RP TU Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

Histidine kinases and response regulators in *Methanosarcina acetivorans*:

Structural and functional insights into archaeal signalling networks

vom Fachbereich Biologie

der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau

zur Verleihung des akademischen Grades Dr. rer. nat. genehmigte

Dissertation

von

Nora F. K. Georgiev, M.Sc., geb. in Speyer

Mündliche Prüfung: 19.12.2024

Dekan:	Prof. Dr. Stefan Kins
Promotionskommissionsvorsitzender:	Prof. Dr. Stefan Kins
Berichterstattende:	Prof. Dr. Nicole Frankenberg-Dinkel
	Prof. Dr. Matthias Hahn

D386

“If you always do what you’ve always done, you’ll always get what you’ve always got.”

— Henry Ford —

Für meine Familie.

Table of contents

Abbreviations

Abbreviations of the International System of Units (SI) are not listed separately. For amino acids the three-letter code or the one-letter code is used.

ADH	alcohol dehydrogenase
AEBSF	4-(2-aminoethyl)benzene-sulfonyl fluoride
AHT	anhydrotetracycline
Amp	ampicillin
Apo	apoferritin
APS	ammonium persulfate
ATP	adenosintriphosphat
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BD	blue dextran
BSA	bovine serum albumin
CA	**C**atalytic **A**TP-binding domain = HATPase
CA	carbonic anhydrase
CHASE	**C**yclases/**H**istidine kinases **A**ssociated **S**ensory **E**xtracellular
Cm	chloramphenicol
CV	column volume ≈ 1 ml
DHp	**D**imerisation and **H**istidine **p**hosphotransfer = HisKA,
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EtOH	ethanol
fwd	forward
GAF	c**G**MP-specific phosphodiesterases, **A**denylyl cyclases and **F**hlA
gDNA	genomic DNA
HAMP	**H**istidine kinases, **A**denylate cyclases, **M**ethyl accepting proteins and **P**hosphatase
HF	high fidelity
HHK	**H**ybrid **H**istidine **K**inase
HK	**H**istidine **K**inase
HPt	**H**istidine **P**hospho**t**ransfer
IPTG	isopropyl β-D-1-thiogalactopyranoside
K_{AV}	gel phase distribution coefficient
kDa	kilodalton

MASE	**M**embrane-**A**ssociated **SE**nsor
MEDS	**ME**thanogen/methylotroph **D**cmR **S**ensory
MeOH	methanol
MWCO	molecular weight cut-off
NBT	nitro blue tetrazolium chloride
OCS	one-component system
OD	optical density
oN	overnight
PAC	**P**AS-**A**ssociated **C**-terminal motif
PAS	**P**er-**A**rnt-**S**IM
pH	potential of hydrogen (latin *potentia hydrogenii*)
PocR	sensory domain found in PocR, with novel variant of the PAS-like fold
PVDF	polyvinylidene fluoride
QC	QuikChange
REC/R	**REC**eiver domain
rev	reverse
RR	**R**esponse **R**egulator
RT	room temperature
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
STKs	Ser/Thr kinases
TCS	two-component system
TEMED	tetramethyl ethylenediamine
TNP-ATP	2'-(3')-O-(2,4,6-trinitrophenyl)-adenosine 5'-triphosphat
Tris	tris(hydroxymethyl)aminomethane
UV/Vis	ultraviolet-visible
β-Amy	β-amylase
ε	molar absorption coefficient
λ	wavelength

The following is an example of terms and spellings used for genes and proteins in this study:

MA_4377 or *rdms*: genes of *M. acetivorans*

MA4377 or RdmS: corresponding protein of the gene MA_4377 or *rdms*

1 Introduction

1.1 Signal transduction allows microorganisms to adapt to their environment

All living organisms must sense and react to cellular and environmental signals and stimuli in order to survive. Microorganisms frequently occur in habitats with rapidly changing conditions, including fluctuations in temperature, pH, salt concentration, and nutrient availability. Furthermore, the interaction with other microorganisms or alterations within the cell can be a crucial factor in microbial survival. With signal transduction systems, microorganisms achieved a simple but effective method to sense and respond quickly to their surroundings (Bhate *et al.*, 2015; Gao & Stock, 2009; Laub & Goulian, 2007; Parkinson, 1993; Stock *et al.*, 2000).

Signal transduction systems are known at different levels of complexity. It can be distinguished between one-component systems (OCSs), two-component systems (TCSs) and phosphorelay systems (Hoch, 2000; Ulrich *et al.*, 2005). OCSs are proteins composed of an input and output domain on a single polypeptide chain (Figure 1.1, A.). An example of an OCS is the single-molecule transcriptional regulator RocR with an HTH (**H**elix-**T**urn-**H**elix) and PAS domain (Calogero *et al.*, 1994; Ulrich *et al.*, 2005). OCSs represent the simplest type of signal transduction systems and are widespread among the three domains of life. The diversity of input and output domains of OCS are much greater than the diversity of these domains in TCSs (Ulrich *et al.*, 2005). A classical TCS is composed of a histidine kinase (HK) (with sensory domain and HK domain) and a response regulator (RR) (with receiver domain and output domain) (Bhate *et al.*, 2015; Capra & Laub, 2012; Laub & Goulian, 2007). A specific signal is sensed by the sensory domain of the HK and leads to an autophosphorylation reaction within the **D**imerisation and **H**istidine **p**hosphotransfer domain (DHp domain) and the **C**atalytic **A**TP-binding domain (CA domain). The γ-phosphate of the bound ATP is transferred to the Nε atom of the conserved histidine (His/H) residue forming a high-energy N-P bond. The phosphoryl group is subsequently transferred to an aspartate (Asp/D) residue in the receiver (REC) domain of the RR (Bhate *et al.*, 2015; Stock *et al.*, 2000). The phosphorylation of the RR results in the activation or deactivation of the output domain, which finally leads to a reaction (Gao *et al.*, 2019; Gao *et al.*, 2007; Perry *et al.*, 2011). HKs can be distinguished regarding the position of their sensory domain. This domain can be either cytoplasmatic (Figure 1.1, B.), membrane-bound (Figure 1.1, C.) or extracellular (Figure 1.1, D.) (Perry *et al.*, 2011; Wolanin *et al.*, 2002). Phosphorelay systems are based on TCS but are more complex, as they involve additional steps (a four-step phosphorelay involving His-Asp-His-Asp). These systems consist of a hybrid HK (HHK) with bound REC domain, an intermediate

Histidine **P**hospho**t**ransfer (HPt) domain or HPt-like protein and a RR with output domain. The HPt component can be either fused to a HHK (Figure 1.1, E.) or can be a stand-alone protein (Figure 1.1, F.) (Gao & Stock, 2009; Singh *et al.*, 2021). Additionally, pseudokinases can also function as HPt components (Kwon *et al.*, 2019). TCSs and phosphorelay systems can also be interconnected, resulting in more complex arrangements, in which case a distinction is made between cross-talk and branched pathways. Cross-talk is defined as communication between two distinct pathways that would leave a distinct and functional pathway if the communication is eliminated (Figure 1.1, G.) (Laub & Goulian, 2007). Additionally, pathways can be branched in two different ways. Either one HK can transfer signals to more than one RR (one-to-many), or several HK can transfer signals to only one RR (many-to-one) (Figure 1.1, H./I.) (Laub & Goulian, 2007). The chemotaxis protein CheA is an example for the one-to-many principle. Both RR CheY and CheB need to be transphosphorylated from the HK CheA to get a functional response. Cross-talk and the involvement of several components into one signal transduction pathway leads to highly complex systems (Laub & Goulian, 2007; Szurmant & Ordal, 2004).

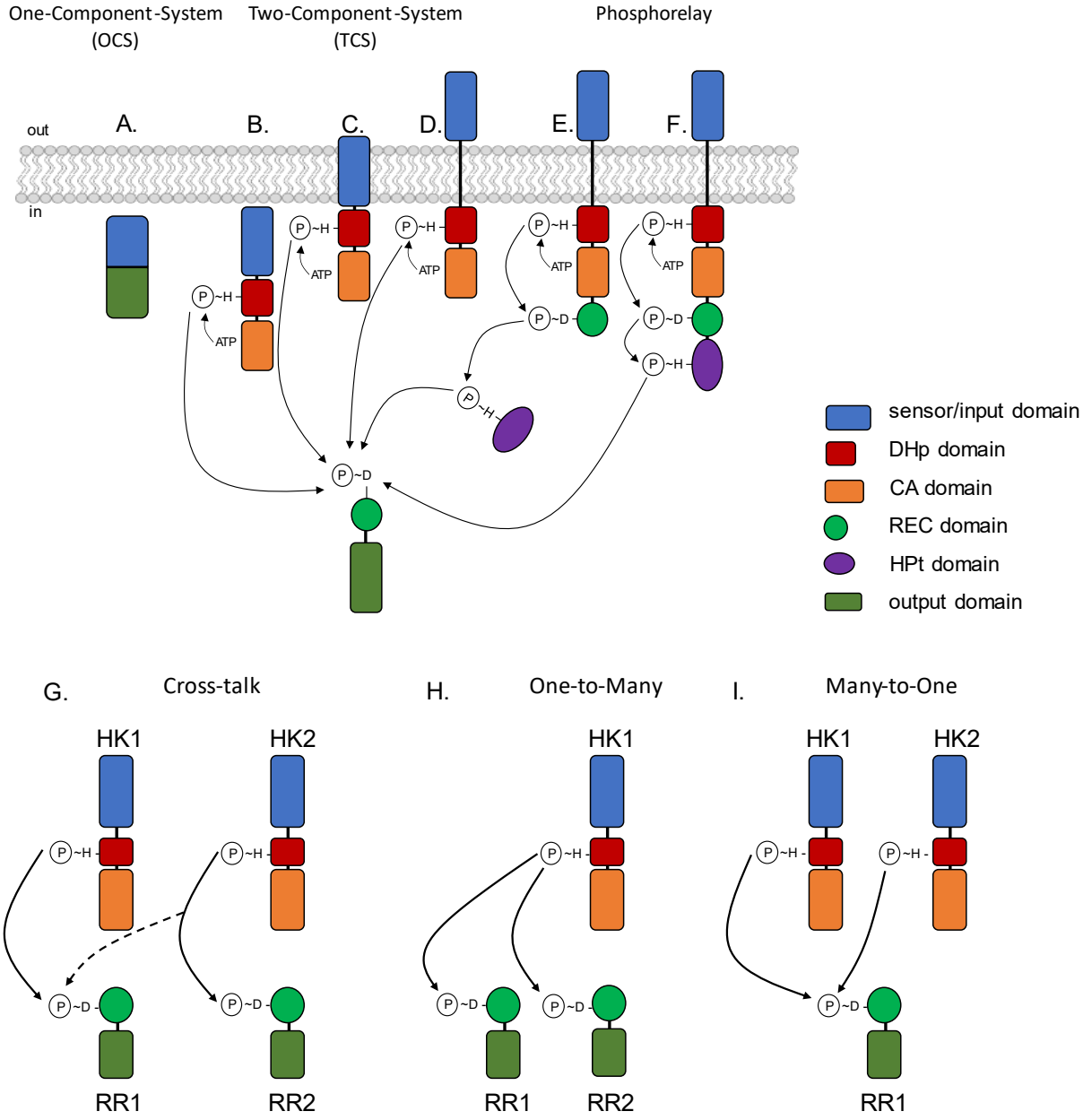

Figure 1.1: Schematic overview of different types of signal transduction pathways and different types of cross regulation. (**A.**) Structure of a one-component system (OCS) with fused input and output domain. (**B.**) Structure of a two-component system (TCS) with cytoplasmatic HK. A signal leads the autophosphorylation of the HK and to transfer of a phosphoryl group to a cognate RR with output domain. (**C.**) Structure of a TCS with membrane-bound HK. (**D.**) Structure of a classical TCS with extracellular HK. (**E.**) Structure of a phosphorelay system with bound REC domain and stand-alone HPt protein. (**F.**) Structure of a phosphorelay system with bound REC and HPt domain. For the sake of simplicity, the phosphorelay system is represented with an extracellular sensor domain, but they also exist with cytoplasmatic and membrane-bound sensor domains. (**G.**) Cross-talk of two TCSs. With the elimination of the cross-talk (dashed line) both TCS would still function as distinct systems. (**H.**) Regulation of a TCS according to the one-to-many principle. The HK1 communicates to RR1 and RR2. (**I.**) Regulation of a TCS according to the many-to-one principle. Both HK1 and HK2 can communicate with RR1. Modified from (Laub & Goulian, 2007).

1.2 Molecular structures give hint on how histidine kinases talk to response regulators

To understand the function of HKs and RRs in more detail it is crucial to look closer into their molecular structure. A typical HK is a homodimeric protein composed of an N-terminal sensor

region (with one or more sensory domains) and a HK domain. The sensory domains can be cytoplasmic (i.e. PAS (**P**er-**A**rnt-**S**IM), PAC (**P**AS-**A**ssociated **C**-terminal motif) and GAF (c**G**MP-specific phosphodiesterases, **a**denylyl cyclases and **F**hlA) domains)), membrane integral or extracellular (i.e. CHASE (**C**yclases/**H**istidine kinases **A**ssociated **S**ensory **E**xtracellular), MEDS (**ME**thanogen/methylotroph **D**cmR **S**ensory) or Cache (**CA**lcium channels and **CHE**motaxis receptors) domains). Additional linker domains (HAMP (**H**istidine kinases, **A**denyl cyclases, **M**ethyl-accepting proteins and **P**hosphatases)) can be situated between sensor domains and the HK domain (Buschiazzo & Trajtenberg, 2019; Gao & Stock, 2009; Zschiedrich *et al.*, 2016). The HK domain consists of a DHp domain and a CA domain. The DHp domain is sometimes designated as HisKA, while the CA is designated as HATPase. The DHp domain drives the homodimer formation which leads to a four-helix bundle of the HK (Ashenberg *et al.*, 2013; Gao & Stock, 2009; Gärtner, 2020; Tomomori *et al.*, 1999). Furthermore, the phosphorylation site (H-box) with the conserved His residue is situated on the α1 helices in the middle of the DHp domain. The CA domain is responsible for the binding and the hydrolysis of the ATP and displays a globular α/β structure. The domain consists of a five-stranded β-sheet, which is packed against three α-helices on one side (Bergerat fold) (Buschiazzo & Trajtenberg, 2019; Casino *et al.*, 2014; Dutta *et al.*, 1999). The HK domain possesses some highly conserved regions, designated as the N, G1, F and G2 boxes. The N, G and G2 boxes are responsible for the formation of the ATP binding pocket, while the loop that includes the F-box acts as the ATP lid (Dutta *et al.*, 1999; Wolanin *et al.*, 2002). Most HK form homodimers after perceiving a signal, but also the existence of a monomeric kinase (Rivera-Cancel *et al.*, 2014; Zschiedrich *et al.*, 2016) and a hexameric kinase (Wojnowska *et al.*, 2013) was reported. The following section will discuss the function of a typical homodimeric kinase. The four-helix bundle structure allows different possible ways how the phosphoryl group gets transferred. In one scenario, the ATP is bound to one monomer, while the phosphoryl group is transferred to the second monomer, resulting in a trans-autophosphorylation (Figure 1.2, A.). Alternatively, the ATP can be bound to one monomer, with phosphoryl group transfer to the His residue of the same monomer, a process known as cis-autophosphorylation (Figure 1.2, B.) (Ashenberg *et al.*, 2013; Gärtner, 2020; Zschiedrich *et al.*, 2016). Both cis- and trans-autophosphorylation are found *in vivo*. The HKs EnvZ, NtrB, AtoS and CheA of *Escherichia coli* (*E. coli)* and the HK AgrC of *Staphylococcus aureus (S. aureus)* autophosphorylate in trans. On the other hand, the HK PhoR of *S. aureus*, ArcB of *E. coli* and HK853 of *Thermotoga maritima* (*T. maritima*) autophosphorylate in cis (Ashenberg *et al.*, 2013). The structural reasons that discriminate between cis- or trans-autophosphorylation were investigated with chimeric HKs (Casino *et al.*, 2014). These experiments showed that the handedness of the loop between the four-helix bundle of the DHp domain and the CA domain is crucial for the autophosphorylation. The four-helix bundle

is formed by two α-helices from each monomer. The loop can either be right-handed or left-handed. A HK with a right-handed loop, like EnvZ, autophosphorylates in trans (Figure 1.2, C.). A HK, like HK853, with a left-handed loop autophosphorylates in cis (Figure 1.2, D.) (Ashenberg *et al.*, 2013; Buschiazzo & Trajtenberg, 2019; Casino *et al.*, 2014; Zschiedrich *et al.*, 2016).

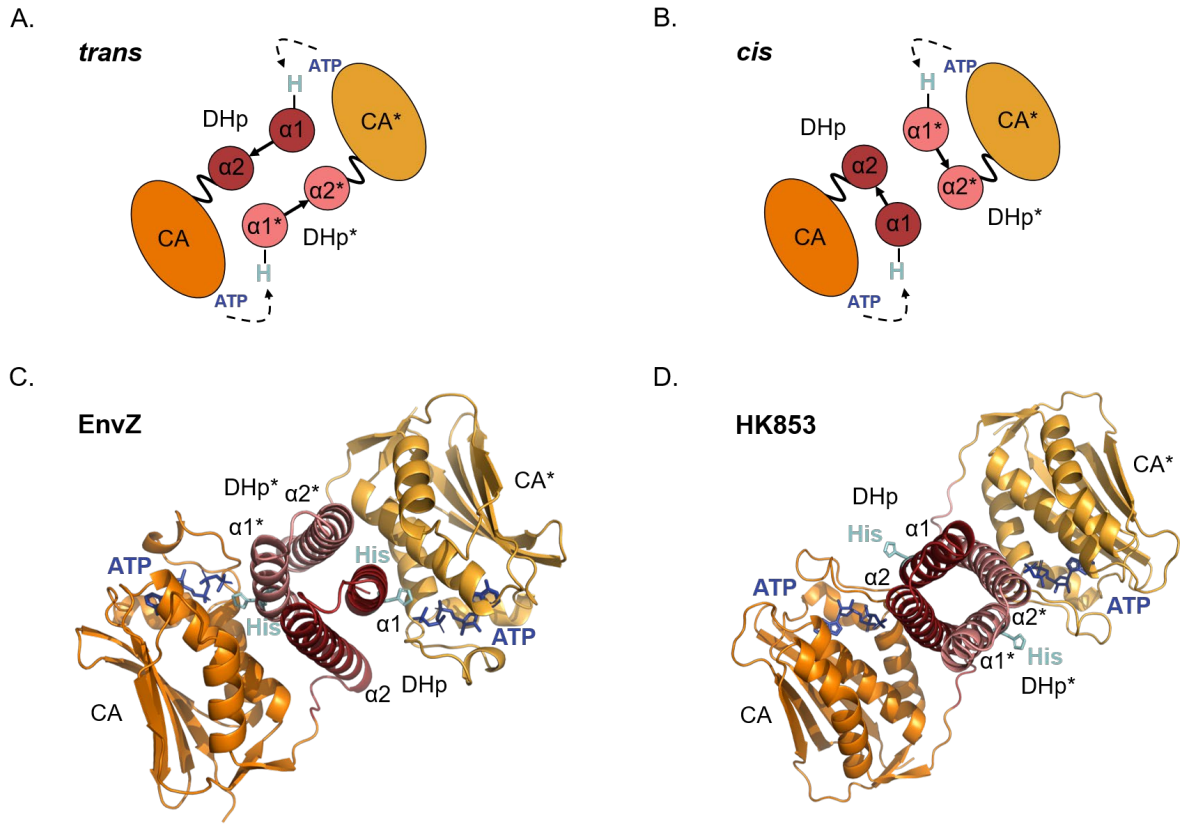

Figure 1.2: Schematic representation of trans- and cis-autophosphorylation and corresponding examples. (**A.**) Schematic representation of the CA (orange) and DHp (red) domain of a HK that autophosphorylates in trans. (**B.**) Schematic representation of the CA (orange) and DHp (red) domain of a HK that autophosphorylates in cis. Whether a HK autophosphorylates in trans or cis depends on the handedness of the loop between the CA and the DHp domain. (**C.**) The displayed structure shows the CA and DHp domain of EnvZ (*E. coli*) predicted with Alphafold3 (Abramson *et al.*, 2024). The HK domain of a HK consists of a CA (orange) and DHp (red) domain. The CA and DHp domain of the second monomer is identified with an asterisk (*). (**D.**) The displayed structure shows the CA and DHp domain of HK853 (*T. maritima*) predicted with AlphaFold3 (Abramson *et al.*, 2024). The HK domain of a HK consists of a CA (orange) and DHp (red) domain. The CA and DHp domain of the second monomer is identified with an asterisk (*). AlphFold3 structures were edited using PyMOL (DeLano, 2020) Modified from (Ashenberg *et al.*, 2013).

A typical RR consists of a REC domain and one or more output domains. The structure of the REC domain in RRs is highly conserved and in the following all known features of the REC domain will be explained (Figure 1.3, A./B.): Five central β-strands are surrounded by two α-helices on one site and three α-helices on the other. The structure can be recognized in the amino acid sequence. Three of the β-strands are located in the middle of the structure, the active site can be found at the C-terminal end of these three β-strands. The strand β1 is followed by two highly conserved Asp residues that are involved in metal ion binding (Mg^{2+}). The highly conserved Asp of the phosphorylation site is situated at the end of β3. Sometimes

β3 is also starting with an Asp residue, but the function is yet unknown. Between β3 and α3 a 180° loop is initiated by a proline (Pro/P) residue. The α3 helices start with a highly conserved glycine (Gly/G) residue and β4 ends in conserved threonine (Thr/T) or serine (Ser/S) residue, followed by alanine (Ala/A), Gly or Thr/Ser. The Thr/Ser residue together with following small residues gives access to the phosphorylation site. Residues of the β5 (less conserved phenylalanine (Phe/F) or tyrosine (Tyr/Y) and highly conserved lysine (Lys/L)) are responsible for phosphorylation-mediated conformational changes (Bourret, 2010; Gao *et al.*, 2019; Gao *et al.*, 2007; Zschiedrich *et al.*, 2016). The AlphaFold3 predicted protein structure of the REC-only RR CheY of *E. coli* was analysed regarding the conserved amino acid residues that are most likely responsible for the function of the RR (Figure 1.3, C./D.).

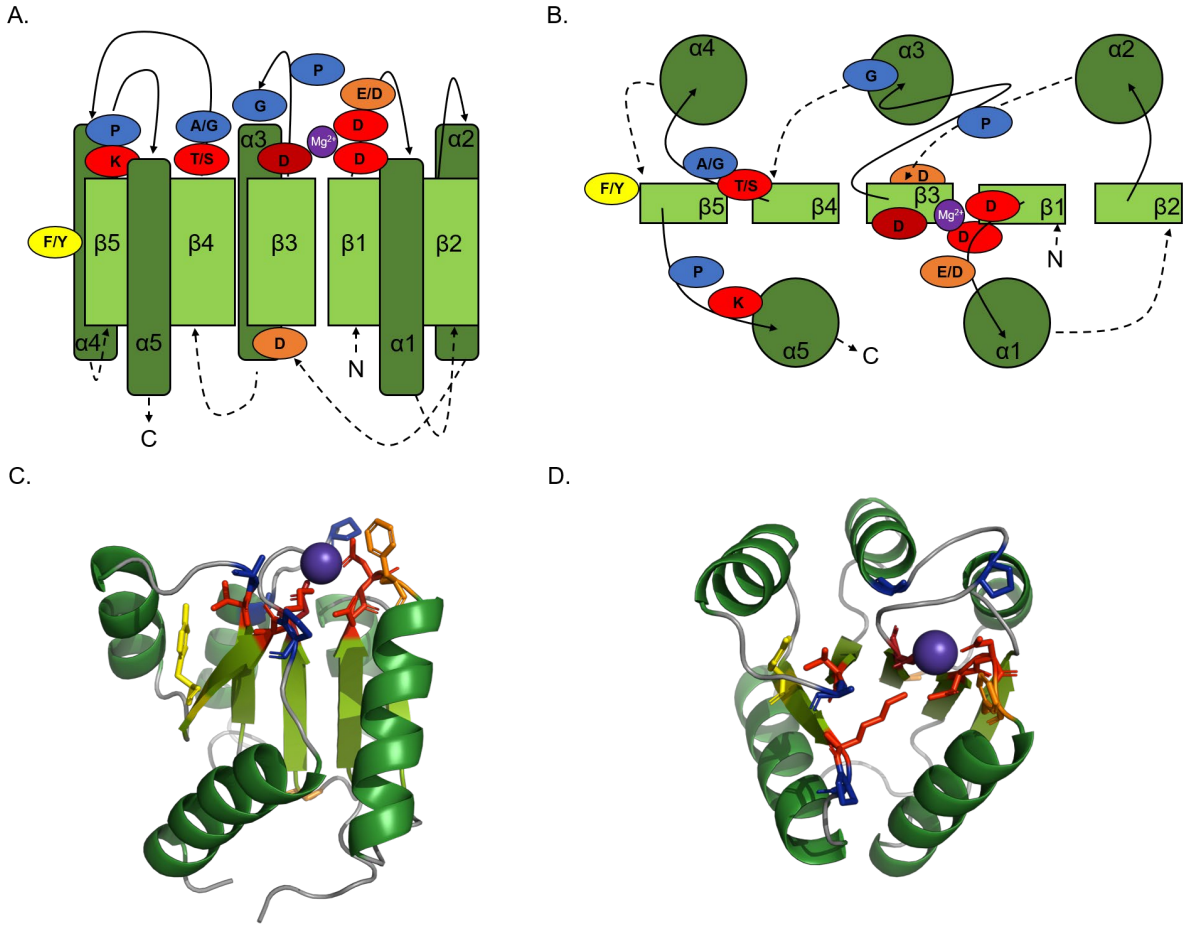

Figure 1.3: Schematic overview of a typical REC domain and protein structure of the REC-only RR CheY of *E. coli*. (**A.**) Schematic representation of a typical REC domain from the side and (**B.**) From the top. Five β-strands (light green) are surrounded by five α-helices (dark green). All functionally relevant amino acids are displayed in different colours. Red indicates highly conserved amino acids, and the conserved Asp residue of the phosphorylation site is coloured in dark red. Less conserved amino acid residues are displayed in yellow. Amino acid residues with suspected structural relevance are displayed in blue and conserved residues with unknown function are displayed in orange. The divalent metal ion is coloured in violet. (**C.**) Protein structure of CheY (*E. coli*) with conserved amino acid residues from the side and (**D.**) From the top. The protein structure was predicted using AlphaFold3 (Abramson *et al.*, 2024) and edited using PyMOL (DeLano, 2020). The residues are displayed in the same colours than in the schematic representation. Modified from (Bourret, 2010).

The conserved nature of the HK domain of HKs and the REC domain of RRs raises the question of how signal transduction systems can be specific. It is known that different TCSs can interact with each other and that substrate competition, but also phosphatases, can prevent cross regulation (Laub & Goulian, 2007). It was shown *in vitro* that the kinase VanS can interact with the RR PhoB, but only in the absence of PhoR. The same was also shown for PhoR to the RR VanR in the absence of VanS. Both kinases are bifunctional, meaning they have kinase and phosphatase activity. If both kinases are present, the phosphatase activity prevents cross-talk (Fisher *et al.*, 1995; Haldimann *et al.*, 1997; Laub & Goulian, 2007).

But how does a HK and a RR decide how to fit together in the first place? The recognition of cognate partners at the molecular level appears to be crucial, and the process is not yet understood. It is known that the four-helix bundle of the HK with the conserved His residue and the α/β structure of the RR with some highly conserved residues is a key point in understanding how the HK talks to the RR. It is known that the RR binds to the DHp domain and catalyses its own phosphorylation. The phosphoryl group situated at the His residue is used as a substrate. The results of several experiments with small molecules, such as acetyl-phosphate, indicate that the RR is catalysing the phosphotransfer actively through the involvement of specific amino acid residues. (Wolfe, 2010).

A study of the EnvZ kinase showed the importance of the DHp domain for transphosphorylation. A cluster of seven residues within the DHp domain was identified as important for the interaction with the RR. In a chimera experiment, this cluster in EnvZ was exchanged with a cluster from a different kinase. The experiment showed that the chimeric kinase transphosphorylates a phosphoryl group only to the cognate RR to which the cluster belongs (Skerker *et al.*, 2008). Two other studies provide the first atomic structure of a complex composed of a RR and an HPt protein (YPD1:SLN1 and Spo0F:Spo0B). Since the structure of the HPt is highly similar to a HK, it can be used to understand the molecular mechanisms between the complex of a HK and a RR. With the help of different studies, some amino acids that are required for binding and also for the catalyses of the phosphotransfer were found (Varughese *et al.*, 2006; Xu *et al.*, 2003; Zschiedrich *et al.*, 2016). However, the knowledge is still limited on how the HK can distinguish between a cognate and non-cognate RR.

1.3 TCSs occur in all three domains of life

Signal transduction systems are known for all three domains of life. HKs can be found in bacteria (Mascher *et al.*, 2006), archaea (Galperin *et al.*, 2018), amoebae (Schaap *et al.*, 2016), plants (Osakabe *et al.*, 2013), fungi (Hérivaux *et al.*, 2016) and viruses/bacteriophages (Delaroque *et al.*, 2001; Hargreaves *et al.*, 2014) and are missing in higher organisms

(Kabbara *et al.*, 2019; Koretke *et al.*, 2000; Wuichet *et al.*, 2010). However, HK activity was also detected in mammals, even though they do not encode for typical HK-like proteins (Attwood, 2013). Besides the fact that all three domains possess signal transduction pathways, there are also distinct differences between the systems of the different domains.

Most common signal transduction systems in prokaryotes are OCSs (Ulrich *et al.*, 2005). TCSs have been identified in both prokaryotic and eukaryotic organisms. In bacteria, a correlation between the number of HK and the size of the genome was demonstrated (Galperin *et al.*, 2010). Moreover, it is proposed that the ecological and environmental niches of microorganisms are also relevant factors in determining the number of HKs in their genome. It has been observed that microorganisms which inhabit extreme environments with rapidly changing conditions, or which coexist with other microorganisms, appear to encode a greater number of HK than parasites or microorganisms which populate environmental niches (Galperin *et al.*, 2010; Ulrich *et al.*, 2005). It is noteworthy that within all three domains, there are species that are characterised by the absence of signal transduction systems. In bacteria some pathogen species (*Mycoplasma spec*.) and endosymbionts (*Amoebophilus spec*.) completely lack HKs (Wuichet *et al.*, 2010). In archaea, TCSs are absent in the phyla *Crenarchaeota*, *Korarchaeota*, *Nanoarchaeota* and Candidatus Nanohaloarchaeota. In contrast, the phyla *Euryarchaeota* and *Thaumarchaeota* display a high number of HKs and RRs (Galperin et al., 2018). TCS are also known in Asgard archaea. This is of particular interest given that Asgard archaea are thought to represent the evolutionary link between bacteria and eukaryotes (Padilla-Vaca *et al.*, 2023). The known TCSs in bacteria are mostly composed of a sensing HK with a RR that directly affects the transcriptional regulation via HTH domains. Moreover, protein-binding and enzymatic reactions are known to be the output of bacterial RRs. A well-studied classical bacterial TCS is the EnvZ/OmpR system from *E. coli*, which is involved in osmosensing (Forst & Roberts, 1994). Besides HKs, also arginine (Arg) kinases, serine (Ser) kinases, tyrosine (Tyr) kinases and Hanks-type Ser/Thr kinases (STKs) are known from bacteria. Hanks-type kinases (Ser/Thr kinases) were first believed to be unique to eukaryotes, however, to date it is known that they are also present in bacteria and archaea (Hunter, 2022; Janczarek *et al.*, 2018; Stancik *et al.*, 2018). A phylogenetic study did not find evidence for horizontal gene transfer of Hanks-type kinase genes between eukaryotes and bacteria, suggesting a monophyletic origin of Hanks-type kinases (Janczarek *et al.*, 2018). HKs are the predominant class of kinases in bacteria, whereas Ser/Thr and Tyr kinases are predominant in eukaryotes (Hunter, 2022). Furthermore, eukaryotic TCSs are mostly composed of an additional module, a histidine-containing phosphotransfer domain that mediates the signal from the HK to the RR. The HK is mostly a HHK with bound REC domains. In eukaryotes, phosphorelay systems are more abundant than classical TCSs. Additionally,

the RRs in eukaryotic TCSs can have an extended function besides the classical role as transcriptional regulator. They can regulate downstream elements like mitogen-activated protein kinase (MAPK) cascades or cyclic adenosine monophosphate (cAMP) signalling systems (Hérivaux *et al.*, 2016; Kabbara *et al.*, 2019). The Sln1/Ydp1/Ssk1 phosphorelay of *Saccharomyces cerevisiae* is one example for a system that activates a MAP kinase pathway (Lu *et al.*, 2003).

Despite the extensive research conducted on bacterial and eukaryotic TCSs in terms of their structural and functional characteristics, the number of known archaeal TCSs remains limited. To date only three pathways have been experimental investigated. The halophilic archaeon *Halobacterium salinarum* contains a CheA/CheY/CheB system that plays a role in chemotaxis and functions in a manner analogous to that observed in bacterial CheA HKs (Rudolph & Oesterhelt, 1995). Furthermore, the FilI/FilR1/FilR2 system from *Methanosaeta harundinacea* is responsible for regulating methanogenesis. The RR FilR1 comprises a REC domain, a DUF1724 domain and a HTH domain, whereas the RR FilR2 is a REC-only protein (Li *et al.*, 2014). The third known archaeal TCS is from *Methanococcoides burtonii*, where the LtrK/LtrR system is responsible for the temperature regulation of the organism (Najnin *et al.*, 2016). All three systems are composed of at least one RR with HTH domain, and they all function similar to their bacterial counterparts. There is a notable distinction between archaea and bacteria with respect to the ratio of HKs to RRs. In bacteria this ratio is approximately 1, whereas archaea typically exhibit a higher number of HK (Galperin *et al.*, 2018). This phenomenon leads to the assumption that archaeal TCSs function in a more complex way with more interconnection between HKs and RRs. In contrast to bacteria, archaea have a limited number of RR with HTH domains, but an elevated number of REC-only RR and novel output domains (Galperin *et al.*, 2018; Krell, 2018). This additionally leads to the assumption that transcriptional regulation is not the main output of archaeal RR (Krell, 2018). Between 73 and 88 % of bacterial HK are proteins with transmembrane regions, which suggests a high importance of extracytoplasmic signal sensing (Wuichet *et al.*, 2010). Cyanobacteria have a lower number of membrane-bound HK and around 62 % of archaeal HK lack transmembrane regions (Galperin *et al.*, 2018). Together with a high number of cytoplasmic sensing domains like PAS and GAF domains, this unveils that cytoplasmatic sensing is more important in archaea than in bacteria (Galperin *et al.*, 2018).

1.4 Lateral gene transfer led to the evolution of TCSs in archaea

The observation of the distribution of TCSs across all three domains of life, along with the identification of differences between the signalling systems present in these organisms, has prompted the question of how these signalling pathways originated and evolved. This is a

challenging question to answer, and our understanding of it is still incomplete. However, significant advances have been made over time.

Signal transduction pathways are believed to have evolved in bacteria after the separation from the last common ancestor. From bacteria they were passed on to archaea and eukaryotes through horizontal gene transfer (Koretke *et al.*, 2000). It is hypothesised that OCSs are the precursor of TCSs and phosphorelay systems and that they became more complex over time. The high diversity of input and output domains in OCSs, their simple composition and widespread distribution supports the idea of OCSs as the precursors of signal transduction systems (Ulrich *et al.*, 2005). Additionally, there are indications to the evolutionary history of HKs. One hypothesis suggests that they may have originated from ATPase-like proteins. The structural similarities between HKs and proteins such as Hsp90, MutL, and topoisomerases, support this theory (Capra & Laub, 2012; Koretke *et al.*, 2000). Additionally, homologues were found for HPts. They could have been originated from two different sources as their only requirement is a His residue that can get phosphorylated. The monomeric HPts most probably evolved *de novo* from other proteins, while the dimeric HPts may have originated from degenerated HKs. However, for RRs these homologies to other proteins are absent (Capra & Laub, 2012).

This leads to the question of what mechanisms are responsible for enabling evolution. The principal factors are gene transfer and gene duplication, in addition to gene loss through deletion or mutation. Gene transfer can occur via phages, plasmids, direct conjugation, competence, or direct uptake (Capra & Laub, 2012). The introduction of a new signal transduction gene into another organism requires a high number of mutations, substitutions, and rearrangements to adapt this gene to a new signalling function. It is established that the sensory and output domains of HKs are less conserved than the kinase catalytic core elements (Capra & Laub, 2012). This indicates that alterations in the input and output domains result in the generation of HK with new functions. The substitution or shuffling of sensory domains can facilitate the rapid generation of HKs that respond to novel signal inputs. Furthermore, small changes in the output domains can easily lead to the regulation of novel target genes. A high number of TCSs are encoded in operons with a HK and a cognate RR. When lateral gene transfer is occurring, the whole operon is most probably transferred from one to another organism. It can also happen that only one gene of an operon gets duplicated, which is an explanation for the existence of orphan TCS parts. However, orphan kinases often keep their function (Capra & Laub, 2012; Koretke *et al.*, 2000; Whitworth & Cock, 2009). An example for this is the sporulation phosphorelay systems of *B. subtilis*. The five HKs KinA/B/C/D/E which might have evolved through duplication can all phosphorylate the same

RR Spo0F (Stephenson & Hoch, 2002). The development of HHKs can be explained with gene fusion events and the mutation of stop codons inside an operon (Capra & Laub, 2012).

