5% nonfat dry milk (only for the anti-phosphotyrosine antibody, 2% BSA instead of 5% nonfat dry milk) in dedicated container for overnight at 4° C. The membrane was washed 2× 5 min in TTBS with gentle shaking and subsequently was incubated with the diluted primary antibodies in TTBS containing 2% nonfat dry milk (for the anti-phosphotyrosine monoclonal antibody, 2% BSA instead of 2% nonfat dry milk was used for all steps) for 2 hrs at RT or overnight at 4° C. In this study, anti-FruA antibodies were 5,000-fold diluted, anti-Hpk8 antibodies were 10,000-fold diluted, anti-FrzCD antibodies were 5,000-fold diluted and anti-phosphotyrosine monomeric antibody (Sigma-Aldrich) was 1,000-fold diluted. The membrane was washed in TTBS 2x 5 min and 1x 15 min and then incubated with diluted secondary antibodies in TTBS containing 2% nonfat dry milk (only for the anti-phosphotyrosine monoclonal antibody, 2% BSA instead of nonfat dry milk was used for all steps) for 1 hr at RT. The membrane was washed in TTBS 5x 5 min in TTBS, detected by chemiluminescence (Perkin Elmer Life Sciences) and exposed on a film.

5.4.2 Over-expression of proteins in *E. coli*

The procedures for protein over-expression in *E. coli* Rosetta 2 (DE3) are as follows. The overexpression construct was transformed into *E. coli* Rosetta 2 (DE3) and a positive colony was inoculated in LB medium with selective antibiotics overnight. Three different induction methods were used for overexpression of individual proteins: 1). Overnight culture was 1,000-fold diluted into ZY medium (1% tryptone, 0.5% yeast extract, 0.025 M (NH₄)₂SO₄, 0.05 M KH₂PO₄, 0.05 M Na₂HPO₄, 0.001 M MgSO₄, 0.5% glycerol, 0.05% glucose and 0.2% α-lactose) with selective antibiotics and was incubated for 3-5 hrs at 37°C. Then the culture was further cultivated at 18°C overnight or over two nights until OD₆₀₀ surpassed 2.5; 2). The culture was 100-fold diluted into 2x YT medium (1.6% Bacto-tryptone, 1% yeast extract, 0.5% NaCl, pH 7.2) and incubated at 37°C until OD₆₀₀ 0.5~0.7. IPTG was added into the culture at a final concentration of 0.1 mM and the culture was further incubated at 18°C for 16 hrs; 3). The culture was 100-fold diluted into 2x YT medium and incubated at

37°C until OD₆₀₀ around 0.3. The ethanol was added to the culture with the final concentration of 2% and the cells were further incubated for 30 min at 37°C. Subsequently, IPTG was added to the culture with the final concentration of 0.1 mM and the culture was induced at 18°C for around 16 hrs.

The procedure for salt induction is as follows. The overexpression construct was transformed into salt induction strain of GJ1158 (Bhandari & Gowrishankar, 1997). A positive colony was picked out and grown in LBON medium (LB medium without NaCl) overnight. The culture was diluted into LBON medium and incubated at 37°C until OD₆₀₀ 0.8. The NaCl solution was added to the culture with a final concentration of 0.3 M NaCl. The culture was further incubated at 37°C for 3 hrs.

5.4.3 The solubility test of over-expressed proteins

10 ml of over-expressed culture was harvested after induction. The cell pellet was resuspended in 1 ml of lysis buffer (for different proteins the buffers are different and the buffer recipes are described in the purification procedures) with protease inhibitor tablet (Roth) and 1 mg/ml lysozyme and incubated on the ice for 30 min. The cell suspension was sonicated 6x 20 s with microtip, output 3 and duty 50%. The lysate was centrifuged at 4,700 rpm for 10 min at 4°C (Heraeus Multifuge 1 S-R). The supernatant was saved as soluble fraction. The pellet was resuspended into lysis buffer and saved as insoluble fraction. The different fractions were examined by SDS-PAGE.

5.4.4 Purification of His-tag protein under native conditions

After induction, 500-1,000 ml culture was spun down and pellet was resuspended in 50 ml lysis buffer (10% glycerol, 50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 10 mM imidazole, pH 8) with protease inhibitor tablet. The cells were lysed by sonication with 4x 5 min, output 5 and duty 50%. The lysis mixture was separated at 40,000 rpm for 1 hr at 4°C in Beckman ultra centrifuge. The supernatant was mixed with 2 ml Ni-NTA slurry (the volume could be variable dependent of the amount of proteins) gently in Labinco LD-79 Rotator at 4°C for 60 min. The lysate-Ni-NTA mixture was load on Poly-Prep Chromatography columns (Bio-Rad). After the resin was settled down, the bottom cap was removed and the flow-through was collected. The column was

washed with 2x 20 ml and 1x 10 ml wash buffer (10% glycerol, 50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 20 mM imidazole, pH 8). The protein was eluted with elution buffer containing different concentrations of imidazole: 50 mM, 100 mM, 150 mM, 200 mM and 3x 250 mM. The different fractions were collected, the concentrations were measured by Bradford assay (Biorad protein assay) and the purity was checked by SDS-PAGE.

5.4.5 Gel filtration chromatography

In this study, the native Trx-His₆-SdeK KD was further purified with gel filtration chromatography. Briefly, Trx-His₆-SdeK KD protein was condensed with Amicon Ultra-15 centrifugal filter 30 kD (Millipore) and changed into running buffer (100 mM Tris-HCl, 150 mM NaCl). The Superose™ 12 10/300 GL Column in ÄKTA™ UNICORN equipment (Amersham Biosciences, München) was equilibrated with 6 column volume (CV) of running buffer (100 mM Tris-HCl, 150 mM NaCl) with a flow speed of 0.3 ml/min. Subsequently, 500 µl of the Trx-His₆-SdeK KD protein was loaded on the column. Then the column was eluted with 3 CV of running buffer. The protein was detected with 280 nm and 254 nm UV detectors. The fractions which contain the protein were collected and checked by SDS-PAGE. The column was further equilibrated with 6 CV running buffer and washed with 3 CV of 20% ethanol.

5.4.6 Purification of His-tag protein under denature conditions

After induction, 500 ml culture was spun down and pellet was resuspended in 50 ml lysis buffer (10% glycerol, 50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 10 mM imidazole, pH 8) with protease inhibitor tablet. The cells were lysed by sonication with 4x 5 min, output 5 and duty 50%. The lysis mix was separated at 40,000 rpm for 1 h at 4°C in Beckman ultra centrifuge. The pellete was resuspended in 25 ml solution A (0.1 M Tris-HCl pH 7.0, 1 mM EDTA) at 4°C using an ultraturrax. 0.5 volume of solution B (60 mM EDTA, 6% Triton X-100, 1.5 M NaCl pH 7) was added to the suspension and incubated for 30 min at 4°C. The inclusion bodies were spun down by centrifugation at 31,000 x g for 10 min, 4°C. The pellet was washed twice with 40 ml solution C (0.1 Tris-HCl pH 7.0, 20 mM EDTA) and then solubilized with 20 ml denaturing buffer (100 mM NaH₂PO₄, 10 mM Tris-HCl and 8 M urea pH 8.0) by stirring the mixture on Labinco LD-79 Rotator at room temperature for 60 min. The lysate was

centrifuged at 10,000 x *g* for 20–30 min at room temperature to pellet the cellular debris. The supernatant was mixed with 3 ml Ni-NTA slurry gently in Labinco LD-79 Rotator at room temperature for 60 min. The lysate-Ni-NTA mixture was load on Poly-Prep Chromatography columns (Bio-Rad). After the resin was settled down, the bottom cap was removed and the flow-through was collected. To purify the desired protein, two following methods could be performed alternatively or consequentially based on the purity of purified protein. One method is to elute protein from the column with a gradient of imidazole concentration. The column was washed with 2x 20 ml and 1x 10 ml wash buffer (10% glycerol, 50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 20 mM imidazole, pH 8). The protein was eluted with elution buffer containing imidazole at gradient concentrations: 50 mM, 100 mM, 150 mM, 200 mM and 3x 250 mM. The other method is eluting protein with pH drop. The column could be washed by 2x buffer C (100 mM NaH₂PO₄, 10 mM Tris-HCl and 8 M urea pH 6.3) and eluted with 5x 1.5 ml buffer D (100 mM NaH₂PO₄, 10 mM Tris-HCl and 8 M urea pH 5.9) and 5x 1.5 ml buffer E (100 mM NaH₂PO₄, 10 mM Tris-HCl and 8 M urea pH 4.5).

5.4.7 Refolding of protein

In this study, the purified denatured Trx-His₆-SdeK KD was refolded with the described protocol (Pollack & Singer, 2001). The purified denatured full length of His₆-SdeK was refolded with following protocol. 4 mg of protein was added to 80 ml of refolding buffer (0.1 M Tris-HCl pH 8.5, 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF and 0.5 M L-Arginine-HCl) drop by drop while stirring. The mixture was stirred at 4°C overnight and subsequently concentrated to a volume of 500 µl at 4°C by using Amicon Ultra-15 with desired buffer. Protein concentration was determined by a standard Bradford assay.

5.4.8 Purification of FruA and FruA_{D59N} with strep-tag

After autoinduction, 500 ml culture was spun down and pellet was resuspended in 50 ml lysis buffer (100 mM Tris-HCl pH 8, 300 mM NaCl, 1 mM EDTA) with protease inhibitor tablet. The cells were lysed by sonication with 4x 5min, output 5 and duty 50%. The lysis mix was separated at 40,000 rpm for 1 h at 4°C in Beckman ultra centrifuge. The supernatant was loaded on the Polypropylene Column (1 ml, Qiagen) filled with 1ml *Strep-Tactin*[®] Superflow[®] 50% suspension

(IBA, the column was equilibrated with 10-20 ml lysis buffer before use). After all of the supernatant passed through, the column was washed with 2x 20 ml and 1x 10 ml lysis buffer. The column was eluted by 6x 1 ml 1x strep tag elution buffer (diluted with lysis buffer from 10x strep tag elution buffer (IBA)). The different fractions and elutes were further analyzed with SDS-PAGE. The strep tag column was washed with 3x 4 ml lysis buffer and regenerated with 3x 4 ml buffer R (IBA, diluted from 10x buffer with lysis buffer). This column could be reused about 5 times.

5.4.9 Autophosphorylation of Hpk8

The autophosphorylation reaction was mixed carefully with following reagents: 50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 10 μ M protein, 50 mM KCl and 20 mM MnCl₂. The reaction is started by adding 1/10 volume of the ATP mixture with 1:1 ratio of 10 mM ATP, [γ -³²P] -ATP (>220 TBq/mmol, Hartmann analytic GmbH) at defined times and a control reaction without ATP mixture was started together with the longest incubation time of reaction. The reactions were incubated at 25°C in Thermomixer (Eppendorf) and stopped at the same time point by adding 3x SDS loading buffer (180 mM Tris-HCl, pH 6.8, 6% SDS, 30% glycerol, 0.015% bromo-phenol-blue, 15 mM EDTA and 0.3 M DTT). All of the reactions were loaded into 12% SDS-PAGE followed by electrophoresis at 150 V for 50 min in Biorad gel system. The gel tank was dissembled and the dye front of the gel was cut to get rid of the signal interruption from free phosphate and ATP. The gel was covered with plastic bag and exposed to the phosphor screen in cassette overnight. After exposure, the phosphor screen was scanned by phosphorimager.

5.4.10 Chemical stability of phosphoryl group

The autophosphorylation reaction was performed as described above. The reaction was stopped by adding 2x SDS loading buffer. 10 μ l of quenched protein was kept at 4°C as a control reaction. Each of 10 μ l quenched protein was added to the tubes respectively containing 2 μ l 1 M NaOH, H₂O, 1 M HCl, H₂O, 99% pyridine or 1 M hydroxylamine and incubated at 42°C for 1hr. Then the reactions were separated by SDS-PAGE and detected by phosphoimaging and pageblue staining.

5.4.11 Phosphotransfer reaction from phosphorylated Hpk8 to FruA

Trx-His₆-Hpk8 KD (10 μM) was autophosphorylated with [γ -³²P] ATP for 18 hrs. 5 μl of the autophosphorylated Trx-His₆-Hpk8 KD (10 μM) was mixed with an equal volume of FruA-strep (10 μM or 20 μM) or FruA_{D59N}-strep (calculated concentration of 20 μM or 26 μM) for each phosphotransfer reaction and incubated at 25°C for a defined time. The reactions were stopped by adding 3x SDS loading buffer, separated by SDS-PAGE and detected by phosphoimaging and pageblue staining.

5.5 Bioinformatics methods

5.5.1 Sequence retrieval and domain analysis of protein structure

All of the protein or gene sequences of *M. xanthus* were retrieved from Tigr database (<http://cmr.tigr.org/tigr-scripts/CMR/GenomePage.cgi?org=gmx>). The proteins from other organisms are from NCBI database (<http://www.ncbi.nlm.nih.gov/sites/gquery>). The domain analyses were performed in SMART database (<http://smart.embl-heidelberg.de/>).