1.5 *Methanosarcina acetivorans:* an archaeal model organism for signal transduction

Methanosarcina acetivorans (*M. acetivorans*) is a mesophilic, obligate anaerobic, methanogenic archaeon. It is classified into the phylum *Euryarchaeota*, the order *Methanosarcinales* and the family *Methanosarcinaea*. It was isolated from marine sediment in a depth of 65 feet from the Sumner branch of Scripps Canyon located near to La Jolla in California (Sowers *et al.*, 1984). *M. acetivorans* has the largest known archaeal genome with 5.75 Mbp and 4524 open reading frames, suggesting a broad ability to adapt to changing environment (Galagan *et al.*, 2002). The organism is conducting methanogenesis for energy conservation and is able to utilize various substrates. This fact makes *M. acetivorans* a good model organism to study metabolism (Galagan *et al.*, 2002; Sowers *et al.*, 1984).

Compared to other archaea, methanogens in general and *M. acetivorans* in particular have a high number of putative signal transduction components (Galperin *et al.*, 2018). In addition to its large genome, the extensive range of substrates on which it can live and the environment from which it was initially isolated makes *M. acetivorans* also a good model organism to study signal transduction in archaea.

1.6 Objectives of this work

This study is focusing on the structure and function of HKs and RRs in the methanogenic archaeon *M. acetivorans*. The organism has one of the largest known archaeal genomes and, with 53 putative HK and 16 putative RR, one of the largest sets of signal transduction systems. To gain insights into the hitherto not very well understood signal transduction in archaea and *M. acetivorans* in particular, the HKs and RRs of the organism were investigated with a combined approach of an *in silico* analysis and a biochemical characterisation of four putative HKs (MA0863, MA2082, MA2013 and MA4377). Employing ATP binding studies and kinase assays the function of four putative HK was determined. This study gives the first indication for the existence of a complex archaeal phosphorelay system in *M. acetivorans*, involving several HKs and RRs.

2 Material and Methods

2.1 Materials and chemicals

2.1.1 Equipment

Table 2.1: List of instruments used in this study.

Instrument	Model	Manufacturer
Agarose gel electrophoresis chamber	MAXI VARIGEL 700-7213	VWR International, Radnor (PA), USA
Anaerobic tent	Type B/Type C (N_2 with 5% H_2)	Coy Laboratory Products Inc., Grass Lake (MA), USA
Autoclave	VX 150	Systec GmbH & Co. KG, Linden, Germany
Blotting system	Trans-Blot SD Semi-Dry Transfer Cell	Bio-Rad Laboratories, Hercules (CA), USA
Centrifuges and rotors		
	Centrifuge 5415 D, rotor: F45-24-11	Eppendorf SE, Hamburg, Germany
	Centrifuge 5430 R, rotor: F-35-6-30	Eppendorf SE, Hamburg, Germany
	Centrifuge 5810 R, rotor: A-4-62	Eppendorf SE, Hamburg, Germany
	Z 32 HK, rotor: 221.17	Hermle Labortechnik GmbH, Wehingen, Germany
	Z 323 K, rotor: 220.78 V02	Hermle Labortechnik GmbH, Wehingen, Germany
	Sorvall Lynx 6000 centrifuge, rotor: T29-8x50 rotor: F9-6x1000 LEX	Thermo Fisher Scientific Inc., Waltham (MA), USA
	Sorvall RC 5B PLUS, rotor: F12S-6x500 LEX	Thermo Fisher Scientific Inc., Waltham (MA), USA
Gel documentation system	Gel iX20 Imager	INTAS Science Imaging Instruments GmbH, Göttingen, Germany
Incubator	Heraeus instruments	Thermo Fisher Scientific Inc., Waltham (MA), USA
Incubator shaker	Innova™ 44 Incubator Shaker Series	New Brunswick Scientific Co., Inc., New Jersey (NJ), USA
Microfluidizer	LM 10 Microfluidizer	Microfluidics International Corporation, Newton (MA), USA
pH meter	pH 50+ DHS	XS Instruments, Carpi, Italy
PhosphorImager	Typhoon FLA 7000	GE Healthcare, Chicago (IL), USA
Pipettes	Research® Plus	Eppendorf SE, Hamburg, Germany
Power supply		
	Power Pac Basic	Bio-Rad Laboratories, Hercules (CA), USA

	Power Source 300 V	VWR International, Radnor (PA), USA
	Power Pac HC	Bio-Rad Laboratories, Hercules (CA), USA
Scales		
	4000/AR	Europe
	Acculab Atilon ATL-6201	Sartorius AG, Göttingen, Germany
	R200D	Sartorius AG, Göttingen, Germany
	R 300 S	Sartorius AG, Göttingen, Germany
SDS-PAGE chamber		
	Mini-PROTEAN Tetra Vertical Electrophoresis Cell	Bio-Rad Laboratories, Hercules (CA), USA
	Mini-PROTEAN	Bio-Rad Laboratories, Hercules (CA), USA
Shaker		
	Polymax 1040	Heidolph Scientific Products GmbH, Schwabach, Germany
	Rocker3D Basic	IKA Werke, Staufen im Breisgau, Germany
	Innova™ 2300 Platform Shaker	New Brunswick Scientific Co., Inc, New Jersey (NJ), USA
Size exclusion chromatography system		
	Äkta Explorer system	GE Healthcare, Chicago (IL), USA
	Äkta Pure system	GE Healthcare, Chicago (IL), USA
	Superdex® 200 Increase 10/300 GL column	GE Healthcare, Chicago (IL), USA
Spectrometer		
	8453 UV-visible Spectroscopy System	Agilent Scientific Technologies, Santa Clara (CA), USA
	Ultrospec 500 pro Photometer	Amersham Biosciences, Amersham, UK
	NanoDrop™Lite Spectrophotometer	Thermo Fisher Scientific Inc., Waltham (MA), USA
	FP-8300 fluorescence spectrometer	Jasco Deutschland GmbH, Pfungstadt, Germany
Sterile bench	Antares	Biohit, Helsinki, Finland
Thermo block	ThermoStat plus	Eppendorf SE, Hamburg, Germany
Thermocycler		
	T gradient	Biometra GmbH, Göttingen, Germany
	T-personal	Biometra GmbH, Göttingen, Germany
	T1 Thermocycler	Biometra GmbH, Göttingen, Germany

	Peqstar 2	Peqlab Biotechnologie GmbH, Erlangen, Germany
Ultrasonic homogenizer		
	Sonoplus	Bandelin electronic GmbH & Co. KG, Berlin, Germany
	Electronic UW 220	Bandelin electronic GmbH & Co. KG, Berlin, Germany
Ultra-pure water system	MilliQ® Integral Water Purification System	Merck Millipore, Burlington (MA), USA
Vacuum pump	Laboport®	KNF DAC GmbH, Hamburg, Germany
Vortex	Vortex	VWR International, Radnor (PA), USA

2.1.2 Chemicals, kits, antibodies

All chemicals used in this study were ACS grade or higher and were purchased from the following companies, unless stated otherwise: Carl Roth GmbH + Co. KG (Karlsruhe, Germany), Merck KGaA (Darmstadt, Germany), AppliChem GmbH (Darmstadt, Germany), Bio-Rad (Hercules (CA), USA), Serva Electrophoresis GmbH (Heidelberg, Germany), Sigma-Aldrich (St. Louis (MO), USA) and Thermo Fisher Scientific (Waltham (MA), USA).

Table 2.2: Special chemicals, antibodies, special materials, and kits used in this study.

Type of product	Name of product	Manufacturer
Special chemicals		
ATP	ATP ≥99 %, 100 mM solution	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Radioactive ATP	[γ-^{32}P]-ATP (SRP-301), 500 µCi	Hartmann Analytic, Braunschweig, Germany
TNP-ATP	2',3'-O-Trinitrophenyl-adenosine-5'-triphosphate, (TNP-ATP)	Jena Bioscience, Jena, Germany
Detergent DDM	Dodecyl-β-d-maltoside	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Nanodisc DIBMA	diisobutylene and maleic acid (ACUSOL™ 460ND Polymer)	Dow Chemical, Midland (MI), USA
Desthiobiotin	D-desthiobiotin	IBA Lifesciences GmbH, Göttingen, Germany
Biotin	D-biotin	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Imidazole	Imidazole	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Enzymes		
Buffers for polymerases		
	5x Phusion HF buffer	New England Biolabs GmbH, Ipswich (MA), USA
Buffers for restriction enzymes		
	CutSmart	New England Biolabs GmbH, Ipswich (MA), USA

	rCutSmart (for HF-enzymes)	New England Biolabs GmbH, Ipswich (MA), USA
DNA ligase	Taq DNA ligase	New England Biolabs GmbH, Ipswich (MA), USA
DNA polymerase	Phusion DNA polymerase	RPTU, Lab collection
	Phusion (2U/µl)	New England Biolabs GmbH, Ipswich (MA), USA
DNaseI	DNaseI	AppliChem GmbH, Darmstadt, Germany
dNTPs	dNTP-Mix (dATP, dCTP, dGTP, dTTP)	Axon Labortechnik, Kaiserslautern, Germany
Lysozyme	Lysozyme, lyophilized	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Restriction enzymes		
	Restriction endonucleases	New England Biolabs GmbH, Ipswich (MA), USA
	High-fidelity (HF®) restriction endonucleases	New England Biolabs GmbH, Ipswich (MA), USA
Antibodies		
Strep-tag II	Strep-Tactin AP conjugate (1:4000)	IBA Lifesciences GmbH, Göttingen, Germany
His-tag	6x-His Tag monoclonal antibody (1:3000)	Invitrogen AG, Walham (MA), USA
Mouse IgG	Goat anti mouse IgG alkaline phosphatase antibody (1:10000)	Sigma-Aldrich, St. Louis (MI), USA
Special material		
Blotting paper	Gel Blotting Papers Whatman® 3MM Thickness 0,34 mm, 13.3 x 10.2 cm	Cytiva, Marlborough (MA), USA
Centrifugal filter units		
	Amicon® Ultra-4, 10 kDa MWCO	Merck Millipore, Burlington (MA), USA
	Amicon® Ultra-4, 50 kDa MWCO	Merck Millipore, Burlington (MA), USA
	Amicon® Ultra-4, 100 kDa MWCO	Merck Millipore, Burlington (MA), USA
	Microspin™G-50 Columns	Cytiva, Marlborough (MA), USA
Colum material for affinity chromatography		
Strep-tag II	Strep-Tactin® Sepharose®	IBA Lifesciences GmbH, Göttingen, Germany
Strep-tag II	Strep-Tactin®XT 4Flow® high capacity resin	IBA Lifesciences GmbH, Göttingen, Germany
His-tag	TALON® Superflow™	Cytiva, Marlborough (MA), USA
Cuvettes		
	Semi-micro cuvettes, 1.6 ml, 10 mm	ratiolab®, Dreieich, Germany
	Quarz SUPRASIL High Precision Cell QS100	Hellma Analytics, Müllheim, Germany
Dialysis tube	SERVAPOR® dialysis tubing, MWCO 12000–14000	SERVA Electrophoresis GmbH, Heidelberg, Germany

DNA loading dye	Gel Loading Dye, Purple (6x)	New England Biolabs GmbH, Ipswich (MA), USA
DNA marker	GeneRuler™ 1 kb Plus DNA Ladder	Thermo Fisher Scientific Inc.,Waltham (MA), USA
Protein marker		
	Unstained Protein Standard, Broad Range	New England Biolabs GmbH, Ipswich (MA), USA
	Color Prestained Protein Standard, Broad Range	New England Biolabs GmbH, Ipswich (MA), USA
	Blue Prestained Protein Standard, Broad Range	New England Biolabs GmbH, Ipswich (MA), USA
Imaging plate	BAS-IP MS 2040	Fujifilm, Tokyo, Japan
Protein standards SEC		
	β-Amylase	Sigma-Aldrich, St. Louis (MI), USA
	Albumin	Sigma-Aldrich, St. Louis (MI), USA
	Apoferritin	Sigma-Aldrich, St. Louis (MI), USA
	Alcoholdehydrogenase	Sigma-Aldrich, St. Louis (MI), USA
	Carboanhydrase	Sigma-Aldrich, St. Louis (MI), USA
	Blue dextran	Sigma-Aldrich, St. Louis (MI), USA
Sterile filter		
	0.22 µm Syringe Filter, PVDF (Sterile), Blue, 33 mm	Starlab International GmbH, Hamburg, Germany
	0.45 µm Syringe Filter, PVDF (Sterile), Yellow, 33 mm	Starlab International GmbH, Hamburg, Germany
Transfer membrane	Transfer membrane ROTI®PVDF 0.45 µm	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Kits		
PCR Clean-up	NucleoSpin® Gel and PCR Clean-up kit	Macherey-Nagel GmbH & Co. KG, Düren, Germany
Plasmid preparation		
	NucleoSpin® Plasmid EasyPure kit	Macherey-Nagel GmbH & Co. KG, Düren, Germany
	E.Z.N.A.® Plasmid DNA Mini Kit I	Omega Bio-Tek Inc, Norcross (GA), USA
Genomic DNA isolation	NucleoSpin® Microbial DNA kit	Macherey-Nagel GmbH & Co. KG, Düren, Germany
Kinase Assay	ADP-Glo™ Kinase Assay	Promega Corporation, Madison (WI), USA

2.2 Bacterial strains, plasmids, and oligonucleotides

2.2.1 Bacterial strains

Table 2.3: Bacterial and archaeal strains used in this study.

Strain	Genotype	Reference
E. coli Strains		
E. coli DH5α	F- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*) U169, *hsd*R17(r_K^- m_K^+), λ–	(Hanahan, 1983)
E. coli BL21 (DE3)	F- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S)	(Studier & Moffatt, 1986)
E. coli Nissle 1917	serotype O6:K5:H1	(Grozdanov *et al.*, 2004)
E. coli C43 (DE3)	F- *omp*T *hsdS*B(r_B^- m_B^-) *gal dcm* (DE3)	(Miroux & Walker, 1996)
E. coli C41 (DE3)	F- *omp*T *gal dcm hsd*SB(r_B^- m_B^-) (DE3)	(Miroux & Walker, 1996)
E. coli TKR 2000	Δ*kdpFABCDE trkA405 trkD1 atp706*	(Kollmann & Altendorf, 1993)
M. acetivorans Strains		
M. acetivorans WWM73	Δ*hpt*::P*mcrB-tetR*-ϕC31-*int*-attP	(Guss *et al.*, 2008)

2.2.2 Plasmids

Table 2.4: Plasmids used in this study.

Plasmid	Characteristic	Reference
pASK-IBA3	Expression vector, heterologous overexpression in *E. coli*, Strep-tag II, *tet* promotor, Amp^R	IBA Lifesciences GmbH
pACYCDuet-1	Expression vector, heterologous overexpression and coexpression in *E. coli*, His-tag and S-tag, T7 promoter, Cm^R	Novogene Co., Ltd.
pET21a(+)	Expression vector, heterologous overexpression in *E. coli*, T7-tag and His-tag, T7 promotor, Amp^R	Novogene Co., Ltd.
pASK-IBA3-MA_4377	pASK-IBA3 derivate, coding region of MA_4377 from *M. acetivorans* at *Sac*II/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	(Sexauer, 2021)
pASK-IBA3-MA_4377-CHPK	pASK-IBA3 derivate, coding region of MA_4377 (CHASE-HK domain) from *M. acetivorans* at *Bam*HI/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pASK-IBA3-MA_4377-PKR1R2	pASK-IBA3 derivate, coding region of truncated cytosolic MA_4377 (HK-R2 domain) from	(Sexauer, 2021)

	M. acetivorans with C-terminal Strep-tag II, *tet* promoter, Amp^R	
pASK-IBA3-MA_4377-$PK_{H497Q}R1_{D818N}R2$	pASK-IBA3-MA_4377-PKR1R2 derivate with amino acid exchange: His497 to Gln and Asp818 to Asn, obtained by site-directed mutagenesis	This study
pASK-IBA3-MA_4377-PKR1	pASK-IBA3 derivate, coding region of truncated cytosolic MA_4377 (HK-R1 domain) from *M. acetivorans* at *Bam*HI/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pASK-IBA3-MA4_377-$PK_{H497Q}R1$	pASK-IBA3-MA4_377-PKR1 derivate with amino acid exchange: His497 to Gln, synthetic gene	This study
pASK-IBA3-MA4_377-PK	pASK-IBA3 derivate, coding region of truncated cytosolic MA_4377 (only HK domain) from *M. acetivorans* with C-terminal Strep-tag II, *tet* promoter, Amp^R	(Sexauer, 2021)
pASK-IBA3-MA4_377-PK_{H497Q}	pASK-IBA3-MA_4377-PK derivate with amino acid exchange: His497 to Gln, synthetic gene	This study
pASK-IBA3-MA_4377-R1	pASK-IBA3 derivate, coding region of truncated cytosolic MA_4377 (only R1 domain) from *M. acetivorans* at *Xba*I/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pASK-IBA3-MA_4377-R2	pASK-IBA3 derivate, coding region of truncated cytosolic MA_4377 (only R2 domain) from *M. acetivorans* at *Xba*I/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pASK-IBA3-MA_0863(rdmS)$_{O216K}$	pASK-IBA3 derivate, coding region of MA_0863 (rdmS) with amino acid exchange (Pyl216 to Lys) from *M. acetivorans* at *Sac*II/*Pst*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	(Kwiatkowski, 2013)
pASK-IBA3-MA_2082	pASK-IBA3 derivate, coding region of MA_2082 from *M. acetivorans* at *Bam*HI/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pASK-IBA3-MA_2013	pASK-IBA3 derivate, coding region of MA_2013 from *M. acetivorans* at *Bam*HI/*Nco*I with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pACYC-duet1-MA_2013-PK	pACYCDuet-1 derivate, coding region of truncated MA_2013 (without R1 and HPt) from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study

pASK-IBA3-MA_2013-R1	pASK-IBA3 derivate, coding region of truncated MA_2013 (only R1) from *M. acetivorans* with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pASK-IBA3-MA_2013-HPT	pASK-IBA3 derivate, coding region of truncated MA_2013 (only HPt) from *M. acetivorans* with C-terminal Strep-tag II, *tet* promoter, Amp^R	This study
pACYC-duet1-MA_0016	pACYCDuet-1 derivate, coding region of MA_0016 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_0018	pACYCDuet-1 derivate, coding region of MA_0018 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_1268	pACYCDuet-1 derivate, coding region of MA_1268 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_1269	pACYCDuet-1 derivate, coding region of MA_1269 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_1366	pACYCDuet-1 derivate, coding region of MA_1366 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_1468	pACYCDuet-1 derivate, coding region of MA_1468 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_1469	pACYCDuet-1 derivate, coding region of MA_1469 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_2445	pACYCDuet-1 derivate, coding region of MA_2445 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_2012	pACYCDuet-1 derivate, coding region of MA_2012 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_2861	pACYCDuet-1 derivate, coding region of MA_2861 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pACYC-duet1-MA_3068	pACYCDuet-1 derivate, coding region of MA_3068 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study

pACYC-duet1-MA_4376 (R3)	pACYCDuet-1 derivate, coding region of MA_4376 from *M. acetivorans* at *Sac*I/*Sal*I with N-terminal His-tag, T7 promoter, Cm^R	(Sexauer, 2021)
pACYC-duet1-MA_4671	pACYCDuet-1 derivate, coding region of MA_4671 from *M. acetivorans* at *Bam*HI/*Nde*I with N-terminal His-tag, T7 promoter, Cm^R	This study
pET21a(+)-MA_4375 (*msrX*)	pET21a(+) derivate, coding region of MA_4375 from *M. acetivorans* with C-terminal His-tag, T7 promoter, Amp^R	(Sexauer, 2021)
pET21a(+)-MA_4384 (*msrC*)	pET21a(+) derivate, coding region of MA_4384 from *M. acetivorans* with C-terminal His-tag, T7 promoter, Amp^R	(Blasius, 2023)

2.2.3 Oligonucleotides

Table 2.5: Oligonucleotides used in this study.

Oligonucleotides	Sequence (5' → 3')
Oligonucleotides for sequencing	
pASK-IBA3-seq-fwd	CGCAGTAGCGGTAAACG
pASK-IBA3-seq-rev	GAGTTATTTTACCACTCCCT
pACYC-duet1-seq-fwd	TCTCCCTTATGCGACTCCTG
pACYC-duet1-seq-rev	GGGTTATGCTAGTTATTGCTCAGC
Oligonucleotides for expression plasmids	
pASK-IBA3-MA_4377-fwd	GCTATCCGCGGCATGAATGTGAGTAGAAAAATTCT
pASK-IBA3-MA_4377-rev	GCATACCATGGCCTCTTCGACAATAAGAATTTCTTTC
pASK-IBA3-MA_4377-CHPK-fwd	ATTCGAGCTCGGTACCCGGGATATGAATGTGAGTAGAAAAATTC
pASK-IBA3-MA_4377-CHPK-rev	GTGGCTCCAAGCGCTGAGACTCTGAGCTTCCCTATTTTC
pASK-IBA3-MA_4377-PKR1R2-fwd	GCTATCCGCGGCAGGTTGAATTCCGATAAGGTTAA
pASK-IBA3-MA_4377-PKR1R2-rev	GCATACCATGGCCTCTTCGACAATAAGAATTTCTTTC
pASK-IBA3-MA_4377-PKR1-fwd	ATTCGAGCTCGGTACCCGGGCTAGGTTGAATTCCGATAAG
pASK-IBA3-MA_4377-PKR1-rev	GTGGCTCCAAGCGCTGAGACGGAAACTGAACTTACCTG
pASK-IBA3-MA_4377-PK-fwd	GCTATCCGCGGCAGGTTGAATTCCGATAAGGTTAA
pASK-IBA3-MA_4377-PK-rev	GCGCCATGGCCTGTGAGGGGAATTG
pASK-IBA3-MA_4377-PK (QC)	ATTCGAGCTCGGTACCCGGGCTATGAGGTTGAATTCCG
pASK-IBA3-MA_4377-PK (QC)	GTGGCTCCAAGCGCTGAGACGCTTTGTGAGGGGAATTG

pASK-IBA-MA_4377-R1-fwd	ATGAATAGTTCGACAAAAATTAGTCTTGTGCTTGTAGTC
pASK-IBA3-MA_4377-R1-rev	GTGGCTCCAAGCGCTGAGACTGAAACTGAACTTACCTG
pASK-IBA3-MA_4377-R2-fwd	ATGAATAGTTCGACAAAAATTGGTAAGTTCAGTTTCGAC
pASK-IBA3-MA_4377-R2-rev	GTGGCTCCAAGCGCTGAGACTTCTTCGACAATAAGAATTTCTTTC
pASK-IBA3-MA_0863(rdmS)$_{O216K}$-fwd	GGACCGCGGAAAAAAGTTCGATATAAATC
pASK-IBA3-MA_0863(rdmS)$_{O216K}$-rev	GGACTGCAGTTATTTTTCGAACTCCGGGTGGCTCCAAGACTCTATTTGATTCATG
pASK-IBA3-MA_2082-fwd	ATTCGAGCTCGGTACCCGGGATATGGAAGGAAAACTGAGAAATTCTGGTATTGAG
pASK-IBA3-MA_2082-rev	GTGGCTCCAAGCGCTGAGACTCTCGACGCCCACAGGGA
pASK-IBA3-MA_2013-fwd	ATTCGAGCTCGGTACCCGGGAAGTGTGCGGTTTAAAAAAAC
pASK-IBA3-MA_2013-rev	GTGGCTCCAAGCGCTGAGACGACAGATCTGTTTTTCCAG
pACYC-duet1-MA_2013-PK-fwd	ACCATCATCACCACAGCCAGATGTGCGGTTTAAAAAAAC
pACYC-duet1-MA_2013-PK-rev	TATCCAATTGAGATCTGCCACTAGCATACTAACAGGCTTTC
pASK-IBA3-MA_2013-R1-fwd	ATTCGAGCTCGGTACCCGGGCGATGGTGAAAAGAGGGG
pASK-IBA3-MA_2013-R1-rev	GTGGCTCCAAGCGCTGAGACCTATACTTTCCAGGTCCAG
pACYC-duet1-MA_0016-fwd	ACCATCATCACCACAGCCAGATGGCAAGAGTAATGATC
pACYC-duet1-MA_0016-rev	TATCCAATTGAGATCTGCCATTAATTTGGTTCAGCAATTTTTTTAATTAC
pACYC-duet1-MA_0018-fwd	ACCATCATCACCACAGCCAGATGCCTGAAATCCTGATC
pACYC-duet1-MA_0018-rev	TATCCAATTGAGATCTGCCATCAACCAGGAGAATTCGTTG
pACYC-duet1-MA_1268-fwd	ACCATCATCACCACAGCCAGATGAAAACAAAGGTAGCTG
pACYC-duet1-MA_1268-rev	TATCCAATTGAGATCTGCCATTATTTTGACGGTAGTTTCAC
pACYC-duet1-MA_1269-fwd	ACCATCATCACCACAGCCAGATGGAAACATGGACGGCATTTAAAC
pACYC-duet1-MA_1269-rev	TATCCAATTGAGATCTGCCATTAAGTAAGATTTACGACTTCAAGCC
pACYC-duet1-MA_1366-fwd	ACCATCATCACCACAGCCAGATGCATAACATCAATCTTGAC
pACYC-duet1-MA_1366-rev	TATCCAATTGAGATCTGCCATCAAATCCGACTACGTCTC
pACYC-duet1-MA_1468-fwd	ACCATCATCACCACAGCCAGATGGACAAAGCAAAGATTC
pACYC-duet1-MA_1468-rev	TATCCAATTGAGATCTGCCATTATTCTTCAGTTACTTTACCC
pACYC-duet1-MA_1469-fwd	ACCATCATCACCACAGCCAGATGAAAAAAGCAAAAATTCTGGTTGTTG
pACYC-duet1-MA_1469-rev	TATCCAATTGAGATCTGCCATCATTCGTCGCTGTCGGAG
pACYC-duet1-MA_2012-fwd	ACCATCATCACCACAGCCAGATGAAAGTTCTGATTGCC
pACYC-duet1-MA_2012-rev	TATCCAATTGAGATCTGCCATTAAAAGCCCAGCTTTTTTC
pACYC-duet1-MA_2445-fwd	ACCATCATCACCACAGCCAGATGTTGAATGAGAAAAACAAAC

pACYC-duet1-MA_2445-rev	TATCCAATTGAGATCTGCCATTATGCTGCTTTTATGATAAATTC
pACYC-duet1-MA_2861-fwd	ACCATCATCACCACAGCCAGATGAAAGTACTGCTTGTAG
pACYC-duet1-MA_2861-rev	TATCCAATTGAGATCTGCCATTAACCTGAACCATTTTTAAC
pACYC-duet1-MA_3068-fwd	ACCATCATCACCACAGCCAGATGGCAAAAGTCTTGATTG
pACYC-duet1-MA_3068-rev	TATCCAATTGAGATCTGCCATTATGCACCGATTACTTTC
pACYC-duet1-MA_4376(R3)-fwd	GCGTCGAGCTCGAAAGAAATTCTTATTGTCG
pACYC-duet1-MA_4376(R3)-rev	CCCTTAGTCGACTTCTCCCAGATATTTATCG
pACYC-duet1-MA_4671-fwd	ACCATCATCACCACAGCCAGATGAAAAGTATACTTGTAGTTGAG
pACYC-duet1-MA_4671-rev	TATCCAATTGAGATCTGCCATCACAATACCTGGTAACC
pET21a(+)-MA_4375(*msrX*)-fwd	CTTTAAGAAGGAGATATACAATGAAAATATCACTTATTGACCTG
pET21a(+)-MA_4375(*msrX*)-rev	AGTGGTGGTGGTGGTGGTGCTTTATTTTGTCCAGCCGC
Oligonucleotides for site directed mutagenesis	
MA_4377-H497Q-fwd	GAGTTCCTTGCCAATATGAGCCAGGAGCTCCGGACGCCTCTG
MA_4377-H497Q-rev	CAGAGGCGTCCGGAGCTCCTGGCTCATATTGGCAAGGAACTC
MA_4377-D818N-fwd	CCTGATGTTATCACTCTGAATGTTTTACTTCCCGATACTAGTGGC
MA_4377-D818N-rev	GCCACTAGTATCGGGAAGTAAAACATTCAGAGTGATAACATCAGG
MA_4377-D941N-fwd	CAGCAGCCGGATATCCTCATCCTCAACCTGCTGATGCCCGAGATAAGT
MA_4377-D941N-rev	ACTTATCTCGGGCATCAGCAGGTTGAGGATGAGGATATCCGGCTGCTG

2.3 Microbiological methods

2.3.1 Sterilisation

Glassware, plastic labware, media and supplements used for cultivation of microbial cultures were sterilised prior to use. Glassware was sterilised by dry heat at 180 °C for 3 h, while heat stable plastic labware and media were autoclaved for 20 min at 121 °C and 1 bar. Heat-sensitive supplements were sterilised by sterile filtration.

2.3.2 Culture media and supplements

E. coli was cultivated in LB medium (Lennox) or on LB agar (Lennox). For LB media/LB agar either granulated media was dissolved in MilliQ H_2O according to the manufacturer's instructions (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) or it was prepared according to Table 2.6.

The cultivation of *M. acetivorans* was carried out under anaerobic conditions in autoclaved HS medium. To prepare 1 l of HS medium, 100 ml of 10x Mix A and 100 ml of 10x Mix B were mixed, subsequently 1 ml 500x trace elements and 1 ml 500x vitamin solution were added (Table 2.6). The solution was filled to a total volume of 1 l with H_2O dest. in a 3 l round-bottom flask. After the addition of 20 mM MOPS, the pH value was adjusted to 7. The solution was boiled for 10–20 min under constant gas flow (N_2) until the colour changed from blue to violet. Afterwards the medium was cooled down for at least 30 min under constant gas flow and introduced into the anaerobic tent (Coy Laboratory Products Inc.). Inside of the anaerobic tent, NH_4Cl and cysteine HCl were added one after the other until the colour of the redox indicator resazurin changed from dark blue to pink. As soon as the colour changed to pink, Na_2S was added to the medium, with this step the media colour changed to a light yellow. If necessary other additives were added to the medium, for standard medium 125 mM Methanol was added. The medium was filled up to 1 l with anaerobic H_2O dest. The medium was aliquoted into serum bottles or anaerobic culture tubes (Balch tubes), sealed with gas-tight butyl plugs and aluminium crimp caps and removed from the anaerobic tent. The gas phase was then adjusted by the addition to N_2/CO_2 (80:20 [v/v]) for 20 seconds.

Table 2.6: Culture media, supplements and antibiotics used in this study.

	Component	**Stock solution**	**Final volume/ concentration per 1 l medium**
LB medium			
	Tryptone		10 g/l
	yeast extract		5 g/l
	NaCl		5 g/l
	Agar-Agar (if appropriate)		15 g/l
HS medium			
10x Mix A	NaCl		400 mM
	KCl		13 mM
	0.1 % (w/v) Resazurin		4 µM
	500 x trace elements		1 ml
	500 x vitamin solution		1 ml
10x Mix B	$MgCl_2$ x 6 H_2O		54 mM
	$CaCl_2$ x 2 H_2O		2 mM
500x trace elements*	Nitrilotriacetic acid	58 mM	58 µM
	$FeSO_4$ x 7 H_2O	20 mM	20 µM
	Na_2SeO_3 x 5 H_2O	11 mM	11 µM
	$CoCl_2$ x 6 H_2O	4 mM	4 µM
	$MnSO_4$ x H_2O	6 mM	6 µM
	Na_2MoO_4 x H_2O	4 mM	4 µM
	Na_2WO_4 x 2 H_2O	3 mM	3 µM
	$ZnSO_4$ x H_2O	3 mM	3 µM

	$NiCl_2$ x 6 H_2O	12 mM	12 µM
	H_3BO_3	1.6 mM	1.6 µM
	$CuSO_4$	0.4 mM	0.4 µM
*500x concentrated in 200 ml H_2O dest., pH 6.5 with KOH adjusted, autoclaved			
500x vitamin solution**	p-Aminobenzoic acid	364.5 µM	729 nM
	Nicotinic acid	406 µM	812 nM
	Ca-Pantothenate	209.5 µM	419 nM
	Pyridoxine-HCl	243 µM	486 nM
	Riboflavin	133 µM	266 nM
	Thiamine-HCl	148 µM	296 nM
	Biotin	102 µM	204 nM
	Folic acid	56.5 µM	113 nM
	a-Lipoic acid	121 µM	242 nM
	Vitamin B12	18.5 µM	37 nM
**500x concentrated in 200 ml H_2O dest., filter sterilized			
Supplements			
Anhydrotetracycline (AHT)		2 mg/ml	200 ng/ml
Isopropyl-β-D-thiogalactopyranosid (IPTG)		1 M	0.5 mM
Antibiotics			
Ampicillin (Amp) in H_2O		100 mg/ml	100 µg/ml
Chloramphenicol (Cm) in Ethanol		34 mg/ml	34 µg/ml

2.3.3 Storage of *E. coli* and *M. acetivorans*

Long term storage of *E. coli* was carried out at -80 °C in glycerol cultures. For glycerol cultures 5 ml of a liquid culture was centrifuged, and the supernatant was discarded. The pellet was subsequently resuspended in fresh LB medium supplemented with an appropriate antibiotic. A volume of 600 µl of this culture was mixed with 400 µl sterile glycerol.

M. acetivorans was stored in 10 ml Balch tubes at room temperature up to one year. Additional glycerol cultures as backup were prepared. For this 1 ml anaerobic culture was mixed with 3 ml of sterile glycerol. The mixture was stored in anaerobic Balch tubes at -80 °C.

2.3.4 Cultivation of *E. coli* and *M. acetivorans*

E. coli cells were cultivated either on solid medium or in liquid medium. For plates, LB Agar was poured in petri dishes, if necessary, antibiotics were added to the agar. *E. coli* cells were spread out on solid medium with an inoculation loop or a drigalski spatula. Plates were incubated over night at 37 °C. For liquid cultivation, pre-cultures were prepared in LB medium. A single colony from a plate was used to inoculate a small amount of LB medium in either a test tube or an Erlenmeyer flask depending on the experiment. Pre-cultures were grown overnight at 37 °C with constant shaking. For the main culture, medium supplemented with

additives was prepared and inoculated 1:100 with the pre-culture. Depending on the experiment, the cultures were grown for some hours or overnight (oN) at 37 °C.

Cultivation of *M. acetivorans* was performed in 10 ml HS medium in a Balch tube or in 50 ml HS medium in serum bottles. Standard HS medium was prepared with 125 mM Methanol. Cultures were inoculated with a sterile cannula, which was rinsed with N_2/CO_2 (80:20 [v/v]) before use. The cultures were incubated at 37 °C without shaking in darkness.

2.3.5 Determination of cell density

To determine the cell density of a microbial culture the optical density (OD) was determined photometrically at a wavelength of 600 nm (OD_{600}). The respective culture medium was used as reference. For *E. coli* the OD was measured with single use cuvettes, for *M. acetivorans,* the optical density was either measured with cuvettes or in Balch tubes.

2.4 Molecular biological methods

2.4.1 Preparation of chemical competent cells

To take up foreign DNA, bacterial cells must be made competent if they do not have natural competence. For *E. coli* cells this can be induced by chemical treatment. All used *E. coli* strains were treated with the following protocol.

100 ml LB medium was inoculated 1:100 with a pre-culture of the respective *E. coli* strain. The culture was incubated at 37 °C under constant shaking at 160 rpm to an OD_{600} of 0.5-0.7. After reaching the desired OD_{600}, 100 ml of the culture was transferred to two sterile 50 ml reaction tubes and then centrifuged for 15 min at 4 °C and 2000 g. After centrifugation, the supernatant was discarded, and the cell pellet was carefully resuspended in 25 ml ice-cold solution I (Table 2.7) and then incubated on ice for one hour. After incubation, the mixture was centrifuged for 10 min at 4 °C and 2000 g. The supernatant was then discarded again, and each cell pellet was resuspended in 2.5 ml ice-cold solution II (Table 2.7). Finally, 200 µl of each of the prepared competent cells were transferred to sterile 1.5 ml reaction tubes and snap-frozen with liquid nitrogen. The competent cells were stored at -80 °C.

Table 2.7: Solution for chemical competent cells.

Solution I	
$CaCl_2$	50 mM
Solution II	
$CaCl_2$	50 mM
Glycerol	15 % (v/v)

2.4.2 Transformation of chemical competent cells

For the transformation of chemical competent *E. coli* cells, aliquots were thawed on ice and mixed gently with 50-100 ng DNA or 15 µl of Gibson assembly reaction mix and incubated on ice for 30 min.

After incubation on ice, the mixture was subjected to a heat shock at 42 °C for 2 min. After the heat shock, the mixture was placed on ice for 2 min. Subsequently, 700 µl of sterile LB medium was added to each sample and the mixture was incubated at 37 °C for 1 h with shaking at 160 rpm. After the incubation time, 100 µl of the mixture was plated with a drigalski spatula on an LB agar plate with the respective antibiotic. In case of a difficult transformation the rest of the mixture was centrifuged (3000 g, RT, 3 min) and a part of the supernatant was discarded. The cell pellet was resuspended in the remaining media (around 100 µl) and plated as aforementioned. As a control, the equal volume of sterile water was mixed with chemical competent cells and treated in the same way. Additionally, a growth control was conducted, for this the chemical competent cells were plated on LB agar without antibiotics. All plates were incubated overnight at 37 °C.

2.4.3 Preparation of plasmid DNA

The NucleoSpin® Plasmid Easy Pure Kit or the E.Z.N.A.® Plasmid DNA Mini Kit I were used to isolate plasmid DNA. The procedure was carried out according to the manufacturer's instructions (Macherey-Nagel GmbH & Co. KG or Omega Bio-Tek Inc.).

2.4.4 Preparation of genomic DNA from *M. acetivorans*

For the isolation of genomic DNA from *M. acetivorans*, 50 ml of a culture with 125 mM methanol was incubated at 37 °C until stationary phase was reached, the cells were collected for 10 min at 3500 g (centrifuge 5415 D, rotor: FS 45-24-11). Subsequently, 0.5 g of the cells were resuspended in 150 µl Tris buffer (50 mM Tris/HCl, pH 7). All further steps were performed with the NucleoSpin® Microbial DNA Mini Kit according to the manufacturer's instructions (Macherey-Nagel GmbH & Co. KG).

2.4.5 Agarose gel electrophoresis

Agarose gel electrophoresis is used to separate DNA fragments according to their size. The negatively charged DNA moves from the cathode to the anode in an electric field. Smaller fragments pass through the gel matrix faster than larger fragments (Flint & Harrington, 1972). To prepare the agarose gels, an appropriate amount of agarose was mixed with an appropriate volume of 1x TAE running buffer, boiled, and subsequently poured into a gel slide. Depending on the purpose, the agarose concentration was adjusted, standard agarose gels were prepared with 0.8 % agarose concentration. The DNA samples were mixed with 6x DNA sample buffer (New England Biolabs GmbH) and then applied to the gel. The DNA was

separated at a constant voltage of 80–120 V. The gel was then incubated in an ethidium bromide solution for 5 min before the DNA signals were visualised by irradiation with UV light (312 nm) in the gel documentation system (INTAS Science Imaging Instruments GmbH). The GeneRuler™1 kb Plus DNA Ladder (Thermo Fisher Scientific Inc.) was used as the size standard.

Table 2.8: TAE buffer for agarose gel electrophoresis.

50x TAE buffer, pH 8.0	
Tris-acetate	2 M
EDTA	50 mM

2.4.6 Restriction digest of DNA

For restriction digest of PCR products or plasmids, restriction endonucleases (New England Biolabs GmbH) were used according to the manufacturer's instructions. The reaction mix for a test restriction with a total volume of 10 µl contained 1 µl of the appropriate buffer (10x Cutsmart or 10x rCutsmart), 1 µl of the respective enzyme and 100-300 ng DNA. For preparative DNA digest, the reaction mix was adjusted to a higher amount of DNA. Incubation time of the reaction mix was dependent on the use of restriction endonuclease. With a standard enzyme the incubation was, 1 h at 37 °C, while for HF-endonucleases the incubation time was 20-30 min. After incubation the enzymes were heat-inactivated if possible. Successful digestion was verified via agarose gel electrophoresis.

2.4.7 Polymerase chain reaction

The polymerase chain reaction (PCR) is a method used to amplify DNA fragments (Mullis *et al.*, 1986). For the reaction the Phusion DNA polymerase was used. The polymerase was purified from the microbiology group (RPTU Kaiserslautern) or purchased from New England Biolabs GmbH. The elongation time was adjusted individually for every reaction according to the size of the DNA fragment. The annealing temperature was adjusted according to the employed primers (approx. T_m – 5 °C).

To verify the successful cloning of expression plasmids (chapter 2.4.11), the colony PCR method was used. Single colonies after a *E. coli* transformation were picked, streaked onto a respective LB agar plate and resuspended in 10 µl MilliQ H_2O. For the PCR reaction 1 µl of this solution was used as the template. To prove the correct insertion of the respective gene, the oligonucleotides used for the Gibson assembly were used. PCR reactions were analysed via agarose gel electrophoresis.

Table 2.9: PCR composition.

PCR (Amplification of DNA)	
Component	**Final concentration**
Template DNA	10–100 ng
Primer fwd (20 pmol/µl)	1.25 µl
Primer rev (20 pmol/µl)	1.25 µl
dNTPs (mix) (10 µM)	1 µl
Phusion HF polymerase	1 µL
5x Phusion HF buffer	1x
MilliQ H_2O	to 50 µL
Colony PCR	
Component	**Final concentration**
Template DNA	10–100 ng
Primer fwd (20 pmol/µl)	1.25 µl
Primer rev (20 pmol/µl)	1.25 µl
dNTPs (mix) (10 µM)	1 µl
Phusion HF polymerase	1 µL
5x Phusion HF buffer	1x
MilliQ H_2O	to 50 µL

Table 2.10: PCR program.

PCR (Amplification of DNA)			
Reaction step	**temperature**	**time**	
Initial denaturation	98 °C	30 sec	
Denaturation	95 °C	30 sec	
Annealing	primer-specific	30 sec	30 cycles
Extension	72 °C	20 sec/kb	
Final extension	72 °C	5 min	
Cooling	4 °C	∞	
Colony PCR			
Reaction step	**temperature**	**time**	
Initial denaturation	95 °C	10 min	
Denaturation	95 °C	30 sec	
Annealing	primer-specific	30 sec	30 cycles
Extension	72 °C	20 sec/kb	

Final extension	72 °C	5 min
Cooling	4 °C	∞

2.4.8 Purification of PCR products

The purification of DNA fragments amplified by PCR was performed using the NucleoSpin® Gel and PCR Purification Kit (Macherey-Nagel GmbH & Co. KG) according to the manufacturer's instructions on the PCR clean-up protocol.

2.4.9 Extraction of DNA fragments from agarose gels

The extraction of DNA fragments amplified by PCR was performed using the NucleoSpin® Gel and PCR Purification Kit (Macherey-Nagel GmbH & Co. KG) according to the manufacturer's instructions on the Gel-clean-up protocol.

2.4.10 Determination of nucleic acid concentrations

To determine the concentration of plasmid or genomic DNA a NanoDrop™ Lite photometer (Thermo Fisher Scientific Inc.) was used. The absorption at 260 nm of 1 µl of a DNA solution was measured. As a reference 1 µl MilliQ H_2O or elution buffer was used. An absorption of 1 corresponds to a DNA concentration of 50 ng/µl for double-stranded DNA. The absorption ratio of $A_{260nm}/A_{280\ nm}$ indicated the purity of the measured sample. A purity of 1.8 was considered as sufficient pure.

As an alternative, the concentration of an DNA sample was measured with an agarose gel (chapter 2.4.5) and an appropriate DNA marker as comparison.

2.4.11 Gibson cloning

The Gibson assembly method is used to assemble up to five DNA fragments at once and is therefore simpler than traditional cloning based on restriction digest. All expression vectors of this study were constructed using the Gibson assembly method (Gibson *et al.*, 2009).

The primers were designed with the NEBuilder online tool (https://nebuilder.neb.com/#!/). The genes of interest were amplified from genomic DNA (*M. acetivorans*) or ordered as synthetic genes from Twist Bioscience (San Francisco (CA), USA). The desired genes were amplified with the appropriate primers (chapter 2.4.7), resulting in a DNA fragment with specific overhangs. In the meantime, the corresponding plasmid was digested with the appropriate restriction endonucleases (chapter 2.4.6). The PCR and restriction digest were validated with an agarose gel electrophoresis (chapter 2.4.5) and the concentration was estimated using the NanoDrop™ Lite photometer (chapter 2.4.10). The mass of insert which was used for the ligation was calculated with the NEBioCalculator (https://nebiocalculator.neb.com/#!/ligation). The ratio of insert:vector was chosen between 3:1 or 5:1. The calculated amount of vector and

insert (5 µl in total) was mixed with 15 µl of Gibson Assembly Master Mix. The Reaction mix was incubated at 50 °C for 1 h, subsequently 10 µl of the reaction mix was used for a transformation in chemical competent *E. coli* DH5α cells.

Table 2.11: Gibson Assembly Master Mix components.

Component	Final volume
5x isothermal reaction buffer (500 mM Tris/HCl pH 7.5, 50 mM $MgCl_2$, 50 mM DTT, 5 mM NAD^+, 1 mM (each) dNTPs, 25 % w/v PEG 8000)	320 µl
Phusion DNA polymerase (2 U/µl)	20 µl
Taq DNA ligase (40 U/µl)	160 µl
T5 exonuclease (10 U/µl)	0.64 µl
MilliQ H_2O	Add 1.2 µl

2.4.12 Site directed mutagenesis

Site-directed mutagenesis (QuikChange) is used to introduce a specific mutation in a DNA sequence to analyse the resulting effect on the structure and/or function of the encoded protein. Specific primers to create base exchanges within the gene were designed manually. The plasmid with the corresponding gene without mutation was used as template for the PCR. The PCR was carried out in two steps: First the reaction mix A and reaction mix B were subjected to a first PCR separately. Afterwards, 25 µl of reaction mix A was mixed with reaction mix B and 1 µl of Phusion polymerase was added. This new reaction mix was then used for the second PCR reaction. Subsequently, the methylated parental plasmid was removed with a restriction digest using *Dpn*I. 10 µl of the digested reaction mix was transformed in *E. coli* DH5α. After transformation the desired plasmid was isolated from single colonies. For some constructs a silent site mutation was inserted, the correct insertion of the mutation were validated in this case via restriction digest. The correct insertion of the mutation was always verified using Sanger sequencing (chapter 2.4.13).

Table 2.12: Reaction mix for site directed mutagenesis PCR.

	Reaction mix A	Reaction mix B
MilliQ H_2O	35 µl	35 µl
5x HF buffer	10 µl	10 µl
Template	1 µl	1 µl
Primer fwd (20pmol/µl)	1.25 µl	/
Primer rev (20pmol/µl)	/	1.25 µl
dNTPs	1 µl	1 µl

Phusion polymerase	1 µl	1 µl
total	50 µl	50 µl

Table 2.13: PCR program for site directed mutagenesis PCR.

PCR 1			
Reaction step	temperature	time	
Initial denaturation	98 °C	30 sec	
Denaturation	98 °C	15 sec	
Annealing	55 °C	30 sec	8 cycles
Extension	72 °C	20 sec/kb	
Final extension	72 °C	5 min	
Cooling	4 °C	∞	
PCR 2			
Reaction step	temperature	time	
Initial denaturation	98 °C	30 sec	
Denaturation	98 °C	15 sec	
Annealing	55 °C	30 sec	18 cycles
Extension	72 °C	20 sec/kb	
Final extension	72 °C	10 min	
Cooling	4 °C	∞	

2.4.13 Sequencing

To verify expression vectors and plasmids with desired mutations generated by site-directed mutagenesis, sequence analyses were carried out using Sanger sequencing (Sanger *et al.*, 1977). The sequencing was carried out by Eurofins Genomics Germany (Ebersberg, Germany). Plasmid DNA or PCR products were mixed with the corresponding oligonucleotides and MilliQ H_2O according to the company's instructions. The sequencing results were analysed using SnapGene® 4.0.8 (San Diego (CA), USA).

2.5 Protein biochemical methods

2.5.1 Test expression of recombinant genes to determine conditions for protein production

To determine the optimal conditions (temperature, quantity of inducer, time, and *E. coli* strain) for protein production in *E. coli,* test expression experiments were conducted for every construct. A main culture with a final volume of 50–250 ml, with the addition of an appropriate

antibiotic, was inoculated 1:100 with an overnight culture. The main culture was incubated at 37 °C under constant shaking at 160 rpm until an OD_{600} of 0.5–0.7 was reached. Once the desired OD was reached, a sample (before induction) was taken. The samples were adjusted to a certain OD, to allow a comparison between the samples from different timepoints. After sampling the culture was cooled down to the appropriate temperature and the gene expression was induced with AHT (200 ng/ml) or IPTG (0.5 mM) depending on the promoter system. To test the solubility and the yield of the protein, the expression of the corresponding gene was tested under different conditions (30 °C for 3 h, 17 °C oN or 25 °C oN). Samples were taken after the induction at different timepoints and handled as mentioned before. Depending on the experimental question, the samples were either directly resuspended with SDS sample buffer or resuspended with lysis buffer and sonicated to get a soluble and insoluble fraction which was afterwards mixed with SDS sample buffer. All samples were heated (95 °C, 10 min), separated by SDS-PAGE (chapter 2.5.8) and analysed via western blot (chapter 2.5.9).

2.5.2 Heterologous production of recombinant proteins in *E. coli*

After the determination of the optimal conditions, a heterologous protein production was performed. For this a pre-culture of the *E. coli* strain harbouring the appropriate expression plasmid was grown in LB media with the respective antibiotic overnight at 37 °C and 160 rpm. The 1 l main culture (LB medium with antibiotics) for protein production was inoculated 1:100 with this pre-culture. The culture was incubated at 37 °C at 100 rpm until an OD_{600nm} of 0.5–0.7. The gene expression was induced by AHT (200 ng/ml) or IPTG (0.5 mM), depending on the used promotor. The incubation time and temperature were selected accordingly to the results of the test expression. After the defined incubation time the cells were harvested and stored at -20 °C or directly used for protein purification.

Table 2.14: Production conditions and characteristics of used proteins. tag = protein tag, Ind. = inducer, AB = antibiotic, Temp. = production temperature, time = time for production, Strep = Strep-tag II, His = His-tag, C = C-terminal, N = N-terminal, AHT = anhydrotetracycline, IPTG = isopropyl β-D-1-thiogalactopyranoside, Amp = ampicillin, Cm = chloramphenicol.

Protein	plasmid	*E. coli* strain	tag	Ind.	AB	Temp	time
MA4377*	pASK-IBA3-MA_4377	BL21(DE3) C43(DE3)	Strep (C)	AHT	Amp	37 °C/17 °C	oN
RdmS	pASK-IBA3-MA_0863(*rdmS*) O216K	Nissle 1917	Strep (C)	AHT	Amp	37 °C/17 °C	oN
MA2082	pASK-IBA3-MA_2082	BL21(DE3)	Strep (C)	AHT	Amp	37 °C/17 °C	oN
MA2013	pASK-IBA3-MA_2013	BL21(DE3)	Strep (C)	AHT	Amp	37 °C/17 °C	oN
	pACYC-duet-MA_2013-PK	BL21(DE3)	His (N)	IPTG	Cm	37 °C/17 °C	oN
	pASK-IBA3-MA_2013-R1/Hpt	BL21(DE3)	Strep (C)	AHT	Amp	37 °C/17 °C	oN

Rec-only proteins**	pACYC-duet1-"Rec only"	BL21(DE3)	His (N)	IPTG	Cm	37 °C/17 °C	oN
MsrX	pET21a(+)-MA_4375 (*msrX*)	BL21(DE3)	His (C)	IPTG	Amp	37 °C/17 °C	oN
MsrC	pET21a(+)-MA_4375 (*msrC*)	BL21(DE3)	His (C)	IPTG	Amp	37 °C/30 °C	3h

* All variants (truncated and site directed mutagenesis) were produced under the same conditions

** "Rec-only" stands for the respective REC only gene. All REC-only proteins were produced under the same conditions

2.5.3 Cell harvesting and cell disruption of *E. coli* cells

Cells were harvested by centrifugation for 15 min and 9,000 g at 4 °C (Sorvall LYNX 6000, rotor: F9-6x1,000). Harvested cells were either used directly or stored at -20 °C. For protein purification the cell pellet was resuspended in ice-cold lysis buffer (Table 2.15) (3 ml buffer per g of cells) supplemented with 0.25 mM 4-(2-aminoethyl) benzene-sulfonyl fluoride (AEBSF), 1 mM 1,4-dithiothreitol (DTT), DNAse (10 µl of 5 mg/ml) and lysozyme. For membrane proteins the pellet was resuspended in ice-cold lysis buffer supplemented with 1 µM Leupeptin, 1 µM E-64, 0.15 µM Aprotinin, 500 µM PMSF and 10 µl DNase (5 mg/ml).

After approximately 30 min incubation time and complete resuspension of the cells, the cells were disrupted using a microfluidizer (Microfluidics International Corporation, Newton (MA), USA). The technology is based on the use of high pressure that is converted to high shear rates. The sample is pumped through the channels of the interaction chamber, where it is subjected to constantly high shear rates and impact forces. For the disruption of *E. coli* cells, a pressure of 15,000 psi for 3–4 rounds was used. The crude extract was further processed depending on the protein.

2.5.4 Purification of recombinant cytosolic proteins

For the purification of cytosolic proteins, the cells were centrifuged for 1 h at 50,000 g (Sorvall LYNX 6000, rotor: T29-8x50) after cell disruption. The crude extract containing Strep-tagged proteins was loaded on a gravity flow column containing 2 ml of Strep-Tactin® Sepharose® resin (IBA Lifesciences GmbH, Göttingen, Germany) and treated according to the instruction of the manufacturer. The washing of the column was conducted with 10 column volumes (CV) of buffer W, the elution was performed using 3 CV of buffer E. The column was stored on buffer R.

The crude extract containing His-tagged proteins was loaded on a gravity flow column containing 2 ml of TALON® Superflow™ resin (Cytiva, Marlborough (MA), USA) and treated according to the instruction of the manufacturer. The washing of the column was conducted with 10 CV of buffer W_{His}, the elution was performed using 3 CV of buffer E_{His}. The column was

stored on 20 % Ethanol. For a subsequent SDS-PAGE, samples were taken at different timepoints (lysate (L), flow-through (FT), washing (W), elution (E)).

Table 2.15: Buffers and solutions used for the purification of recombinant cytosolic proteins.

Lysis buffer (buffer W_{Cyt}), pH 8.0	
Tris/HCl	100 mM
NaCl	300 mM
EDTA	1 mM
Lysis buffer (buffer W_{His}), pH 7.5	
Tris/HCl	100 mM
NaCl	300 mM
Imidazole	10 mM
Elution buffer (buffer E_{cyt})	
Tris/HCl	100 mM
NaCl	300 mM
EDTA	1 mM
Desthiobiotin	2.5 mM
Elution buffer (buffer E_{His}), pH 7.5	
Tris/HCl	100 mM
NaCl	300 mM
Imidazole	500 mM
Regeneration buffer (buffer R_{Cyt}), pH 8.0	
Tris/HCl	100 mM
NaCl	300 mM
EDTA	1 mM
HABA (hydroxy-azophenyl-benzoic acid)	1 mM
Regeneration buffer (buffer R_{His})	
Ethanol	20 %

2.5.5 Purification of recombinant membrane proteins

For the purification of membrane proteins, the cells were centrifuged for 1 h at 50,000 g (Sorvall LYNX 6000, rotor: T29-8x50) after cell disruption. Samples were taken from the supernatant (S1) and the pellet (P1). The pellet was resuspended in the same volume of buffer W_{memb}. The supernatant (S1) was centrifuged again for 1 h at 80,000 g and 4 °C (Sorvall Lynx 6000 centrifuge; rotor: T29-8x50). A sample of the supernatant (S2) was taken. The pellet (P2) generated by the ultracentrifugation was resuspended in buffer W_{memb} to give a pellet

concentration of 150 mg/ml. A sample (P2) was taken. In a next step the solution was solubilised with either the detergent dodecyl-β-D-maltoside (DDM) or the copolymer nanodisc Diisobutylene-maleic acid (DIBMA).

Solubilisation with detergent DDM

For solubilisation, DDM was added to the solution to a final concentration of 1 % DDM and 50 mg/ml membrane particles. Solubilisation was then carried out at 4 °C under constant stirring for 2 h. After solubilisation, a sample was taken (SO). Thereafter, an ultracentrifugation was performed at 4 °C and 70,000 g for 1h. After ultracentrifugation, a sample of the supernatant (S3) and the pellet (P3) were taken. P3 was resuspended in buffer W_{memb} as the pellet samples aforementioned.

The supernatant (S3) was mixed with 2 ml of Strep-Tactin®XT 4Flow® high capacity resin (IBA Lifesciences GmbH, Göttingen, Germany) and incubated at 4 °C under constant rotation overnight. On the next day, the mixture was loaded on a gravity flow column and treated according to the instruction of the manufacturer. Purification was carried out at 4 °C. Samples were taken from all purification steps (flow-through (FT), washing (W), elution (E)). The column material was washed with 10 CV buffer W_{memb}, supplemented with 0.25 % DDM. The elution was performed using 10 CV of buffer E_{memb}, supplemented with 0.25 % DDM. To regenerate and store the column material for reuse, the column was first washed with at least 15 CV of 3 M $MgCl_2$. The column was then washed with 8 CV of buffer W_{memb} and finally stored in 20 % ethanol until reuse.

Solubilisation with nanodisc DIBMA

For the solubilisation with DIBMA a 10 % DIBMA solution was prepared in buffer W_{memb} and the pH was adjusted to 8.0. The 10 % DIBMA solution was added to the solution (P2) with a final concentration of 5 % DIBMA and 50 mg/ml membrane particles. Solubilisation was then carried out at 4 °C under constant stirring oN. After solubilisation, a sample was taken (SO). Thereafter, an ultracentrifugation was performed at 4 °C and 80,000 g for 1 h. After ultracentrifugation, a sample of the supernatant (S3) and the pellet (P3) were taken. P3 was resuspended in buffer W_{memb} as the pellet samples before.

The supernatant (S3) was mixed with 2 ml of Strep-Tactin®XT 4Flow® high capacity resin (IBA Lifesciences GmbH, Göttingen, Germany) and incubated at 4 °C under constant rotation overnight. On the next day the mixture was loaded on a gravity flow column and treated according to the instruction of the manufacturer. Purification was carried out at 4 °C. Samples were taken from all purification steps (flow-through (FT), washing (W), elution (E)). The column material was washed with 10 CV buffer W_{memb}. The elution was performed using 10 CV of buffer $E_{memb.}$ To regenerate and store the column material for reuse, the column was first

washed with at least 15 CV of 3 M $MgCl_2$. The column was then washed with 8 CV of buffer W_{memb} and finally stored in 20 % ethanol until reuse.

Table 2.16: Buffers and solutions used for the purification of recombinant membrane proteins.

Lysis buffer (buffer W_{memb}), pH 8.0	
Tris/HCl	100 mM
NaCl	300 mM
Lysis buffer with DDM (buffer W_{memb} DDM), pH 8.0	
Tris/HCl	100 mM
NaCl	300 mM
DDM	0.25 %
Elution buffer (buffer E_{memb}), pH 8.0	
Tris/HCl	100 mM
NaCl	300 mM
D-Biotin	50 mM
Elution buffer with DDM (buffer E_{memb}), pH 8.0	
Tris/HCl	100 mM
NaCl	300 mM
D-biotin	50 mM
DDM	0.25 %
Regeneration buffer (buffer R_{memb})	
$MgCl_2$	3 M in H_2O
Storage buffer (buffer S_{memb})	
Ethanol	20 %
Protease-inhibitors	
Aprotinin	1 mM
E-64	10 mM
Leupeptin	10 mM
Phenylmethylsulfonyl fluoride (PMSF)	100 mM
5 % DDM stock solution	
buffer W_{memb}	
DDM	5 %
10 % DIBMA stock solution	
buffer W_{memb}	
DIBMA	10 %

2.5.6 Dialysis and concentration of purified proteins

To increase the protein stability and to remove desthiobiotin, biotin or imidazole from the purified proteins the solution was dialysed in kinase buffer (Table 2.22) oN at 4 °C under constant stirring. The dialysis tube used had a cut-off of 12–14 kDa (SERVAPOR® dialysis tubing, MWCO 12000-14000). The volume of buffer for the dialysis was 100 times the volume of the pooled elution fractions. Proteins were concentrated by several centrifugation steps at 4 °C and 3,500 g for 4 min (centrifuge 5810 R, rotor: A-4-62) using a centrifugal concentrator (Amicon® Ultra-4).

2.5.7 Estimation of protein concentration

The concentration of purified proteins was determined photometrically based on the absorbance at 280 nm and the use of the Lambert-Beer law. The molecular extinction coefficient was calculated using the calculation tool from Protein Calculator v3.4 (https://protcalc.sourceforge.net/) based on the calculation of Gill and van Hippel.

$$c\ [\mu M] = \frac{A_{280\,nm} \cdot DF}{\varepsilon_{280\,nm} \cdot d}$$

$$c\ \left[\frac{mg}{ml}\right] = \frac{M_r \cdot A_{280\,nm} \cdot DF}{\varepsilon_{280\,nm} \cdot d}$$

c = protein concentration [µM or mg/ml]
$A_{280\,nm}$ = absorption at 280 nm
DF = dilution factor
$\varepsilon_{280\,nm}$ = molar absorption coefficient
d = optical path length [cm]
M_r = molecular mass [Da or g/mol]

Table 2.17: Molecular mass and molar absorption coefficients (ε_{280nm}) used for the calculation of protein concentrations.

Protein	Molecular mass [kDa]	Molar absorption coefficient ε_{280nm} [$M^{-1}cm^{-1}$]
RdmS	113	115080
MA2082	83	81940
MA2013	116	91110
MA2013-PK	80	77620
MA2013-R1	25	10930
MA2013-HPt	21	9530
MA4377	116	50810
CHPK	87	50810
PKR1R2 *	77	42890
PKR1 *	61	40210

PK *	42	31840
R1 *	15	13940
R2 *	17	2680
MA0016	15	1280
MA0018	15	1520
MA1268	18	8250
MA1269	18	15220
MA1366	21	16500
MA1468	16	3840
MA1469	17	5120
MA2445	33	20700
MA2012	16	27880
MA2861	16	10810
MA3068	14	2560
MA4376 (R3) *	18	3840
MA4671	15	3960
MsrX	32	24300
MsrC	30	32790

*For the respective protein variants (site directed mutagenesis) the same values are valid

2.5.8 SDS-PAGE

To separate and analyse proteins according to their size and molecular weight, denaturating and discontinue SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) was used. The principle of SDS-PAGE is based on the denaturation of the proteins and the coating with the detergent SDS that leads to a negative charge proportional to the molecular weight. Applying electric current, the negatively charged proteins migrate through the gel according to their size. Large proteins migrate slower than small proteins. Depending on the protein size, gels with different acrylamide concentration can be used (Laemmli, 1970). To separate proteins, two different types of polyacrylamide gels were prepared. Standard gels consisted of a stacking gel (5.25 %), a resolving gel (12 % or 15 %) and a stopping gel (Table 2.19). Gradient gels were prepared with 4 % - 20 % acrylamide gradient (Table 2.20).

Prior to gel application, samples were mixed with SDS sample buffer, incubated for 5–10 min at 95°C and centrifuged (centrifuge 5415 D, rotor: F45-24-11; 10000 g, 5 min). Each gel pocket was loaded with 10–20 µl of protein sample. A volume of 5 µl protein marker (Unstained Protein Standard, Broad Range or Color Prestained Protein Standard, Broad Range, New England Biolabs GmbH) was used as reference for protein size estimation. Electrophoresis was performed applying a constant voltage of 180 V for ~ 60 min in SDS running buffer.

Polyacrylamide gels were stained in staining solution (Table 2.18) for 10–30 min and destained with destaining solution (Table 2.18) or water under constant shaking until the signals were clearly visible. The protein staining is based on the binding of coomassie brilliant blue (G250) to amino acids with aromatic or basic side chains.

Table 2.18: Buffers and solutions used for the SDS-PAGE.

Stacking gel buffer (4x)	
Tris-HCl pH 6.8	0.5 M
SDS	4 % (w/v)
Resolving gel buffer (4x)	
Tris-HCl pH 8.8	1.5 M
SDS	0.4 % (w/v)
SDS running buffer (10x)	
Tris-HCl pH 8.8	0.25 M
Glycine	1.92 M
SDS	1 % (w/v)
SDS sample buffer (4x)	
Tris-HCl pH 6.8	0.1 M
SDS	8 % (w/v)
Glycerol	40 % (v/v)
β-mercaptoethanol	10 % (v/v)
Bromophenol blue	1 % (w/v)
Staining solution	
Acetic acid	10 % (v/v)
Ethanol	30 % (v/v)
Coomassie brilliant blue	0.25 % (w/v)
Destaining solution	
Acetic acid	10 % (v/v)
Ethanol	30 % (v/v)

Table 2.19: Composition of stacking, resolving, and stopping gel for SDS-PAGE.

	Stacking gel	**Resolving gel**		**Stopping gel**
	5.25 %	**12 %**	**15 %**	
Acrylamide 30%	1.1 ml	8 ml	10 ml	420 µl
Stacking gel buffer (4x)	2.5 ml	/	/	/
Resolving gel buffer (4x)		5 ml	5 ml	250 µl

H_2O	6.1 ml	6.8 ml	4.8 ml	333 µl
10% SDS	50 µl	100 µl	100 µl	/
APS	30 µl	100 µl	100 µl	13 µl
TEMED	20 µl	10 µl	10 µl	2 µl

Table 2.20: Composition of gradient gels for SDS-PAGE.

	4 % gradient gel (12x0.75mm)	**20 % gradient gel(12x0.75mm)**
Acrylamid 30%	5.3 ml	26.7 ml
Resolving gel buffer (4x)	10 ml	10 ml
H_2O	24.7 ml	3.3 ml
APS	220 µl	220 µl
TEMED	22 µl	22 µl

2.5.9 Western blot

Western blot analyses were performed for the immunological detection of tagged proteins after separation by SDS-PAGE.

Specific antibodies directed against the respective fusion tags (His-tag, Strep-tag II) of the proteins can recognise them. The antibodies are conjugated with an alkaline phosphatase, whose activity is used for detection. The protein samples were first separated using SDS-PAGE before transferring to a PVDF membrane. The PVDF membrane was first incubated in methanol for 15 min before being equilibrated in Towbin buffer together with the SDS-PAGE gel and the blotting papers (Whatman® 3 mm). The proteins were transferred using the semi-dry blot method for 18 min at a constant voltage of 15 V. For smaller proteins (~15 kDa) the time was reduced to 15 min. The membrane was first incubated for 1 h at RT or oN at 4 °C in 10 ml blocking solution to saturate unspecific binding sites on the membrane. After three washing steps with TBS-T the membrane was incubated with the primary antibody for 1 h. For the detection of the Strep-tag II, 2 µg/ml avidin dissolved in TBS-T was added prior to incubation with the antibody to counteract the detection of endogenous biotin-carboxyl carrying proteins present in *E. coli* cells. This was followed by three washing steps with TBS-T for 5 min each to remove unbound antibodies. In case of the His-tag the membrane was afterwards incubated for 1 h with the secondary antibody, which binds to the primary antibody together with the alkaline phosphatase. After washing again 3 times with TBS-T for 5 min and TBS for 1 min each time, the membrane was equilibrated in 10 ml AP buffer. The transferred proteins were detected by adding 66 µl 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and 33 µl nitro blue tetrazolium chloride (NBT). This results in a colour reaction caused by BCIP, NBT and the alkaline phosphatase coupled to the first (in the case of detection against Strep-tag II) or

second antibody (in the case of detection against His-tag). The colourimetric reaction was stopped by washing the membrane with H_2O dest.

Table 2.21: Buffers and solutions for the western blot.

TBS (pH 7,6)	
Tris/HCl	50 mM
NaCl	140 mM
TBS-T (pH 7,6)	
Tris/HCl	50 mM
NaCl	140 mM
Tween-20	0,05 %
Towbin buffer (pH 8,3)	
Tris	25 mM
Glycin	192 mM
AP buffer (pH 9,5)	
Tris/HCl	100 mM
NaCl	100 mM
$MgCl_2$	5 mM
Blocking solution	
in TBS-T Puffer	3 % BSA
Avidin solution	
in TBS-T	2 µg/ml
NBT solution	
in 70% DMF	100 mg/ml
BCIP solution	
in 100% DMF	50 mg/ml

2.5.10 Size exclusion chromatography

Size exclusion chromatography (SEC) is used to analyse the oligomerisation state of proteins (Moore, 1964). With the help of a special gel matrix filled column, the proteins are separated based on their hydrodynamic radius. Small molecules can diffuse into the pores and are thus retained. Large molecules, on the other hand, are less able to penetrate the hydrophilic pores of the gel matrix and can therefore pass through more quickly. The smaller a molecule is, the lower its migration speed in the column and the larger its elution volume (V_e). By comparing the elution volumes with those of protein standards, the relative molecular mass of an unknown protein can be calculated. SEC was performed using an ÄKTA Pure or ÄKTA Explorer with a

Superdex® 200 Increase 10/300 GL column (GE Healthcare) with a column volume of approx. 24 ml each. The flow rate was 0.2 ml/min. Before the protein sample was injected, the column was first equilibrated with H_2O dest. and afterwards with kinases buffer. The elution behaviour of the sample was monitored spectroscopically at 280 nm. The unknown protein samples were fractionated in 0.5 ml aliquots and analysed via SDS-PAGE to determine which absorbance peak is corresponding to the protein. As protein standards, carbonic anhydrase (CA; 29 kDa), albumin (Alb; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), β-amylase (β-Amy; 200 kDa) and apoferritin (Apo; 443 kDa) were loaded onto the column and their elution volumes at 280 nm were compared with that of the unknown protein sample. Blue dextran (BD; 2000 kDa) was used to determine the void volume (V_0) and the column volume (V_c) was measured using acetone.

$$K_{AV} = \frac{V_e - V_0}{V_c - V_0}$$

V_e = protein elution volume
V_0 = column void volume
V_c = measured column volume
K_{AV} = gel phase distribution coefficient

2.5.11 TNP-ATP Assay

To investigate the ATP binding affinity of the putative HK, the ATP derivative 2',3'-O-trinitrophenyl adenosine 5'-triphosphate (TNP-ATP) was used. TNP-ATP has the advantage over native ATP that it can be analysed using fluorescence spectroscopy, as the fluorescence in aqueous solution is very low, but increases with binding to the nucleotide binding site of proteins (Hiratsuka, 2003). To investigate the binding of TNP-ATP, an increasing concentration of TNP-ATP (0–80 µM) was mixed with the putative HKs (2.5 µM, 5 µM and 7.5 µM) in kinase buffer. A fluorescence emission spectrum in the range of 450 - 650 nm was recorded using a fluorescence spectrometer (FP-8300 fluorescence spectrometer, Jasco) and a quartz cuvette (SUPRASIL® cuvette, 3x3 mm, Hellma Analytics). The excitation wavelength was 410 nm, and the slit width was 5 nm. The fluorescence of TNP-ATP was determined as a maximum at 541 nm. The measurements were carried out in triplicates and the fluorescence emission values were corrected by subtracting the buffer emission values and the protein emission values determined under the same conditions. To determine the binding affinity, the recorded emission maxima were plotted against the TNP-ATP concentrations (Figure 2.1). The K_d value was determined using the SOLVER function of Microsoft Excel.

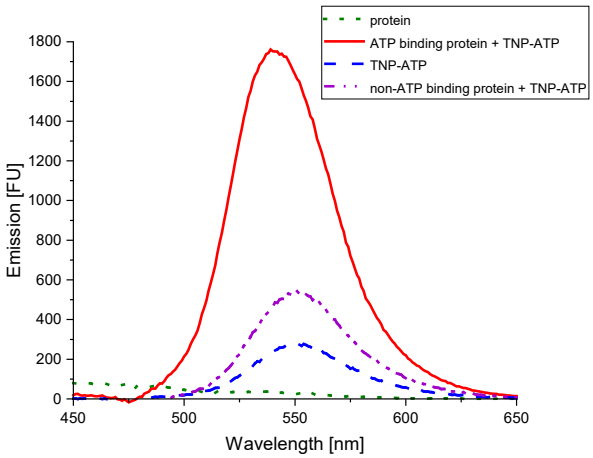

Figure 2.1: Overview of the TNP-ATP assay. The addition of TNP-ATP to a protein solution containing an ATP-binding protein results in a binding of TNP-ATP to the protein further eliciting an increased fluorescence peak. The protein itself shows almost no background fluorescence. TNP-ATP shows a low background fluorescence. The mix of TNP-ATP with a non-ATP binding shows clearly that the signal is not increased.

2.5.12 ADP-Glo™ Kinase Assay

To investigate whether the putative HKs can hydrolyse ATP, the ADP-Glo™ Kinase Assay (Promega) was performed according to the manufacturer's instructions (Figure 2.2).

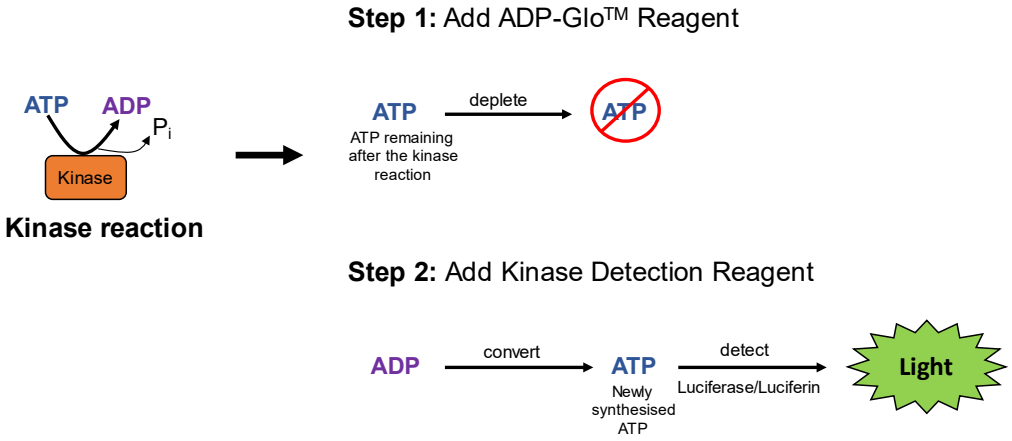

Figure 2.2: ADP-Glo™ Kinase Assay principle. The assay is composed of two different steps and an initial kinase reaction. In a kinase reaction the employed kinase is hydrolysing ATP to ADP and P_i. Afterwards the first step is to add the ADP-Glo™ Reagent that terminates the kinase reaction and depletes the remaining ATP. In the second step the addition of Kinase Detection Reagent converts ADP to ATP and generates light from the newly synthesised ATP using a luciferase/luciferin reaction. The generated light is proportional to the ADP present and therefore also to the kinase reaction.