5.5.2 Identification and phylogenetic analysis of TCS proteins.

TCS proteins can generally be divided into three groups: HPK, which contain either a HisKA, HisKA₂, HisKA₃ (here after these three domains will be collectively referred to as the HisKA domain) or Hpt domain and the conserved HATPase_c domain; HPK-like proteins, which contain the HisKA domain and lack the HATPase_c domain or *vice versa*; and, RR which contain the conserved receiver domain. HPK, HPK-like proteins and receiver domain-containing proteins were identified using the TIGR Comprehensive Microbial Resource database (<http://cmr.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi>) and HMMER (HMMER 2.3, <http://hmmer.wustl.edu>) searches (these two searches done by Stuart Huntley) on the *M. xanthus* proteome using the pfam HisKA, HisKA₂, HisKA₃, Hpt, and HATPase_c matrices for domains in HPKs and the Response_{reg} matrix for receiver domains. An E-value cutoff of 0.02 was set for HisKA, Hpt and Response_{reg} domains and an E-value cutoff of 1e-5 was set for the HATPase_c domain. Sequences from proteins containing one or more of the homologous matrices of HisKA, Hpt, HATPase_c and Response_{reg} were extracted based on boundaries identified using the HMMER results, SMART analysis results (Letunic *et al.*, 2004), and manual

curation using the domain boundary definitions described by Grebe and Stock (Grebe & Stock, 1999).

To generate phylogenetic trees (done by Stuart Huntley), the *M. xanthus* TCS domains were compared to the domain sets used by Grebe and Stock (Grebe & Stock, 1999) by aligning all HPK domains and, separately, all receiver domains using CLUSTALW 1.83 alignment software (gap open penalty = 50, gap extension penalty = 0.5, Gonnet protein matrix) (<http://www.ebi.ac.uk/clustalw>) (Thompson *et al.*, 1994). The resulting alignments were improved by manual curation and were then used to generate phylogenetic trees using the neighbor joining method (Saitou & Nei, 1987) (1,000 bootstrapping replications implemented).

To identify output domains in RR (done by Stuart Huntley), HMMER searches were performed on all *M. xanthus* proteins bearing receiver domains using all matrices provided in the pfam database (<http://www.sanger.ac.uk/Software/Pfam>) (Finn *et al.*, 2006). Results from the searches were inspected and, based on the annotations provided for each pfam matrix; all hits from known output domains were noted. For those proteins for which the above search failed to identify an output domain, possible cryptic, conserved domains were searched as follows: Peptide sequences leading and trailing the receiver domains from these proteins were collected and subjected to a BLAST search of the NCBI non-redundant protein database. Results were manually analyzed, but no new domains were identified.

Trans-membrane helices were predicted using the TMHMM 2.0c software package with the provided default model and options (<http://www.cbs.dtu.dk/services/TMHMM>) (Sonnhammer *et al.*, 1998) (done by Stuart Huntley). If a protein was predicted to contain a single trans-membrane helix, the location of the trans-membrane helix was inspected. If the helix was located within the first 20 amino acid residues and would, thus, overlap with the signal peptide, the HPK was categorized as cytoplasmic.

5.5.3 Classification of TCS proteins in *M. xanthus* based on genetic organization

Based on the genetic organization of TCS genes, a set of criteria was developed to classify these genes into three groups (Figure 10) using the

following criteria. Two adjacent genes encoding an HPK and a RR, or an HPK-like protein and a RR and transcribed in the same direction were grouped into paired genes. TCS genes in gene clusters containing two or more RR genes, two or more HPK or HPK-like genes, or three or more TCS genes independently of their transcriptional direction, were grouped into complex genes. All other gene organizations of TCS genes were grouped into orphan genes. Three *che* gene clusters are flanked by three orphan genes (*MXAN6953*, *MXAN6955* and *MXAN6966*) encoding HPKs and one gene (*MXAN2687*) encoding an HPK-like protein that are not CheA-like and these genes are classified as orphan.

6 Supplementary data

6.1 TCS proteins in *M. xanthus*

Table S1. Characteristics of histidine protein kinase genes and proteins in *M. xanthus*.

TIGR_MXAN (Gene symbol)	Ref.	Max. expression ratio (microarrays) /Induction time (hrs) ^{a,b}	Max. expression ratio (qRT-PCR) /Induction time (hrs) ^a	Effect on growth	Effect on motility ^c	Effect on development	TM helices	Grebe & Stock classification ^d	Domain organization ^e	Genetic organization ^f
Che clusters										
<i>MXAN2686</i> (<i>cheA4</i>)	(Vlamakis <i>et al.</i> , 2004)	NR		No	S	ND ^h	0	HPK9	Hpt-HATPase_c-RR	# # # # +cheW4 +cheR4 +mcp4 +cheY4 +cheW4 +cheA4 -2687 # # # #
<i>MXAN4140</i> (<i>frzE</i>)	(Bustamante <i>et al.</i> , 2004)	NR		No	A, S	Yes	0	HPK9	Hpt-HATPase_c-RR	# # # # -frzF -frzG -frzE -frzCD -frzB -frzA +frzZ # # # # +4149
<i>MXAN4758</i> (<i>cheA8</i>)	J. Kirby	NOA					0	HPK9	Hpt-HATPase_c	# # # # -cheY -cheB -cheR # # -cheW -cheW -cheA -cheY # # # #
<i>MXAN5147</i> (<i>cheA3</i>)	(Kirby & Zusman, 2003)	NR		No	No	Yes	0	HPK9	Hpt-HATPase_c-RR	# # # # -cheR3 -cheB3 # -cheA3 -mcp -mcp # cheW3 # +5153 # # # #
<i>MXAN6029</i> (<i>cheA5</i>)	J. Kirby	NR					0	HPK9	Hpt-HATPase_c-RR	# # # # -mcp -cheB -cheA # -cheR -cheW -cheY # # # #
<i>MXAN6692</i> (<i>difE</i>)	(Yang <i>et al.</i> , 1998)	NR		No	S	Yes	0	HPK9	Hpt-HATPase_c	# # # # -difG -difE -difD -difC -difB -difA # # # #
<i>MXAN6951</i> (<i>cheA6</i>)	J. Kirby	NR					0	HPK9	Hpt-HATPase_c-RR	# # # # +cheW +cheR +cheW +mcp +cheA +cheB +6953 # +6955 -6956 # #
<i>MXAN6964</i> (<i>cheA7</i>)	J. Kirby	NR					0	HPK9	Hpt-HATPase_c	# +6955 -6956 # # -cheB -cheR # -mcp -cheW -cheA -cheY -6966 # +6968 #

Complex clusters										
MXAN0195		NR					0	HPK3g	HisKA-HATPase_c-RR	##### -0195 +0196 -0197 ###
MXAN0196		NOA					0	HPK3f	HisKA-HATPase_c	##### -0195 +0196 -0197 ###
MXAN0197		NR					8	HPK3f	HisKA- HATPase_c	##### -0195 +0196 -0197 ###
MXAN0229		NR					0	HPK1a	RR-HisKA-HATPase_c-RR	##### -0229 -0230 #####
MXAN0459 (redC)	(Higgs et al., 2005)	1.9xup /6-9		No	No	Yes	1	HPK4	HisKA- HATPase_c	##### +0459 +0460 +0461 +0462 #####
MXAN4244		NR					0	HPK4	HisKA-HATPase_c-RR	# +4240 ### -4244 -4245 -4246 ##### +4251
MXAN4246		NR					0	HPK4	HisKA- HATPase_c	# +4240 ### -4244 -4245 -4246 ##### +4251
MXAN4251		NR					3	HPK4	HisKA- HATPase_c	-4246 ### +4251 +4252 +4253 ### +4257 #
MXAN4444		NR					2	HPK3g	HisKA-HATPase_c	##### -4444 -4445 #####
MXAN4445		NR					0	HPK1a	RR-HisKA-HATPase_c-RR-RR	##### -4444 -4445 #####
MXAN4786		NR					2	HPK4	HisKA- HATPase_c	##### -4785 -4786 +4787 ### ##
MXAN5365 (hsfB)	(Ueki & Inouye, 2002)	NR		Essential	NA ^h	NA ^h	0	HPK3g	RR-HisKA-HATPase_c	##### -5364 -5365 -5366 ### ##
MXAN6734		2.7xup /4-6					0	HPK3f	RR-HisKA-HATPase_c	##### -6734 -6735 #####
MXAN6735		NR					0	HPK1b	HisKA-HATPase_c-RR-RR-RR	##### -6734 -6735 #####
MXAN6865	This study	NR		No	No	No	0	HPK4	HisKA-HATPase_c	##### -6865 -6866 #####
MXAN7002	This study	2.5xup /0-2		No	No	No	0	HPK3f	HisKA- HATPase_c	- ##### +7001 -7002 +7003 ### ##
MXAN7003	This study	NR		No	No	No	7	HPK3f	HisKA-HATPase_c	- ##### +7001 -7002 +7003 ### ##
MXAN7363		NR					2	HPK1a	HisKA-HATPase_c-RR-RR	##### -7362 -7363 -7364 ## +7368 #
Orphan genes										
MXAN0060	This study	NOA	19.3xup /0-6	No	No	No	2	HPK3f	HisKA-HATPase_c	##### +0060 #####
MXAN0095		NR					0	HPK3g	HisKA-HATPase_c-RR	##### -0095 #####
MXAN0168		NOA	11.0xdown /0-6				1	HPK1a	HisKA-HATPase_c	##### +0168 ### - #

MXAN0176	This study	NOA	14.8xup /6-12	No	No	No	0	HPK3f	HisKA-HATPase_c	# -0172 # # # +0176 # # # # #
MXAN0245		NOA	8.9xdown /0-6				0	HPK3f	HisKA-HATPase_c	# # # # # +0245 # # # # #
MXAN0304		NR					3	HPK3f	HisKA-HATPase_c	# # # # # -0304 # # # # #
MXAN0314		NR					0	HPK3g	HisKA-HATPase_c-RR	# # -0311 # # +0314 # # # # #
MXAN0336		NR					0	HPK3f	RR-HisKA-HATPase_c	# # # # # -0336 # # # -0340 #
MXAN0340	This study	2.7xup /4-6		No	No	No	0	HPK3f	HisKA-HATPase_c	-0336 # # # -0340 # # # # #
MXAN0347	This study	NR		No	No	No	7	HPK3f	HisKA-HATPase_c	# # # # # -0347 # # # # #
MXAN0399		NR					0	HPK4	HisKA-HATPase_c	# # # # # -0399 # # # # #
MXAN0571	This study	1.9xup /2-4		No	No	No	0	HPK3f	HisKA-HATPase_c	# # # # # +0571 # # # # #
MXAN0612		NR					0	HPK3g	HisKA-HATPase_c	# # # # # +0612 # # # # #
MXAN0643		NR					0	HPK3f	HisKA-HATPase_c	# # # # # -0643 # # # # #
MXAN0706	This study	NOA	3.4xup /0-6	No	No	No	0	HPK3f	HisKA-HATPase_c	# # # # # -0706 # # # +0710 #
MXAN0712	This study	1.6xup /12-15		No	No	Yes	0	HPK1b	HisKA-HATPase_c-RR-RR	# # # +0710 # +0712 # # +0715 # #
MXAN0720	This study	NR		No	No	No	0	HPK3f	HisKA-HATPase_c	+0715 # # # # # +0720 # # # # #
MXAN0736	This study	2.2xup /0-2	128xup /0-6	No	No	No	2	HPK3f	HisKA-HATPase_c	# -0732 -0733 # # -0736 # # # # #
MXAN0928	This study	10.9xup /2-4	29.7xup /0-6	No	No	No	0	HPK3g	HisKA-HATPase_c	# # # # # -0928 # # +0931 # #
MXAN0931 (espA)	(Cho & Zusman, 1999b)	4.4xup /0-2		No	No	Yes	0	HPK3g	HisKA-HATPase_c-RR	# # -928 # # +0931 # # # +935 #
MXAN0993		NR					0	HPK3f	RR-HisKA-HATPase_c	# # # # # +0993 # # # # #
MXAN1014 (sdeK)	(Garza et al., 1998, Pollack & Singer, 2001)	2.6xup /0-2		No	No	Yes	0	HPK3f	HisKA-HATPase_c	# # # # # +1014 # # # # #
MXAN1249 (sass)	(Yang & Kaplan, 1997)	NR		No	No	Yes	2	HPK4	HisKA-HATPase_c	# -1245 # # # -1249 # # # # #
MXAN2317	This study	NR	NR	No	No	No	0	HPK1b	HisKA-HATPase_c-RR-RR-Hpt	# # # # # +2317 # # # # #
MXAN2368		NR					0	HPK3f	HisKA-HATPase_c	# # # # # +2368 # # # # #