The kinase reaction was carried out in white 96-well plates (Greiner LUMITRACTM 200) and detection was done using a luminometer (FLUOstar® Omega, BMG Labtech). All measurements were carried out in triplicates and the luminescence values determined were corrected by subtracting the values determined under the same conditions for a blank sample. A standard curve was created to quantify the amount of ATP converted during the kinase reaction.

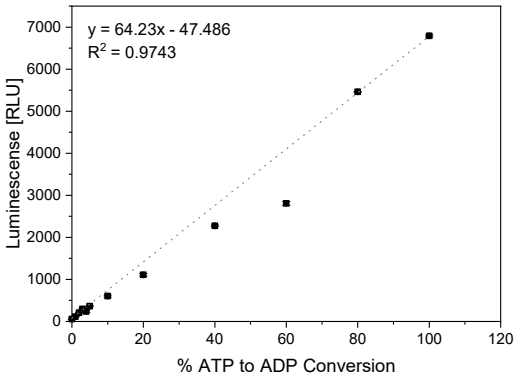

Figure 2.3: Calibration curve the ADP-Glo™ Kinase Assay. The ATP to ADP conversion curves were prepared according to the manufacturer's instructions with a 10 µM ATP concentration and can be used to calculate the conversion of ATP (%) during the kinase reaction. RLU: relative light unit

2.5.13 *In vitro* protein kinase assays

For autophosphorylation assays, 10 µM of the purified HK in kinase buffer was incubated with the addition of 200 µM ATP and 3.3 µM [γ-^{32}P]-ATP (3000 Ci/mmol and 10 mCi/ml; Hartmann Analytic) in a final volume of 12.5 µl to start the reaction. For transphosphorylation assays the kinase was autophosphorylated as described before and excessive ATP was removed using illustra™ MicroSpin™ G-25 Columns (GE healthcare). Afterwards the phosphate acceptor protein (10 µM) was added to the reaction mixture. Samples were taken at different timepoints. To stop the reaction, 12.5 µl of the reaction mixture was added to 2.5 µl of 4 x SDS-sample buffer (with β-mercaptoethanol). Samples were separated by SDS-PAGE gels (percentage depending on the size of the proteins) (chapter 2.5.8) and imaged using a Typhoon FLA 7000 PhosphorImager (GE Healthcare).

Table 2.22: Composition of kinase buffer.

Kinase buffer, pH 8.0	
Tris/HCl	50 mM
NaCl	150 mM
KCl	50 mM
$MgCl_2$	5 mM
Glycerol	10 %

2.6 Computational analysis

2.6.1 Data evaluation

Raw data were analysed using Microsoft Excel (Microsoft 365 Apps, Version 2105; Redmond (WA), USA) or OriginLab (Origin 2022, 9.9.0.225, Northampton, (MA), USA). Plots were generated with OriginLab. Images were edited with CorelDraw (CorelDRAW GraphicsSuite 2017, Version 19.1.0.448; Ottawa, Canada), Microsoft PowerPoint (Microsoft 365 Apps,

Version 2105; Redmond (WA), USA) and Inkscape (Inkscape 1.3.2, 2023). Schematic representative figures were created using Microsoft PowerPoint. Sequencing data were analysed with SnapGene (SnapGene® 4.0.8; San Diego (CA), USA)

2.6.2 Identification and analysis of nucleotide and amino acid sequences

Nucleotide sequences for the cloning of new constructs were obtained from the NCBI database (https://www.ncbi.nlm.nih.gov/nucleotide/). The analysis of the genomic distribution of signal transduction genes in *M. acetivorans* was done with the help of the free online tool PROKSEE (https://proksee.ca/).

Amino acid sequences for multiple sequence alignments were obtained from UniProt (https://www.uniprot.org/). The list of all used amino acid sequences is found in the appendix with accession number (Table S1). The databases SMART (http://smart.embl-heidelberg.de/) (Letunic *et al.*, 2020) and InterPro (https://www.ebi.ac.uk/interpro/) (Paysan-Lafosse *et al.*, 2023) were used to compare and classify protein sequences and protein domains. Transmembrane domains were analysed using the DeepTMHMM 1.0 server (https://services.healthtech.dtu.dk/services/DeepTMHMM-1.0/) (Hallgren *et al.*, 2022). The Microbial Signal Transduction Database 4.0 (MiST 4.0) (https://mistdb.com/mist/genomes) (Gumerov *et al.*, 2023) and the Procaryotic 2-Component System database (P2CS) (http://www.p2cs.org/) (Ortet *et al.*, 2015) were used to find signal transduction proteins of *M. acetivorans*. Multiple sequence alignments were generated using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) (Sievers *et al.*, 2011) and edited using ESPript v.3.0 (https://espript.ibcp.fr/ESPript/ESPript/) or Microsoft PowerPoint.

2.6.3 Prediction of protein structures

The predicted three-dimensional protein structures of the RR were generated using AlphFold2 with ColabFold by utilising the AlphaFold2 advanced notebook (Jumper *et al.*, 2021; Mirdita *et al.*, 2022). The structure of the 4 HKs, R3 and the homodimer structures were generated using AlphFold3 (https://alphafoldserver.com/about) using the default settings (Abramson *et al.*, 2024). All structures were analysed using the molecular visualisation software PyMOL (DeLano, 2020).

2.6.4 Phylogenetic analysis

Accession numbers of the used sequences are found in the Appendix (Table S1). The amino acid sequence alignment for the phylogenetic tree was created with Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo). The maximum-likelihood phylogenetic tree was constructed using PhyML 3.0 (Guindon *et al.*, 2010) with the WAG substitution model for amino acids (Whelan & Goldman, 2001) and a bootstrap replication of 100, bootstrap values were calculated using the online tool BOOSTER (Lemoine *et al.*, 2018). Bootstrap values

(above 0.8) are displayed as light blue circles in different sizes. To generate the tree, only the HK domain of the HKs was considered. A second phylogenetic tree was constructed with additional archaeal, bacterial, and eukaryotic HK sequences. The phylogenetic trees were afterwards modified with the programs iTOL (https://itol.embl.de/itol.cgi) and Inkscape (Inkscape 1.3.2, 2023).

3 Results

Signal transduction pathways are crucial for all living organisms. However, the knowledge about signal transduction systems in archaea is still limited. This thesis aims to get insights in the structure and function of HKs and RRs of the methanogenic archaeon *M. acetivorans*. An *in silico* analysis gives a broad overview of signal transduction systems in *M. acetivorans* and a biochemical approach examines the detailed function of selected HKs.

Chapter 1: *In silico* analysis of putative HKs and RRs encoded in the genome of *M. acetivorans*

3.1 Genomic distribution of HKs and RRs displays a 3:1 ratio

To get a first overview on the total number and the distribution of putative signal transduction components, the genome of the methanogenic archaeon *M. acetivorans* (NC_003552.1) was analysed using the online tool PROKSEE (https://proksee.ca/) together with data from the MiST 4.0 database (https://mistdb.com/mist/genomes), the P2CS database (http://www.p2cs.org/) and data provided by Galperin *et al.* 2018 (https://www.ncbi.nlm.nih.gov/Complete_Genomes/TCSarchaea.html).

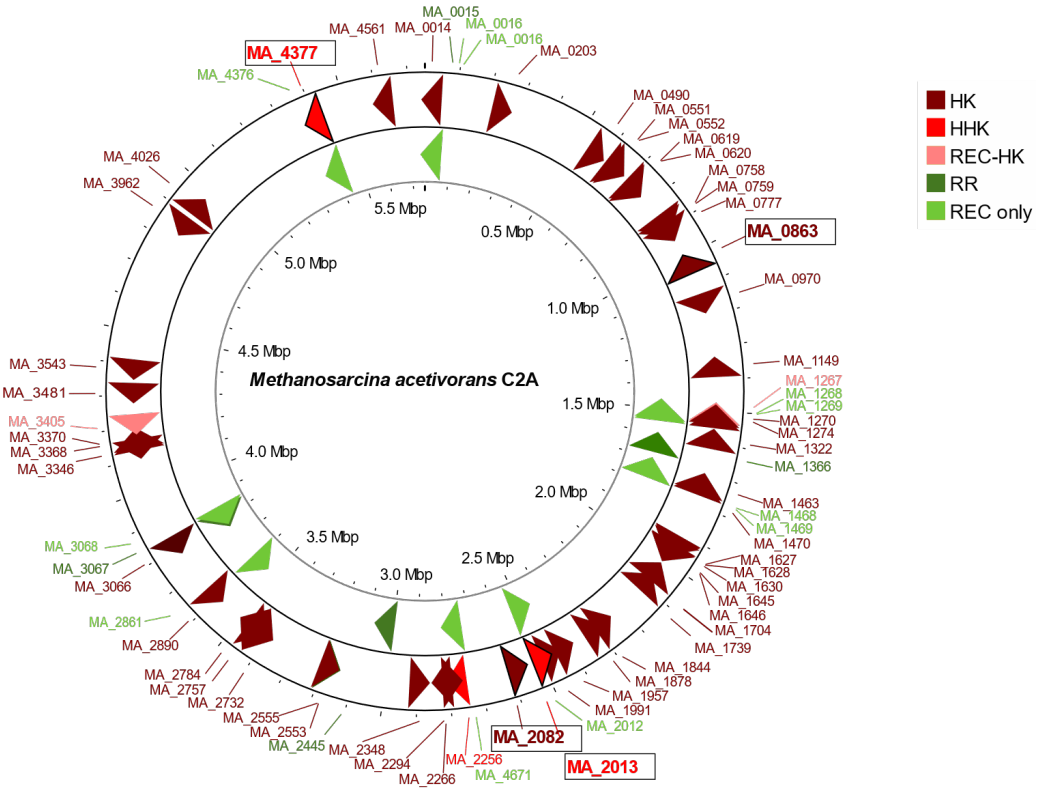

Figure 3.1*: Genomic distribution of HKs and RRs in the genome of *M. acetivorans*. Genomic distribution of histidine kinases (HK = Histidine Kinase, HHK = hybrid histidine kinase, REC-HK = histidine kinase with N-terminal REC domain) and response regulators (RR = response regulator with output domain, REC only = response regulator with only REC domain) in the genome of *M. acetivorans*. The genome of *M. acetivorans* (NC_003552.1) was obtained from the NCBI database and analysed using the online tool PROKSEE (https://proksee.ca/).

M. acetivorans has a genome size of around 5.7 Mbp and encodes for 53 putative HKs and 15 putative RRs. The ratio of HKs to RRs is around 3:1, in contrast to most bacteria which have a 1:1 ratio. Based on their structure, most of the putative HKs are classical HKs. Three putative HKs are predicted to be HHKs (MA_2013, MA_2256 and MA_4377). Two putative HKs (REC-HK) have a N-terminal REC domain (MA_1267, MA_3405). Looking into their genome localisation, they group in orphan kinases and kinases with genomic proximity to RRs. The genes of seven putative HKs (MA_0014, MA_1270, MA_1463, MA_2013, MA_2256, MA_3066 and MA_4377) are encoded in operons with one or more RR, while the other HKs are orphan. The 15 RRs can be divided into classical RRs and RRs with only a REC domain (REC-only). Most of the RR/REC-only genes are located in proximity of a putative HK, while two are orphan RR (MA_1366, MA_2861) (Figure 3.1).

3.2 HKs of *M. acetivorans* can be categorised into four groups

The protein sequences of the 53 putative HKs were obtained from the UniProt database (https://www.uniprot.org/), and the domain structure was analysed using the protein databases SMART (http://smart.embl-heidelberg.de/) and InterPro (https://www.ebi.ac.uk/interpro/). The 53 putative HKs are multidomain proteins with the characteristic HK domain, consisting of the DHp domain and the CA domain. Furthermore, the examined HK reveal different N-terminal sensing domains, mainly PAS, PAC, and GAF domains. Moreover, CHASE4, novel MEDS, MASE3 (**M**embrane-**A**ssociated **SE**nsor) and PocR (sensory domain found in PocR, with novel variant of the PAS-like fold) domains. Four putative HKs (MA1739, MA2348, MA2553, MA2555, MA4377) have a HAMP domain which is known to be a signal transduction element between membrane and cytosolic sensing parts. Furthermore, five putative HK display REC domains, either N-terminal (MA1267, MA3405) or C-terminal (MA2013, MA2256, MA4377). Three putative HK (MA0014, MA2013, MA3066) have HPt domains. The proteins MA0014 and MA3066, annotated as CheA homologs have a C-terminal CheW-like domain and a N-terminal HPt domain, probably serving as the phosphorylation site. On the other hand, the protein MA2013 has a C-terminal HPt domain. Among the 53 HKs, eleven HKs are membrane bound while the others are cytosolic proteins (Figure 3.2).

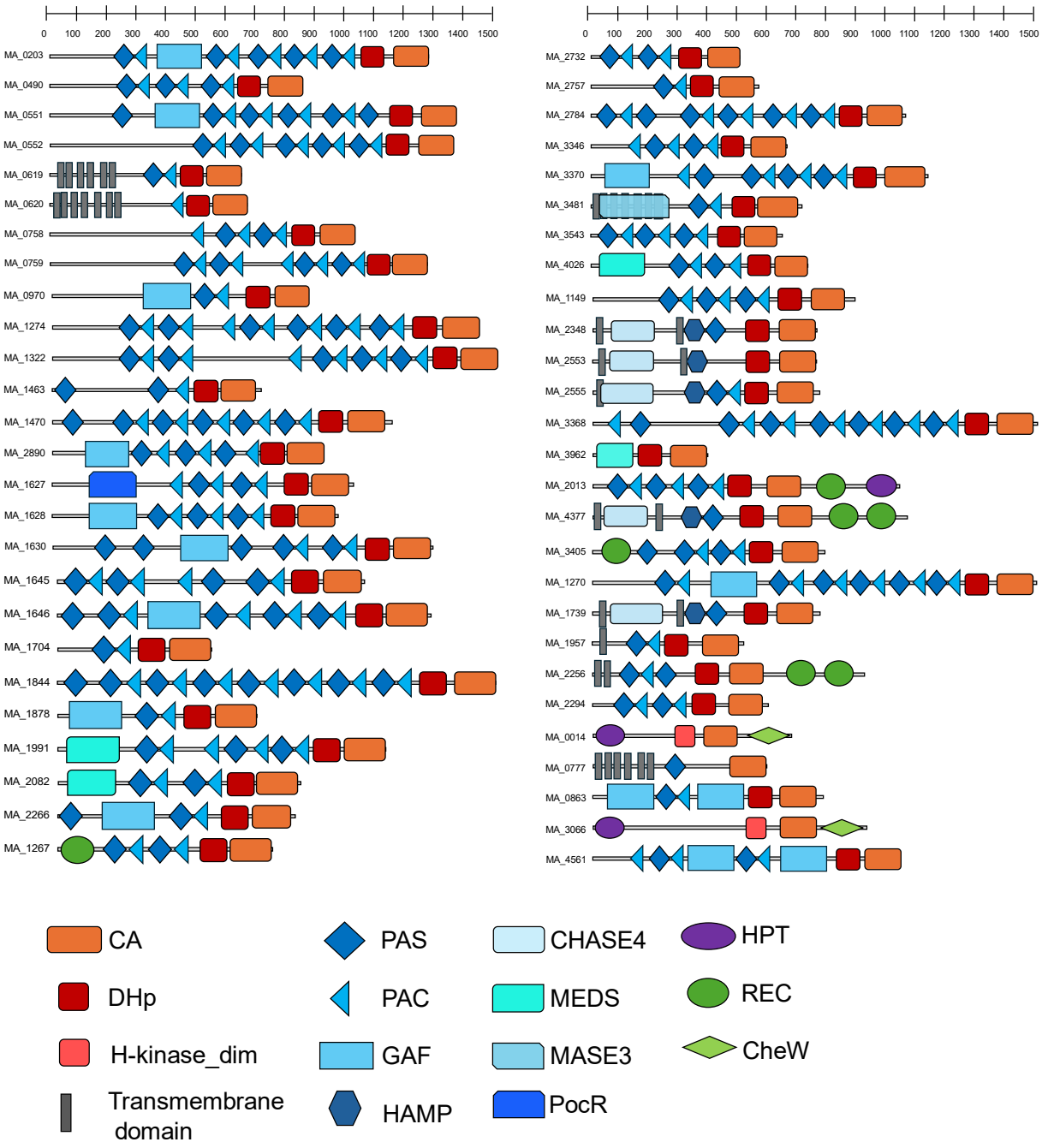

Figure 3.2: Domain organisation of putative HK of *M. acetivorans*. Domain structure of 53 putative HKs of *M. acetivorans*. The domain structures of the proteins were analysed using the programs SMART (Letunic *et al.*, 2020) and InterPro (Paysan-Lafosse *et al.*, 2023). PAS domain (Per-Arnt-SIM), PAC domain (PAS-Associated C-terminal motif), GAF domain (cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA), CHASE4 domain (Cyclases/Histidine kinases Associated Sensory Extracellular), HAMP domain (Histidine kinases, Adenylate cyclases, Methyl accepting proteins and Phosphatases), MEDS domain (MEthanogen/methylotroph DcmR Sensory), DHp domain (Dimerisation and Histidine phosphotransfer), CA domain (Catalytic ATP-binding), R (Receiver domain), HPt domain (Histidine Phosphotransfer), CheW.

In the light of the study of Kim and Forst, bacterial HKs can be categorised into five different groups (Type I, Type II, Type III, Type IV and CheA) according to the structure of their H-box (Kim & Forst, 2001). Indeed, an amino acid sequence alignment and a phylogenetic tree of the HK domain of all 53 putative HKs of *M. acetivorans* showed that they group in four distinct groups depending on the amino acid sequence of the H-box region (Figure 3.3, Figure 3.4).

Figure 3.3: Partial amino acid sequence alignment of HK domain of all putative HKs of *M. acetivorans*. Acession numbers of the employed sequences are listed in Table S1. For the alignment only the amino acid sequences of the HK domain were considered. The alignment was constructed using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) and modified using Microsoft PowerPoint. The conserved His residue is highlighted in bold with gray backgound. The H-box region groups into four distinct groups (Ma_type 1A/B/C, Ma_type 2, Ma_type 3 and MA_CheA). MA_type 1 is similar to the Type I HKs of bacteria. As the amino acid arrangement differs within the group, the subgroups A, B and C were created. MA_type 2 kinases are characterised by the absence of an H-box and MA_type 3 kinases cannot be grouped with a known type of bacterial kinases. MA_CheA are highly similar to bacterial CheA proteins.

As only one of these groups belongs to the known bacterial kinase types, new groups were designated for *M. acetivorans*. They were named MA_type 1A/B/C, MA_type 2, MA_CheA and MA_type 3. MA_type 1 is similar to the Type I HKs of bacteria (Kim & Forst, 2001). Since the amino acid arrangement differs within the group, the subgroups A, B and C were created. MA_type 2 kinases are characterised by the absence of the H-box. Both MA_type 2 and MA_type 3 kinases cannot be grouped with a known type of bacterial kinases. Additionally, the MA_type 3 kinases lack the F-box region which functions as the ATP lid (Appendix, Figure S1).

MA_CheA are highly similar to bacterial CheA proteins. The protein MA0777 cannot be grouped in any of the designated types.

A phylogenetic tree constructed with the HK domain of 53 putative HK supports the findings from the amino acid sequence alignment. A second phylogenetic tree constructed with the HK domain of *M. acetivorans* kinases and additional HKs from bacteria and eukaryotes reveal that MA_type 1 kinases groups within bacterial Type I kinases. MA_CheA kinases clearly group with bacterial CheA representatives. The HHK MA2013 groups within the eukaryotic and bacterial kinases. MA_type 2, with the lacking H-box, build their own group. The MA_type 3 representatives form a distinct group with only two representatives of the archaeon *Methanothermobacter thermoautothrophicus*.

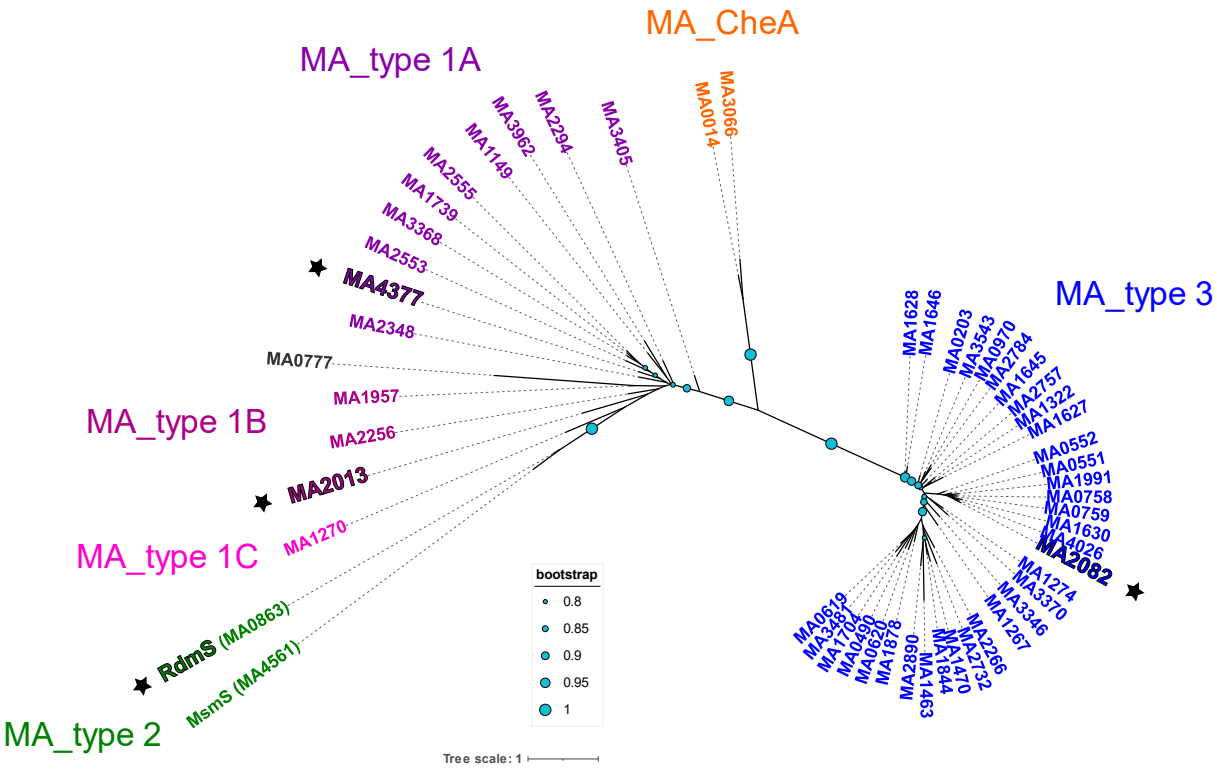

Figure 3.4: Unrooted phylogenetic tree depicting the HK domain of 53 putative HKs from M. acetivorans. Phylogenetic analysis of the HK domain from *M. acetivorans*. The maximum-likelihood phylogenetic tree was constructed using PhyML 3.0 (Guindon *et al.*, 2010), bootstrap values were calculated using the online tool BOOSTER (Lemoine *et al.*, 2018). Bootstrap values are displayed as light blue circles in different sizes. To generate the tree, only the HK domain of the HKs was considered. The phylogenetic tree was afterwards adjusted with the programs iTOL (https://itol.embl.de/itol.cgi) and Inkscape (Inkscape 1.3.2, 2023). Labels at branches are coloured according to the HK type. Accession numbers of the used sequences are found in the Appendix (Table S1).

3.3 RRs of *M. acetivorans* are characterised by an elevated number of REC-only RR and novel output domains

The analyses of the putative RRs revealed that eleven of them are so-called REC-only proteins only bearing a REC domain and completely lacking an output domain (Figure 3.5). Four of the RRs show the characteristic domain organisation with a REC domain and a fused output

domain. Two of them have a fused CheB domain (MA_0015 and MA_3067), one a HTH domain (MA_1366) and one a methanogen output domain 1 (MetOD1 domain) (MA_2445). The two proteins MA_0015 and MA_3067, harbouring a CheB-like output domain (methylesterase), are associated with the putative MA_CheA proteins MA_0014 and MA_3066. MA_1366 looks like a typical bacterial RR with a DNA binding HTH domain. The MetOD1 domain is a specific domain of methanogens with yet unknown function. Twelve RRs are encoded in proximity to a putative HK, while three are not encoded in proximity to a HK (MA_1366, MA_2445 and MA_2861).

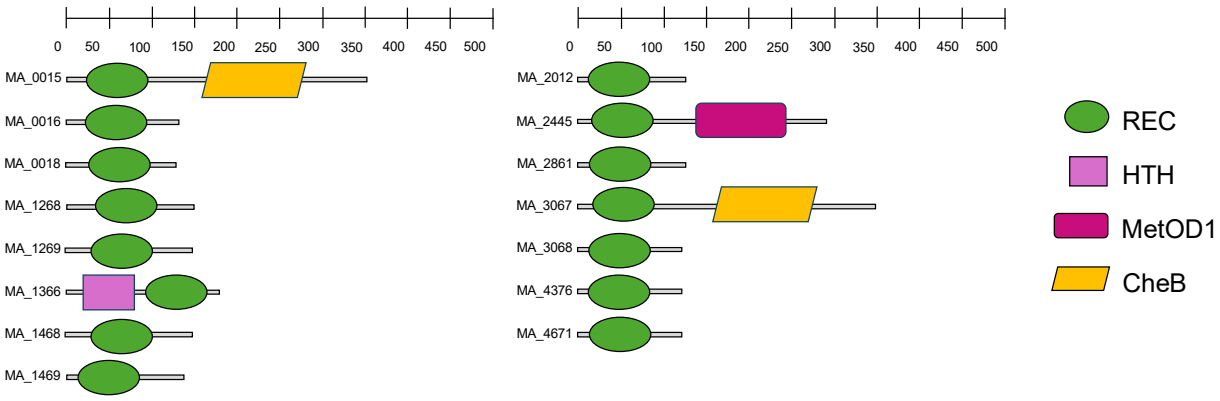

Figure 3.5: Domain organisation of putative RRs of *M. acetivorans*. Domain structure of 15 putative RRs of *M. acetivorans*. The domain structures of the proteins were analysed using the programs SMART (http://smart.embl-heidelberg.de/) and InterPro (https://www.ebi.ac.uk/interpro/). REC domain (RECeiver domain), CheB (methylesterase), HTH domain (Helix-Turn-Helix, DNA binding), MetOD1 domain (Methanogen output domain 1).

An amino acid sequence alignment (Figure 3.6) and a structural overlay (Figure 3.7) of the REC domain of 15 RRs were generated to analyse the similarities between the REC domains. The structural overlay demonstrated a high degree of similarity among the analysed REC domains, which all exhibited a $(\beta/\alpha)_5$ structure. The amino acids highlighted in the alignment are known to be important for the transphosphorylation activity (chapter 1.2, Figure 1.3). Nevertheless, the conserved amino acid residues exhibit variation. Only the highly conserved phosphorylation site (D) and a highly conserved glycine residue (G) are consistent within all REC domains. The remaining amino acid residues display greater variability, thereby conferring a unique character to each REC domain.

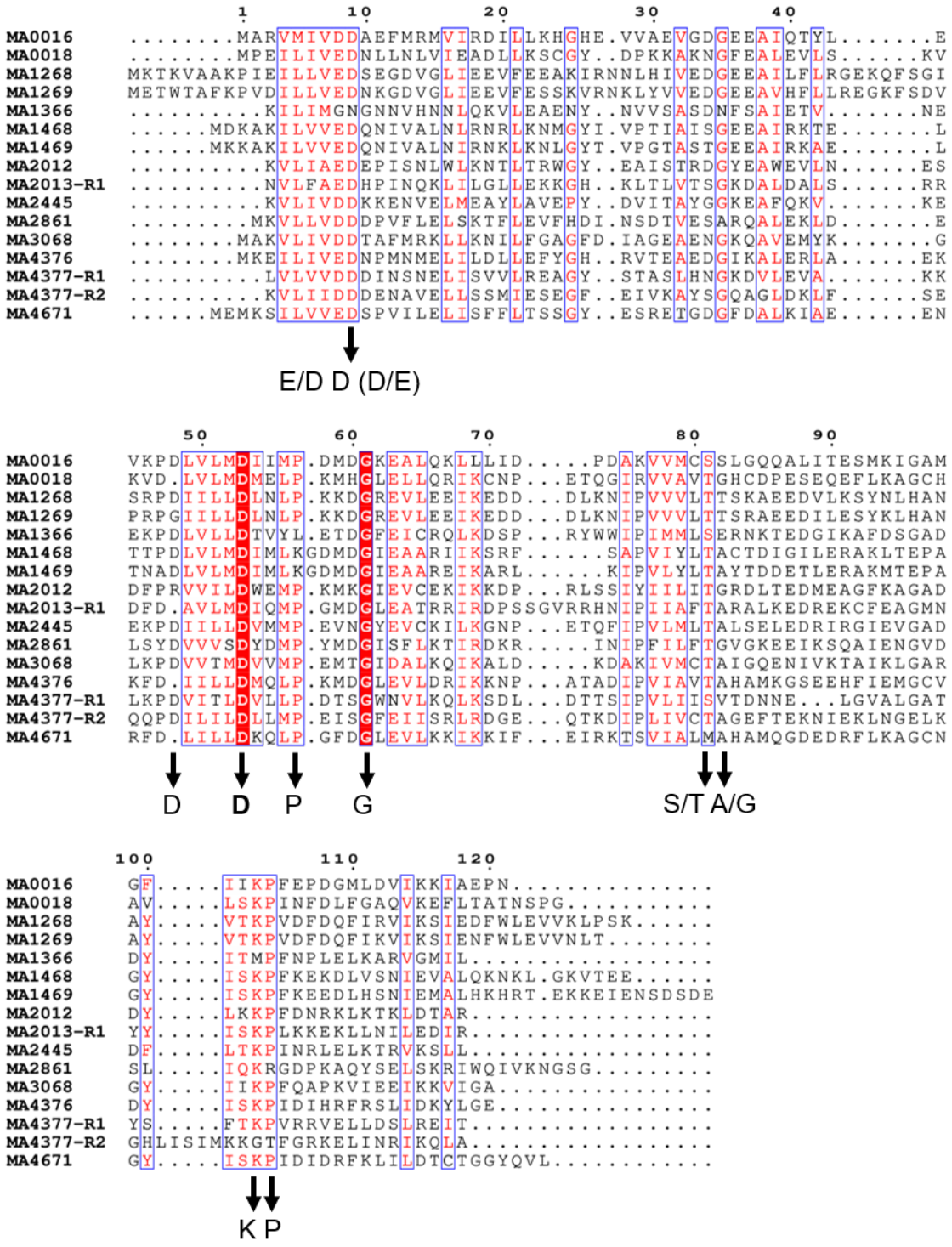

Figure 3.6: Amino acid sequence alignment of REC domains of all putative RRs of *M. acetivorans*. Amino acid sequence alignment of REC domains. Acession numbers of the employed sequences are listed in Table S1, for the alignment only the amino acid sequences of the REC domains were considered. The alignment was constructed using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) and modified using ESPript v.3.0 (https://espript.ibcp.fr/ESPript/ESPript/). Important conserved residues are indicated with arrows, the conserved Asp (D) residue is highlighted in bold.

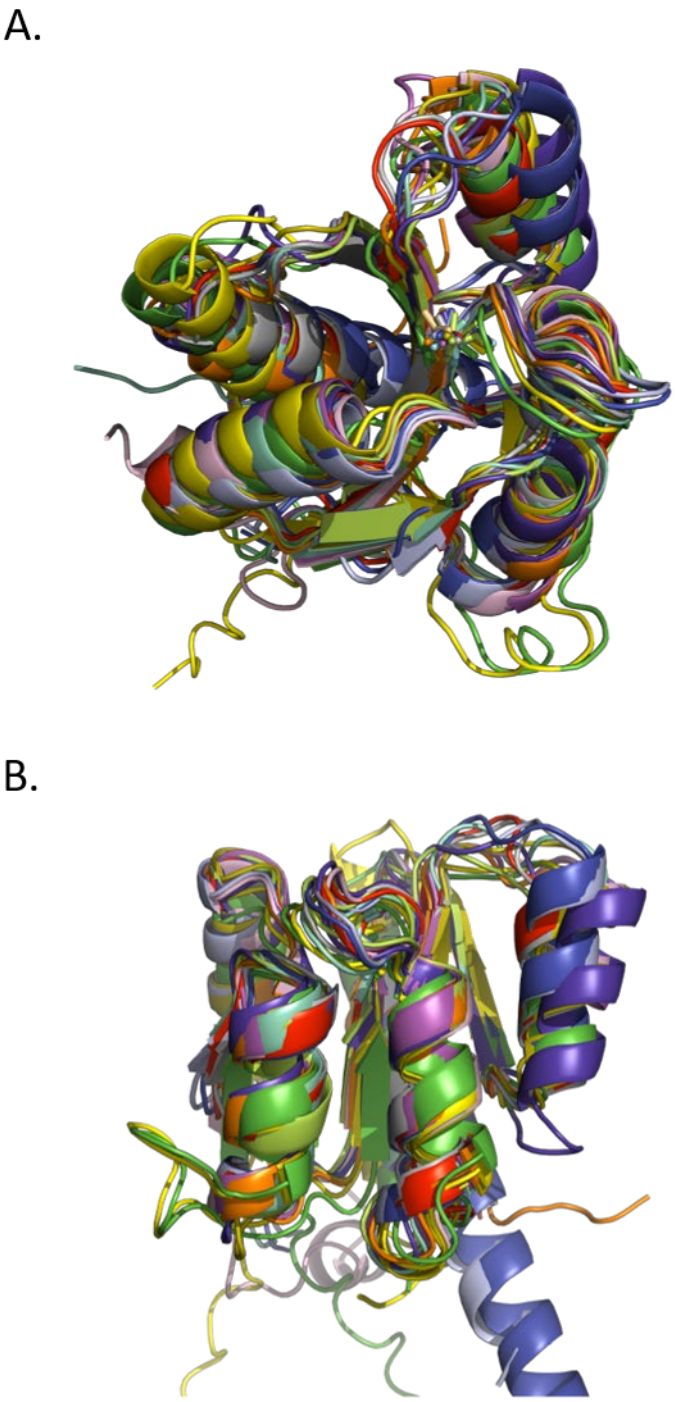

Figure 3.7: Overlayed AlphaFold2 structures of REC domains of all putative RRs of *M. acetivorans*. Predicted AlphaFold2 protein structures of the REC domains. Structures were displayed in cartoon representation and overlayed using PyMOL (DeLano, 2020). (**A.**) Represents the structures from the top. (**B.**) Represents the structures from the side.

Chapter 2: Biochemical analysis of four putative HKs of *M. acetivorans*

The biochemical analysis presented in this chapter has the aim of providing more insights into the mechanism of archaeal signal transduction. One member of the each newly designated HK group was chosen for further investigations (Figure 3.3, Figure 3.4). The MA_CheA HKs were excluded from further studies, since CheA proteins have already been extensively studied. For MA_type 1, a cytosolic (MA2013) and a membrane bound representative (MA4377) were selected. Both proteins are putative HHKs with bound REC-domains. MA2013 additionally has a HPt domain at the C-terminus. As a representative for MA_type 2 kinases, the orphan, cytoplasmic putative HK RdmS (MA0863), which is lacking the H-box, was selected. MA2082 represents the MA_type 3 kinases. The protein is an orphan, cytoplasmic putative HK, possessing a methanogen-specific MEDS domain.

3.4 All examined HKs bind and hydrolyse ATP

The function of HKs relies on the binding of ATP to an ATP binding site in the CA domain and its subsequent hydrolysis in the DHp domain, which results in a binding of a phosphoryl group to the conserved His residue in the H-box. In order to test the ATP binding and hydrolysis, all four putative HK were recombinantly produced in *E. coli* and purified using affinity chromatography.

ATP binding was tested with the ATP-derivate 2'-(3')-O-(2,4,6-trinitrophenyl)-adenosine 5'-triphosphat (TNP-ATP) (chapter 2.5.11). With this substance the binding of proteins to ATP can be tested via fluorescence spectroscopy. In water, the TNP-ATP only shows a low background fluorescence. Bound to an ATP binding protein, the fluorescence signal is strongly increasing (Figure 2.1, Figure 3.8).

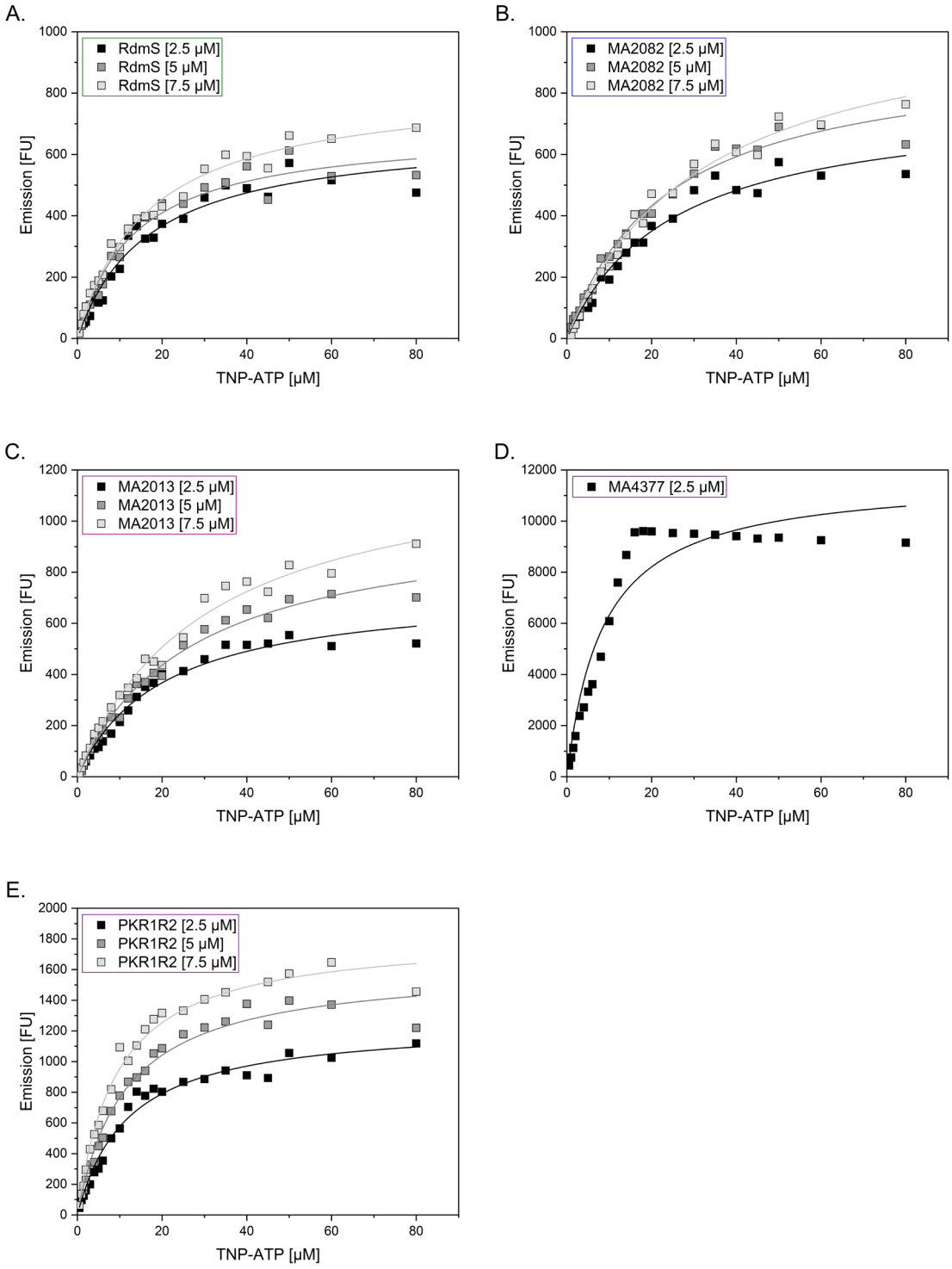

Figure 3.8: TNP-ATP binding kinetics of putative HKs. Constant protein concentration (2.5 µM, 5 µM or 7.5 µM) was mixed with variable TNP-ATP concentrations (0–80 µM) in kinase buffer. A fluorescence emission spectrum in the range of 450-650 nm was recorded using a fluorescence spectrometer (FP-8300 fluorescence spectrometer, Jasco) and a quartz cuvette (SUPRASIL® cuvette, 3x3 mm, Hellma Analytics). The excitation wavelength was 410 nm, and the slit width was set to 5 nm. The fluorescence of TNP-ATP was determined as a maximum at 541 nm. The plots were generated using OriginLab. (**A.**) TNP-ATP binding kinetics of RdmS. (**B.**) TNP-ATP binding kinetics of MA2082. (**C.**) TNP-ATP binding kinetics of MA2013. (**D.**) TNP-ATP binding kinetics of MA4377. (**E.**) TNP-ATP binding kinetics of PKR1R2. FU: fluorescence unit

The dissociation constant (K_d) for TNP-ATP towards the putative HKs was calculated using the Excel SOLVER function (Figure 3.9, A.). All putative HKs displayed affinity towards TNP-ATP with K_d values in the micromolar range. MA4377 and the variant PKR1R2 show the highest affinity (K_d ~11 µM), while the affinity of RdmS (K_d ~17 µM), MA2013 (K_d ~28 µM) and MA2082 (K_d ~29 µM) suggest a lower binding affinity towards TNP-ATP.

A.

protein	K_d TNP-ATP [µM]
RdmS	17.0 ± 5.7
MA2082	29.6 ± 4.4
MA2013	28.2 ± 1.3
MA4377	11.7
PKR1R2	11.4 ± 2.6

B.

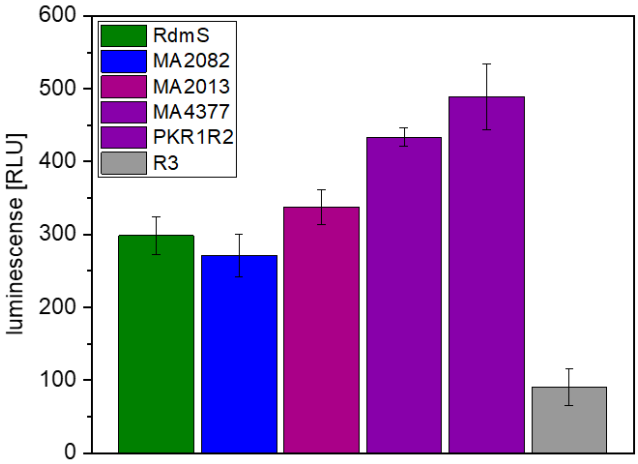

Figure 3.9: Analysis of ATP binding and hydrolysis. (**A.**) K_d values of different putative HKs determined for TNP-ATP. Data were generated using the TNP-ATP assay. Emission spectra were recorded for different protein concentrations (2.5 µM, 5 µM, 7.5 µM) and TNP-ATP concentrations (0–80 µM) in kinase buffer. The fluorescence maximum at 541 nm was plotted against the TNP-ATP concentration for every protein concentration. K_d values were examined using the SOLVER function of Microsoft Excel. The results were averaged over three independent measurements. (**B.**) ADP-Glo™ Kinase Assay to examine the hydrolysis of ATP for different HKs. R3, the REC-only RR in vicinity to MA4377 was used as negative control since it is a non-ATP binding protein. The mean values and standard deviations of three independent kinase reactions are shown. RLU: relative light unit.

ATP binding was shown for all four putative HKs. A next step was to investigate their ability to hydrolyse ATP. For this the ADP-Glo™ Kinase Assay was used (chapter 2.5.12, Figure 2.2). The principle of this assay is based on the measurement of luminescence that correlates with the kinase activity. The results show an increased luminescent signal for all tested putative HKs compared to the negative control (R3). This suggests that the kinases can also hydrolyse ATP. The luminescent signal for MA3477 and PKR1R2 is higher than for RdmS, MA2082 and MA2013. Using the calibration curve (Figure 2.3), the ATP to ADP conversion rate was calculated, all putative HKs showed a conversion rate of around 3–5 %. Knowing that the four putative kinases can bind and hydrolyse ATP, in a next step their kinase activity had to be tested via radioactive autokinase assays.

3.5 RdmS is an unusual kinase lacking autokinase activity

The protein RdmS (**R**edox **d**ependent **m**ethyltransferase-associated **S**ensor) is a MA_type 2 kinase with alternating PAS/PAC and GAF domains and a HK domain (Figure 3.10, A.).

M. acetivorans harbours two MA_type 2 kinases (RdmS and MsmS (**M**ethyl **s**ulfide **m**ethyltransferase-associated **S**ensor)) which are structurally similar and characterised by the absence of an H-box. In the protein database UniProt (https://www.uniprot.org/) RdmS is annotated as a HK and, besides the lacking H-box, it has all relevant components of a HK.

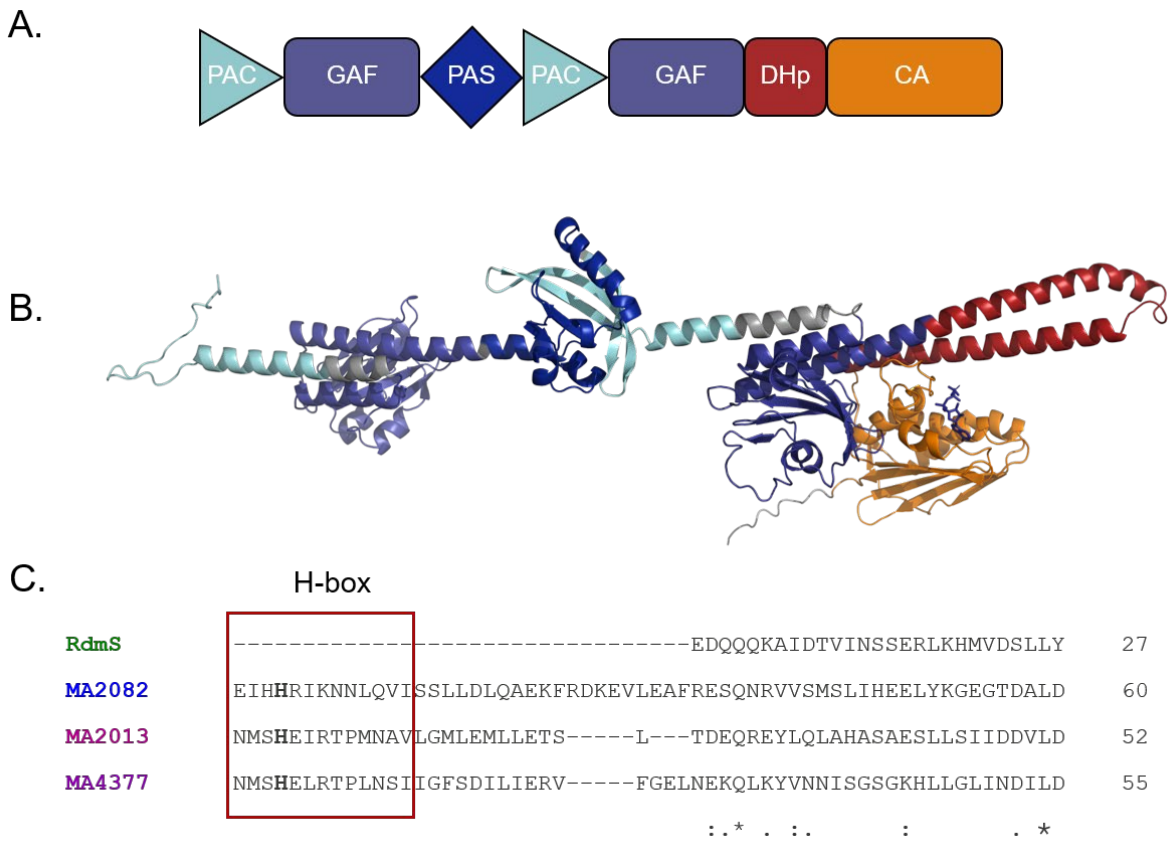

Figure 3.10: Schematic domain representation and predicted protein structure of the protein RdmS. (**A.**) Schematic representation of the domain organisation of RdmS. The domain structures were analysed using the tools InterPro (https://www.ebi.ac.uk/interpro/) and SMART (http://smart.embl-heidelberg.de/). RdmS is a cytosolic protein displaying several PAS (**P**er-**A**rnt-**S**IM) /PAC (**P**AS-**A**ssociated **C**-terminal motif), GAF (c**G**MP-specific phosphodiesterases, **a**denylyl cyclases and **F**hlA) and a DHp (**D**imerisation and **H**istidine **p**hosphotransfer) and CA (**C**atalytic **A**TP-binding) domain. (**B.**) Protein structure of RdmS predicted with Alphafold3 (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured referring to the domain structure with PyMOL (DeLano, 2020). (**C.**) Partial amino acid sequence alignment of the HK domain of representatives of all four examined kinase types. RdmS clearly lacks the H-box.

RdmS was purified via affinity chromatography and subsequently analysed by SDS-PAGE (Figure 3.11, A.). The coomassie stained SDS-PAGE revealed a protein with relative molecular weight of ~115 kDa, which is consistent with the predicted molecular weight of ~113 kDa calculated from the amino acid sequence (chapter 2.5.7).

To investigate the oligomerisation state of RdmS in kinase buffer a SEC using a Superdex® 200 Increase 10/300 GL column (Figure S4) was conducted. Two major A_{280} peaks were observed with a retention volume of 7.8 ml and 9.7 ml. Peak 1 at 7.8 ml most likely represents aggregates since blue dextran (2000 kDa) elutes at the same retention volume. Blue dextran was used to determine the void volume. Peak 2 at 9.7 ml corresponds to a relative molecular

weight of 560 kDa and presumably represents RdmS. To verify this, fractions of the eluted sample were collected and analysed via SDS-PAGE. The SDS-PAGE revealed a ~113 kDa protein in the fractions B10-B4, with the highest protein amount in fraction B6/B5 corresponding to peak 2. Furthermore, RdmS was identified in the fractions corresponding to the void volume, indicating that the protein is likely partially aggregated. The calculated molecular weight of peak 2 corresponds to a tetramer. It is not likely that RdmS represents a tetramer since most HKs form dimers. The oligomerisation state of RdmS was already determined in a previous study, with a similar result (505 kDa) (Fiege, 2019). The author of this study stated that the protein elutes as a long-stretched dimer to explain the high molecular weight. A second study analysing the oligomerisation state of a RdmS homolog of *Methanosarcina siciliae* revealed that this protein forms a dimer with a molecular weight of 230 kDa. In this study, the Superdex® 200 Increase 10/300 GL column and the Superose™ 6 Increase 10/300 GL column were used to confirm the result. Important to note is, that the protein concentration that was applied on the SEC was very low (Köhler, 2021). Further experiments are required to determine whether this molecular weight can be explained with the long-stretched structure of the protein or the fusion of two dimers that occurs due to a high protein concentration.

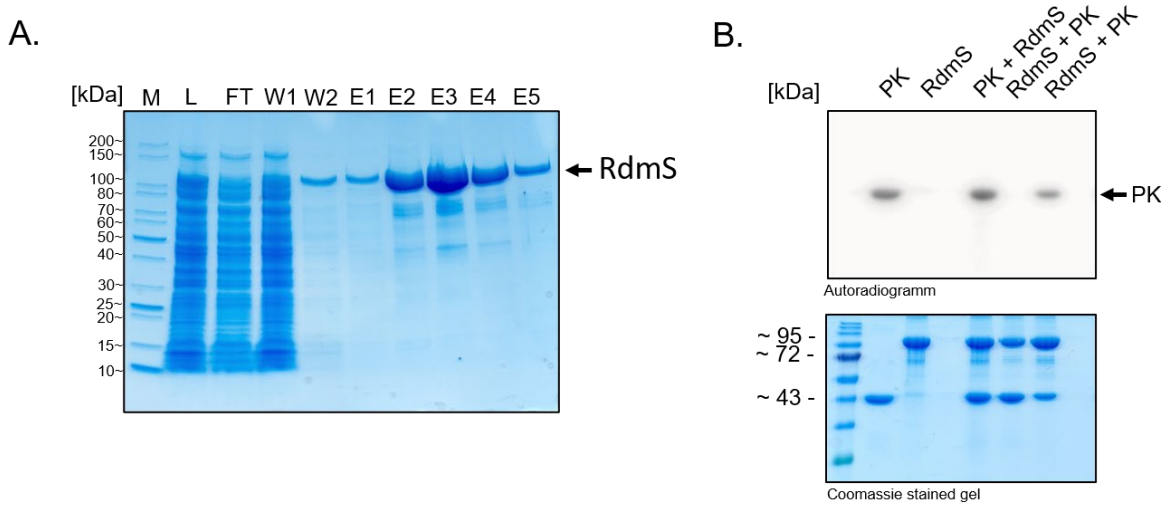

Figure 3.11: Purification of the protein RdmS and kinase assay of RdmS with the control protein PK. (**A.**) Coomassie stained gel of RdmS after recombinant production and affinity chromatography purification with Strep-Tactin® Sepharose®. Sample application of lysate after cell lysis (L), flow-through (FT), wash fractions (W) and protein elution fractions (E). As reference the colour-stained protein marker (M) from NEB was used. RdmS has a predicted molecular weight of ~116 kDa. (**B.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation/transphosphorylation assay. First radiolabel [γ-^{32}P]-ATP was added to RdmS and PK (both 10 µM) and the reaction was terminated with SDS sample buffer (with BME) to test the autokinase activity. For the second part of the assay either the protein PK or RdmS were autophosphorylated and the second protein was added in equimolar concentration to analyse a potential transphosphorylation. Additionally, the radiolabeled [γ-^{32}P]-ATP was added to a mix of RdmS and PK (both 10 µM). The coomassie stained gel shows that the same protein concentration was used. The autoradiogram displays autokinase activity only for the control protein PK. As reference the unstained/color prestained protein standard (M) from NEB was used.

Previous work indicated that MsmS and RdmS are redox-dependent kinases with autophosphorylation under oxidising conditions likely at a tyrosine residue (Fiege, 2019; Fiege

& Frankenberg-Dinkel, 2019; Molitor, 2013; Molitor *et al.*, 2013). Further investigations of this phenomenon however revealed that autokinase activity under oxidising conditions was only observed when the reaction was terminated with a SDS sample buffer lacking a reducing agent (i.e. β-mercaptoethanol (BME)) (Blasius, 2023). Furthermore, this autophosphorylation was dependent on the presence of cysteine residues in the protein and appears to be an unspecific phosphorylation phenomenon which also occurs in non-kinase proteins (Blasius, 2023). While it was confirmed that RdmS binds and hydrolyses ATP (chapter 3.4), the investigation of the autokinase activity was still not completely resolved. Therefore, recombinant RdmS was re-investigated in autophosphorylation assays together with the control protein PK (classical HK). Compared to the control protein, RdmS is not showing autokinase activity under the standard conditions, which clearly demonstrates the lack of activity for RdmS (Figure 3.11, B.). To test whether RdmS is phosphorylated by another HK, or if RdmS can transfer a phosphoryl group to a classical HK, a radioactive transphosphorylation assay was conducted. This assay shows no phosphoryl group transfer from RdmS to the control protein PK or from PK to RdmS (Figure 3.11, B.). This result leads to the assumption that RdmS is a kinase with an unknown substrate.

3.6 MA2082 is a cytosolic orphan HK with autokinase activity

The putative HK MA2082 is a cytosolic representative of MA_type 3 kinases possessing a methanogen-specific MEDS domain, two PAS/PAC domains and a HK domain consisting of a DHp and CA domain. The protein has an H-box with a putative phospho-accepting His residue (Figure 3.12).

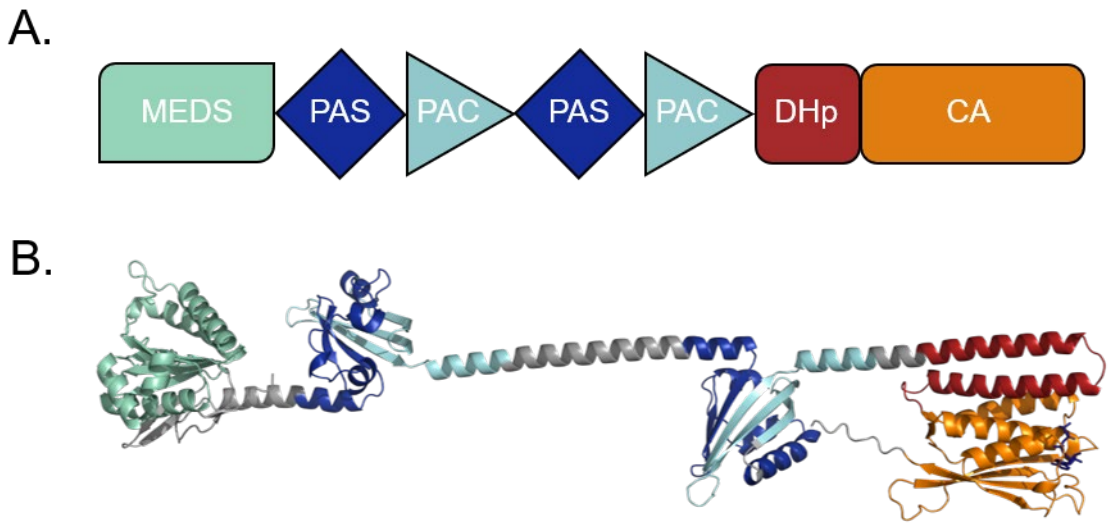

Figure 3.12: Schematic domain representation and predicted protein structure of the putative HK MA2082. (**A.**) Schematic representation of the domain organisation of MA2082. The domain structures were analysed using the tools InterPro (https://www.ebi.ac.uk/interpro/) and SMART (http://smart.embl-heidelberg.de/). MA2082 is a cytosolic protein displaying a MEDS (**ME**thanogen/methylotroph **D**cmR **S**ensory) domain, a PAS (**P**er-**A**rnt-**S**IM), PAC (**P**AS-**A**ssociated **C**-terminal motif) and a DHp (**D**imerisation and **H**istidine **p**hosphotransfer) and CA (**C**atalytic **A**TP-binding) domain. (**B.**) Protein structure of MA2082 predicted with AlphaFold3 (Abramson *et al.*, 2024) Displayed in cartoon representation and coloured referring the domain structure with PyMOL (DeLano, 2020).

MA2082 was purified via affinity chromatography and subsequently analysed by SDS-PAGE (Figure 3.13). The coomassie stained SDS-PAGE revealed a protein with relative molecular weight of ~80 kDa, which is consistent with the predicted molecular weight of ~83 kDa calculated from the amino acid sequence (chapter 2.5.7).

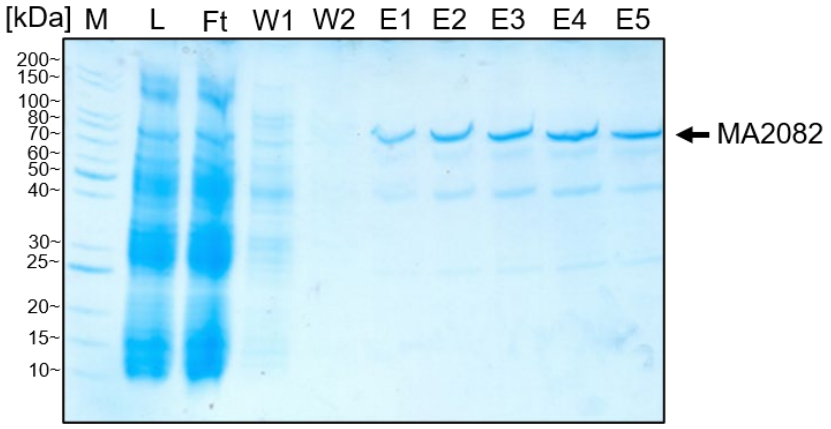

Figure 3.13: Purification of the protein MA2082. Coomassie stained gel of MA2082 after recombinant production and affinity chromatography purification with Strep-Tactin® Sepharose®. Sample application of lysate after cell lysis (L), flow-through (FT), wash fractions (W) and protein elution fractions (E). As reference the colour-stained protein marker (M) from NEB was used. MA2082 has a predicted molecular weight of ~83 kDa.

HKs predominantly form dimers. To investigate the oligomerisation state of the putative HK MA2082 in kinase buffer (Table 2.22) a SEC using a Superdex® 200 Increase 10/300 GL column (Figure S5) was conducted. Two major A_{280} peaks were observed with a retention volume of 7.6 ml and 10.3 ml. Peak 1 at 7.6 ml most likely represents aggregates since blue dextran (2000 kDa) elutes at the same retention volume. Peak 2 at 10.3 ml corresponds to a relative molecular weight of 420 kDa and presumably represents MA2082. To verify this, fractions of the eluted sample were collected and loaded on an SDS-PAGE to analyse the presence of MA2082. The SDS-PAGE revealed a ~80 kDa protein in the fractions B10-B4, with the highest protein amount in fraction B5/B4 corresponding to peak 2. Additionally, MA2082 was identified in the fractions corresponding to the void volume, indicating that the protein is likely partially aggregated. The calculated molecular weight of peak 2 corresponds to a tetramer. It is not likely that MA2082 represents a tetramer. Considering the SEC data of RdmS, further experiments are required to determine whether this molecular weight can be explained with the long-stretched structure of the protein or the fusion of two dimers that occurs due to a high protein concentration.

Furthermore, the autokinase activity of MA2082 was investigated with a radioactive autophosphorylation assay (Figure 3.14).

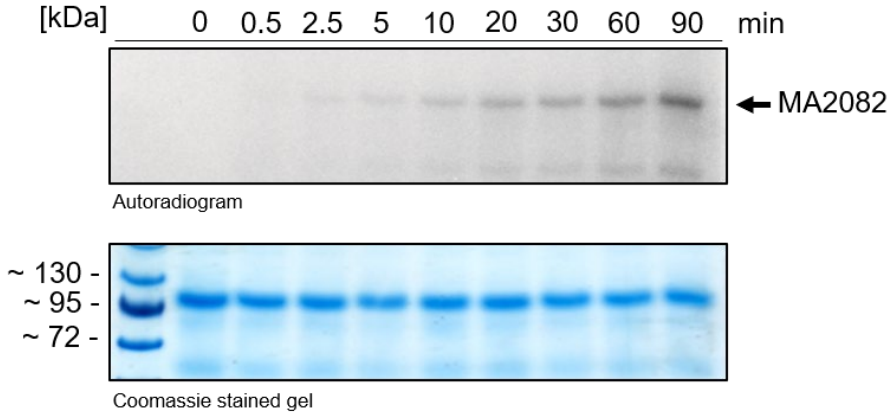

Figure 3.14: Autokinase activity of heterologus produced MA2082. Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM MA2082 and the reaction was terminated after different timepoints with SDS sample buffer (with BME). The coomassie stained gel shows that the same protein concentration was loaded. The autoradiogram displays an increasing signal starting from 2.5 min. As reference the color prestained protein standard (M) from NEB was used.

A time-depending increase of the radioactive signal with the highest intensity at 90 minutes was observed for the recombinant MA2082 protein. The autophosphorylation assay provides evidence that MA2082 is a classical HK with autokinase activity.

3.7 MA2013 is a hybrid kinase with bound REC and HPt domain

MA2013 is a cytosolic representative of MA_type 1 kinases. The putative HHK possesses three PAS/PAC domains, a HK domain, a C-terminal REC domain and a HPt domain. A REC-only RR (MA2012) is encoded in proximity of MA2013. The structure with a bound REC domain, a HPt domain and a REC-only RR suggests a putative role in a phosphorelay system.

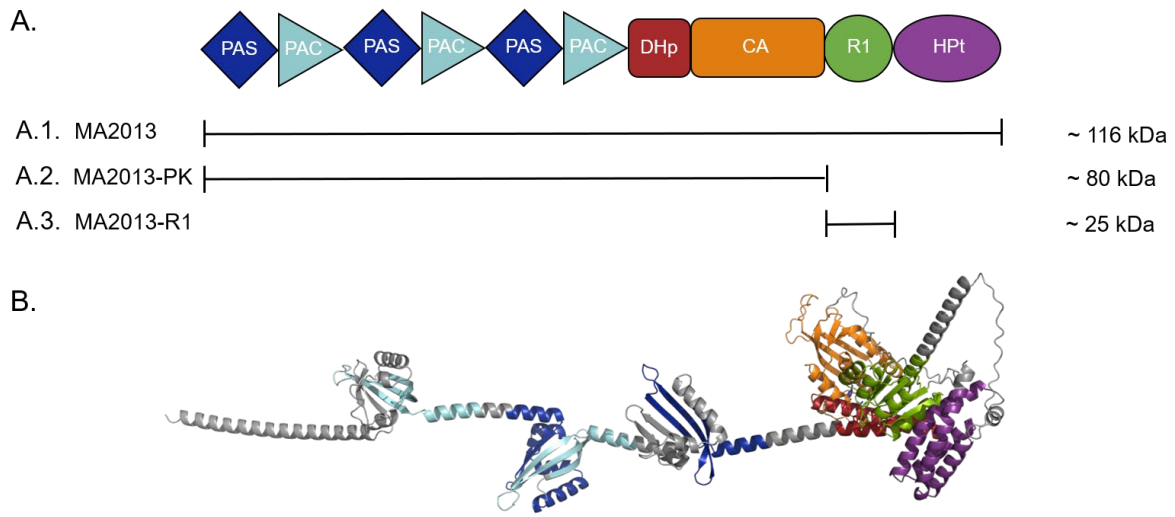

Figure 3.15: Schematic domain representation and predicted protein structure of the putative HHK MA2013. (**A.**) Schematic representation of the domain organisation and overview of truncated variants of the protein MA2013. All truncated variants of the protein MA2013 used in this study are shown with their calculated molecular weight. The domain structures were analysed using the tools InterPro (https://www.ebi.ac.uk/interpro/) and SMART

(http://smart.embl-heidelberg.de/). MA2013 is a cytosolic protein displaying several PAS (**P**er-**A**rnt-**S**IM), PAC (**P**AS-**A**ssociated **C**-terminal motif) domains, a DHp (**D**imerisation and **H**istidine **p**hosphotransfer) and CA (**C**atalytic **A**TP-binding) domain and a REC (**R**eceiver domain) and HPt (**H**istidine **P**hospho**t**ransfer) domain. (**A.1.**) full-length MA2013 (**A.2.**) truncated MA2013-PK variant with only the sensing domains and the HK domain. (**A.3.**) only the REC domain of MA2013 (MA2013-R1). (**B.**) Protein structure of MA2013 predicted with AlphaFold3 (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

MA2013 and the truncated variants of the protein were purified via affinity chromatography and subsequently analysed by SDS-PAGE (Figure 3.16).

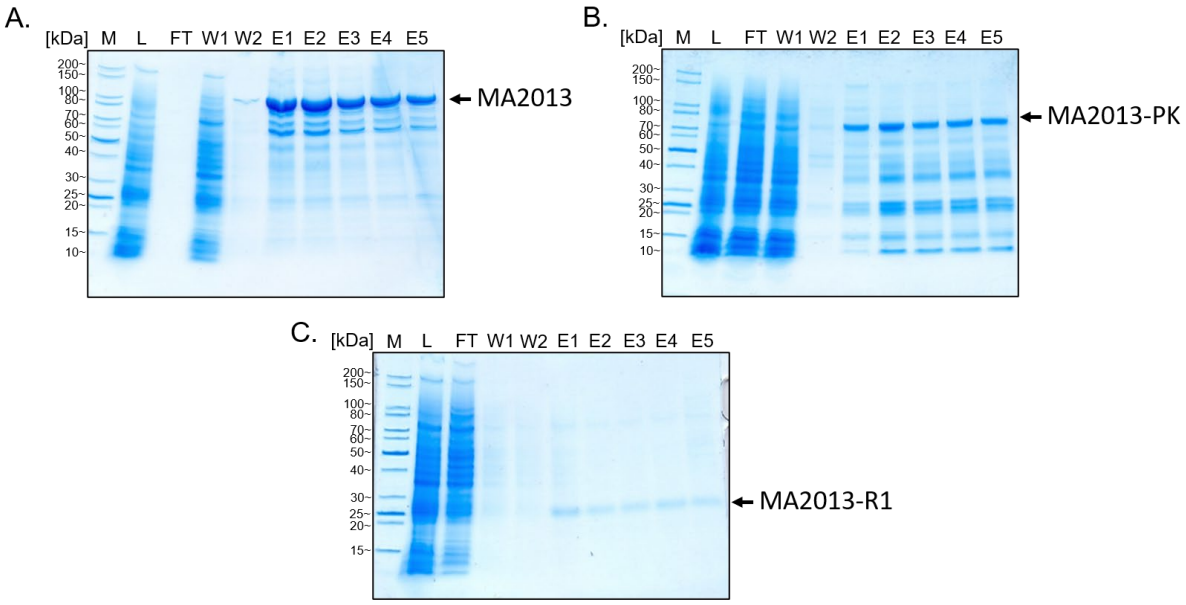

Figure 3.16: Purification of the MA2013 protein variants after recombinant production and affinity chromatography purification with Strep-Tactin® Sepharose®. Sample application of lysate after cell lysis (L), flow-through (FT), wash fractions (W) and protein elution fractions (E). (**A.**) Coomassie stained gel (4 %-20 % gradient) of MA2013. MA2013 has a predicted molecular weight of ~116 kDa. (**B.**) Coomassie stained gel (4 %-20 % gradient) of MA2013-PK. MA2013-PK has a predicted molecular weight of ~80 kDa. (**C.**) Coomassie stained gel (4 %-20 % gradient) of MA2013-R1. MA2013-R1 has a predicted molecular weight of ~25 kDa. As reference the unstained protein standard (M) from NEB was used.

The coomassie stained SDS-PAGE of the corresponding purification revealed a protein with a relative molecular weight of ~110 kDa (MA2013), ~80 kDa (MA2013-PK) and ~25 kDa (MA2013-R1). The molecular weight of the purified proteins is consistent with the predicted molecular weight of ~116 kDa (MA2013), ~80 kDa (MA2013-PK) and ~25 kDa (MA2013-R1) calculated from the amino acid sequence (chapter 2.5.7).

To determine the oligomerisation state of the full-length protein MA2013 a SEC using a Superdex® 200 Increase 10/300 GL column (Figure S6) was conducted. Two major A_{280} peaks were observed with a retention volume of 7.6 ml and 9.5 ml. The first peak (7.6 ml) represents most likely aggregates. The peak at 9.5 ml corresponds to a relative molecular weight of 505 kDa and presumably represents MA2013. The calculated molecular weight of peak 2 corresponds to a tetramer. It is not likely that MA2013 represents a tetramer. Further

experiments are required to determine whether this molecular weight can be explained with the long-stretched structure of the protein or the aggregation due to high protein concentration.

3.8 MA2013 shows autokinase and transphosphorylation activity

The autokinase activity of the two recombinant protein variants MA2013 and MA2013-PK were investigated with radioactive autophosphorylation assays (Figure 3.17, A./B.). The increasing radioactive signal on the autoradiogram corresponds to autokinase activity of the proteins MA2013 and MA2013-PK. It is noticeably that the autokinase activity of MA2013-PK decreases earlier than for MA2013. The differences in the autokinase activity between the two MA2013 variants might be due to the bound REC and HPt domain of the full-length protein. After the autokinase reaction, the phosphoryl group is likely intramolecular transferred to the REC and HPt domain and therefore the signal on the autoradiogram remains visible for a longer time.

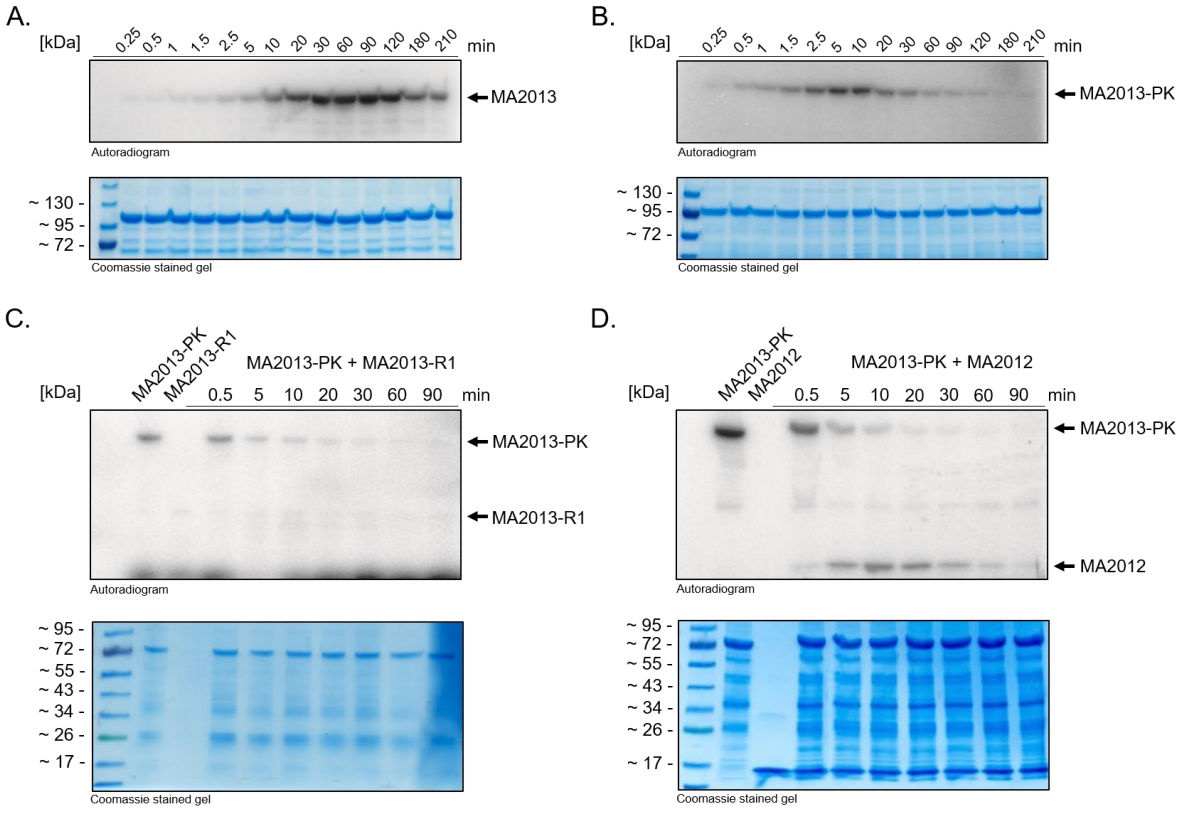

Figure 3.17: Autokinase activity and transphosphorylation activity of MA2013. (**A.**) Autoradiogram and coomassie stained gel (4-20% gradient) after in vitro autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM MA2013 and the reaction was terminated after different timepoints. (**B.**) Autoradiogram and coomassie stained gel (4-20% gradient) after in vitro autophosphorylation assay of MA2013-PK, a truncated protein variant without the REC and HPt domain. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM MA2013-PK and the reaction was terminated after different timepoints (**C.**) Transphosphorylation of MA2013-PK to the bound REC domain MA2013-R1. Purified recombinant MA2013-PK (10 µM) was phosphorylated with [γ-^{32}P]-ATP, after removing excess ATP with Microspin™G-50 Columns, the second protein (MA2013-R1) was added in equimolar amount and the reaction was terminated after different timepoints. (**D.**) Transphosphorylation of MA2013-PK to REC only RR MA2012. Purified recombinant MA2013-PK (10 µM) was phosphorylated with [γ-^{32}P]-ATP, after removing excessive ATP with Microspin™G-50 Columns, the second protein (MA2012) was added in equimolar amount and the reaction was terminated after different timepoints. The coomassie stained gel shows that protein concentration was the same for every time point. As reference the color prestained protein standard (M) from NEB was used.