MXAN2386		NR					0	HPK3g	HisKA-HATPase_c-RR	##### +2386 #####
MXAN2606		NR					0	HPK1a	HisKA-HATPase_c-RR-HisKA-HATPase_c	##### +2606 #####
MXAN2763	This study	2.4xup /0-2		No	No	No	0	HPK3g	HisKA-HATPase_c-RR	##### +2763 #####
MXAN2785		NR					0	HPK1a	HisKA-HATPase_c-RR	##### -2785 #####
MXAN3036	This study	2.1xup /4-6	29.8xup /0-6	Essential ?	NA ^h	NA ^h	3	HPK4	HisKA-HATPase_c	##### +3036 #####
MXAN3098	This study	1.9xup /0-2	4.1xup /0-6	No	No	No	0	HPK3f	HisKA-HATPase_c	##### -3098 #####
MXAN3290	This study	2.8xup /0-2	4.4xup /0-6	No	No	Yes	0	HPK3f	HisKA-HATPase_c-RR	##### -3290 #####
MXAN3343	This study	NOA	3.7xdown /18-24	No	No	No	0	HPK4	HisKA-HATPase_c	##### -3343 #####
MXAN3879		NOA	16.6xdown /0-6				0	HPK4	HisKA-HATPase_c-RR-RR	##### +3879 #####
MXAN3974		NR					0	HPK3g	HisKA-HATPase_c	##### -3974 #####
MXAN4053		NR					0	HPK3g	HisKA-HATPase_c	# -4049 ### +4053 #####
MXAN4465	This study	2.1xup /4-6		No	No	Yes	0	HPK4	HisKA-HATPase_c-RR	# -4461 # +4463 # +4465 # -4468 #
MXAN4640 (sgmT)	(Youderian & Hartzell, 2006)	NR		ND ^h	S	ND ^h	0	HPK1a	HisKA-HATPase_c-RR	##### -4640 ##### -4645
MXAN4988	This study	1.6xup /0-2	2.8xup /0-6	Essential ?	NA ^h	NA ^h	8	HPK4	HisKA-HATPase_c	##### -4988 #####
MXAN5034		NR					2	HPK4	HisKA-HATPase_c	##### +5034 #####
MXAN5184	This study	NR		No	No	No	2	HPK4	HisKA-HATPase_c	##### +5184 ##### +5189
MXAN5483	This study	2.2xup /4-6		No	No	No	0	HPK3f	HisKA-HATPase_c	##### -5483 #####
MXAN5704	This study	NOA	14.0xdown /0-6	No	No	No	5	HPK1a	HisKA-HATPase_c	##### +5704 #####
MXAN5990		NR					0	HPK3f	HisKA-HATPase_c	##### -5990 ##### +5995
MXAN6015		NR					0	HPK4	HisKA-HATPase_c	## +6012 ## +6015 #####
MXAN6053	This study	NR		No	No	No	2	HPK3f	HisKA-HATPase_c	##### -6053 #####
MXAN6117		NOA	1.9xdown /0-6				0	HPK3f	HisKA-HATPase_c	##### +6117 #####
MXAN6315	This study	4.2xup /0-2	36.5xup /0-6	No	No	No	0	HPK3f	HisKA-HATPase_c-RR	##### -6315 #####

MXAN6335		NR					0	HPK3f	HisKA-HATPase_c-RR	##### -6335 #####
MXAN6586		NR					0	HPK3f	HisKA-HATPase_c	##### +6586 #####
MXAN6702		NR					0	HPK1a	HisKA-HATPase_c	##### -6702 #####
MXAN6847		NR					0	HPK3f	RR-HisKA-HATPase_c	##### -6847 #####
MXAN6855 (<i>espC</i>)	(Lee <i>et al.</i> , 2005)	1.7xup /12-15		No	No	Yes/No ^g	8	HPK3g	HisKA-HATPase_c-RR	##### -6855 #####
MXAN6941	This study	NOA	7.5xup /6-12	No	No	No	0	HPK3f	HisKA-HATPase_c	##### +6941 #####
MXAN6953 (<i>socD</i>)	L. Shimkets	NR		ND ^h	ND ^h	ND ^h	0	HPK3f	HisKA-HATPase_c	+cheR +cheW +mcp +cheA +cheB +6953 # 6955 -6956 # #
MXAN6955 (<i>todK</i>)	(Rasmusen & Sogaard-Andersen, 2003)	NR		No	No	Yes	0	HPK3f	HisKA-HATPase_c	+mcp +cheA +cheB +6953 # +6955 -6956 # # -cheB -cheR
MXAN6966		NR					0	HPK3f	RR-HisKA-HATPase_c	# -mcp -cheW -cheA -cheY -6966 # +6968 # # -6971
MXAN6971		NR					0	HPK3f	HisKA-HATPase_c	-6966 # +6968 # # -6971 #####
MXAN6994	This study	4.4xup /2-4	76.4xup /0-6	No	No	No	0	HPK3f	HisKA-HATPase_c	##### -6994 # -6996 # # #
MXAN6996 (<i>asgD</i>)	(Cho & Zusman, 1999a)	1.8xup /12-15		No	No	Yes	0	HPK3f	RR-HisKA-HATPase_c	### -6994 # -6996 # # # # +7001
MXAN7027		NOA	2.1xdown /0-6				0	HPK3f	HisKA-HATPase_c	## -7024 # # +7027 #####
MXAN7059		NR					0	HPK3f	HisKA-HATPase_c	##### +7059 #####
MXAN7123	This study	4.0xup /2-4	143xup /0-6	No	No	No	0	HPK3f	HisKA-HATPase_c	##### +7123 #####
MXAN7180		NR					2	HPK3f	HisKA-HATPase_c	### +7178 # -7180 #####
MXAN7206 (<i>mokA</i>)	(Kimura <i>et al.</i> , 2001)	NOA	9.2xup /0-6	No	No	Yes/No ^g	0	HPK3g	HisKA-HATPase_c-RR	##### -7206 #####
MXAN7368		NR					0	HPK3f	HisKA-HATPase_c	-7362 -7363 -7364 # # +7368 # # # # #
MXAN7398	This study	NOA	3.8xup /0-6	No	No	No	0	HPK3f	HisKA-HATPase_c	### -7396 # -7398 #####
MXAN7444	This study	NR		No	No	No	0	HPK3f	RR-HisKA-HATPase_c	+7439 +7440 # # # +7444 # # # # #
Paired genes										
MXAN0733	(Rasmusen)	NR		No	No	Yes	0	HPK1b	HisKA-HATPase_c-RR-RR-RR	# -0728 # # # -0732 -0733 # # -0736

(rodK)	sen <i>et al.</i> , 2005, Rasmussen <i>et al.</i> , 2006)									##
MXAN0938		NR					0	HPK3g	HisKA-HATPase_c	## +0935 # -0937 -0938 #####
MXAN1077 (spdS)	(Hager <i>et al.</i> , 2001)	NOA		ND ^h	ND ^h	ND ^h	0	HPK4	HisKA-HATPase_c	##### +1077 +1078 #####
MXAN1129 (frgB)	(Cho <i>et al.</i> , 2000)	NR		No	No	No	9	HPK4	HisKA-HATPase_c	##### -1128 -1129 #####
MXAN1166		NR					0	HPK3h	HisKA-HATPase_c	##### +1166 +1167 #####
MXAN1190		NOA					2	HPK4	HisKA-HATPase_c	##### -1189 -1190 #####
MXAN1350		NR					2	HPK1a	HisKA-HATPase_c	##### -1349 -1350 #####
MXAN1553	This study	2.5xup /12-15		No	No	No	0	HPK1b	HisKA-HATPase_c	##### -1552 -1553 #####
MXAN1679		NR					2	HPK2a	HisKA-HATPase_c	##### -1679 -1680 #####
MXAN2779 (phoR2)	(Moraleda-Munoz <i>et al.</i> , 2003)	NR		No	No	Yes	2	HPK2a	HisKA-HATPase_c	##### -2778 -2779 #####
MXAN3419		NOA					2	HPK3g	HisKA-HATPase_c	##### -3418 -3419 #####
MXAN3451		NR					2	HPK2a	HisKA-HATPase_c	##### +3450 +3451 #####
MXAN3812		NR					5	HPK4	HisKA-HATPase_c	##### -3811 -3812 #####
MXAN4071		NR					5	HPK7	HisKA_3- HATPase_c	##### + 4071 +4072 #####
MXAN4165		NR					2	HPK3B	HisKA-HATPase_c	##### +4164 +4165 #####
MXAN4197		1.8xup /2-4					2	HPK4	HisKA-HATPase_c	##### - 4196 -4197 ##### -4202
MXAN4262		NR					2	HPK4	HisKA-HATPase_c	# + 4257 ### -4261 -4262 #####
MXAN4579		NR					2	HPK4	HisKA-HATPase_c	##### + 4579 +4580 #####
MXAN4778 (phoR1)	(Carrero-Lerida <i>et al.</i> , 2005)	NR		No	No	Yes	2	HPK1a	HisKA-HATPase_c	##### + 4777 +4778 #####
MXAN5082		NR					3	HPK1a	HisKA-HATPase_c	##### +5082 +5083 #####
MXAN5211		NR					0	HPK2a	HisKA-HATPase_c	##### -5211 -5212 #####
MXAN5314		2.1xup /12-15					1	HPK2a	HisKA-HATPase_c	##### +5313 +5314 #####
MXAN5778 (pilS2)	(Wu <i>et al.</i> , 1998)	NR		ND ^h	ND ^h	ND ^h	0	HPK3g	HisKA-HATPase_c	##### -5777 -5778 #####

<i>MXAN5785</i> (<i>pilS</i>)	(Wu & Kaiser, 1995)	NOA		No	No	ND ^h	5	HPK4	HisKA-HATPase_c	##### -5784 -5785 #####
<i>MXAN5852</i>		NR					12	HPK4	HisKA-HATPase_c	##### + 5852 +5853 #####
<i>MXAN5996</i>		NOA					2	HPK1a	HisKA-HATPase_c	-5990 ##### +5995 +5996 #####
<i>MXAN6150</i>		NR					0	HPK7	HisKA_3- HATPase_c	##### -6149 -6150 #####
<i>MXAN6223</i>		NR					5	HPK3e	HisKA-HATPase_c	##### +6223 +6224 #####
<i>MXAN6414</i> (<i>phoR3</i>)	(Moralet al., 2003)	NR		No	No	Yes	2	HPK2a	HisKA-HATPase_c	##### -6413 -6414 #####
<i>MXAN6979</i>		NR					4	HPK3e	HisKA-HATPase_c	##### +6979 +6980 #####
<i>MXAN7142</i>		NR					2	HPK4	HisKA-HATPase_c	##### -7142 -7143 #####
<i>MXAN7439</i>		NR					2	HPK4	HisKA-HATPase_c	##### +7439 +7440 ### -7444 #

^a Expression ratios were calculated as the expression in developing cells over the expression in vegetative cells. Maximum expression ratios indicate the maximum expression ratio for a particular gene during all time points tested using DNA microarrays (0, 2, 4, 6, 9, 12, 15, 18 and 24 hrs) and qRT-PCR (0, 6, 12, 18, 24 hrs). Induction time indicates the time interval in which the expression ratio began to change relative to that observed in vegetative cells.

^b NOA: Gene not represented or not detected on DNA microarray ; NR: Gene not significantly regulated in DNA microarray and/or qRT-PCR analyses.

^c S indicates an effect on the Type IV pili dependent S-motility system and A indicates an effect on the A-motility system.

^d See Grebe and Stock (Grebe & Stock, 1999) for a definition of different classes of HPK and RR. For hybrid HPKs only the classification of the kinase domain is included.

^e All domains present in a protein are indicated except for domains in the sensor domain.

^f Genetic organization: HPK, RR and HPK-like protein encoding genes are indicated in red, blue and lavender, respectively. Genes are indicated by their MXAN numbers except in the cases of *che* gene clusters for which the *che* or established nomenclature was adopted. Transcription using the lower or upper strand as template is indicated by + and -, respectively. # indicates a gene not encoding a TCS protein. With the exception of *che* gene clusters, gene organization is shown for 5 genes upstream and downstream from the query (indicated in italic).

^g *espC* (MXAN6855) and *mokA* (MXAN7206) were reported to be important for development in the DZ2 wild-type strain (Lee *et al.*, 2005, Kimura *et al.*, 2001). However, no developmental defects were observed in the DK1622 wild-type background used in the present study.

^h NA, not applicable; ND, not determined.

Table S2. Characteristics of histidine protein kinase-like genes and proteins in *M. xanthus*.