In a next step the transphosphorylation activity of MA2013-PK to the bound REC-domain (MA2013-R1) and to the REC-only RR (MA2012), which is encoded in genomic proximity to MA2013, was investigated (Figure 3.17, C./D.). For this, the protein MA2013-PK was autophosphorylated and the phosphoryl group transfer to MA2013-R1 and MA2012 were examined. A time-dependent decrease of the radioactive signal of the protein MA2013-PK on the autoradiogram and a simultaneous appearance of a signal for the protein MA2013-R1 or MA2012 corresponds to transphosphorylation activity. The autokinase activity and the transphosphorylation activity towards MA2013-R1 and MA2012 suggested that MA2013 is likely involved in a phosphorelay. The function of the HPt domain has still to be investigated.

3.9 The full-length membrane hybrid kinase MA4377

The putative HHK MA4377 is a membrane-bound representative of the MA_type 1 kinases with a CHASE4 domain which is linked to a cytoplasmic HAMP and PAS domain. The central HK domain, consisting of a DHp and CA domain, is C-terminally fused to two REC domains (R1 and R2) (Figure 3.18). In addition, a REC-only RR (R3) is encoded downstream (MA_4376), overlapping with MA_4377 by 29 nucleotides. Just like MA2013 the complexity of MA4377 suggests a putative role in a phosphorelay system. Since MA4377 is the most complex kinase of the four selected kinases, its function and phosphorylation cascade were investigated in more detail.

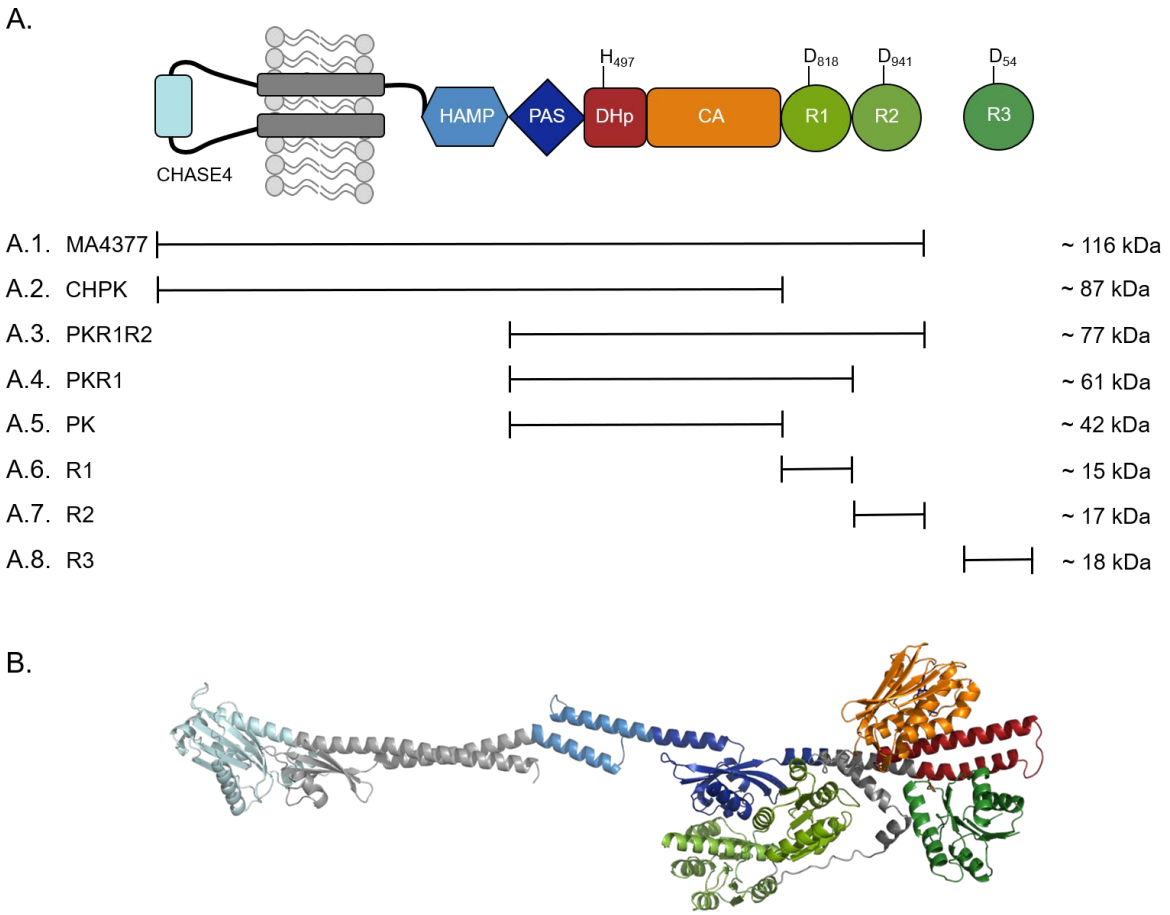

Figure 3.18: Schematic domain representation and predicted protein structure of the putative HK MA4377. (**A.**) Schematic representation of the domain structure and overview of truncated variants of the protein. All variants of the protein MA4377 used in this study are shown with their molecular weight. (**A.1.**) Full length membrane protein MA4377, (**A.2.**) MA4377 without C-terminal REC domains, (**A.3.**) MA4377 without the CHASE and HAMP domain, (**A.4.**) MA4377 without the CHASE4, HAMP and R2 domains, (**A.5.**) MA4377 without CHASE, HAMP, R1 and R2 domains, (**A.6.**) Only the R1 domain of MA4377, (**A.7.**) Only the R2 domain of MA4377. (**A.8.**) REC only RR MA4376 (R3). (**B.**) Protein structure of MA4377 predicted with AlphaFold3 (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

Since the overproduction of membrane proteins can be toxic to microorganisms, different *E. coli* strains were examined with a test expression to find the best strain to express the gene MA_4377. The first tested strain was *E. coli* BL21(DE3), which is still the gold standard for protein production. The strain *E. coli* C41 is based on *E. coli* BL21(DE3) and has at least one mutation that prevents negative effects of toxic proteins. *E. coli* C43 was obtained from *E. coli* C41 by selecting for resistance to a toxic protein. Additionally, the strain *E. coli* Nissle 1917 was tested. The test expression revealed that *E. coli* Nissle 1917 did not lead to the expression of the gene MA_4377 (Figure 3.19). The SDS-PAGE and the western blot showed no difference between the three other strains (Figure 3.19). For further experiments the gene was heterologously overexpressed in *E. coli* BL21(DE3) and *E. coli* C43 to produce MA4377. The protein yield shows no noticeable difference (data not shown).

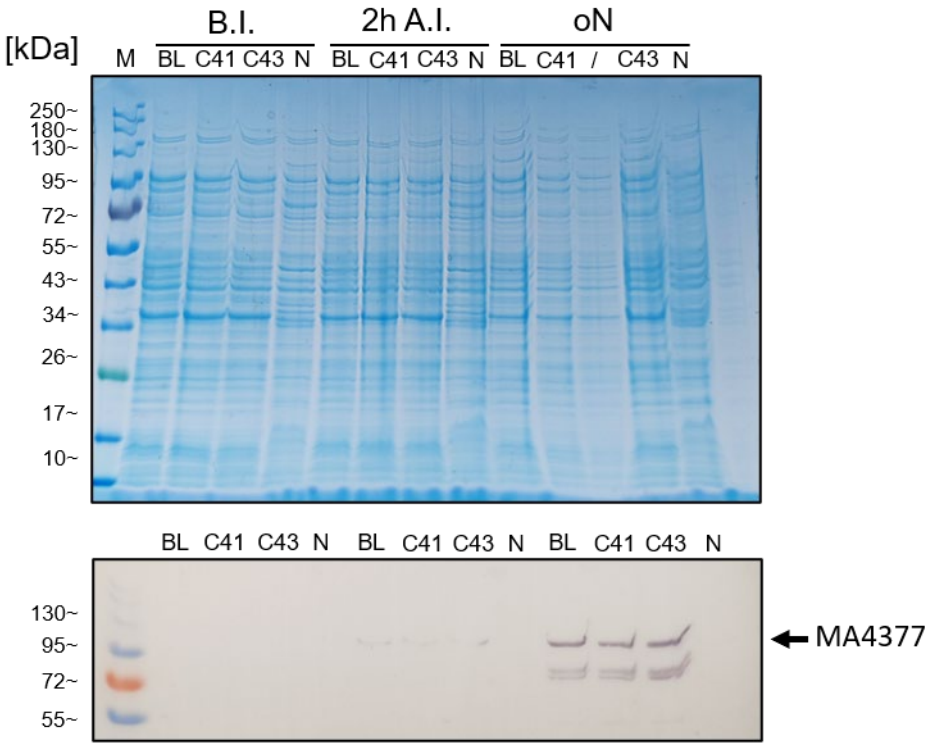

Figure 3.19: Test expression of MA4377 employing four different *E. coli* strains. Coomassie stained gel (4 %-20 % gradient) and western blot of MA4377 after test expression with the strains *E. coli* BL21(DE3) (BL), *E. coli* C41 (C41), *E. coli* C43 (C43) and *E. coli* Nissle 1917 (N). Sample application of before induction (B.I.), 2 h after induction (2h A.I.) and over night (oN). As reference the color prestained protein standard (M) from NEB was used. MA4377 has a predicted molecular weight of ~116 kDa.

MA4377 and the membrane protein variant without bound REC domains (CHPK) were purified applying two different membrane extraction protocols. First the proteins MA4377 and CHPK were purified after the solubilisation with the detergent dodecyl-β-D-maltoside (DDM). Using detergents to purify membrane proteins is still a widely used method, but it has the downside that the detergent has always to be present in the buffers after the purification. Furthermore, the protein is removed from the membrane and packed into a protein-detergent complex, which might influence the folding and therefore also the function of the protein (Thoma & Burmann, 2020). The second method used for the purification of MA4377 was the membrane extraction with the nanodisc diisobutylene-maleic acid (DIBMA). DIBMA works like a cookie cutter, that cuts out the protein and parts of the surrounding membrane (Thoma & Burmann, 2020). The protein MA4377 was successfully purified with both methods (Figure 3.20). The coomassie stained SDS-PAGE and a western blot after the purification of the full-length protein MA4377 revealed a protein with a relative molecular weight of ~110 kDa, which is consistent with the predicted molecular weight of ~116 kDa based on the amino acid sequence (chapter 2.5.7).

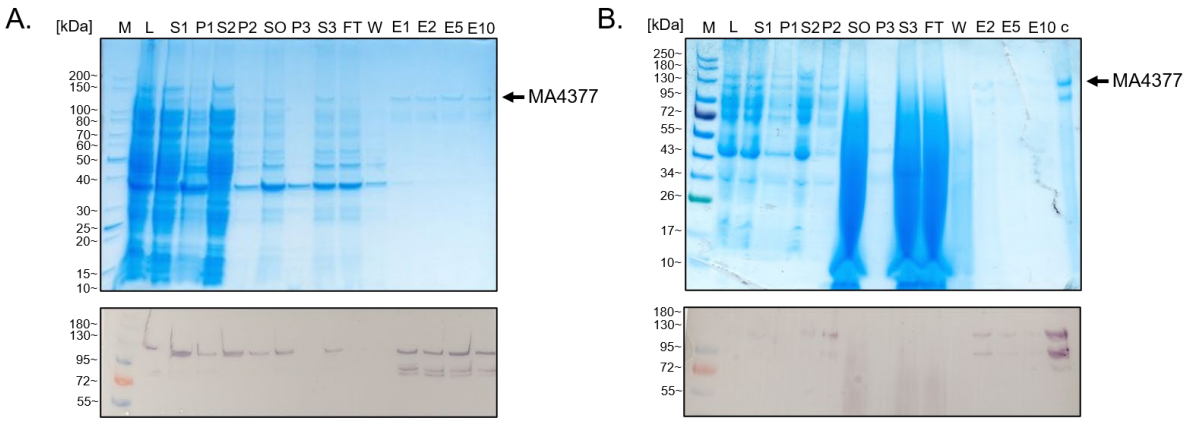

Figure 3.20: Purification of the membrane protein MA4377. (**A.**) Coomassie stained gel and western blot of MA4377 after recombinant production, solubilisation with DDM and affinity chromatography purification with Strep-Tactin®XT 4Flow® high capacity. Sample application of lysate after cell lysis (L), flow-through (FT), wash fractions (W) and protein elution fractions (E). (**B.**) Coomassie stained gel and western blot of MA4377 after recombinant production, solubilisation with DIBMA and affinity chromatography purification with Strep-Tactin®XT 4Flow® high capacity. Sample application of lysate after cell lysis (L), flow-through (FT), wash fractions (W), protein elution fractions (E) and concentrated protein (c). As reference the unstained/color prestained protein standard (M) from NEB was used. MA4377 has a predicted molecular weight of ~116 kDa.

HKs typically function as homodimers. To analyse the oligomerisation state of the full-length protein MA4377 in kinase buffer (purified with DDM) a SEC was conducted, using a Superdex® 200 Increase 10/300 GL column (Figure 3.21). Four A_{280} peaks were observed with a retention volume of 11.7 ml, 14.1 ml, 17.1 ml and 18.8 ml. Peak 1 at 11.7 ml corresponds to a relative molecular weight of 230 kDa and presumably represents MA4377. To verify this, fractions of the eluted sample were analysed with an SDS-PAGE. The SDS-PAGE revealed a ~116 kDa protein in the fractions 9–15, with the highest protein amount in fraction 13/14 corresponding to the peak 1 at 11.7 ml. A western blot confirmed the data of the SDS-PAGE (Figure 3.21, C.). The calculated molecular weight corresponds to a dimer. To get an impression how a homodimer of MA4377 would look like a three-dimensional protein structure was predicted with AlphaFold3 (Abramson *et al.*, 2024) (Figure 3.21, D.). The predicted homodimer of MA4377 is a long-stretched protein. The grey part between the CHASE and the HAMP domain represents the transmembrane helices. The HK domain shows the typical structure of the CA and DHp domain and both REC domains are positioned close to the HK domain.

Considering the SEC data for the other putative HKs, which suggest a tetrameric structure, the results of MA4377 provide further evidence that the applied protein concentration may be the reason that the putative HKs do not elute as dimers. The protein concentration used for the SEC of the membrane protein was clearly lower than for the cytosolic proteins. This is consistent with the results from Köhler (2021). This study used very low concentrations of a RdmS homolog, resulting in a dimeric structure (Köhler, 2021). Moreover, MA4377 exhibits a similar long-stretched structure compared to the other putative HKs, indicating that the

structural characteristics do not noticeable impact the determination of the oligomerisation state.

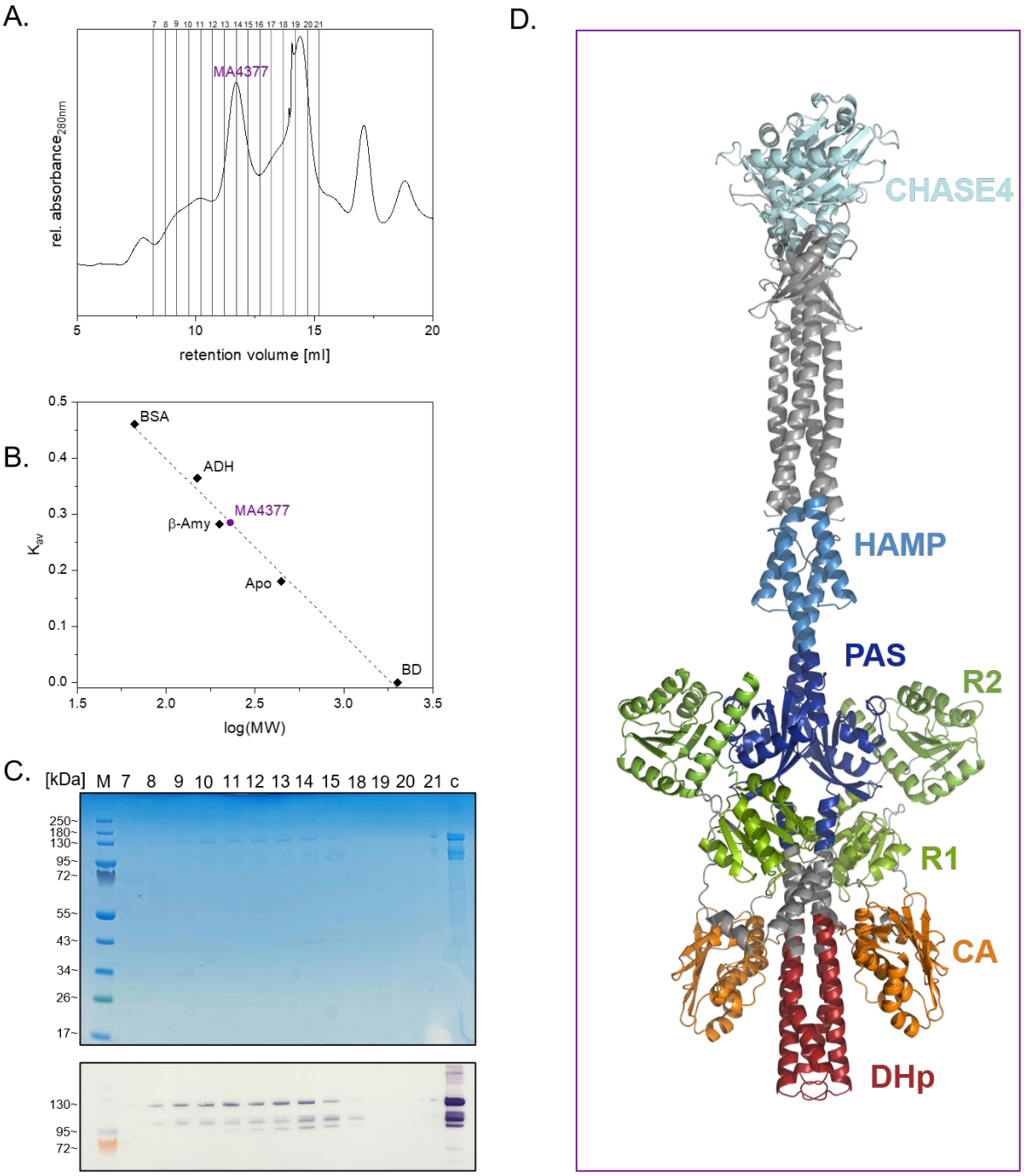

Figure 3.21: Determination of oligomerisation state and predicted homodimer structure of MA4377. (**A.**) Elution profile of MA4377 solubilised with DDM. Size exclusion chromatography was performed with a Superdex® 200 Increase 10/300 GL column equilibrated with kinase buffer, pH 8.0 containing 5 % glycerol. The relative absorbance was detected at 280 nm and plotted against the retention volume. Sample was fractionated in 0.5 ml steps. Fractionated samples are displayed with vertical lines and numbers. (**B.**) Calibration curve of the Superdex® 200 Increase 10/300 GL column. Marker proteins: carbonic anhydrase (CA; 29 kDa), bovine serum albumin (BSA; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), β-amylase (β-Amy; 200 kDa) and apoferritin (Apo; 443 kDa), blue dextran (BD; 2000 kDa). The void volume was determined using blue dextran (BD). The actual column volume was determined with acetone. The calibration curve was obtained by plotting the K_{av} of the standard proteins against the logarithm of the molecular weight (black squares). The fitting was determined using linear regression in OriginLab. MA4377 is indicated on the calibration curve (purple dot). (**C.**) Coomassie stained gel and western plot of the fractionated samples to determine which peak belongs to the protein MA4377. As reference the color prestained protein standard (M) from NEB was used. Monomeric MA4377 has a predicted molecular weight of

~116 kDa. (**D.**) AlphaFold3 prediction of MA4377 as a homodimer (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020)

3.10 Recombinant full-length MA4377 shows no autokinase activity

After successful purification of the full-length protein MA4377, the activity of the putative HK was tested with *in vitro* kinase assays. To examine the autokinase activity, the recombinant protein MA4377 was incubated with [γ-^{32}P]-ATP. A first autophosphorylation assay showed no autokinase activity for the full-length membrane protein purified with DDM (data not shown). Subsequently, different factors that might affect the activity of the membrane protein were systematically investigated.

MA4377 was purified using two different membrane isolation methods. Besides to MA4377 purified with the detergent DDM, also the protein purified with the nanodisc DIMBA was tested with a kinase assay. Compared to the purification with DDM, DIBMA allows the extraction of the protein surrounded by an intact membrane, and it is not connected to a potential detergent-like activity interference. However, the protein MA4377 purified with both methods has no autokinase activity. A next step was to test the autokinase activity of the protein variant CHPK. This protein variant lacks the C-terminal REC domains. The REC domains could lead to a lack of autokinase activity signal if the phosphate transfer and additional dephosphorylation at the Asp residue would occur in a short time (Immormino *et al.*, 2016). The assay revealed no autokinase activity of CHPK, which leads to the assumption that the REC domains have no effect on this reaction (Wojnowska *et al.*, 2013) (Figure 3.22, A.).

Furthermore, different *E. coli* strains were used to produce MA4377. Besides the common strain *E. coli* BL21 (DE) also the strain *E. coli* C43 (DE) was used. *E. coli* C43 (DE) is optimized for the overproduction of toxic proteins (Dumon-Seignovert *et al.*, 2004; Miroux & Walker, 1996; Studier & Moffatt, 1986). However, neither the protein yield, nor the activity of the protein differed between the two strains. Additionally, the strain *E. coli* TKR 2000 was used as an expression host. *E. coli* TKR 2000 is lacking the F_1F_0 ATPase activity and prevents the fast depletion of ATP, which could lead to a false negative result in autokinase assays (Jung *et al.*, 1997; Kollmann & Altendorf, 1993). However, the assay revealed no autokinase activity for MA4377, while the control (PK) showed a clear signal on the autoradiogram (Figure 3.22, A.).

After testing the activity of MA4377 purified with different methods and produced in different *E. coli* strains without getting an active protein, the question arose if a missing input signal can lead to a lacking autokinase activity of MA4377. MA4377 has a membrane bound CHASE4 domain. This domain might be responsible for sensing signals from cell surroundings. A literature search revealed that CHASE domains can sense different signals and are divided

into different subgroups (CHASE2, CHASE3, CHASE4) (Zhulin *et al.*, 2003). Several studies report that CHASE domains are responsible to sense cytokinins (Wang *et al.*, 2017). A multiple sequence alignment with common CHASE domains of different organisms revealed that the CHASE4 domain of MA4377 is highly similar to other CHASE4 domains of *M. acetivorans* and to the CHASE4 domain of the protein FilI of *Methanosaeta harundinacea* (Supplement Figure S3). The CHASE domain of PcrK, a HHK similar to MA4377, differs from the CHASE4 domain of MA4377. The HHK PcrK of *Xanthomonas campestris* is involved in cytokinin sensing, while the TCS (FilI/FilR) is known to sense N-acyl homoserine lactones (AHLs) associated with quorum sensing (Li *et al.*, 2014; Wang *et al.*, 2017). With an *in vitro* autokinase assay, the activity of the protein MA4377 was examined in the presence of different potential input signals. These included three cytokinins, two AHLs and acetate as a metabolic substrate of *M. acetivorans*. The effect of the different potential input signals was analysed in comparison to pure MA4377 and a positive control (PKR1R2).

Table 3.1: Potential input signal for the activity of MA4377.

Potential input signal	Abbreviation	Concentration
N6-(2-Isopentenyl) adenine	2iP	10 µM
adenine sulfate	A	10 µM
trans-zeatin	Z	10 µM
N-dodecanoyl-L-Homoserine lactone	12-HSL	10 µM
N-(Ketocaproyl)-DL-homoserine lactone	DL-HSL	10 µM
acetate	Ac	10 µM, 20 µM

The results lead to the assumption that none of the tested substances leads to autokinase activity of MA4377 (Figure 3.22, B.). Further experiments are required to elucidate whether the selected substances have no effect on the protein, whether the concentration was not optimal, or if the protein remains inactive for other reasons.

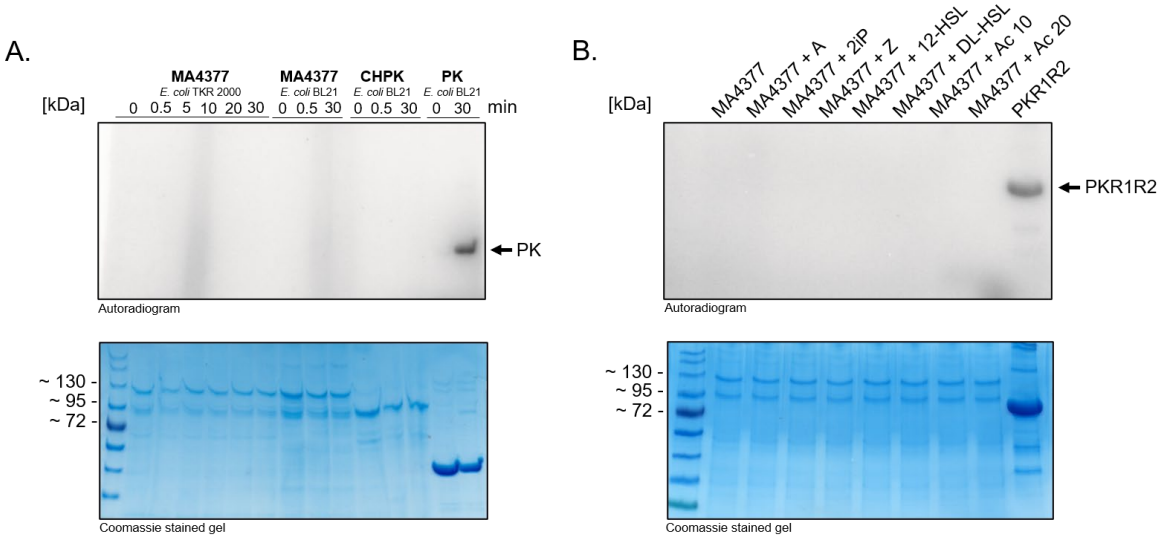

Figure 3.22: Autokinase activity of heterologously produced membrane MA4377 variants. (**A.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to MA4377 variants and the reaction was terminated after different timepoints. The coomassie stained gel indicate the loaded protein. The autoradiogram displays a signal for PK after 30 min. The protein MA4377, produced with the strains *E. coli* TKR2000 and *E. coli* BL21(DE3) and additionally the protein CHPK *E. coli* BL21(DE3) reveal no signal on the autoradiogram. (**B.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP and potential input signal were added to MA4377 (equimolar) and the reaction was terminated after 20 min. N6-(2-Isopentenyl)adenine (2iP), adenine sulfate (A), trans-zeatin (Z), N-dodecanoyl-L-Homoserine lactone (12-HSL), N-(Ketocaproyl)-DL-homoserine lactone (DL-HSL), Acetate (Ac), 10 = 10 µM, 20 = 20 µM. As reference the color prestained protein standard (M) from NEB was used.

In a next step the cytosolic truncated variants of MA4377 were investigated regarding their autokinase activity.

3.11 Purification of the truncated cytosolic MA4377 variants and R3

The cytosolic truncated variants of the protein MA4377 were purified via affinity chromatography and subsequently analysed by SDS-PAGE. The coomassie stained SDS-PAGE of the corresponding purification revealed proteins with relative molecular weight of ~72 kDa (PKR1R2), ~60 kDa (PKR1), ~40 kDa (PK), ~15 kDa (R1), ~18 kDa (R2), ~17 kDa (R3) (Figure 3.23). The molecular weight of the purified proteins is consistent with the predicted molecular weight calculated from the amino acid sequence (chapter 2.5.7). The oligomerisation state of the truncated protein variant PKR1R2 and R3 were determined via SEC using a Superdex® 200 Increase 10/300 GL column (Figure S7, Figure S8). For PKR1R2 three A_{280} peaks were observed with a retention volume of 7.7 ml, 9.2 ml and 10.7 ml. The peak at 7.7 ml elutes in the void volume and most likely represents aggregates. The peak at 9.4 ml corresponds to a relative molecular weight of 740 kDa and the peak at 10.7 ml corresponds to a relative molecular weight of 320 ml. To analyse which peak represents the protein PKR1R2, fractions of the eluted sample were collected and loaded on an SDS-PAGE. The SDS-PAGE revealed a ~72 kDa protein in the fractions B10-B2, with the highest protein amount in fraction B4/B3. This reveals that PKR1R2 is present in both peaks but with a higher

concentration in the peak at 10.7 ml. The molecular weight of 320 kDa corresponds to a tetrameric structure of the protein variant. Considering the results of the SEC data of the previous chapters, it is likely that PKR1R2 elutes as a tetramer due to the high protein concertation. The predicted three-dimensional structure of PKR1R2 reveals a more globular structure of the protein, which additionally provides confirmation that the long-stretched structure observed in other proteins is likely not the reason for the high molecular weight. For R3 one major peak was observed with a retention volume of 15.9 ml corresponding to a molecular weight of 18.5 kDa most likely representing R3 as a monomer. To verify this, fractions of the eluted sample were collected and analysed via SDS-PAGE. The SDS-PAGE revealed a ~18 kDa protein in the fractions B7 and C6, confirming that R3 is a monomer.

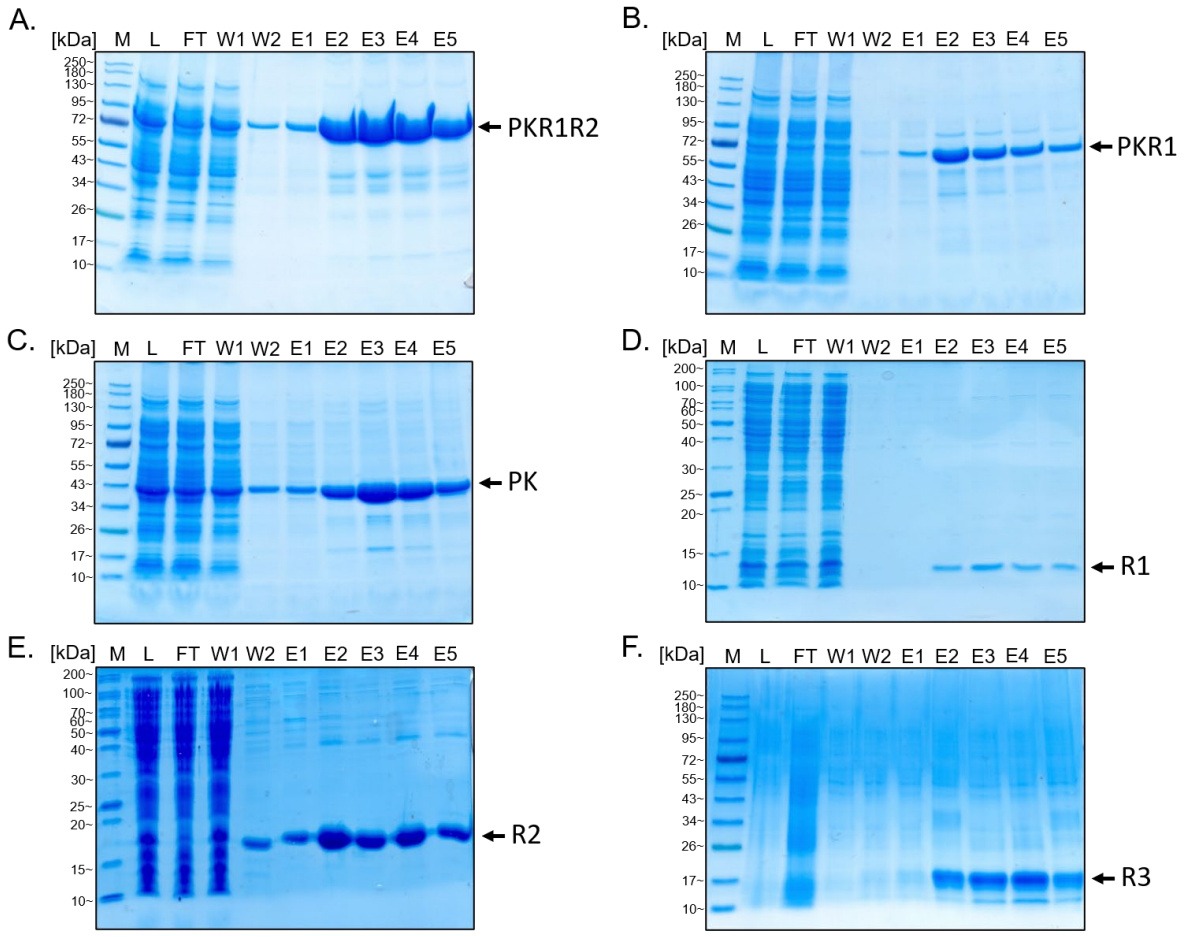

Figure 3.23: Purification of the MA4377 protein variants after recombinant production and affinity chromatography purification with Strep-Tactin® Sepharose®. Sample application of lysate after cell lysis (L), flow-through (FT), wash fractions (W) and protein elution fractions (E). (**A.**) Coomassie stained gel (4 %-20 % gradient) of PKR1R2. PKR1R2 has a predicted molecular weight of ~77 kDa. (**B.**) Coomassie stained gel (4 %-20 % gradient) of PKR1. PKR1 has a predicted molecular weight of ~61 kDa. (**C.**) Coomassie stained gel (4 %-20 % gradient) of PK. PK has a predicted molecular weight of ~42 kDa. (**D.**) Coomassie stained gel (15 %) of R1. R1 has a predicted molecular weight of ~15 kDa. (**E.**) Coomassie stained gel (15 %) of R2. R2 has a predicted molecular weight of ~17 kDa. (**F.**) Coomassie stained gel (4 %-20 % gradient) of R3. R3 has a predicted molecular weight of ~18 kDa. As reference the unstained/color prestained protein standard (M) from NEB was used.

3.12 Truncated cytosolic MA4377 variants show autokinase activity

Since the full-length membrane protein MA4377 did not show autokinase activity, it was subsequently tested if truncated cytosolic variants of MA4377 are active. If the membrane domain of MA4377 is responsible for the lacking autokinase activity cytosolic variants of the protein could still be active. A literature search revealed that most authors worked with truncated cytosolic variants of membrane-bound kinases and were able to show autokinase activity (Barr *et al.*, 2023; Francis *et al.*, 2018; Li *et al.*, 2014; Najnin *et al.*, 2016; Vega-Baray *et al.*, 2022). Three different truncated cytosolic variants were investigated. The variant PKR1R2 is lacking the CHASE4 and HAMP domain, while PKR1 is additionally lacking the second REC domain (R2) and the variant PK consists of only the PAS and the HK domain. Interestingly, the recombinant protein variants PKR1R2, PKR1 and PK showed clear autokinase activity after incubation with [γ-^{32}P]-ATP. For all three variants a clear time-dependent increase of autokinase activity was revealed (Figure 3.24, A./B./C.). Based on the amino acid sequence alignment His 497 is the phosphorylation site within MA4377 (Figure 3.3). To confirm H497 as the conserved histidine residue, the protein variant PK_{H497Q} was generated, in which the His residue H497 was replaced by glutamine (Q) using site-directed mutagenesis. This protein variant is binding ATP, but not showing autokinase activity (Figure 3.24, F.). Moreover, the protein variants $PK_{H497Q}R1$, $PK_{H497Q}R1_{D818N}R2$ and $PK_{H497Q}R1R2$ were generated. Here also the H497 residue was replaced by glutamine (Q) and for one construct the D818 residue was replaced by asparagine (N) via site-directed mutagenesis. All protein variants were tested with autophosphorylation assays, and it was confirmed that H497 is crucial for the autokinase activity of the protein (Figure 3.24, D./E./F.).

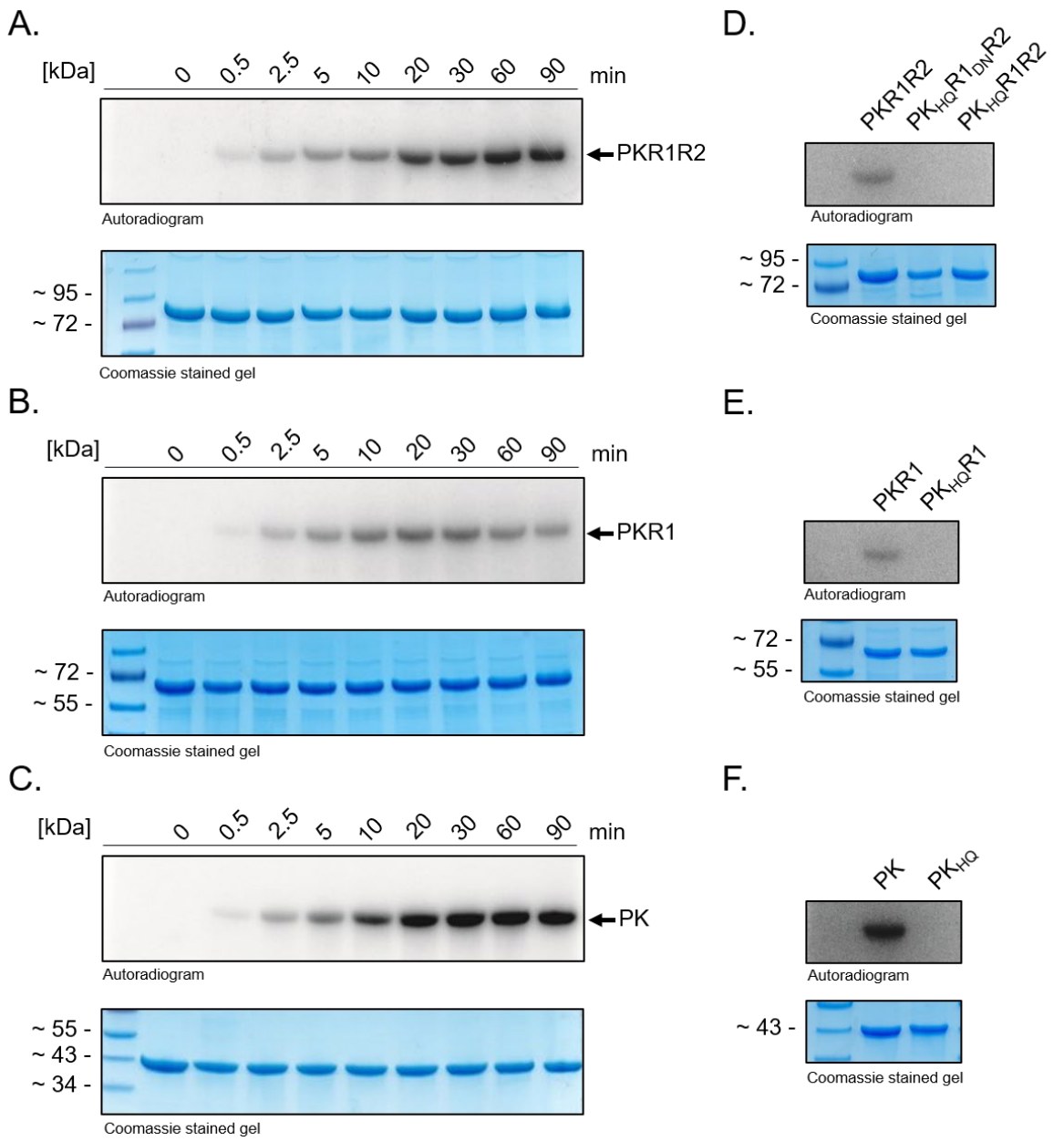

Figure 3.24: Autokinase activity of heterologously produced cytosolic MA4377 variants. (**A.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM PKR1R2 and the reaction was terminated after different timepoints with SDS sample buffer (with BME). The coomassie stained gel shows the loaded protein. The autoradiogram displays an increasing signal over 90 min, corresponding to autokinase activity. (**B.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM PKR1 and the reaction was terminated after different timepoints with SDS sample buffer (with BME). The coomassie stained gel shows the loaded protein. The autoradiogram displays an increasing signal over 90 min, corresponding to autokinase activity. (**C.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM PK and the reaction was terminated after different timepoints with SDS sample buffer (with BME). The coomassie stained gel shows the loaded protein. The autoradiogram displays an increasing signal over 90 min, corresponding to autokinase activity. (**D.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM PKR1R2, PK$_{H497Q}$R1$_{D818N}$R2 and PK$_{H497Q}$R1R2 and the reaction was terminated after 20 min with SDS sample buffer (with BME). The coomassie stained gel shows the loaded protein. The autoradiogram displays a signal only if the residue H497 is present. (**E.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM PKR1 and PK$_{H497Q}$R1 and the reaction was terminated after 20 min with SDS sample buffer (with BME). The coomassie stained gel shows the loaded protein. The autoradiogram displays a signal only if the residue H497 is present. (**F.**) Autoradiogram and coomassie stained gel (4-20% gradient) after *in vitro* autophosphorylation assay. Radiolabeled [γ-^{32}P]-ATP was added to 10 µM PK and PK$_{H497Q}$ and the reaction was terminated after 20 min with SDS sample buffer (with BME). The coomassie

stained gel shows the loaded protein. The autoradiogram displays a signal only if the residue H497 is present. As reference the color prestained protein standard (M) from NEB was used.

3.13 Transphosphorylation to the bound REC domains and to R3

HHK are often involved in complex multistep phosphorelays. MA4377 is a hybrid kinase with two C-terminal REC domains (R1 & R2) and a REC-only RR (R3) in genomic proximity (Figure 3.18). The question arose if the fused REC domains and R3 are involved in a phosphorelay reaction of MA4377. To investigate the transphosphorylation activity of MA4377, the protein variant PK was autophosphorylated and the transphosphorylation towards R1, R2 and R3 was explored.

First the protein variant PK was autophosphorylated with [γ-^{32}P] ATP. After removing excessive ATP, the protein variants $PK_{H497Q}R1$ or $PK_{H497Q}R1_{D818N}R2$ were added in equimolar amounts (Figure 3.25, A./B.). The same experiment was also performed for R1 and R2 as individual proteins (Figure 3.25, C./D.).

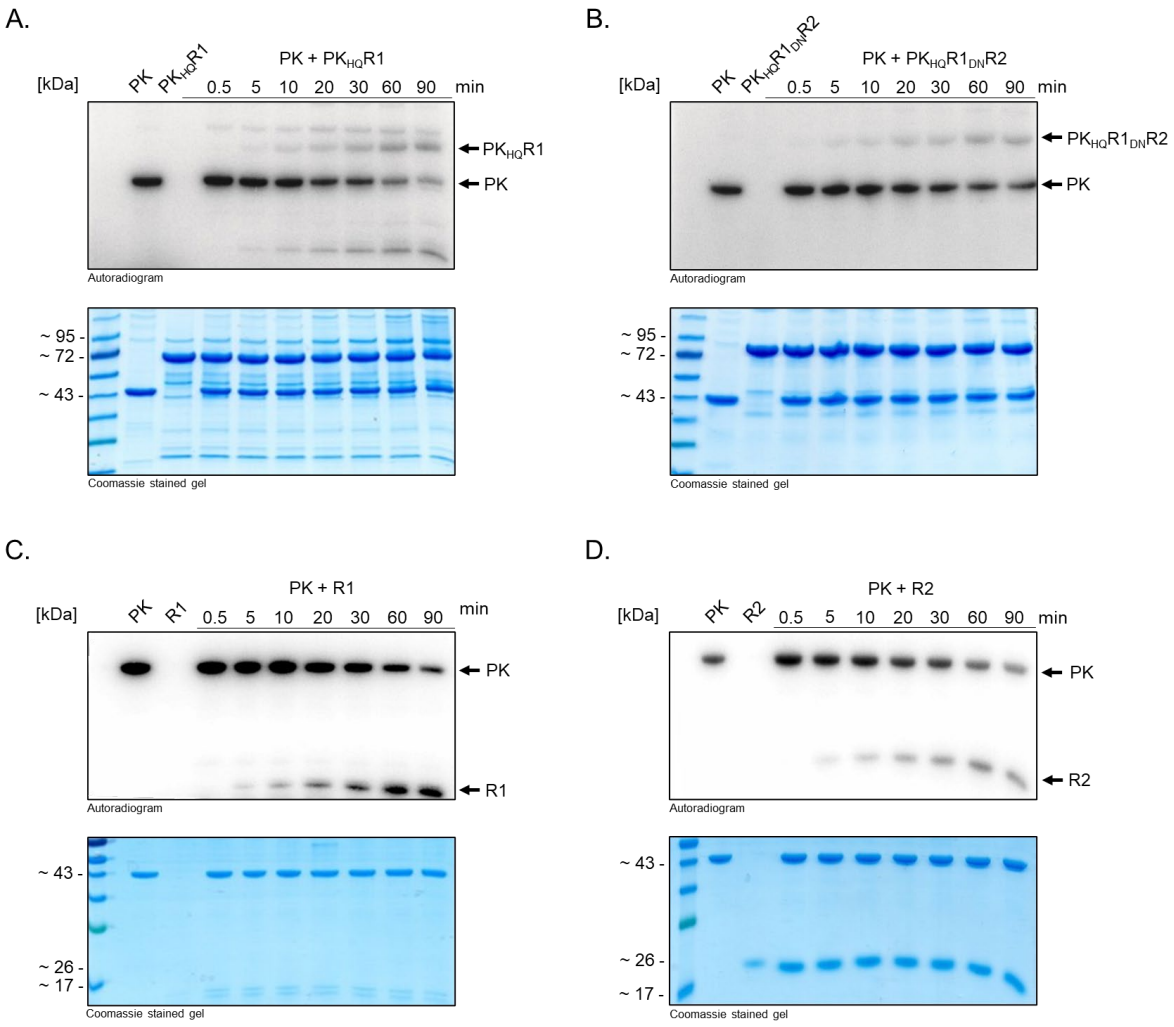

Figure 3.25: Transphosphorylation assays of the protein variant PK. Autoradiogram and corresponding SDS-PAGE (Coomassie stained gel) of the transphosphorylation reaction of different protein variants. Purified

recombinant PK (10 µM) was phosphorylated with [γ-^{32}P]-ATP, after removing excessive ATP with Microspin™G-50 Columns, the second protein was added in equimolar amount. (**A.**) Transphosphorylation assay with PK and $PK_{H497Q}R1$. (**B.**) Transphosphorylation assay of PK with $PK_{H497Q}R1_{D818N}R2$. (**C.**) Transphosphorylation assay with PK and R1. (**D.**) Transphosphorylation assay with PK and R2. As reference the color prestained protein standard (M) from NEB was used.

The transphosphorylation assays revealed that the radioactive autokinase signal of the variant PK was decreased over time, while a signal of the added protein variants ($PK_{H497Q}R1$, $PK_{H497Q}R1_{D818N}R2$, R1 or R2) appeared on the autoradiogram. This verified transphosphorylation activity to both bound REC domains (Figure 3.25).

In a next step the transphosphorylation activity to the REC-only RR R3 was investigated with the same experimental setup. Immediately after the addition of R3 to the autophosphorylated protein PK, a fast transfer of the radioactive signal was observed, while the R3 protein itself showed no autophosphorylation activity (Figure 3.26, A.). The transphosphorylation of PK to R3 seems to be faster than to the bound REC domains R1 and R2.

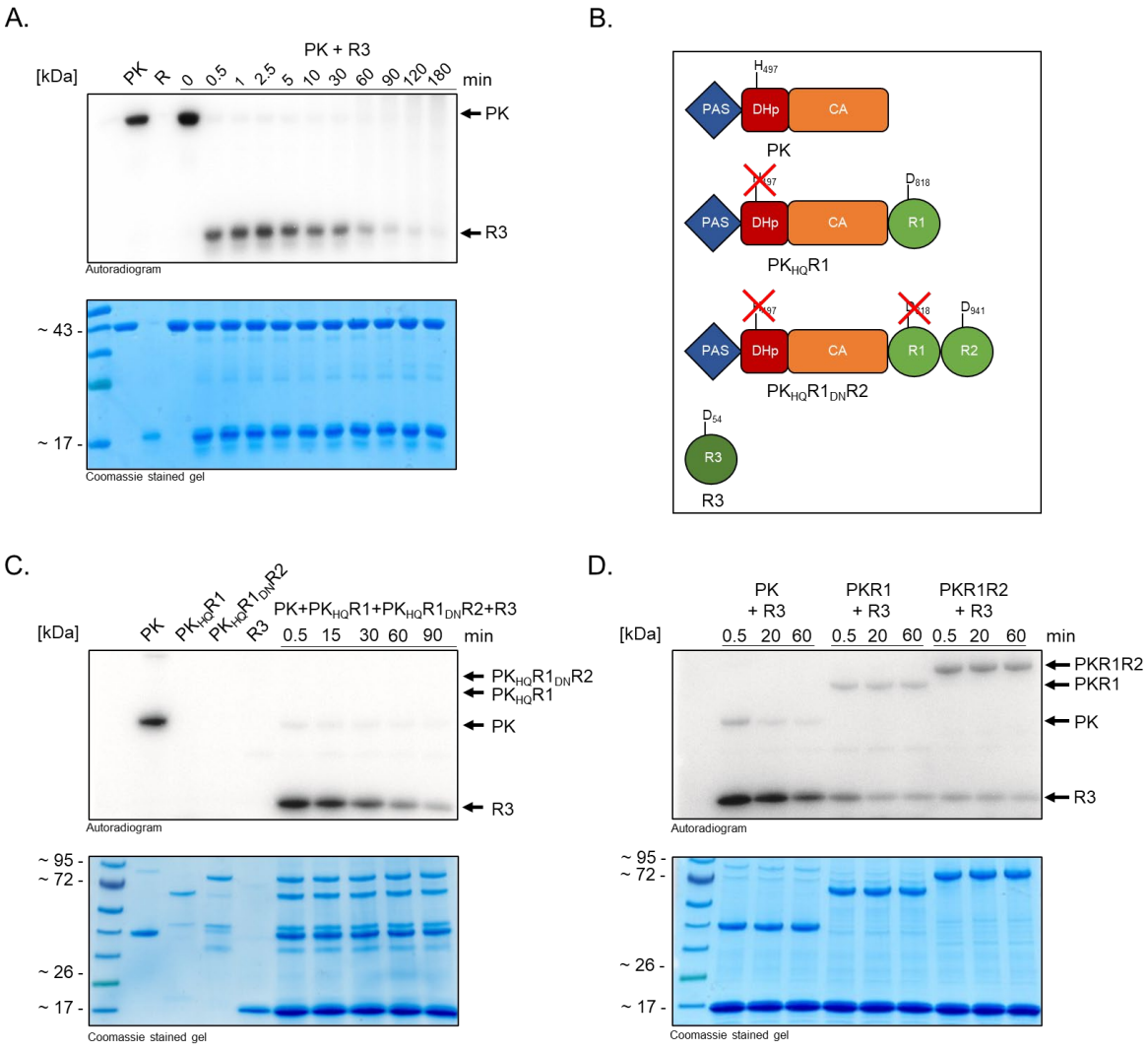

Figure 3.26: Transphosphorylation assays of different cytosolic MA4377 variants. Autoradiogram and corresponding SDS-PAGE (Coomassie stained gel) of the transphosphorylation reaction of different protein

variants. (**A.**) Purified recombinant PK (10 µM) was phosphorylated with [γ-^{32}P]-ATP, after removing excessive ATP with Microspin™G-50 Columns, the second protein (R3) was added in equimolar amount. (**B.**). Schematic representation of applied protein variants (**C.**) Transphosphorylation assay with PK and $PK_{H497Q}R1$, $PK_{H497Q}R1_{D818N}R2$, R3. PK was autophosphorylated and all constructs were added simultaneously. The H497 and D818 residues were replaced by glutamine (Q) and asparagine (N) via site directed mutagenesis to prevent autophosphorylation of the constructs. In case of $PK_{H497Q}R1_{D818N}R2$ to additionally prevent phosphoryl group transfer to R1. (**D.**) Transphosphorylation assay with different MA4377 variants (PK, PKR1, PKR1R2). The signal that corresponds to the autophosphorylated kinase and the REC/RR is indicated by an arrow. As reference the color prestained protein standard (M) from NEB was used.

After observing phosphotransfer to R1, R2 and R3 the question arose whether there is a preference for the transfer when all three proteins are present simultaneously. For this the protein variant PK was again autophosphorylated with [γ-^{32}P]-ATP. After removing excessive ATP, the protein variants $PK_{H497Q}R1$, $PK_{H497Q}R1_{D818N}R2$ and R3 were simultaneously added in equimolar amount (Figure 3.26, B.). Interestingly, R3 seems to outcompete R1 and R2 as it is the only protein being phosphorylated by PK (Figure 3.26, C.).

Together with the previous results, this leads to the assumption that the kinase preferable transfers phosphate to R3. Another experiment with cytosolic MA4377 variants with intact bound REC domains showed a similar result. For this the variants PKR1R2, PKR1 and PK were autophosphorylated with [γ-^{32}P]-ATP. After removing excessive ATP, the protein R3 was added in equimolar amount to each of the variants. While PK shows almost complete phosphotransfer to R3, the variant PKR1 still shows a radioactive signal and the variant PKR1R2 shows a stronger radioactive signal additional to the signal of R3 (Figure 3.26, D.). This leads to the assumption that if R1 or R1R2 are present that the signal is transferred to the bound REC domains and to R3. All together it can be summarized that R1 as well as R2 and R3 are involved in a phosphorelay of MA4377, but R3 seems to be the preferred cognate RR.

3.14 Phosphotransfer profiles provide hints for cross regulation of HKs and RRs in *M. acetivorans*

M. acetivorans encodes a total of 15 RRs, 11 of them are so called REC-only RRs, only consisting of a REC domain, and four are having a fused output domain. As REC domains typically display a high degree of structural similarity, the question arose on whether the observed transfer to R1, R2, and R3 is specific or if other RRs are also involved in the phosphorelay of MA4377 (Figure 3.26, Figure 3.27). Furthermore, it was investigated if the proteins MA2013 and MA2082 have transphosphorylation activity towards the RRs of *M. acetivorans*. To investigate the role of the RRs a phosphotransfer profile was conducted. For this, a kinase is autophosphorylated and different RRs are added as individual proteins. If in addition to the signal from the HK, a second signal is visible on the autoradiogram or if the signal of the HK disappeared, a transfer to the RR is verified (Laub *et al.*, 2007).

For 13 RRs, phosphotransfer profiling assays were performed to investigate whether they are involved in the phosphorelay mediated by MA4377. The remaining two RR were excluded as they are associated with CheA type HKs. The protein variants PK and PKR1R2 were autophosphorylated and, after removing excessive ATP, the different RRs were added. Fast transphosphorylation from phosphorylated PK to MA2445 and R3 was observed within 30 seconds (Figure 3.27, A.). After 60 minutes, transphosphorylation to MA2861, MA4671, R1 and R2 was additionally visible (Figure 3.27, B.). In a second experiment, the autophosphorylated variant PKR1R2 was used. This experiment verified the transphosphorylation of MA2445 and R3 (Figure 3.27, C.). For MA2861 and MA4671 the transfer is ambiguous, as the signals are just slightly visible after 60 minutes (Figure 3.27, D.).

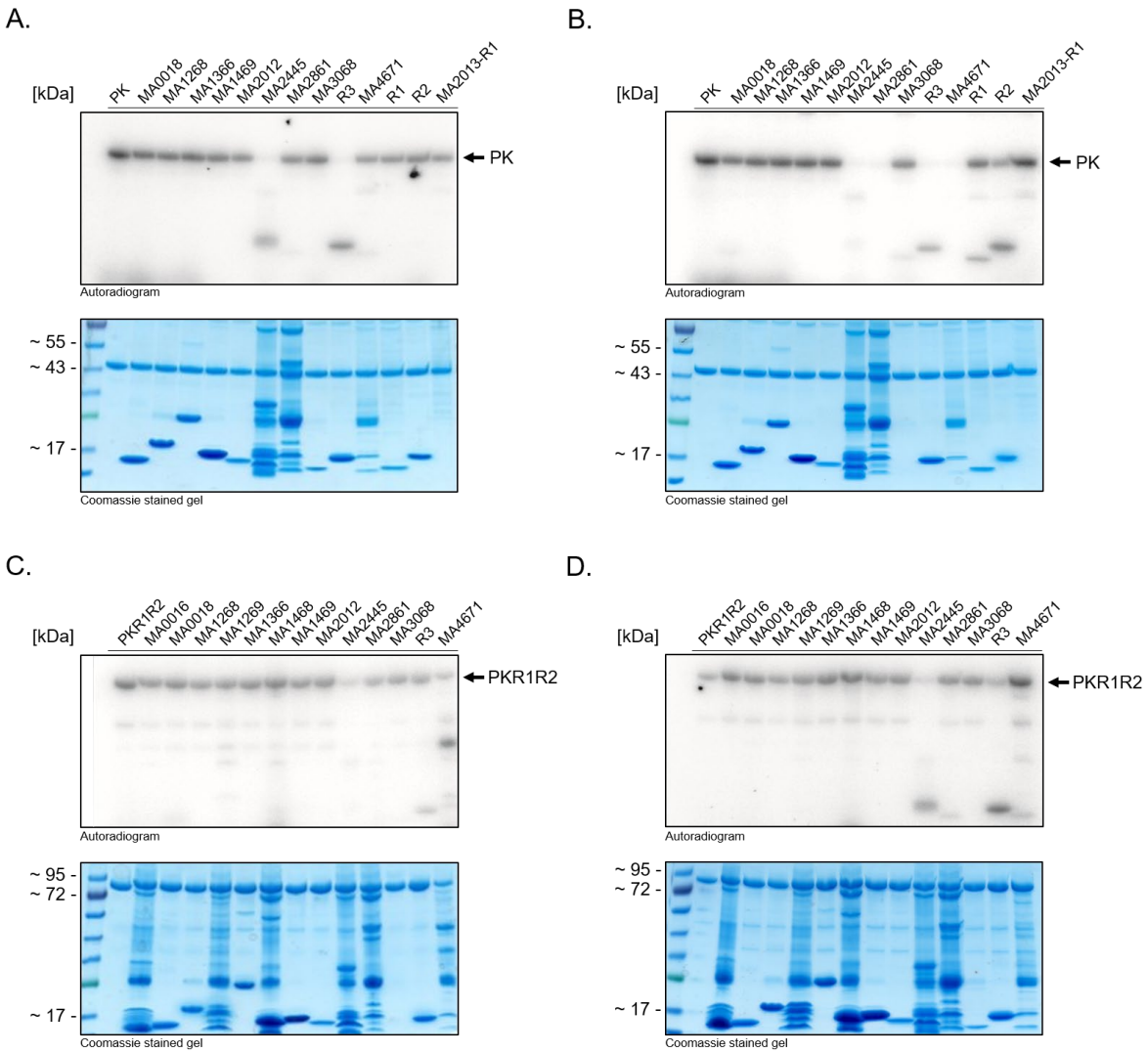

Figure 3.27: Phosphotransfer profiling of PK and PKR1R2. The purified 10 µM of the HKs PK and PKR1R2 were incubated with radioactive labelled [γ-^{32}P]-ATP, after removing excessive ATP with Microspin™G-50 Columns, the second protein (RR) was added in equimolar amount and the reaction was stopped after 30 seconds or 60 minutes using 4x SDS sample buffer with BME. Samples were separated by SDS-PAGE and the radioactive signals detected by phosphorimaging (Autoradiogram). The same SDS gel was stained with coomassie (Coomassie stained gel) to visualise the loaded protein (**A.**) Phosphotransfer profile of PK after 30 seconds. (**B.**) Phosphotransfer profile after 60 minutes. (**C.**) Phosphotransfer profile of PKR1R2 after 30 seconds. (**D.**) Phosphotransfer profile after 60 minutes. The signal that corresponds to the autophosphorylated kinase is indicated by an arrow.

For the protein MA2013 it was already shown that phosphate gets transferred to the bound REC domain (MA2013-R1) and the REC-only RR (MA2012) in genomic proximity (Figure 3.17). The phosphotransfer profiling for MA2013 verifies MA2012 as the cognate RR. After 30 seconds only MA2012 gets transphosphorylated (Figure 3.28, A.). After 60 minutes however additional vague signals for MA1269 and MA2445 were visible (Figure 3.28, B.). The phosphotransfer profiling assay of MA2082, an orphan kinase, reveals a vague signal for MA1366 and MA1469 after 30 seconds (Figure 3.28, C.) and for MA2012 and MA1468 after 60 minutes (Figure 3.28, D.). The transfer is ambiguous, as the signals are just slightly visible which leads to the conclusion that MA2082 has no clear cognate RR.

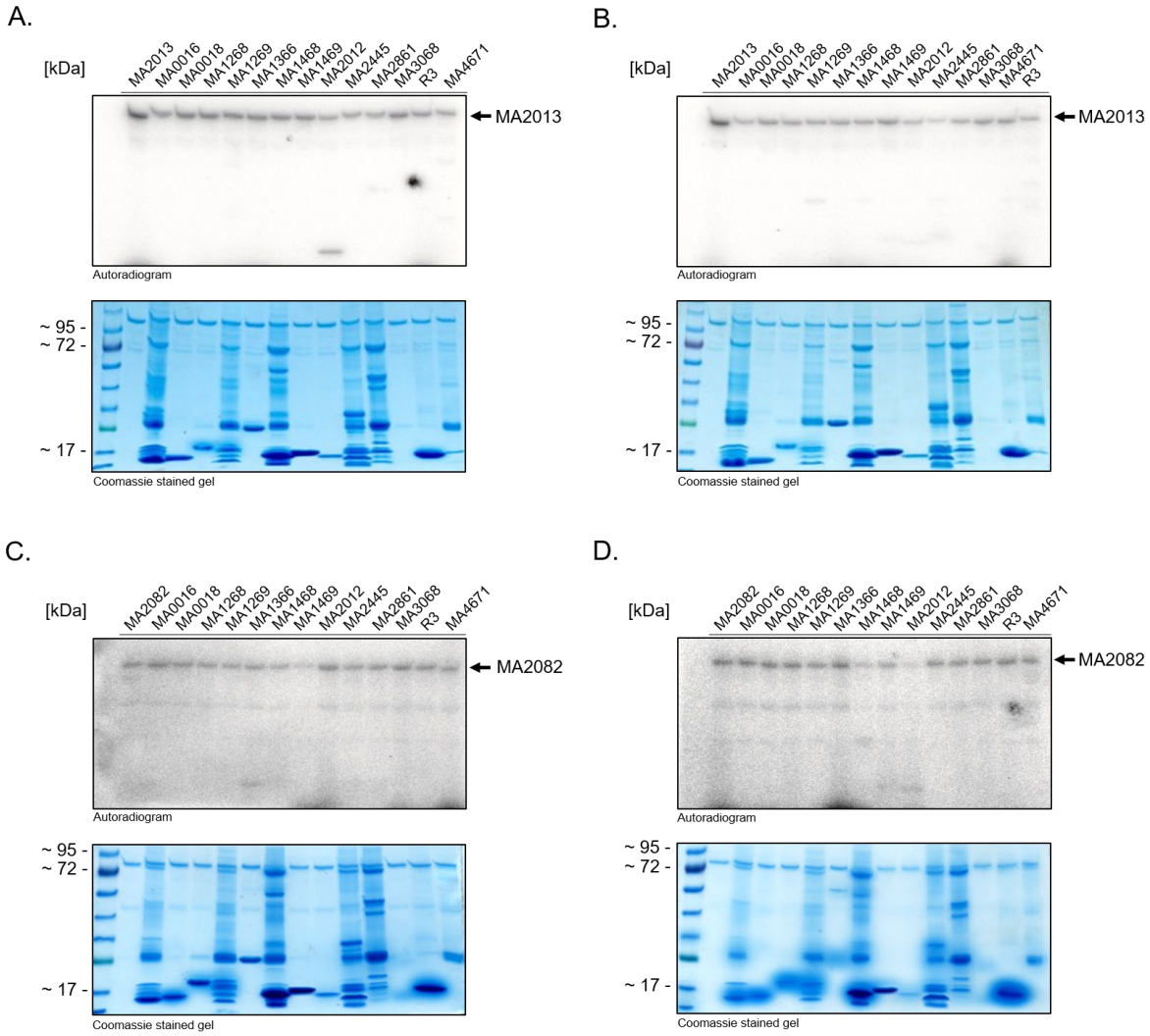

Figure 3.28: Phosphotransfer profiling of MA2013 and MA2082. The purified HKs MA2013 and MA2082 are incubated with radioactive labelled ATP and then incubated with each of the purified RRs. The reactions were analysed by SDS-PAGE and phosphorimaging. (**A.**) Phosphotransfer profile of MA2013 after 30 seconds. Transfer is visible for the REC-only RR MA2012. (**B.**) Phosphotransfer profile after 60 minutes. Transfer of phosphate was not clearly identified, possible transfer to MA2445 and MA1268. A signal around 110 kDa corresponds to the autophosphorylated kinase, a second signal corresponds to the transphosphorylated RR. (**C.**) Phosphotransfer profile of MA2082 after 30 seconds. (**D.**) Phosphotransfer profile after 60 minutes. A signal around 80 kDa corresponds to the autophosphorylated kinase, a second signal corresponds to the transphosphorylated RR. Transfer is not clearly identified, MA1468, MA1469, MA2012 and MA1366 might be involved in the phosphorelay.

3.15 Possible further steps of the phosphorelay of MA4377

A classical phosphorelay system consists of a HHK with bound REC domains, a HPt domain/protein and a RR with output domain. The phosphoryl group from the His residue of the kinase is transferred to an Asp residue in the REC domain. This then targets a His residue of a HPt domain or a HPt-like protein, finally ending up at an Asp of a RR with output domain.

For the putative phosphorelay system of MA4377, transphosphorylation to several RR was shown. One of the identified RR is MA2445, a protein containing a novel output domain, which could lead to a direct output reaction. However, most of potential RR do not possess an output domain, suggesting that further steps are needed to create a final output signal. Based on a classical four step phosphorelay system, the next signalling partner to any of the identified RRs should be a HPt domain/HPt-like protein.

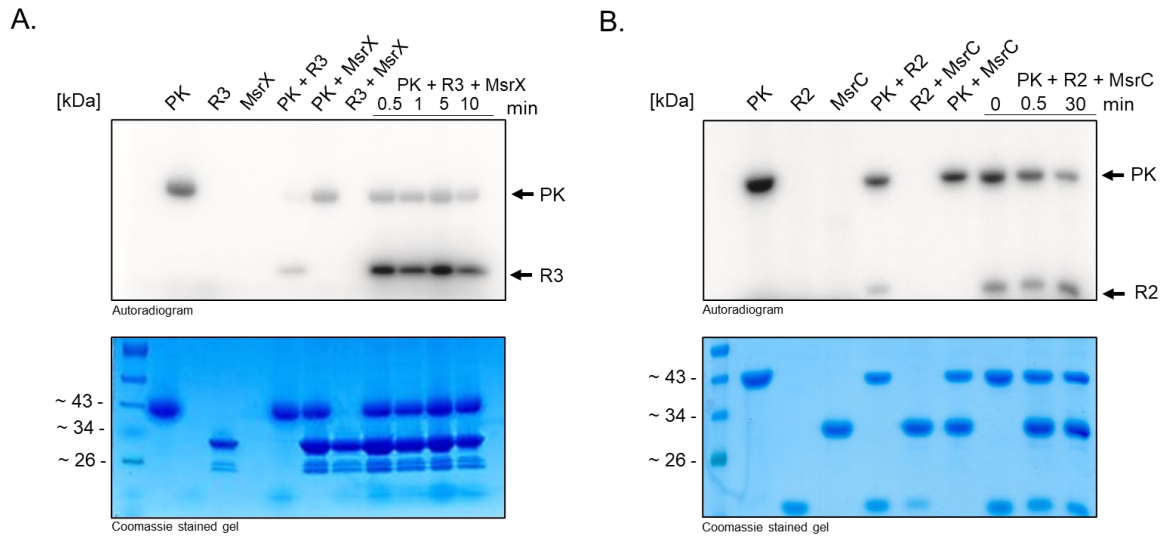

Figure 3.29: Putative further components of the MA4377 phosphorelay. (**A.**) Autoradiogram and coomassie stained gel (4-20% gradient) after in vitro autophosphorylation/transphosphorylation assay. First radiolabeled [γ-^{32}P]-ATP was added to PK, R3 and MsrX (10 µM) and the reaction was terminated with SDS sample buffer (with BME) to test the autokinase activity. For the second part of the assay the proteins PK or R3 were autophosphorylated and the second protein (R3 or MsrX) was added after removing excessive ATP with Microspin™G-50 Columns. For the third part of the assay the protein PK (10 µM) was autophosphorylated, excessive ATP with Microspin™G-50 Columns was removed and first the protein R3 and afterwards the protein MsrX were added in equimolar concentration. The coomassie stained gel shows that the same protein concentration was used. The autoradiogram displays autokinase activity only for the control protein PK and transphosphorylation activity to R3. (**B.**) Autoradiogram and coomassie stained gel (4-20% gradient) after in vitro autophosphorylation/ transphosphorylation assay. First radiolabel [γ-^{32}P]-ATP was added to PK, R2 and MsrC (10 µM) and the reaction was terminated with SDS sample buffer (with BME) to test the autokinase activity. For the second part of the assay the proteins PK or R2 were autophosphorylated and the second protein (R2 or MsrC) was added after removing excessive ATP with Microspin™G-50 Columns. For the third part of the assay the protein PK (10 µM) was autophosphorylated, excessive ATP with Microspin™G-50 Columns was removed and first the protein R2 and afterwards the protein MsrC were added in equimolar concentration. The coomassie stained gel shows that the same protein concentration was used. The autoradiogram displays autokinase activity only for the control protein PK and transphosphorylation activity to R2. As reference the color prestained protein standard (M) from NEB was used.

Therefore, the genome of *M. acetivorans* was screened for HPt domain containing proteins. Overall, three such proteins (MA0014, MA3066 and MA2013) were identified. Two of them are CheA-like chemotaxis proteins (MA0014 and MA3066) with an N-terminal HPt domain and are

therefore likely not involved in the MA4377 phosphorelay. The third one (MA2013) contains a C-terminal HPt domain and is therefore a candidate for the phosphorelay. Since the number of HPt domain containing proteins is low, regulatory proteins encoded in the genomic vicinity of MA4377 were included in the search. Two transcriptional regulators (MsrC and MsrX) with a HTH domain and a DUF1724 domain were identified. The preliminary data indicated that MsrC and MsrX are not involved in a transphosphorylation reaction with MA4377 and therefore are unlikely to be associated with in the phosphorelay (Figure 3.29).

Given the time constrains, further investigation of the phosphorelay was not conducted. Consequently, future experiments must be designed to elucidate the role of the MA2013 HPt domain and additional downstream signalling components

Due to time constrains, the phosphorelay was not investigated further. Therefore, future experiments have to be designed to investigate the role of the MA2013 HPt domain and further downstream signalling components.

4 Discussion

Sensing and reacting to environmental stimuli via signal transduction systems is a key factor for microorganisms to survive. Several examples of bacterial and eukaryotic signal transduction systems have been investigated experimentally in the last decades. However, the knowledge of archaeal signal transduction is still limited. Currently there are only a few publications that experimentally examined the structure and function of archaeal signal transduction. With an increasing number of fully sequenced archaeal genomes there is the possibility to examine archaeal signal transduction systems in more detail. This study provides a first glimpse at the complexity of signal transduction in the methanogenic archaeon *M. acetivorans*.

4.1 HKs and RRs of *M. acetivorans* show distinct differences to bacterial signal transduction systems

A closer look at the number and distribution of HK and RR in the genome of *M. acetivorans* reveals that HKs outnumber RR proteins by at least three times. In addition, only six of the HKs are encoded in an operon with a corresponding RR, resulting in a high number of orphan kinases. This is almost opposite to the situation of TCSs in *E. coli*. There, 89 % of the HK are encoded in operons with their corresponding RR (Capra & Laub, 2012). The high number of HKs suggests that signal transduction in *M. acetivorans* does not work in strict pairs and that cross regulation exists (Galperin *et al.*, 2018; Krell, 2018). Alternatively, *M. acetivorans* might use unknown signalling partners for its orphan kinases. Another interesting result from the *in silico* analysis is the number of predicted membrane-bound HKs. Most bacterial HKs are membrane bound (around 73 %) (Wuichet *et al.*, 2010), compared to archaea that exhibit around 38 % membrane-bound HKs (Galperin *et al.*, 2018). *M. acetivorans* only has about 23 % predicted membrane HKs. This observation already gives a hint on the putative signals to be sensed. Membrane-bound kinases are often needed to detect signals that cannot easily enter the cell, usually external signals, whereas cytosolic signal input domains are used to detect internal signals or those that can easily pass through the membrane (e.g. light, gases). Despite the differences in membrane-bound HKs, the overall domain organisation is largely similar to bacterial HKs with PAS and GAF domains being highly prevalent (Galperin *et al.*, 2018; Koretke *et al.*, 2000; Krell, 2018).

Looking closer into the RRs of *M. acetivorans*, it is noticeable that they differ from bacterial RRs. In bacterial signal transduction, the signal receiving REC domain is often linked to DNA binding domains (such as HTH motifs) which play a role in transcriptional regulation (Galperin, 2006; Hoch & Varughese, 2001). Overall, there is only one RR with HTH domain in the genome of *M. acetivorans* (MA1366). On the other hand, a high number of REC-only RRs, without any

output domain are present. This is not only the case for the methanogenic archaeon investigated in this study but an overall phenomenon within the archaea suggesting that other signal transduction mechanisms have evolved (Galperin, 2006, 2010). Additional novel output domains occur that are structural and functional undescribed (halobacterial output domain (HalOD) and methanogen output domain (MetOD) (Galperin *et al.*, 2018). This leads to the assumption that transcriptional regulation might not be the primary function of archaeal signal transduction systems (Krell, 2018). Besides transcriptional regulation, bacterial RRs are known to contribute to ligand binding, protein binding or enzymatic reactions (Galperin, 2006, 2010). REC-only RR can produce an output directly through a conformational change in the REC domain (Koshland, 1977; Stewart *et al.*, 2000). A prominent example for a REC-only RR is CheY that works together with CheA and CheB in the chemotaxis signalling in *E. coli*. Phosphorylated CheY binds to the flagella motor protein FliM which controls the direction of flagellar rotation (Muok *et al.*, 2020; Stock *et al.*, 1988). Furthermore, REC-only proteins are also known as phosphate sinks and thereby can serve as buffers in complex phosphorelay systems (Galperin, 2006; Szurmant & Ordal, 2004).