TIGR_MXAN (Gene symbol)	Ref.	Max. expression ratio (microarrays) /Induction time (hrs) ^{a, b}	Max. expression ratio (qRT-PCR) /Induction time (hrs) ^a	Effect on growth	Effect on motility ^c	Effect on development	TM helices	Domain organization ^d	Genetic organization ^e
Complex clusters									
<i>MXAN0230</i>	This study	NOA		No	No	No	0	HATPase_c-RR-RR	##### -0229 -0230 #####
<i>MXAN0461</i> (<i>redE</i>)	(Higgs <i>et al.</i> , 2005)	NR		No	No	Yes	0	HATPase_c	##### +0459 +0460 +0461 +0462 #####
<i>MXAN6866</i>		NR					0	HisKA	##### -6865 -6866 #####
Orphan genes									
<i>MXAN0799</i>		NR					0	HisKA	##### +0799 #####
<i>MXAN2687</i>		NR					0	HisKA	+cheR4 +mcp4 +cheY4 +cheW4 +cheA4 -2687 #####
<i>MXAN2961</i>		NR					0	HisKA	##### -2961 +2962 #####
<i>MXAN4203</i>		NR					4	HATPase_c	##### -4202 +4203 #####
<i>MXAN4432</i>	This study	NOA		No	No	No	0	RR-HATPase_c	##### -4432 #####
<i>MXAN5715</i>	This study	5.7xup /0-2	4.0xup /0-6	No	No	No	0	RR-HATPase_c	##### +5715 #####
Paired genes									
<i>MXAN1280</i>		NOA					4	HisKA	##### -1279 -1280 #####
<i>MXAN2670</i> (<i>asgA</i>)	(Plamann <i>et al.</i> , 1995)	4.2xup /0-2		Yes	No	Yes	0	HATPase_c-RR	##### +2670 +2671 #####
<i>MXAN3606</i>		NR					0	HATPase_c	##### -3605 -3606 #####
<i>MXAN4043</i>		NOA					0	HisKA	##### -4042 -4043 ##### -4049
<i>MXAN5123</i> (<i>mrpA</i>)	(Sun & Shi, 2001a, Sun & Shi, 2001b)	NOA		No	No	Yes	0	HisKA	##### +5123 + 5124 #####

^a Expression ratios were calculated as the expression in developing cells over the expression in vegetative cells. The maximum expression ratios indicate the maximum ratio for a particular gene at all time points tested using DNA microarrays (0, 2, 4, 6, 9, 12, 15, 18 and 24 hrs) and qRT-PCR (0, 6, 12, 18, 24 hrs). The induction time indicates the time interval in which the expression ratio began to change relative to that observed in vegetative cells.

^b NOA: Gene not represented or not detected on DNA microarray ; NR: Gene not significantly regulated in DNA microarray and/or qRT-PCR analyses.

^c S indicates an effect on the Type IV pili dependent S-motility system and A indicates an effect on the A-motility system.

^d Only TCS domains present in a protein are indicated.

^e Genetic organization: HPK, RR and HPK-like protein encoding genes are indicated in red, blue and lavender, respectively. Genes are indicated by their MXAN numbers except in the cases of che gene clusters for which the che or established nomenclature was adopted. Transcription using the lower or upper strand as template is indicated by + and -, respectively. # indicates a gene not encoding a TCS protein. With the exception of che gene clusters, gene organization is shown for 5 genes upstream and downstream from the query (indicated in italic).

Table S3. Characteristics of response regulator genes and proteins in *M. xanthus*.

TIGR_MXAN (Gene symbol)	Ref.	Max. expression ratio (microarrays) /Induction time (hrs) ^a	Effect on growth	Effect on motility ^b	Effect on development	Grebe & Stock classification ^c	Domain organization ^d	Genetic organization ^e
Che clusters								
<i>MXAN2684 (cheY4)</i>	(Vlamakis <i>et al.</i> , 2004)	NR	No	S	ND ^f	C1	RR	# # # # +cheW4 +cheR4 +mcp4 +cheY4 +cheW4 +cheA4 -2687 # # # # #
<i>MXAN4144 (frzZ)</i>	(Trudeau <i>et al.</i> , 1996)	NR	No	A, S	Yes	U	RR-RR	# # # # -frzF -frzG -frzE -frzCD -frzB -frzA +frzZ # # # # +4149
<i>MXAN4751 (cheY8a)</i>	J. Kirby	NR				A3	RR	# # # # -cheY -cheB -cheR # # -cheW -cheW -cheA -cheY # # # # #
<i>MXAN4752 (cheB8)</i>	J. Kirby	NOA				C5	CheB-RR	# # # # -cheY -cheB -cheR # # -cheW -cheW -cheA -cheY # # # # #
<i>MXAN4759 (cheY8b)</i>	J. Kirby	3.1xup /2-4				A3	RR	# # # # -cheY -cheB -cheR # # -cheW -cheW -cheA -cheY # # # # #
<i>MXAN5145 (cheB3)</i>	(Kirby & Zusman, 2003)	2.1xup /12-15	ND ^f	ND ^f	ND ^f	C5	CheB-RR	# # # # -cheR3 -cheB3 # -cheA3 -mcp -mcp # cheW3 # +5153 # # # # #
<i>MXAN6028 (cheB5)</i>	J. Kirby	NR				C5	CheB-RR	# # # # -mcp -cheB -cheA # -cheR -cheW -cheY # # # # #
<i>MXAN6032</i>		NR				F	RR-CheW	# # # # -mcp -cheB -cheA # -cheR -cheW -cheY # # # # #

MXAN6033 (<i>cheY5</i>)	J. Kirby	NR				A2	RR	##### -mcp - <i>cheB</i> - <i>cheA</i> # - <i>cheR</i> - <i>cheW</i> - <i>cheY</i> #####
MXAN6693 (<i>difD</i>)	(Yang <i>et al.</i> , 1998, Black & Yang, 2004)	4.0xdown /0-2	No	S	Yes	C2	RR	##### - <i>difG</i> - <i>difE</i> - <i>difD</i> - <i>difC</i> - <i>difB</i> - <i>difA</i> #####
MXAN6952 (<i>cheB6</i>)	J. Kirby	NR				C5	CheB-RR	##### + <i>cheW</i> + <i>cheR</i> + <i>cheW</i> +mcp + <i>cheA</i> + <i>cheB</i> +6953 # +6955 -6956 #
MXAN6959 (<i>cheB7</i>)	J. Kirby	NOA				C5	CheB-RR	# +6955 -6956 ## - <i>cheB</i> - <i>cheR</i> # -mcp - <i>cheW</i> - <i>cheA</i> - <i>cheY</i> -6966 # +6968 # #
MXAN6965 (<i>cheY7</i>)	J. Kirby	NR				A3	RR	# # -mcp - <i>cheW</i> - <i>cheA</i> - <i>cheY</i> -6966 # +6968 # #
Complex clusters								
MXAN0460 (<i>redD</i>)	(Higgs <i>et al.</i> , 2005)	NR	No	No	Yes	A4	RR-RR	##### +0459 +0460 +0461 +0462 #####
MXAN0462 (<i>redF</i>)	(Higgs <i>et al.</i> , 2005)	NOA	No	No	Yes	U	RR	##### +0459 +0460 +0461 +0462 #####
MXAN3213 (<i>actA</i>)	(Gronewold & Kaiser, 2001)	NOA	No	No	Yes	U	RR-GGDEF	##### +3213 +3214 #####
MXAN3214 (<i>actB</i>)	(Gronewold & Kaiser, 2001)	NR	No	No	Yes	F	RR-Sigma54-HTH_8	##### +3213 +3214 #####
MXAN3734		NR				B1	DUF-RR ⁹	##### -3734 -3735 # +3738 # # #
MXAN3735		NOA				C1	RR-GGDEF	##### -3734 -3735 # +3738 # # #
MXAN4245		NOA				U	RR-RR	# +4240 # # # -4244 -4245 -4246 # # # # +4251
MXAN4252 (<i>nla20</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	-4246 # # # # +4251 +4252 +4253 # # # +4257 #
MXAN4253		NR				U	RR	-4246 # # # # +4251 +4252 +4253 # # # +4257 #
MXAN4785 (<i>nla3</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A1	RR-Sigma54-HTH_8	##### -4785 -4786 +4787 # # # # #
MXAN4787 (<i>phoB</i>)	(Goldman <i>et al.</i> , 2006)	NR	ND ^f	ND ^f	ND ^f	A1	RR-Trans_reg_C	##### -4785 -4786 +4787 # # # # #
MXAN5052		NR				B1	RR-DUF ⁹	##### -5052 -5053 # # # # #
MXAN5053		NOA				B1	RR-GAF-GGDEF	##### -5052 -5053 # # # # #
MXAN5364 (<i>hsfA</i>)	(Ueki & Inouye, 2002)	3.8xup /0-2	Essential	NA ^f	NA ^f	A4	RR-Sigma54-HTH_8	##### -5364 -5365 -5366 # # # # #
MXAN5366		NOA				U	RR-GGDEF	##### -5364 -5365 -5366 # # # # #

MXAN7001		5.8xup /0-2				F	RR	- # # # # +7001 -7002 +7003 # # # # #
MXAN7362		NOA				B1	RR-Hpt-RR-RR-GGDEF	# # # # # -7362 -7363 -7364 # # +7368 # #
MXAN7364		NOA				A2	RR	# # # # # -7362 -7363 -7364 # # +7368 # #
Orphan genes								
MXAN0172		NR				A4	RR-Sigma54-HTH_8	# +0168 # # # -0172 # # # +0176 #
MXAN0259		NR				F	RR	# # # # # +0259 # # # # #
MXAN0311		NOA				B1	RR-GerE	# # # # # -0311 # # +0314 # #
MXAN0524		9.1xup /2-4				B1	RR	# # # # # +0524 # # # # #
MXAN0710		NOA				F	RR	# -0706 # # # +0710 # +0712 # # +0715
MXAN0715		NR				A2	RR-DUF ^g	+0710 # +0712 # # +0715 # # # # +0720
MXAN0726		NOA				U	RR	# # # # # -0726 # # # -0732 -0733
MXAN0763		2.0xup /12-15				U	RR	# # # # # -0763 # # # # #
MXAN0935		NR				A2	RR	# -0931 # # # +0935 # -0937 -0938 # #
MXAN1087		1.6xup /12-15				A1	RR-PilZ	# # # # # +1087 # # # # #
MXAN1093		2.4xup /0-2h				B1	MerR-RR	# # # # # -1093 # # # # #
MXAN1245 (<i>sasR</i>)	(Guo <i>et al.</i> , 2000)	1.5xup /6-9h	ND ^f	ND ^f	ND ^f	A4	RR-Sigma54-HTH_8	# # # # # -1245 # # # -1249 #
MXAN1378		2.9xup /4-6				F	RR	# # # # # -1378 # # # # #
MXAN1429		NOA				U	RR-DUF ^g	# # # # # +1429 # # # # #
MXAN2021		NR				U	RR	# # # # # +2021 # # # # #
MXAN2050		NR				A2	RR-DUF ^g	# # # # # +2050 # # # # #
MXAN2516 (<i>nla4</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	Yes	A4	RR-Sigma54-HTH_8	# # # # # -2516 # # # # #
MXAN2807		NOA				B3	GSPII_E_N-HDc-RR	# # # # # -2807 # # # # #
MXAN2962		NOA				F	RR	# # # # # -2961 +2962 # # # # #
MXAN2991 (<i>aglZ</i>)	(Yang <i>et al.</i> , 2004)	2.0xup /0-2	No	A	Yes	U	RR-DUF ^g	# # # # # -2992 # # # # #
MXAN3117 (<i>fruA</i>)	(Ogawa <i>et al.</i> , 1996, Ellehaug <i>et</i>	NOA	No	No	Yes	U	RR-GerE	# # # # # +3117 # # # # #

	<i>al.</i> , 1998)							
MXAN3555		NR				A4	RR-Sigma54-HTH_8	##### -3555 #####
MXAN3711		NR				C1	RR-GerE	##### +3711 #####
MXAN3738		NR				U	RR	## -3734 -3735 # +3738 #####
MXAN4049		NR				F	Pkin-RR- Guanylate_cyc	-4043 ##### -4049 ### +4053 #
MXAN4149 (<i>frzS</i>)	(Ward <i>et al.</i> , 2000)	NR	No	S	No	U	RR-DUF ⁹	+4144 ##### +4149 #####
MXAN4202		6.6xdown /2-4				F	RR-HTH_8	##### -4202 +4203 #####
MXAN4232		1.7xup /0-2				B1	RR-HDc	##### +4232 #####
MXAN4240 (<i>nla22</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	##### +4240 ### -4244 -4245
MXAN4257		NR				U	RR-GAF-GGDEF	+4252 +4253 ### +4257 ### -4261 -4262
MXAN4438		NR				B3	RR-DUF ⁹	##### +4438 #####
MXAN4461 (<i>romR</i>)	(Leonardy <i>et al.</i> , 2007)	NR	No	A	Yes	A1	RR-DUF ⁹	##### -4461 # +4463 # +4465 #
MXAN4463		NR				A4	RR-GGDEF	##### -4461 # +4463 # +4465 ## -4468
MXAN4468		NOA				U	RR	+4463 # +4465 ## -4468 ### ##
MXAN4645		NOA				U	RR	-4640 ##### -4645 #####
MXAN4675		NR				F	RR-HDc	##### +4675 #####
MXAN4717		NOA				U	RR-DNAJ-TPR_2	##### -4717 #####
MXAN4794		NOA				F	RR	##### +4794 #####
MXAN4975		NR				U	RR-DUF ⁹	##### -4975 # -4977 ###
MXAN4977		NR				A4	RR-Sigma54-HTH_8	### -4975 # -4977 #####
<i>crdA</i> (<i>nla26</i>)	(Caberoy <i>et al.</i> , 2003, Kirby & Zusman, 2003)	NR	No	No	Yes	A4	RR-Sigma54-HTH_8	##### -cheR3 -cheB3 # -cheA3 -mcp -mcp # cheW3 # +5153 ## ###
MXAN5189		NOA				F	RR	+5184 ##### +5189 #####
MXAN5340		NR				U	RR-GGDEF	##### -5340 #####
MXAN5505		NR				F	RR-DUF ⁹	##### +5505 #####
MXAN5592 (<i>digR</i> / <i>sgmW</i>)	(Overgaard <i>et al.</i> , 2006, Youderian & Hartzell,	NR	No	A, S	Yes	F	RR-HTH_3	##### +5592 #####

	2006)							
MXAN5656		NR				A1	RR	##### -5656 #####
MXAN5688		NR				U	RR	##### +5688 #####
MXAN5791		1.8xup /6-9				U	RR-GGDEF	##### +5791 #####
MXAN5889		NR				U	RR	##### +5889 #####
MXAN6012		2.5xup /0-2				F	RR	##### +6012 ## +6015 ##
MXAN6046		NR				A3	RR	##### +6046 #####
MXAN6099		NR				U	RR	##### -6099 #####
MXAN6296		1.8xdown /4-6				U	RR	##### +6296 #####
MXAN6620		NR				F	RR	##### +6620 #####
MXAN6627 (<i>sgnC</i>)	(Youderian & Hartzell, 2006)	NR	No	S	ND ^f	B3	RR-DUF ^g	##### -6627 #####
MXAN6926		1.8xup /2-4				F	RR	##### +6926 #####
MXAN6956 (<i>dotR</i>)	(Rasmussen & Sogaard- Andersen, 2003)	2.6xup/ 9-12	No	No	No	U	RR	+cheA +cheB +6953 # +6955 -6956 ## -cheB -cheR #
MXAN6968		NR				F	RR	-cheW -cheA -cheY -6966 # +6968 ## -6971 ##
MXAN7024		2.1xup /4-6				U	RR-PilZ	##### -7024 ## +7027 ##
MXAN7033		NR				B1	RR	##### -7033 #####
MXAN7150		1.9xup /12-15				F	RR	##### +7150 #####
MXAN7178		NR				F	RR	##### +7178 # -7180 ## #
MXAN7396		NR				U	RR-GGDEF	##### -7396 # -7398 ## #
MXAN7420		NR				B3	RR	##### +7420 #####
Paired genes								
MXAN0732		2.1xdown /0-2				U	RR	# -0728 ### -0732 -0733 ## -0736 ##
MXAN0937 (<i>nla7</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	## +0935 # -0937 -0938 #####
MXAN1078 (<i>spdR/nla19</i>)	(Hager <i>et al.</i> , 2001, Caberoy <i>et</i>	NR	No	S	Yes	A4	RR-Sigma54-HTH_8	##### +1077 +1078 #####

	<i>al.</i> , 2003)							
MXAN1128 (<i>frgC/nla25</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	##### -1128 -1129 #####
MXAN1167 (<i>nla28</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	Yes	A4	RR-Sigma54-HTH_8	##### +1166 +1167 #####
MXAN1189		NR				B1	RR-Sigma54-HTH_8	##### -1189 -1190 #####
MXAN1279		1.8xup /2-4				C3	RR-LytTR	##### -1279 -1280 #####
MXAN1349		NR				A1	RR-Trans_reg_C	##### -1349 -1350 #####
MXAN1552		1.7xup /6-9				B1	RR	##### -1552 -1553 #####
MXAN1680		NOA				A1	RR-Trans_reg_C	##### -1679 -1680 #####
MXAN2671		NR				F	RR-DUF ⁹	##### +2670 +2671 #####
MXAN2778 (<i>phoP2</i>)	(Moraleda-Munoz <i>et al.</i> , 2003)	NR	No	No	Yes	A1	RR-Trans_reg_C	##### -2778 -2779 #####
MXAN3418 (<i>nla17</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	##### -3418 -3419 #####
MXAN3450		1.6xdown /4-6				A1	RR-Trans_reg_C	##### +3450 +3451 #####
MXAN3605		NOA				B1	RR-DUF ⁹	##### -3605 -3606 #####
MXAN3811 (<i>nla13</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	##### -3811 -3812 #####
MXAN4042 (<i>nla6</i>)	(Caberoy <i>et al.</i> , 2003)	2.8xup /6-9	No	No	Yes	A4	RR-Sigma54-HTH_8	##### -4042 -4043 #####
MXAN4072		NR				C1	RR-GerE	##### +4071 +4072 #####
MXAN4164		NOA				A1	RR-Trans_reg_C	##### +4164 +4165 #####
MXAN4196		NR				A4	RR-Sigma54-HTH_8	##### -4196 -4197 ##### -4202
MXAN4261		NR				A4	RR-Sigma54-HTH_8	##### +4257 ##### -4261 -4262 #####
MXAN4580 (<i>nla8</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	No	No	A4	RR-Sigma54-HTH_8	##### +4579 +4580 #####
MXAN4777 (<i>phoP1</i>)	(Carrero-Lerida <i>et al.</i> , 2005)	NR	No	No	Yes	A1	RR-Trans_reg_C	##### +4777 +4778 #####
MXAN5083		NOA				C3	RR-LytTR	##### +5082 +5083 #####
MXAN5124 (<i>mrpB</i>)	(Sun & Shi, 2001a, Sun & Shi, 2001b)	NOA	No	No	Yes	F	RR-Sigma54-HTH_8	##### +5123 +5124 #####
MXAN5212		NR				A1	RR-Trans_reg_C	##### -5211 -5212 #####

MXAN5313		NR				A1	RR-Trans_reg_C	##### +5313 +5314 #####
MXAN5777 (<i>pilR2</i> , <i>nla23</i>)	(Caberoy <i>et al.</i> , 2003)	NOA	No	S	Yes	F	RR-Sigma54-HTH_8	##### -5777 -5778 #####
MXAN5784 (<i>pilR</i>)	(Wu & Kaiser, 1995)	NR	No	S	ND ^f	A4	RR-Sigma54-HTH_8	##### -5784 -5785 #####
MXAN5853 (<i>nla1</i>)	(Caberoy <i>et al.</i> , 2003)	NR	No	S	Yes	A4	RR-Sigma54-HTH_8	##### + 5852 +5853 #####
MXAN5995		NR				A1	RR-Trans_reg_C	-5990 ##### +5995 +5996 #####
MXAN6149		NR				C1	RR-GerE	##### -6149 -6150 #####
MXAN6224		NOA				A4	RR-HTH_8	##### +6223 +6224 #####
MXAN6413 (<i>phoP3</i>)	(Moraleda-Munoz <i>et al.</i> , 2001)	NR	No	No	Yes	A1	RR-Trans_reg_C	##### -6413 -6414 #####
MXAN6980		NR				A4	RR	##### +6979 +6980 #####
MXAN7143		NR				A4	RR-Sigma54-HTH_8	##### -7142 -7143 #####
MXAN7440 (<i>nla24</i>)	(Caberoy <i>et al.</i> , 2003, Lancero <i>et al.</i> , 2004)	NR	No	A, S	Yes	A4	RR-Sigma54-HTH_8	##### +7439 +7440 ### -7444 #

^a Expression ratios were calculated as the expression in developing cells over the expression in vegetative cells. The maximum expression ratios indicate the maximum expression ratio for a particular gene at all time points tested (0, 2, 4, 6, 9, 12, 15, 18 and 24 hrs). The induction time indicates the time interval in which the expression ratio began to change relative to that observed in vegetative cells.

^b NOA: Gene not represented or not detected on DNA microarray ; NR: Gene not significantly regulated in DNA microarray analyses.

^c See Grebe and Stock (Grebe & Stock, 1999) for a definition of different classes of HPK and RR.

^d All domains present in a protein are indicated. DUF indicates a domain of unknown function.

^e Genetic organization: HPK, RR and HPK-like protein encoding genes are indicated in red, blue and lavender, respectively. Genes are indicated by their MXAN numbers except in the cases of the gene clusters for which the che or established nomenclature was adopted. Transcription using the lower or upper strand as template is indicated by + and -, respectively. # indicates a gene not encoding a TCS protein. With the exception of the che gene clusters, gene organization is shown for 5 genes upstream and downstream from the query (indicated in italic).

^f NA, not applicable; ND, not determined.

^g For RR with DUF domains the presence of this domain in other species is indicated.

7 References

- Alm, E., K. Huang & A. Arkin, (2006) The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS Comp. Biol.* 2: e143.
- Alves, R. & M. A. Savageau, (2003) Comparative analysis of prototype two-component systems with either bifunctional or monofunctional sensors: differences in molecular structure and physiological function. *Mol Microbiol* 48: 25-51.
- Aravind, L. & C. P. Ponting, (1997) The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem Sci* 22: 458-459.
- Besant, P. G. & P. V. Attwood, (2005) Mammalian histidine kinases. *Biochim Biophys Acta* 1754: 281-290.
- Bhandari, P. & J. Gowrishankar, (1997) An Escherichia coli host strain useful for efficient overproduction of cloned gene products with NaCl as the inducer. *J Bacteriol* 179: 4403-4406.
- Bibikov, S. I., R. Biran, K. E. Rudd & J. S. Parkinson, (1997) A signal transducer for aerotaxis in Escherichia coli. *J Bacteriol* 179: 4075-4079.
- Bilwes, A. M., L. A. Alex, B. R. Crane & M. I. Simon, (1999) Structure of CheA, a signal-transducing histidine kinase. *Cell* 96: 131-141.
- Biondi, E. G., S. J. Reisinger, J. M. Skerker, M. Arif, B. S. Perchuk, K. R. Ryan & M. T. Laub, (2007) Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* 444: 899 - 904.
- Birck, C., M. Malfois, D. Svergun & J. Samama, (2002) Insights into signal transduction revealed by the low resolution structure of the FixJ response regulator. *J Mol Biol* 321: 447-457.
- Birck, C., L. Mourey, P. Gouet, B. Fabry, J. Schumacher, P. Rousseau, D. Kahn & J. P. Samama, (1999) Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure* 7: 1505-1515.
- Black, W. P. & Z. Yang, (2004) Myxococcus xanthus chemotaxis homologs DifD and DifG negatively regulate fibril polysaccharide production. *J Bacteriol* 186: 1001-1008.
- Boysen, A., E. Ellehauge, B. Julien & L. Sogaard-Andersen, (2002) The DevT protein stimulates synthesis of FruA, a signal transduction protein required for fruiting body morphogenesis in Myxococcus xanthus. *J Bacteriol* 184: 1540-1546.
- Bruder, S., J. U. Linder, S. E. Martinez, N. Zheng, J. A. Beavo & J. E. Schultz, (2005) The cyanobacterial tandem GAF domains from the cyaB2 adenylyl cyclase signal via both cAMP-binding sites. *Proc Natl Acad Sci U S A* 102: 3088-3092.
- Burbulys, D., K. A. Trach & J. A. Hoch, (1991) Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay. *Cell* 64: 545-552.
- Bustamante, V. H., I. Martinez-Flores, H. C. Vlamakis & D. R. Zusman, (2004) Analysis of the Frz signal transduction system of Myxococcus xanthus shows the importance of the conserved C-terminal region of the cytoplasmic chemoreceptor FrzCD in sensing signals. *Mol Microbiol* 53: 1501-1513.
- Caberoy, N. B., R. D. Welch, J. S. Jakobsen, S. C. Slater & A. G. Garza, (2003) Global mutational analysis of NtrC-like activators in Myxococcus xanthus: identifying activator mutants defective for motility and fruiting body development. *J Bacteriol* 185: 6083-6094.
- Carrero-Lerida, J., A. Moraleda-Munoz, R. Garcia-Hernandez, J. Perez & J. Munoz-Dorado, (2005) PhoR1-PhoP1, a third two-component system of the family PhoRP from Myxococcus xanthus: role in development. *J Bacteriol* 187: 4976-4983.

- Cheng, Y. & D. Kaiser, (1989) Dsg, a Gene Required for Cell-Cell Interaction Early in Myxococcus Development. *Journal of Bacteriology* 171: 3719-3726.
- Cho, K., A. Treuner-Lange, K. A. O'Connor & D. R. Zusman, (2000) Developmental aggregation of Myxococcus xanthus requires frgA, an frz-related gene. *J Bacteriol* 182: 6614-6621.
- Cho, K. & D. R. Zusman, (1999a) AsgD, a new two-component regulator required for A-signalling and nutrient sensing during early development of Myxococcus xanthus. *Mol Microbiol* 34: 268-281.
- Cho, K. & D. R. Zusman, (1999b) Sporulation timing in Myxococcus xanthus is controlled by the espAB locus. *Mol Microbiol* 34: 714-725.
- Christie, J. M., M. Salomon, K. Nozue, M. Wada & W. R. Briggs, (1999) LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. *Proc Natl Acad Sci U S A* 96: 8779-8783.
- D'Autreaux, B., N. P. Tucker, R. Dixon & S. Spiro, (2005) A non-haem iron centre in the transcription factor NorR senses nitric oxide. *Nature* 437: 769-772.
- Downard, J., S. V. Ramaswamy & K. S. Kil, (1993) Identification of esg, a genetic locus involved in cell-cell signaling during Myxococcus xanthus development. *J Bacteriol* 175: 7762-7770.
- Duclos, B., S. Marcandier & A. J. Cozzone, (1991) Chemical properties and separation of phosphoamino acids by thin-layer chromatography and/or electrophoresis. *Methods Enzymol* 201: 10-21.
- Dutta, R., L. Qin & M. Inouye, (1999) Histidine kinases: diversity of domain organization. *Mol Microbiol* 34: 633-640.
- Dworkin, M., (1996) Recent advances in the social and developmental biology of the myxobacteria. *Microbiological Reviews* 60: 70-&.
- Ellehaug, E., (1999) Identification and characterisation of FruA: a central component in the C-signalling pathway in Myxococcus xanthus. *PhD thesis*.
- Ellehaug, E., M. Norregaard-Madsen & L. Sogaard-Andersen, (1998) The FruA signal transduction protein provides a checkpoint for the temporal co-ordination of intercellular signals in Myxococcus xanthus development. *Mol Microbiol* 30: 807-817.
- Fabret, C., V. A. Feher & J. A. Hoch, (1999) Two-component signal transduction in Bacillus subtilis: how one organism sees its world. *J Bacteriol* 181: 1975-1983.
- Falke, J. J., R. B. Bass, S. L. Butler, S. A. Chervitz & M. A. Danielson, (1997) The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu Rev Cell Dev Biol* 13: 457-512.
- Finn, R., J. Mistry, B. Schuster-Böckler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. L. Sonnhammer & A. Bateman, (2006) Pfam: clans, web tools and services. *Nucl. Acids Res. Database Issue* 34: D247-D251.
- Fisher, S. L., W. Jiang, B. L. Wanner & C. T. Walsh, (1995) Cross-talk between the histidine protein kinase VanS and the response regulator PhoB. Characterization and identification of a VanS domain that inhibits activation of PhoB. *J Biol Chem* 270: 23143-23149.
- Fisher, S. L., S. K. Kim, B. L. Wanner & C. T. Walsh, (1996) Kinetic comparison of the specificity of the vancomycin resistance VanS for two response regulators, VanR and PhoB. *Biochemistry* 35: 4732-4740.
- Galperin, M. Y., (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. *BMC Microbiol* 5: 35.
- Galperin, M. Y., (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J Bacteriol* 188: 4169-4182.

- Garza, A. G., J. S. Pollack, B. Z. Harris, A. Lee, I. M. Keseler, E. F. Licking & M. Singer, (1998) SdeK is required for early fruiting body development in *Myxococcus xanthus*. *J Bacteriol* 180: 4628-4637.
- Gill, R. E. & M. G. Cull, (1986) Control of developmental gene expression by cell-to-cell interactions in *Myxococcus xanthus*. *J Bacteriol* 168: 341-347.
- Goldman, B. S., W. C. Nierman, D. Kaiser, S. C. Slater, A. S. Durkin, J. A. Eisen, C. M. Ronning, W. B. Barbazuk, M. Blanchard, C. Field, C. Halling, G. Hinkle, O. Iartchuk, H. S. Kim, C. Mackenzie, R. Madupu, N. Miller, A. Shvartsbeyn, S. A. Sullivan, M. Vaudin, R. Wiegand & H. B. Kaplan, (2006) Evolution of sensory complexity recorded in a myxobacterial genome. *Proc Natl Acad Sci U S A* 103: 15200-15205.
- Gong, W., B. Hao, S. S. Mansy, G. Gonzalez, M. A. Gilles-Gonzalez & M. K. Chan, (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc Natl Acad Sci U S A* 95: 15177-15182.
- Gouet, P., B. Fabry, V. Guillet, C. Birck, L. Mourey, D. Kahn & J. P. Samama, (1999) Structural transitions in the FixJ receiver domain. *Structure* 7: 1517-1526.
- Grebe, T. W. & J. B. Stock, (1999) The histidine protein kinase superfamily. *Adv Microb Physiol* 41: 139-227.
- Grimshaw, C. E., S. Huang, C. G. Hanstein, M. A. Strauch, D. Burbulys, L. Wang, J. A. Hoch & J. M. Whiteley, (1998) Synergistic kinetic interactions between components of the phosphorelay controlling sporulation in *Bacillus subtilis*. *Biochemistry* 37: 1365-1375.
- Gronewold, T. M. & D. Kaiser, (2001) The act operon controls the level and time of C-signal production for *Myxococcus xanthus* development. *Mol Microbiol* 40: 744-756.
- Guo, D., Y. Wu & H. B. Kaplan, (2000) Identification and characterization of genes required for early *Myxococcus xanthus* developmental gene expression. *J Bacteriol* 182: 4564-4571.
- Hagen, D. C., A. P. Bretscher & D. Kaiser, (1978) Synergism between morphogenetic mutants of *Myxococcus xanthus*. *Dev Biol* 64: 284-296.
- Hagen, K. D. & J. C. Meeks, (1999) Biochemical and genetic evidence for participation of DevR in a phosphorelay signal transduction pathway essential for heterocyst maturation in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* 181: 4430-4434.
- Hager, E., H. Tse & R. E. Gill, (2001) Identification and characterization of spdR mutations that bypass the BsgA protease-dependent regulation of developmental gene expression in *Myxococcus xanthus*. *Mol Microbiol* 39: 765-780.
- Haldimann, A., S. L. Fisher, L. L. Daniels, C. T. Walsh & B. L. Wanner, (1997) Transcriptional regulation of the *Enterococcus faecium* BM4147 vancomycin resistance gene cluster by the VanS-VanR two-component regulatory system in *Escherichia coli* K-12. *J Bacteriol* 179: 5903-5913.
- Haldimann, A., M. K. Prahalad, S. L. Fisher, S. K. Kim, C. T. Walsh & B. L. Wanner, (1996) Altered recognition mutants of the response regulator PhoB: a new genetic strategy for studying protein-protein interactions. *Proc Natl Acad Sci U S A* 93: 14361-14366.
- Harris, B. Z., D. Kaiser & M. Singer, (1998) The guanosine nucleotide (p)ppGpp initiates development and A-factor production in *myxococcus xanthus*. *Genes Dev* 12: 1022-1035.
- Henke, J. M. & B. L. Bassler, (2004) Three parallel quorum-sensing systems regulate gene expression in *Vibrio harveyi*. *J Bacteriol* 186: 6902-6914.
- Higgs, P. I., K. Cho, D. E. Whitworth, L. S. Evans & D. R. Zusman, (2005) Four unusual two-component signal transduction homologs, RedC to RedF, are necessary for timely development in *Myxococcus xanthus*. *J Bacteriol* 187: 8191-8195.
- Ho, Y. S., L. M. Burden & J. H. Hurley, (2000) Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *EMBO J* 19: 5288-5299.

- Hodgkin, J. & D. Kaiser, (1977) Cell-to-cell stimulation of movement in nonmotile mutants of *Myxococcus*. *Proc Natl Acad Sci U S A* 74: 2938-2942.
- Hodgkin, J. & D. Kaiser, (1979) Genetics of Gliding Motility in *Myxococcus-Xanthus* (Myxobacteriales) - 2 Gene Systems Control Movement. *Molecular & General Genetics* 171: 177-191.
- Horiuchi, T., T. Akiyama, S. Inouye & T. Komano, (2002a) Analysis of *dofA*, a *fruA*-dependent developmental gene, and its homologue, *dofB*, in *Myxococcus xanthus*. *J Bacteriol* 184: 6803-6810.
- Horiuchi, T., M. Taoka, T. Isobe, T. Komano & S. Inouye, (2002b) Role of *fruA* and *csgA* genes in gene expression during development of *Myxococcus xanthus*. Analysis by two-dimensional gel electrophoresis. *J Biol Chem* 277: 26753-26760.
- Hulko, M., F. Berndt, M. Gruber, J. U. Linder, V. Truffault, A. Schultz, J. Martin, J. E. Schultz, A. N. Lupas & M. Coles, (2006) The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell* 126: 929-940.
- Igo, M. M., A. J. Ninfa, J. B. Stock & T. J. Silhavy, (1989) Phosphorylation and dephosphorylation of a bacterial transcriptional activator by a transmembrane receptor. *Genes Dev* 3: 1725-1734.
- Inclan, Y. F., S. Laurent & D. R. Zusman, (2008) The receiver domain of FrzE, a CheA-CheY fusion protein, regulates the CheA histidine kinase activity and downstream signalling to the A- and S-motility systems of *Myxococcus xanthus*. *Mol Microbiol* 68: 1328-1339.
- Inclan, Y. F., H. C. Vlamakis & D. R. Zusman, (2007) FrzZ, a dual CheY-like response regulator, functions as an output for the Frz chemosensory pathway of *Myxococcus xanthus*. *Mol Microbiol* 65: 90-102.
- Ishizuka, T., T. Shimada, K. Okajima, S. Yoshihara, Y. Ochiai, M. Katayama & M. Ikeuchi, (2006) Characterization of cyanobacteriochrome TePixJ from a thermophilic cyanobacterium *Thermosynechococcus elongatus* strain BP-1. *Plant Cell Physiol* 47: 1251-1261.
- Jelsbak, L. & L. Sogaard-Andersen, (1999) The cell surface-associated intercellular C-signal induces behavioral changes in individual *Myxococcus xanthus* cells during fruiting body morphogenesis. *Proc Natl Acad Sci U S A* 96: 5031-5036.
- Jelsbak, L. & L. Sogaard-Andersen, (2000) Pattern formation: fruiting body morphogenesis in *Myxococcus xanthus*. *Curr Opin Microbiol* 3: 637-642.
- Jiang, M., W. Shao, M. Perego & J. A. Hoch, (2000) Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol Microbiol* 38: 535-542.
- Julien, B., A. D. Kaiser & A. Garza, (2000) Spatial control of cell differentiation in *Myxococcus xanthus*. *Proc Natl Acad Sci U S A* 97: 9098-9103.
- Kaiser, D., (1979) Social gliding is correlated with the presence of pili in *Myxococcus xanthus*. *Proc Natl Acad Sci U S A* 76: 5952-5956.
- Kashefi, K. & P. L. Hartzell, (1995) Genetic suppression and phenotypic masking of a *Myxococcus xanthus frzF* defect. *Mol. Microbiol.* 15: 483-494.
- Kim, S. K. & D. Kaiser, (1990) Cell motility is required for the transmission of C-factor, an intercellular signal that coordinates fruiting body morphogenesis of *Myxococcus xanthus*. *Genes Dev* 4: 896-904.
- Kimura, Y., H. Nakano, H. Terasaka & K. Takegawa, (2001) *Myxococcus xanthus mokA* encodes a histidine kinase-response regulator hybrid sensor required for development and osmotic tolerance. *J Bacteriol* 183: 1140-1146.
- Kirby, J. R. & D. R. Zusman, (2003) Chemosensory regulation of developmental gene expression in *Myxococcus xanthus*. *Proc Natl Acad Sci U S A* 100: 2008-2013.

- Kroos, L. & D. Kaiser, (1987) Expression of many developmentally regulated genes in *Myxococcus* depends on a sequence of cell interactions. *Genes Dev* 1: 840-854.
- Kruse, T., S. Lobedanz, N. M. Berthelsen & L. Sogaard-Andersen, (2001) C-signal: a cell surface-associated morphogen that induces and co-ordinates multicellular fruiting body morphogenesis and sporulation in *Myxococcus xanthus*. *Mol Microbiol* 40: 156-168.
- Kuner, J. M. & D. Kaiser, (1982) Fruiting body morphogenesis in submerged cultures of *Myxococcus xanthus*. *J Bacteriol* 151: 458-461.
- Kuspa, A., L. Kroos & D. Kaiser, (1986) Intercellular signaling is required for developmental gene expression in *Myxococcus xanthus*. *Dev Biol* 117: 267-276.
- Kuspa, A., L. Plamann & D. Kaiser, (1992) A-Signaling and the Cell-Density Requirement for *Myxococcus-Xanthus* Development. *Journal of Bacteriology* 174: 7360-7369.
- Lancero, H., N. B. Caberoy, S. Castaneda, Y. Li, A. Lu, D. Dutton, X. Y. Duan, H. B. Kaplan, W. Shi & A. G. Garza, (2004) Characterization of a *Myxococcus xanthus* mutant that is defective for adventurous motility and social motility. *Microbiology* 150: 4085-4093.
- Laub, M. T. & M. Goulian, (2007) Specificity in two-component signal transduction pathways. *Annu Rev Genet* 41: 121-145.
- Lee, B., P. I. Higgs, D. R. Zusman & K. Cho, (2005) EspC is involved in controlling the timing of development in *Myxococcus xanthus*. *J Bacteriol* 187: 5029-5031.
- Leonardy, S., G. Freymark, S. Hebenner, E. Ellehaug & L. Sogaard-Andersen, (2007) Coupling of protein localization and cell movements by a dynamically localized response regulator in *Myxococcus xanthus*. *Embo J* 26: 4433-4444.
- Letunic, I., R. R. Copley, S. Schmidt, F. D. Ciccarelli, T. Doerks, J. Schultz, C. P. Ponting & P. Bork, (2004) SMART 4.0: towards genomic data integration. 32 Database issue: D142-D144.
- Li, J., R. V. Swanson, M. I. Simon & R. M. Weis, (1995) The response regulators CheB and CheY exhibit competitive binding to the kinase CheA. *Biochemistry* 34: 14626-14636.
- Li, S., A. Ault, C. L. Malone, D. Raitt, S. Dean, L. H. Johnston, R. J. Deschenes & J. S. Fassler, (1998) The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. *EMBO J* 17: 6952-6962.
- Li, Y., V. H. Bustamante, R. Lux, D. Zusman & W. Shi, (2005) Divergent regulatory pathways control A and S motility in *Myxococcus xanthus* through FrzE, a CheA-CheY fusion protein. *J Bacteriol* 187: 1716-1723.
- Licking, E., L. Gorski & D. Kaiser, (2000) A common step for changing cell shape in fruiting body and starvation-independent sporulation of *Myxococcus xanthus*. *J Bacteriol* 182: 3553-3558.
- Little, R. & R. Dixon, (2003) The amino-terminal GAF domain of *Azotobacter vinelandii* NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions. *J Biol Chem* 278: 28711-28718.
- Lobedanz, S. & L. Sogaard-Andersen, (2003) Identification of the C-signal, a contact-dependent morphogen coordinating multiple developmental responses in *Myxococcus xanthus*. *Genes Dev* 17: 2151-2161.
- Lukat, G. S., B. H. Lee, J. M. Mottonen, A. M. Stock & J. B. Stock, (1991) Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. *J Biol Chem* 266: 8348-8354.
- Maeda, S., C. Sugita, M. Sugita & T. Omata, (2006) A new class of signal transducer in His-Asp phosphorelay systems. *J Biol Chem* 281: 37868-37876.

- Manoil, C. & D. Kaiser, (1980) Guanosine pentaphosphate and guanosine tetraphosphate accumulation and induction of *Myxococcus xanthus* fruiting body development. *J Bacteriol* 141: 305-315.
- Marina, A., C. Mott, A. Auyzenberg, W. A. Hendrickson & C. D. Waldburger, (2001) Structural and mutational analysis of the PhoQ histidine kinase catalytic domain. Insight into the reaction mechanism. *J Biol Chem* 276: 41182-41190.
- Marina, A., C. D. Waldburger & W. A. Hendrickson, (2005) Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. *EMBO J* 24: 4247-4259.
- McCleary, W. R., M. J. McBride & D. R. Zusman, (1990) Developmental sensory transduction in *Myxococcus xanthus* involves methylation and demethylation of FrzCD. *J Bacteriol* 172: 4877-4887.
- McCleary, W. R. & D. R. Zusman, (1990) FrzE of *Myxococcus xanthus* is homologous to both CheA and CheY of *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* 87: 5898-5902.
- Mignot, T., J. P. Merlie, Jr. & D. R. Zusman, (2005) Regulated pole-to-pole oscillations of a bacterial gliding motility protein. *Science* 310: 855-857.
- Mignot, T., J. W. Shaevitz, P. L. Hartzell & D. R. Zusman, (2007) Evidence that focal adhesion complexes power bacterial gliding motility. *Science* 315: 853-856.
- Morais Cabral, J. H., A. Lee, S. L. Cohen, B. T. Chait, M. Li & R. Mackinnon, (1998) Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain. *Cell* 95: 649-655.
- Moraleda-Munoz, A., J. Carrero-Lerida, A. L. Extremera, J. M. Arias & J. Munoz-Dorado, (2001) Glycerol 3-phosphate inhibits swarming and aggregation of *Myxococcus xanthus*. *J Bacteriol* 183: 6135-6139.
- Moraleda-Munoz, A., J. Carrero-Lerida, J. Perez & J. Munoz-Dorado, (2003) Role of two novel two-component regulatory systems in development and phosphatase expression in *Myxococcus xanthus*. *J Bacteriol* 185: 1376-1383.
- Nariya, H. & S. Inouye, (2005) Modulating factors for the Pkn4 kinase cascade in regulating 6-phosphofructokinase in *Myxococcus xanthus*. *Mol Microbiol* 56: 1314-1328.
- Nariya, H. & S. Inouye, (2006) A protein Ser/Thr kinase cascade negatively regulates the DNA-binding activity of MrpC, a smaller form of which may be necessary for the *Myxococcus xanthus* development. *Mol Microbiol* 60: 1205-1217.
- Ogawa, M., S. Fujitani, X. Mao, S. Inouye & T. Komano, (1996) FruA, a putative transcription factor essential for the development of *Myxococcus xanthus*. *Mol Microbiol* 22: 757-767.
- Overgaard, M., S. Wegener-Feldbrugge & L. Sogaard-Andersen, (2006) The orphan response regulator DigR is required for synthesis of extracellular matrix fibrils in *Myxococcus xanthus*. *J Bacteriol* 188: 4384-4394.
- Pellequer, J. L., K. A. Wager-Smith, S. A. Kay & E. D. Getzoff, (1998) Photoactive yellow protein: a structural prototype for the three-dimensional fold of the PAS domain superfamily. *Proc Natl Acad Sci U S A* 95: 5884-5890.
- Perraud, A. L., B. Kimmel, V. Weiss & R. Gross, (1998) Specificity of the BvgAS and EvgAS phosphorelay is mediated by the C-terminal HPT domains of the sensor proteins. *Mol Microbiol* 27: 875-887.
- Plamann, L., Y. Li, B. Cantwell & J. Mayor, (1995) The *Myxococcus xanthus* *asgA* gene encodes a novel signal transduction protein required for multicellular development. *J Bacteriol* 177: 2014-2020.
- Pollack, J. S. & M. Singer, (2001) SdeK, a histidine kinase required for *Myxococcus xanthus* development. *J Bacteriol* 183: 3589-3596.

- Porter, S. L., G. H. Wadhams & J. P. Armitage, (2007) In vivo and in vitro analysis of the Rhodobacter sphaeroides chemotaxis signaling complexes. *Methods Enzymol* 423: 392-413.
- Rasmussen, A. A., (1998) An analysis of the importance of a surface-exposed eight amino acid loop located downstream from Asp59 in FruA for the activity of the protein in vivo in *M. xanthus*. *Bachelor thesis*.
- Rasmussen, A. A., S. L. Porter, J. P. Armitage & L. Sogaard-Andersen, (2005) Coupling of multicellular morphogenesis and cellular differentiation by an unusual hybrid histidine protein kinase in *Myxococcus xanthus*. *Mol Microbiol* 56: 1358-1372.
- Rasmussen, A. A. & L. Sogaard-Andersen, (2003) TodK, a putative histidine protein kinase, regulates timing of fruiting body morphogenesis in *Myxococcus xanthus*. *J Bacteriol* 185: 5452-5464.
- Rasmussen, A. A., S. Wegener-Feldbrugge, S. L. Porter, J. P. Armitage & L. Sogaard-Andersen, (2006) Four signalling domains in the hybrid histidine protein kinase RodK of *Myxococcus xanthus* are required for activity. *Mol Microbiol* 60: 525-534.
- Repik, A., A. Rebbapragada, M. S. Johnson, J. O. Haznedar, I. B. Zhulin & B. L. Taylor, (2000) PAS domain residues involved in signal transduction by the Aer redox sensor of *Escherichia coli*. *Mol Microbiol* 36: 806-816.
- Rodrigue, A., Y. Quentin, A. Lazdunski, V. Mejean & M. Foglino, (2000) Two-component systems in *Pseudomonas aeruginosa*: why so many? *Trends Microbiol* 8: 498-504.
- Rolbetzki, A., M. Ammon, V. Jakovljevic, A. Konovalova & L. & Sogaard-Andersen, (2008) Regulated secretion of a protease activates intercellular signaling during fruiting body formation in *M. xanthus*. *Dev. Cell*. In press.
- Romling, U. & D. Amikam, (2006) Cyclic di-GMP as a second messenger. *Curr Opin Microbiol* 9: 218-228.
- Rybalkin, S. D., I. G. Rybalkina, M. Shimizu-Albergine, X. B. Tang & J. A. Beavo, (2003) PDE5 is converted to an activated state upon cGMP binding to the GAF A domain. *EMBO J* 22: 469-478.
- Ryjenkov, D. A., R. Simm, U. Romling & M. Gomelsky, (2006) The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. *J Biol Chem* 281: 30310-30314.
- Saitou, N. & M. Nei, (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425.
- Sambrook, J. & D. Russell, (2000) Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*.
- Sardiwal, S., S. L. Kendall, F. Movahedzadeh, S. C. Rison, N. G. Stoker & S. Djordjevic, (2005) A GAF domain in the hypoxia/NO-inducible Mycobacterium tuberculosis DosS protein binds haem. *J Mol Biol* 353: 929-936.
- Scott, A. E., E. Simon, S. K. Park, P. Andrews & D. R. Zusman, (2008) Site-specific receptor methylation of FrzCD in *Myxococcus xanthus* is controlled by a tetra-trico peptide repeat (TPR) containing regulatory domain of the FrzF methyltransferase. *Mol Microbiol*.
- Shi, X., S. Wegener-Feldbrugge, S. Huntley, N. Hamann, R. Hedderich & L. Sogaard-Andersen, (2008) Bioinformatics and experimental analysis of proteins of two-component systems in *Myxococcus xanthus*. *J Bacteriol* 190: 613-624.
- Shimkets, L. J., (1990) Social and Developmental Biology of the Myxobacteria. *Microbiological Reviews* 54: 473-501.
- Shimkets, L. J. & D. Kaiser, (1982) Induction of coordinated movement of *Myxococcus xanthus* cells. *J Bacteriol* 152: 451-461.

- Skerker, J. M., B. S. Perchuk, A. Siryaporn, E. A. Lubin, O. Ashenberg, M. Goulian & M. T. Laub, (2008) Rewiring the specificity of two-component signal transduction systems. *Cell* 133: 1043-1054.
- Skerker, J. M., M. S. Prasol, B. S. Perchuk, E. G. Biondi & M. T. Laub, (2005) Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol* 3: e334.
- Sogaard-Andersen, L. & D. Kaiser, (1996) C factor, a cell-surface-associated intercellular signaling protein, stimulates the cytoplasmic Frz signal transduction system in *Myxococcus xanthus*. *Proc Natl Acad Sci U S A* 93: 2675-2679.
- Sogaard-Andersen, L., F. J. Slack, H. Kimsey & D. Kaiser, (1996) Intercellular C-signaling in *Myxococcus xanthus* involves a branched signal transduction pathway. *Genes Dev* 10: 740-754.
- Sonnhammer, E. L., G. von Heijne & A. Krogh, (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 6: 175-182.
- Spratt, B. G., P. J. Hedge, S. te Heesen, A. Edelman & J. K. Broome-Smith, (1986) Kanamycin-resistant vectors that are analogues of plasmids pUC8, pUC9, pEMBL8 and pEMBL9. *Gene* 41: 337-342.
- Stock, A. M., V. L. Robinson & P. N. Goudreau, (2000) Two-component signal transduction. *Annu Rev Biochem* 69: 183-215.
- Sun, H. & W. Shi, (2001a) Genetic studies of *mrp*, a locus essential for cellular aggregation and sporulation of *Myxococcus xanthus*. *J Bacteriol* 183: 4786-4795.
- Sun, H. & W. Y. Shi, (2001b) Analyses of *mrp* genes during *Myxococcus xanthus* development. *Journal of Bacteriology* 183: 6733-6739.
- Szurmant, H., R. A. White & J. A. Hoch, (2007) Sensor complexes regulating two-component signal transduction. *Curr Opin Struct Biol* 17: 706-715.
- Taylor, B. L., (2007) Aer on the inside looking out: paradigm for a PAS-HAMP role in sensing oxygen, redox and energy. *Mol Microbiol* 65: 1415-1424.
- Terwilliger, T. C., E. Bogonez, E. A. Wang & D. E. Koshland, Jr., (1983) Sites of methyl esterification on the aspartate receptor involved in bacterial chemotaxis. *J. Biol. Chem.* 258: 9608-9611.
- Terwilliger, T. C. & D. E. Koshland, Jr., (1984) Sites of methyl esterification and deamination on the aspartate receptor involved in chemotaxis. *J. Biol. Chem.* 259: 7719-7725.
- Terwilliger, T. C., J. Y. Wang & D. E. Koshland, Jr., (1986) Kinetics of receptor modification. The multiply methylated aspartate receptors involved in bacterial chemotaxis. *J. Biol. Chem.* 261: 10814-10820.
- Thomasson, B., J. Link, A. G. Stassinopoulos, N. Burke, L. Plamann & P. L. Hartzell, (2002) MglA, a small GTPase, interacts with a tyrosine kinase to control type IV pili-mediated motility and development of *Myxococcus xanthus*. *Mol Microbiol* 46: 1399-1413.
- Thompson, J. D., D. G. Higgins & T. J. Gibson, (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22: 4673-4680.
- Thony-Meyer, L. & D. Kaiser, (1993) *devRS*, an autoregulated and essential genetic locus for fruiting body development in *Myxococcus xanthus*. *J Bacteriol* 175: 7450-7462.
- Tomomori, C., T. Tanaka, R. Dutta, H. Park, S. K. Saha, Y. Zhu, R. Ishima, D. Liu, K. I. Tong, H. Kurokawa, H. Qian, M. Inouye & M. Ikura, (1999) Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. *Nat Struct Biol* 6: 729-734.

- Trudeau, K. G., M. J. Ward & D. R. Zusman, (1996) Identification and characterization of FrzZ, a novel response regulator necessary for swarming and fruiting-body formation in *Myxococcus xanthus*. *Mol Microbiol* 20: 645-655.
- Ueki, T. & S. Inouye, (2002) Transcriptional activation of a heat-shock gene, lonD, of *Myxococcus xanthus* by a two component histidine-aspartate phosphorelay system. *J Biol Chem* 277: 6170-6177.
- Ueki, T. & S. Inouye, (2003) Identification of an activator protein required for the induction of fruA, a gene essential for fruiting body development in *Myxococcus xanthus*. *Proc Natl Acad Sci U S A* 100: 8782-8787.
- Ueki, T. & S. Inouye, (2005a) Activation of a development-specific gene, dofA, by FruA, an essential transcription factor for development of *Myxococcus xanthus*. *Journal of Bacteriology* 187: 8504-8506.
- Ueki, T. & S. Inouye, (2005b) Identification of a gene involved in polysaccharide export as a transcription target of FruA, an essential factor for *Myxococcus xanthus* development. *J Biol Chem* 280: 32279-32284.
- Ueki, T. & S. Inouye, (2006) A novel regulation on developmental gene expression of fruiting body formation in Myxobacteria. *Appl Microbiol Biot* 72: 21-29.
- Utsumi, R., R. E. Brissette, A. Rampersaud, S. A. Forst, K. Oosawa & M. Inouye, (1989) Activation of bacterial porin gene expression by a chimeric signal transducer in response to aspartate. *Science* 245: 1246-1249.
- Viswanathan, P., K. Murphy, B. Julien, A. G. Garza & L. Kroos, (2007a) Regulation of dev, an Operon That Includes Genes Essential for *Myxococcus xanthus* Development and CRISPR-Associated Genes and Repeats. *J Bacteriol* 189: 3738-3750.
- Viswanathan, P., T. Ueki, S. Inouye & L. Kroos, (2007b) Combinatorial regulation of genes essential for *Myxococcus xanthus* development involves a response regulator and a LysR-type regulator. *Proc Natl Acad Sci U S A*.
- Vlamakis, H. C., J. R. Kirby & D. R. Zusman, (2004) The Che4 pathway of *Myxococcus xanthus* regulates type IV pilus-mediated motility. *Mol Microbiol* 52: 1799-1811.
- Wadhams, G. H. & J. P. Armitage, (2004) Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol* 5: 1024-1037.
- Ward, M. J., H. Lew & D. R. Zusman, (2000) Social motility in *Myxococcus xanthus* requires FrzS, a protein with an extensive coiled-coil domain. *Mol Microbiol* 37: 1357-1371.
- Ward, M. J. & D. R. Zusman, (1999) Motility in *Myxococcus xanthus* and its role in developmental aggregation. *Curr Opin Microbiol* 2: 624-629.
- West, A. H. & A. M. Stock, (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* 26: 369-376.
- White, R. A., H. Szurmant, J. A. Hoch & T. Hwa, (2007) Features of protein-protein interactions in two-component signaling deduced from genomic libraries. *Methods Enzymol* 422: 75-101.
- Whitworth, D. E. & P. J. Cock, (2008) Two-component systems of the myxobacteria: structure, diversity and evolutionary relationships. *Microbiology* 154: 360-372.
- Whitworth, D. E., A. Millard, D. A. Hodgson & P. F. Hawkins, (2008) Protein-protein interactions between two-component system transmitter and receiver domains of *Myxococcus xanthus*. *Proteomics* 8: 1839-1842.
- Wolgemuth, C., E. Hoiczky, D. Kaiser & G. Oster, (2002) How myxobacteria glide. *Curr Biol* 12: 369-377.

- Wu, J., N. Ohta, J. L. Zhao & A. Newton, (1999) A novel bacterial tyrosine kinase essential for cell division and differentiation. *Proc Natl Acad Sci U S A* 96: 13068-13073.
- Wu, S. S. & D. Kaiser, (1995) Genetic and functional evidence that Type IV pili are required for social gliding motility in *Myxococcus xanthus*. *Mol Microbiol* 18: 547-558.
- Wu, S. S., J. Wu, Y. L. Cheng & D. Kaiser, (1998) The pilH gene encodes an ABC transporter homologue required for type IV pilus biogenesis and social gliding motility in *Myxococcus xanthus*. *Mol Microbiol* 29: 1249-1261.
- Yamada, S., S. Akiyama, H. Sugimoto, H. Kumita, K. Ito, T. Fujisawa, H. Nakamura & Y. Shiro, (2006) The signaling pathway in histidine kinase and the response regulator complex revealed by X-ray crystallography and solution scattering. *J Mol Biol* 362: 123-139.
- Yamamoto, K., K. Hirao, T. Oshima, H. Aiba, R. Utsumi & A. Ishihama, (2005) Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *J Biol Chem* 280: 1448-1456.
- Yang, C. & H. B. Kaplan, (1997) *Myxococcus xanthus* sasS encodes a sensor histidine kinase required for early developmental gene expression. *J Bacteriol* 179: 7759-7767.
- Yang, R., S. Bartle, R. Otto, A. Stassinopoulos, M. Rogers, L. Plamann & P. Hartzell, (2004) AglZ is a filament-forming coiled-coil protein required for adventurous gliding motility of *Myxococcus xanthus*. *J Bacteriol* 186: 6168-6178.
- Yang, Z., Y. Geng, D. Xu, H. B. Kaplan & W. Shi, (1998) A new set of chemotaxis homologues is essential for *Myxococcus xanthus* social motility. *Mol Microbiol* 30: 1123-1130.
- Yoder-Himes, D. R. & L. Kroos, (2006) Regulation of the *Myxococcus xanthus* C-signal-dependent Omega4400 promoter by the essential developmental protein FruA. *J Bacteriol* 188: 5167-5176.
- Yoshihara, S., M. Katayama, X. Geng & M. Ikeuchi, (2004) Cyanobacterial phytochrome-like PixJ1 holoprotein shows novel reversible photoconversion between blue- and green-absorbing forms. *Plant Cell Physiol* 45: 1729-1737.
- Youderian, P. & P. L. Hartzell, (2006) Transposon insertions of magellan-4 that impair social gliding motility in *Myxococcus xanthus*. *Genetics* 172: 1397-1410.
- Zapf, J., U. Sen, Madhusudan, J. A. Hoch & K. I. Varughese, (2000) A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure* 8: 851-862.
- Zhou, R. & C. P. Wolk, (2003) A two-component system mediates developmental regulation of biosynthesis of a heterocyst polysaccharide. *J Biol Chem* 278: 19939-19946.
- Zusman, D. R., A. E. Scott, Z. Yang & J. R. Kirby, (2007) Chemosensory pathways, motility and development in *Myxococcus xanthus*. *Nat Rev Microbiol* 5: 862-872.

Acknowledgments

This research project would not have been possible without the support of many people.

First of all, I would like to express my profound gratitude to my advisor, Prof. Dr. MD Lotte Sogaard-Andersen, who offered guidance and suggestions throughout my PhD study.

I am also much indebted to my thesis and defense committee members, Prof. Dr. Erhard Bremer, Dr. Penelope Higgs, Prof. Dr. Lars-Oliver Essen, and Prof. Dr. Hans-Ulrich Mösch for their valuable advice from which I have benefited, especially to Prof. Dr. Erhard Bremer for his time and effort in reviewing this work.

I would like to make a grateful acknowledgement to Dr. Sigrun Wegener-Feldbrügge for crucial comments on my thesis. A special obligation is due to Dr. Ronald Brudler and Meike Ammon for translating my abstract into German.

To our collaborator Dr. Robert A. White, thanks for the prediction of the connectivity of TCS proteins in *M. xanthus*, and to Associate Prof. Dr. Mitchell Singer, thanks for presenting a strain.

My sincere thanks go to Dr. Stuart Huntley for all of the bioinformatic analyses, and to Dr. Sigrun Wegener-Feldbrügge, Dr. Nils Hamann and PD Dr. Reiner Hedderich for sharing microarray data. I am thankful to Bongsoo Lee, Dr. Penelope Higgs for providing strains and to Tobias Petters for sharing the phenotypes of three mutants. I appreciate the aid from Dr. Barbara Clement, Ambhighainath Ganesan in some constructs. I would also like to acknowledge Steffi Lindow for some Y2H constructs, Jörg Kahnt for MS Maldi-Tof analyses, Susanne Kneip for testing Hpk37 antibodies and Dr. Manfred Irmeler for all of the computational support. I am obliged to all of the persons in Dept. of Ecophysiology, especially the persons in our lab including the previous lab members, who gave me much assistance and enjoyable time during my study here.

It is a pleasure to convey thanks to International Max Planck Research School for Environmental, Cellular and Molecular Microbiology for providing financial aid. I am grateful to Susanne Rommel, Dr. Ronald Brudler, Dr. Juliane Dörr and Christian Bengelsdorff for taking care of all of the documents for my daily life.

Words fail me to express my deepest love and gratitude to my mother, who loves me more than herself. Hope what I have done makes her smile in heaven.

Finally, I take this opportunity to thank my other devoted family members including my sister, my brother and my father for their understanding and endless love, as well as all of my friends for their moral support, particularly beloved Shouqiang, who has been always waiting for me and offering help.

Curriculum Vitae

Personal data

Name Xingqi Shi
Date of Birth November 24, 1980
Place of Birth Anhui, China

Education

09/1997-06/2001 Bachelor of Science
Anhui University, Heifei, China
Bachelor thesis: Tissue culture and industrialization of cultivating seedlings of *Amorphophallus rivieri Durieu*

09/2001-07/2004 Master of Science
Nankai University, Tianjin, China
Master thesis: The study of genes involved in flagellar motility of *Pseudomonas aeruginosa*

10/2005-10/2008 PhD (Dr. rer. nat.)
Philipps-Universität Marburg
Max-Planck-Institute for Terrestrial Microbiology, Germany
PhD thesis: An analysis of two-component regulatory systems in *Myxococcus xanthus*

List of publications

Shi, X., S. Wegener-Feldbrugge, S. Huntley, N. Hamann, R. Hedderich & L. Sogaard-Andersen, (2008) Bioinformatics and experimental analysis of proteins of two-component systems in *Myxococcus xanthus*. *J Bacteriol* 190: 613-624.

Shan, Z., M. Qiao, **X. Shi**, H. Xu, H. Yao, Y. Bai, X. Zhang, & C. Gao, (2006) The study of genes involved in swimming motility of *Pseudomonas aeruginosa*. *Nan Kai Da Xue Xue Bao (Zi Ran Ke Xue Ban)* 39: 20-25. (In Chinese)

Shan, Z., H. Xu, **X. Shi**, Y. Yu, H. Yao, X. Zhang, Y. Bai, C. Gao, P. E. Saris & M. Qiao, (2004) Identification of two new genes involved in twitching motility in *Pseudomonas aeruginosa*. *Microbiology* 150: 2653-2661.

Shan, Z., H. Xu, **X. Shi**, Z. Nie, Y. Yu, X. Zhang, Y. Bai, M. Qiao & C. Gao, (2004) Study on Genes Involved in Twitching Motility of *Pseudomonas aeruginosa*. *Acta Microbiologica Sinica*. 44:319-323. (In Chinese)

Shan, Z., H. Xu, **X. Shi**, Z. Nie, Y. Yu, X. Zhang, Y. Bai, M. Qiao & C. Gao, (2004) The study of optimal conditions of electroporation in *Pseudomonas aeruginosa*. *Yi Chuan Xue Bao* 31: 311-316. (In Chinese)

Erklärung

Ich erkläre, daß ich meine Dissertation

An analysis of two-component regulatory systems in *Myxococcus xanthus*

selbständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 14.10.2008

Xingqi Shi