4.2 Representatives of type MA_type 1 and MA_type 3 kinases are autokinases

Based on the phylogeny and the amino acid sequence structure of the H-box region of the HK domain, four new types of HKs in *M. acetivorans* were defined. The experimental data of this study support the conclusion that all examined kinases are true autokinases with the exception of the MA_type 2 representative (RdmS).

ATP binding assays with the ATP analogue TNP-ATP were used to determine the dissociation constant (K_d) towards TNP-ATP. All putative HKs displayed affinity towards TNP-ATP with K_d values in the micromolar range (11–30 µM). Several studies have investigated the K_d values towards ATP analogues of different bacterial kinases. The highest affinity was reported for CheA (0.5/1.7 µM) (Stewart *et al.*, 1998), FixL (0.8 µM) (Sousa *et al.*, 2005) and EnvZ (1.0 µM) (Plesniak *et al.*, 2002) of *E. coli* and for HK1 (1.3 µM) of *Sinorhizobium meliloti* (Shrivastava *et al.*, 2007). The protein PhoQ of *Salmonella typhimurium* was reported to have a K_d value of around 300 µM (Guarnieri *et al.*, 2011), which is a noticeable lower affinity than for the other reported kinases. Considering that each kinase has a unique nucleotide binding pocket and specific interacting residues at the ATP lid, it seems likely that they have different binding affinities despite the overall structural similarity (Guarnieri *et al.*, 2011). A typical bacterial cell has an intracellular ATP concentration of around 1–3 mM (Buckstein *et al.*, 2008). *M. acetivorans* however is a methanogenic archaeon that produces energy through methanogenesis. Methanogenesis is known to produce less than one molecule of ATP for each molecule of methane. Nevertheless, methanogens are perfectly adapted to live with a

low energy level and use several methods to conserve energy (Valentine, 2007). The intracellular ATP concentration of an archaeal cell is unknown. Given the diversity of archaeal organisms, it is also not easy to assume whether the ATP concentration is similar or lower than in bacteria. However, even if the ATP concentration of *M. acetivorans* is lower than that of a typical bacterial cell the values from the ATP binding study still seem reasonable.

The ADP-Glo™ Kinase Assay data further supported the conclusion that the putative HKs are kinases since they can hydrolyse ATP. Autophosphorylation assays confirmed that MA2013, MA2082 and cytosolic variants of MA4377 are true autokinases. Most likely the three autokinases are HKs as they all possess a conserved His residue within the H-box. For the kinase MA4377, the His residue 497 was confirmed as the phosphorylation site. MA_type 2 kinases completely lack the H-box region but are also able to bind and hydrolyse ATP, suggesting that they might phosphorylate another signalling partner, possibly a protein partner similar to what is realized in MAP kinase signalling in Eukarya (Blasius, 2023; Pearson *et al.*, 2001). However, more experimental data will be required to proof this hypothesis.

4.3 What is the reason for the missing autokinase activity of the full-length membrane protein MA4377?

Truncated cytosolic variants of the protein MA4377 confirmed its function as a HK with autokinase activity. However, the autokinase activity of the full-length membrane protein was not observed in this study. A systematic investigation was conducted to elucidate the influence of various factors on the autokinase activity of MA4377. Different expression strains, a truncated membrane variant without the bound REC domains and different purification methods did not lead to an active protein. Looking into publications that characterize membrane-bound kinases, a high number of authors worked with cytosolic variants of the respective kinase (Barr *et al.*, 2023; Francis *et al.*, 2018; Li *et al.*, 2014; Najnin *et al.*, 2016; Vega-Baray *et al.*, 2022). Other authors succeeded to work with active full-length membrane proteins (Jung *et al.*, 1997; Wang *et al.*, 2017). In this case a successful isolation of active membrane proteins was either conducted via inverted membrane isolation or the reconstitution into liposomes. Whereas the solubilisation of the membrane protein with detergent removes the protein out of the native membrane and therefore likely loses the correct folding and activity of the protein. A subsequent reconstitution of the protein into liposomes can lead to an active protein.

In this study also the solubilisation with a nanodiscs was carried out. This method was more promising than solubilisation with a detergent, as the nanodiscs assist in the extraction of the protein with the surrounding membrane, potentially leading to a more native membrane environment. Nevertheless, neither the solubilisation with detergent, nor the solubilisation with

the nanodisc led to autokinase activity of MA4377. An additional reconstitution of the protein into *E. coli* liposomes was not successful (data not shown).

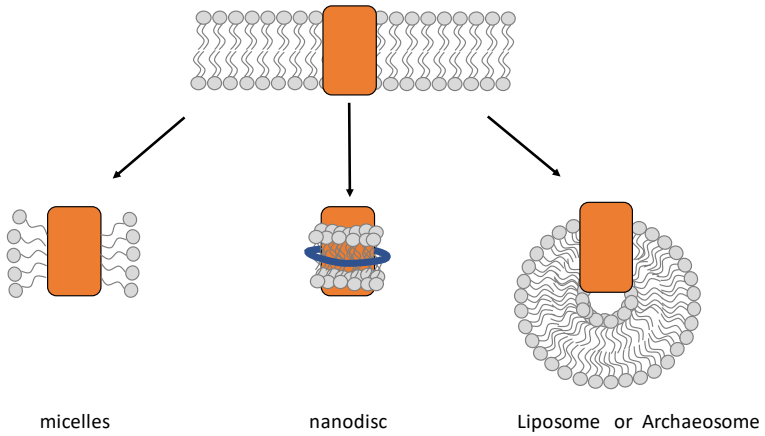

Figure 4.1: Schematic representation of different membrane isolation methods. The solubilisation of a membrane protein with detergent leads to the formation of micelles. Nanodiscs extract the protein with the surrounding native membrane. The reconstitution into liposomes or Archaeosomes can help to get a more native folding of the protein. Modified from (Thoma & Burmann, 2020).

For the HK MA4377 it must be considered that the protein is of archaeal origin but was heterologously produced in a bacterial host. Given that the protein MA4377 is membrane-bound, it can be hypothesised that the membrane of the host organism may play a crucial role in the production, folding and activity of the protein. For this a closer look must be taken to the membrane composition of bacteria and archaea. The chemical composition of the membrane lipids of bacteria and archaea exhibits notable differences. The glycerophosphate backbone of bacterial phospholipids is a glycerol-3-phosphate, whereas archaeal phospholipids contain a glycerol-1-phosphate. Furthermore, archaea possess branched isoprenoid chains (phytanyl sidechain) that are linked via ether bonds, whereas bacteria have ester-linked fatty acids. Additionally, bacterial phospholipids form a bilayer, whereas archaeal phospholipids are predominantly monolayer (Caforio & Driessen, 2017). The structural distinctions in the membrane composition may be a contributing factor to the observed absence of autokinase activity in the full-length protein. Two potential approaches can be considered to address this issue:

Recent publications have reported the use of so-called Archaeosomes. Archaeosomes are vesicles made from polar lipids extracted from archaea or synthetic archaeal lipids. Due to their ether linkages, Archaeosomes are more stable than bacterial liposomes and therefore well suited for drug delivery and biomedical applications. Heterologous produced MA4377 solubilised with detergent could be reconstituted into Archaeosomes to mimic a more natural membrane environment (González-Paredes *et al.*, 2010; Santhosh & Genova, 2023; Vidakovic *et al.*, 2024). A second potential approach would be the production of MA4377 in an archaeal host. Given the current limitations in genetic tools for *M. acetivorans* and the suboptimal growth

parameters for protein production, a homologous production in this organism is challenging. However, heterologous production in a more suitable archaeal host organism may offer a strategy for obtaining an active protein.

In addition to the technical factors that contribute to the observed deficiency in autokinase activity, it is essential to consider the potential absence of the essential input signal. It has been demonstrated that the membrane-bound HK MA4377 is capable of binding and hydrolysing ATP with values that are almost identical to those observed for the truncated cytosolic variant PKR1R2. Nevertheless, autokinase activity was not observed. Furthermore, the investigation of different potential input signals did not result in any effects on the protein. It can be concluded that the appropriate signal may still be absent. One hypothesis is that the protein is in an inactive state when the signal is absent, similar to the bacterial phytochrome BphP of *Pseudomonas aeruginosa* (*P. aeruginosa*). For BphP, it has been reported that autokinase activity can be finetuned through the presence of different light conditions. A high autokinase activity was observed for the protein under far-red light conditions, while darkness led to an almost lacking autokinase activity (Huber, 2024; Huber *et al.*, 2024). It can be postulated that the protein MA4377 is in an inactive state in the absence of the appropriate input signal, and that it can only be transferred to an active state in the presence of the required input signal.

Moreover, the HAMP domain of MA4377 may potentially influence the activity of the protein. HAMP domains are small protein modules with a four-helix bundle structure, situated between the membrane-bound input domains and the cytoplasmic part of a signalling protein. It is hypothesised that they mediate signals between input and output domains via rotations of their α-helices into two different states. A study of HAMP domains in chemoreceptors indicated that they may function as a receptor on/off switch (Hulko *et al.*, 2006; Schultz *et al.*, 2015; Swain & Falke, 2007; Swain *et al.*, 2009). This illustrates the complexity of signalling proteins. It is possible that the HAMP domain of MA4377 may also be a contributing factor to the observed lack of autokinase activity.

Further investigation is required to determine why the full-length membrane protein MA4377 exhibits no autokinase activity despite its ability to bind and hydrolyse ATP.

4.4 HHKs are part of a phosphorelay system in *M. acetivorans*

This study provides the first experimental evidence of a phosphorelay system in *M. acetivorans*. The HHK MA4377 was shown to be a HK with autokinase activity and to be involved in intra- and intermolecular transphosphorylation of the bound REC domains and the downstream encoded REC-only RR (R3). R3 seems to be the preferred phosphate acceptor. However, it was also shown that intramolecular transphosphorylation still takes place and R1

and R2 are still being phosphorylated, even if the reaction seems to be slower than to R3. Such preferences are also known from bacterial phosphorelay systems which show diverse options for the preference and involvement of RECs/RR. For instance, the chemotaxis system of *Rhizobium meliloti* (CheA/CheY1/CheY2) is using one of the two REC-only CheY RRs as a phosphate sink. But only the contribution of both CheY proteins leads to a functional reaction (Sourjik & Schmitt, 1998). The kinase RodK from *Mycococcus xanthus* and PcrK from *Xanthomonas campestris* are examples that not all present REC domains are actively involved in the phosphorelay (Wang *et al.*, 2017; Wegener-Feldbrügge & Søgaard-Andersen, 2009). For the HHK MA2013 it was also demonstrated that it is an autokinase with transphosphorylation activity towards the bound REC domain (MA2013-R1) and the REC-only RR (MA2012) in genomic proximity (Figure 3.17).

Further steps of the MA4377 phosphorelay systems and the MA2013 phosphorelay systems have not yet been elucidated. A classical four-step phosphorelay system is transferring phosphate from an autophosphorylated HHK to a bound REC domain, subsequently to an Hpt domain/protein/HPt-like component and finally to a RR with output domain. *M. acetivorans* encode genes for three proteins with HPt domain, two of them are CheA related proteins with a N-terminal HPt domain that most likely is used as the phosphorylation site. For this reason, they were excluded from further investigations. The HPt domain situated at the C-terminus of the protein MA2013 is the only annotated HPt domain in the genome of *M. acetivorans* that might function as a classical HPt domain. The function of the MA2013-HPt domain must be examined with further experiments. Additionally, to classical HPt domains also other HPt-like components could be involved in a phosphorelay. To date no HPt-like component was found to be involved in the putative MA4377 phosphorelay of *M. acetivorans*. Also, the transcriptional regulators (MsrX and MsrC), which are encoded in vicinity of MA_4377, were found to be not involved in the putative phosphorelay.

As REC domains typically display a high degree of structural similarity, the question arose as to whether the observed transfer to the bound REC domains and the nearby encoded RRs is specific and whether other RRs are also involved in the phosphorelay. *M. acetivorans* encodes a total of 15 RRs, 11 of them are so called REC-only RRs, only consisting of a REC domain, and four are fused to output domains. Phosphotransfer profiling assays were performed with 13 RRs to investigate whether they are involved in a phosphorelay mediated by MA4377 or MA2013. The remaining two RR were excluded as they are associated with CheA type HK.

The Phosphotransfer profiling method enables the systematic identification of signal transduction pathways. This *in vitro* technique allows the examination of a HK for its ability to transfer a phosphoryl group to all RRs encoded within the genome of an organism of interest (Laub *et al.*, 2007; Skerker *et al.*, 2005).

The phosphotransfer profiling assay of MA4377 demonstrated that in addition to R1, R2 and R3 also MA2445, MA2861 and MA4671 may also be involved in a phosphorelay (Figure 3.27). The proteins MA2861 and MA4671 are identified as REC-only RR, whereas MA2445 is characterised as a RR with a MetOD1 output domain. The protein MA4671 is encoded in an operon with the HHK MA2256 and therefore likely constitutes a signalling system. In contrast, MA2445 and MA2861 are orphan RRs. The proteins R3 and MA2445 received phosphoryl groups already after 30 seconds. This leads to the assumption that this transfer may be of greater relevance than the transfer towards the two other proteins. Overall, the data indicates that MA4377 is part of a complex phosphorelay system that involves multiple RR proteins. In addition to MA4377 a phosphotransfer profiling assay was conducted for the HHK MA2013. The data obtained from this assay suggests that MA2013 is likely to be involved in complex phosphorelay systems comprising multiple components. It has previously been demonstrated that phosphate transfer occurs to the bound REC domain (MA2013-R1) and the REC-only RR (MA2012) (Figure 3.17). The phosphotransfer profiling for MA2013 identifies MA2012 as the cognate RR. After 30 seconds only MA2012 exhibited evidence of transphosphorylation. After 60 minutes, additionally a signal for MA2445 and a vague signal for MA1269 was observed. Furthermore, also the orphan kinase MA2082 was analysed with a phosphotransfer profiling assay. However, the data for MA2082 are less conclusive. The assays revealed a vague signal for MA1366 and MA1469 after 30 seconds and for MA2012 and MA1468 after 60 minutes. The transfer is ambiguous, as the signals are only slightly visible which leads to the conclusion that MA2082 has no clear cognate RR (Figure 3.28).

The phosphotransfer profiling method, as described by Skerker *et al.* (2005), represents a robust tool for the identification of the cognate RR of a HK based on a biochemical approach. In a comprehensive study, the authors were able to successfully identify the cognate RR of several *E. coli* and *Caulobacter crescentus* (*C. crescentus*) HKs. In the case of classical TCS, the HK and RR are frequently encoded in an operon, allowing for a straightforward identification based on their genomic localisation (Skerker *et al.*, 2005). However, this is more challenging in the case of orphan HKs or complex pathways with cross-regulation. The study demonstrated that examination of the phosphoryl group transfer after ten seconds is an effective method to verify the cognate RR of a HK. The transphosphorylation activity of the kinase EnvZ towards the cognate RR OmpR was demonstrated after 10 seconds, while after 60 minutes, other RRs were also phosphorylated. The same result was observed for the orphan kinases DivJ and PleC of *C. crescentus*. After 10 seconds, both RRs PleD and DivK were phosphorylated, which are known to be the cognate partners of DivJ and PleC. The present study demonstrated that an *in vitro* method can be employed to identify *in vivo* cognate partners of signal transduction pathways. (Skerker *et al.*, 2005).

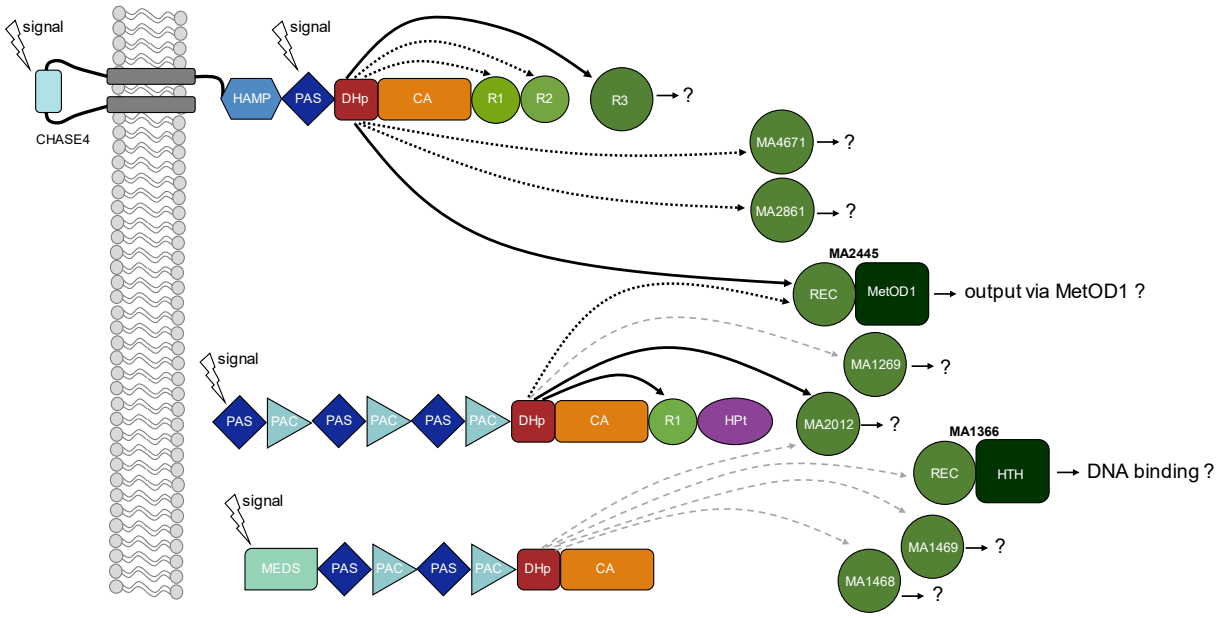

Figure 4.2: Summary of the phosphotransfer profiling data of the three putative HKs MA4377, MA2013 and MA2082 of *M. acetivorans*. Signals can be sensed from the outside and the inside of the membrane with different potential sensing domains. Phosphoryl group transfer is represented with different lines. (Black lines = verified transfer after 30 sec, dotted lines = transfer after 60 min and grey dashed lines = ambiguous transfer).

In conclusion, the results of the phosphotransfer profiling demonstrate that MA4377 is involved in a phosphorelay system, which might involve multiple REC domains and RR proteins. The phosphoryl group transfer to R3 and MA2445 after 30 seconds, suggests these RRs as the cognate *in vivo* partners of MA4377. Data of the transphosphorylation assays of MA4377 also support this hypothesis. When the bound REC domains (R1, R2) and the REC-only RR R3 were present at the same time only R3 was phosphorylated (Figure 3.26, C.). It seems likely that the protein MA2013 is also involved in a phosphorelay system. Nevertheless, only the transphosphorylation activity towards the bound REC domain (MA2013-R1) and to the REC-only RR (MA2012) has been demonstrated with certainty after 30 seconds. To gain further insight into the full function of the system, it is necessary to examine the function of the HPt domain in future experiments. It is possible that the RRs identified after 60 minutes are also part of the *in vivo* phosphorelay system for MA4377 and MA2013. However, the phosphotransfer profiling method is unable to confirm this with certainty.

The function of the orphan HK MA2082 remains unclear. To confirm the ambiguous phosphoryl group transfer, additional experiments are required. However, the initial preliminary results have indicated that four different RRs may be involved in a reaction with MA2082 (Figure 4.2).

4.5 Evidence for a complex signalling network in M. acetivorans

Based on the obtained data, it can be postulated that signal transduction of *M. acetivorans* functions as a complex signalling network.

This hypothesis can be supported by the observation that *M. acetivorans* exhibit a higher number of HKs than RRs (3:1). Suggesting that signal transduction pathways do not function in strict pairs. Furthermore, the majority of HKs in *M. acetivorans* are orphan proteins. It is possible that some of them have a cognate RR that is not encoded in the same operon, similar to the phytochrome BphP and the RR AlgB of *P. aeruginosa* (Huber, 2024; Mukherjee *et al.*, 2019). However, it is also reasonable to consider the possibility that several orphan HKs may interact with the same RR. A prominent example for a pathway that involves several orphan kinases is the KinA-E/Spo0F/Spo0B/Spo0A pathway of *Bacillus subtilis* (*B. subtilis*) (Francis & Porter, 2019; Jiang *et al.*, 2000). The correct function of all five kinases is crucial for the function of the system and the regulation of *B. subtilis* sporulation (Jiang *et al.,* 2000). Another aspect that supports the idea of a complex signalling network in *M. acetivorans* is the high number of REC-only RR, which is common for archaea. This suggests that those proteins are the key players in archaeal signal transduction and a coordination hub for processing and forwarding signals. The specific function of these proteins in *M. acetivorans* remains unclear. Given the role of the REC-only RR CheY in binding to the flagella motor protein FliM when it is phosphorylated (Muok *et al.*, 2020; Stock *et al.*, 1988), it seems likely that protein binding represents one function of the REC-only RRs. Another function can be the forwarding of signals, similar to the REC-only RR Spo0F, which is part of the sporulation network of *B. subtilis*. Spo0F gets phosphorylated from various HKs, subsequently transferring phosphate to an HPt protein (Spo0B), which then phosphorylates a RR with an output domain (Spo0A) (Jiang *et al.*, 2000). Involving several REC-only RR in the same phosphorelay reaction can increases the complexity of the pathway.

The hypothesis for MA4377 is that it can sense different input signals and, accordingly, react with transfer of phosphoryl groups to different RRs. MA4377 contains different sensing domains, the extracellular CHASE4 domain could sense a signal from outside the cell, while the PAS domain could sense a cytoplasmic signal. One way could be a TCS of MA4377 and MA2445 with a direct output of the MetOD1 domain. Another way could involve R1, R2 and R3 with either an unknown HPt-like component and a final output, or a direct reaction through a conformational change. Furthermore, one could think of protein:protein interaction directly influencing the function of the protein partner (i.e. an enzymatic reaction). MA2013 might be involved in a more classical four-step phosphorelay since it harbours an HPt domain. However, also the function of this phosphorelay system is not fully elucidated. Similar to MA4377, different pathways seem possible for MA2013. One way could be a four-step phosphorelay involving the bound REC domain, the HPt domain and MA2012 with an unknown output. On the other hand, MA2013 could also interact directly with MA2012 with a different unknown output respond.

For bacteria several studies confirm the existence of highly complex signalling networks. For example, the previous discussed sporulation network of *B. subitlis*. A similar complex system is the quorum sensing network of *Vibrio barveyi* that also involves several HKs. A signalling system with multiple interactions between the single parts is the cell division network of *C. crescentus*. It is likely that the majority of signal transduction pathways do not function in classical TCSs, but rather show cross regulation. The known TCSs are often encoded in operons and were therefore easier to identify. It is likely that the future research will reveal an extended function of already known TCSs. It is imaginable that a TCS can function as such but also has an extended function as a part of a more complex pathway.

These examples already show how complex the interaction of signal transduction pathways can be. Combining the knowledge of obtained data of this study and the example of bacterial signal transduction networks makes it possible to postulate a hypothesis for the function of both HKs MA4377 and MA2013 in a signalling network.

The kinases GacS and RetS of *P. aeruginosa*, are two HHK, with a similar domain organisation than MA2013 and MA4377 based on data of the InterPro database. The proteins GacS and RetS interact via three distinct mechanisms, but GacS can also directly phosphorylate the RR GacA. A first interaction mechanism is that autophosphorylated GacS can transfer a phosphoryl group to the REC2 domain of RetS. Secondly, the RetS has phosphatase activity, that leads to a faster dephosphorylation of the phosphorylated REC domain of GacS. Finally, the kinase domain of RetS can inhibit the autophosphorylation of GacS. Moreover, this pathway is integrated in an even more complex network with other signalling partners (Francis & Porter, 2019; Francis *et al.*, 2018).

Further experiments are needed to investigate whether MA2013 might interact with MA4377 in a similar manner (Figure 4.3). So far, both systems were only considered as distinct pathways. The phosphotransfer profiles suggest that both HK might interact with the RR MA2445.

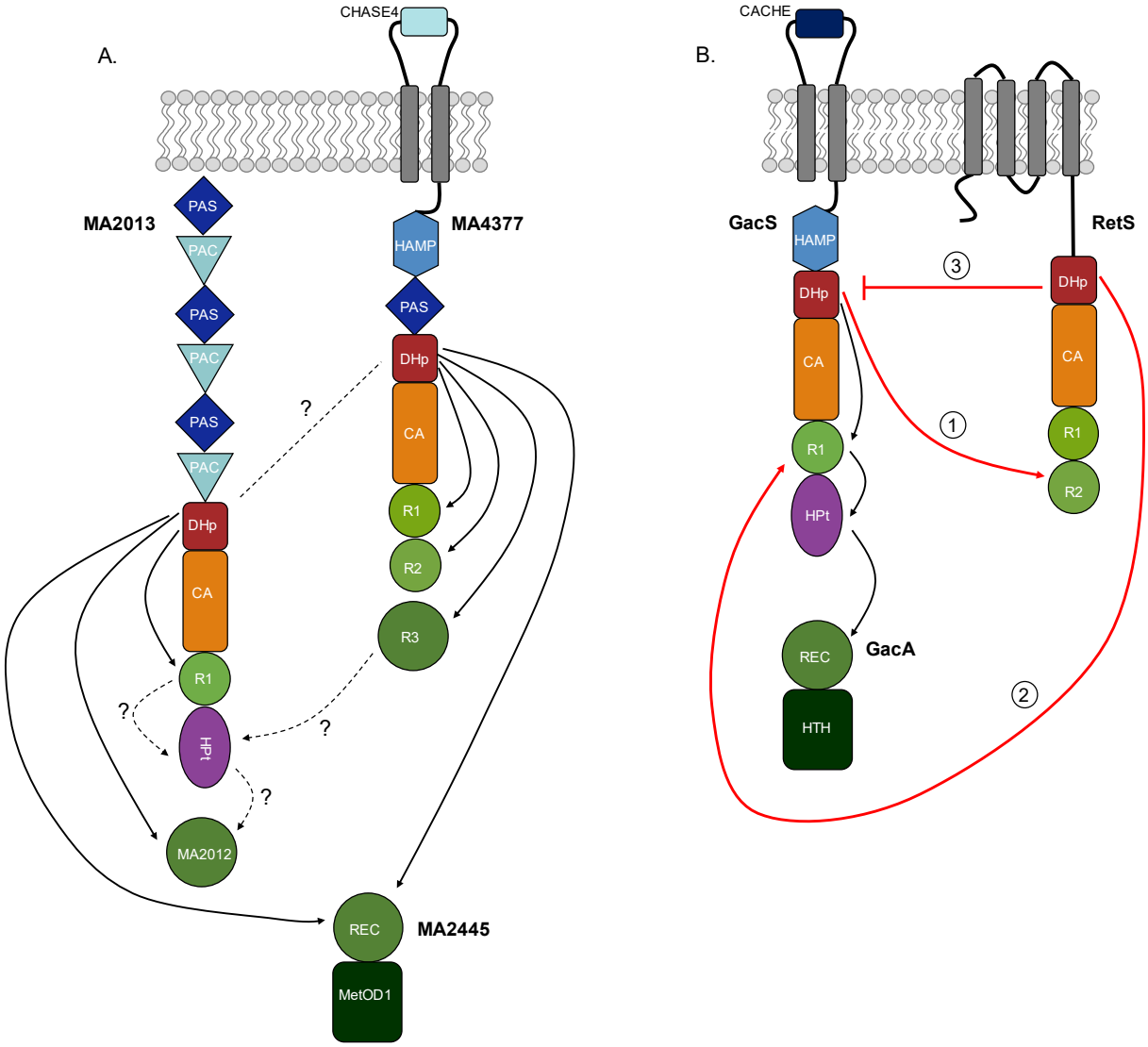

Figure 4.3: Model of the interaction of MA4377 with MA2013 based on the communication mechanisms of the *P. aeruginosa* kinases GacS and RetS. Phosphoryl group transfer is represented with black lines. Dotted lines represent phosphoryl group transfer or direct interaction that must be confirmed. Red lines represent interaction between the *P. aeruginosa* proteins.

Overall, this study only provides a small glimpse into the complexity of signal transduction in *M. acetivorans* and archaea in general. More experimental data including genetic studies are needed to understand the complexity of archaeal signal transduction in detail. Nevertheless, this study provides the first evidence for the existence of phosphorelay systems and complex signalling networks in archaea.

5 Summary

The ability of microorganisms to sense and react to environmental stimuli via signal transduction systems is a crucial factor in their survival. A variety of bacterial and eukaryotic signal transduction systems have been the subject of experimental investigations over the past few decades. Nevertheless, our understanding of archaeal signal transduction remains incomplete. At present, there are only a few publications that have experimentally investigated the structure and function of archaeal signalling. However, with the increasing number of fully sequenced archaeal genomes, there is an opportunity to examine these systems in greater detail. The methanogenic archaeon *M. acetivorans* has one of the largest known archaeal genomes and, with 53 putative HK and 16 putative RR, one of the largest sets of signal transduction systems.

In a comprehensive study the histidine kinases and response regulators of the methanogenic archaeon *M. acetivorans* were first analysed *in silico*. This revealed that the HKs can be classified into four different groups with respect to their HK domain (MA_type 1 A/B/C, MA_type 2, MA_type 3 and MA_CheA). Afterwards four selected putative HK (RdmS, MA2082, MA2013 and MA4377) were studied in detail using a biochemical approach. The recombinant proteins were examined regarding their ATP binding affinity and their auto- and transphosphorylation activity was tested using radioactive autokinase assays. This revealed that three of the four proteins assayed are true autokinases and that RdmS is an unusual kinase lacking the H-box and autokinase activity.A phosphotransfer profiling assay of the three autokinases with all RRs encoded in the *M. acetivorans* genome revealed that the HK may interact with several RRs. In addition, MA4377 and MA2013 were identified as HHKs involved in phosphorelay systems. This study provided the first evidence for the existence of a complex signalling network in *M. acetivorans* involving multiple HKs and RRs.

6 Zusammenfassung

Die Fähigkeit von Mikroorganismen, Umweltreize über Signaltransduktionssysteme wahrzunehmen und darauf zu reagieren, ist ein entscheidender Faktor für ihr Überleben. Eine Vielzahl von bakteriellen und eukaryotischen Signaltransduktionswegen war in den letzten Jahrzehnten Gegenstand von experimentellen Untersuchungen. Dennoch ist das Verständnis der Signaltransduktion bei Archaeen noch unvollständig. Derzeit gibt es nur wenige Veröffentlichungen, die die Struktur und Funktion archaeller Signalübertragung experimentell untersucht haben. Mit der zunehmenden Zahl an vollständig sequenzierten archaeller Genomen besteht jedoch die Möglichkeit, diese Systeme genauer zu untersuchen. Das methanogene Archaeon *M. acetivorans* verfügt über eines der größten archaellen Genome, sowie über eine große Anzahl an Signaltransduktionssystemen, welches 53 potenzielle Histidin Kinasen und 16 potenzielle Antwort Regulatoren umfasst.

In einer umfassenden Studie wurden die Histidin Kinasen und Antwort Regulatoren des methanogenen Archaeons *M. acetivorans* zunächst *in silico* analysiert. Dabei zeigte sich, dass sich die Histidin Kinasen hinsichtlich ihrer HK Domäne in vier verschiedene Gruppen einteilen lassen (MA_type 1 A/B/C, MA_type 2, MA_type 3 und MA_CheA). Anschließend wurden vier ausgewählte potenzielle Histidin Kinasen (RdmS, MA2082, MA2013 und MA4377) mit Hilfe eines biochemischen Ansatzes im Detail untersucht. Die rekombinanten Proteine wurden auf ihre ATP-Bindungsaffinität untersucht und ihre Auto- und Transphosphorylierungsaktivität wurde mit Hilfe von radioaktiven Autokinase Assays getestet. Dabei zeigte sich, dass drei der vier untersuchten Proteine echte Autokinasen sind und dass RdmS eine ungewöhnliche Kinase ist, der die H-Box und die Autokinase-Aktivität fehlen. Ein Phosphotransfer-Profiling Assay der drei Autokinasen mit allen im Genome von *M. acetivorans* kodierten Antwort Regulatoren ergab, dass die Histidin Kinasen mit mehreren Antwort Regulatoren interagieren können. Darüber hinaus wurden MA4377 und MA2013 als Hybrid Histidin Kinasen identifiziert, die an Phosphorelay Systemen beteiligt sind. Diese Studie lieferte den ersten Nachweis für die Existenz eines komplexen Signalnetzwerks in *M. acetivorans*, an dem eine Vielzahl an Histidin Kinasen und Antwort Regulatoren beteiligt sind.

References

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung C-C, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers Al, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, & Jumper JM. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature,* 630(8016), 493-500.

Ashenberg O, Keating AE, & Laub MT. (2013). Helix bundle loops determine whether histidine kinases autophosphorylate in cis or in trans. *Journal of Molecular Biology,* 425(7), 1198-1209.

Attwood PV. (2013). Histidine kinases from bacteria to humans. *Biochemical Society Transactions,* 41(4), 1023-1028.

Barr SA, Kennedy EN, McKay LS, Johnson RM, Ohr RJ, Cotter PA, & Bourret RB. (2023). Phosphorylation chemistry of the *Bordetella* PlrSR TCS and its contribution to bacterial persistence in the lower respiratory tract. *Molecular Microbiology,* 119(2), 174-190.

Bhate MP, Molnar KS, Goulian M, & DeGrado WF. (2015). Signal transduction in histidine kinases: insights from new structures. *Structure,* 23(6), 981-994.

Blasius L. (2023). Untersuchungen zu atypischen Signaltransduktionssystemen im methanogenen Archaeon *Methanosarcina acetivorans*. doctoral thesis. *Technische Universität Kaiserslautern*.

Bourret RB. (2010). Receiver domain structure and function in response regulator proteins. *Current Opinion in Microbiology,* 13(2), 142-149.

Buckstein MH, He J, & Rubin H. (2008). Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *Journal of Bacteriology,* 190(2), 718-726.

Buschiazzo A, & Trajtenberg F. (2019). Two-component sensing and regulation: How do histidine kinases talk with response regulators at the molecular level? *Annual Review of Microbiology,* 73, 507-528.

Caforio A, & Driessen AJM. (2017). Archaeal phospholipids: Structural properties and biosynthesis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids,* 1862(11), 1325-1339.

Calogero S, Gardan R, Glaser P, Schweizer J, Rapoport G, & Debarbouille M. (1994). RocR, a novel regulatory protein controlling arginine utilization in *Bacillus subtilis*, belongs to the NtrC/NifA family of transcriptional activators. *Journal of Bacteriology,* 176(5), 1234-1241.

Capra EJ, & Laub MT. (2012). Evolution of two-component signal transduction systems. *Annual Review of Microbiology,* 66, 325-347.

Casino P, Miguel-Romero L, & Marina A. (2014). Visualizing autophosphorylation in histidine kinases. *Nature Communications,* 5(1), 3258.

DeLano WL. **(2020)**. The PyMOL Molecular Graphics System, Version 2.3.

Delaroque N, Müller DG, Bothe G, Pohl T, Knippers R, & Boland W. (2001). The complete DNA sequence of the *Ectocarpus siliculosus* Virus EsV-1 genome. *Virology,* 287(1), 112-132.

Dumon-Seignovert L, Cariot G, & Vuillard L. (2004). The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification,* 37(1), 203-206.

Dutta R, Qin L, & Inouye M. (1999). Histidine kinases: diversity of domain organization. *Molecular Microbiology,* 34(4), 633-640.

Fiege K. (2019). Function of two redox sensing kinases from the methanogenic archaeon *Methanosarcina acetivorans*. doctoral thesis. *Technische Universität Kaiserslautern*.

Fiege K, & Frankenberg-Dinkel N. (2019). Thiol-based redox sensing in the methyltransferase associated sensor kinase RdmS in *Methanosarcina acetivorans*. *Environmental Microbiology,* 21(5), 1597-1610.

Fisher SL, Jiang W, Wanner BL, & Walsh CT. (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. *The Journal of Biological Chemistry,* 270(39), 23143-23149.

Flint DH, & Harrington RE. (1972). Gel electrophoresis of deoxyribonucleic acid. *Biochemistry,* 11(25), 4858-4864.

Forst SA, & Roberts DL. (1994). Signal transduction by the EnvZ-OmpR phosphotransfer system in bacteria. *Research in Microbiology,* 145(5), 363-373.

Francis VI, & Porter SL. (2019). Multikinase networks: Two-component signaling networks integrating multiple stimuli. *Annual Review of Microbiology,* 73, 199-223.

Francis VI, Waters EM, Finton-James SE, Gori A, Kadioglu A, Brown AR, & Porter SL. (2018). Multiple communication mechanisms between sensor kinases are crucial for virulence in *Pseudomonas aeruginosa*. *Nature Communications,* 9(1), 2219.

Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, DeArellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, Conway de Macario E, Ferry JG, Jarrell KF, Jing H, Macario AJ, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WW, & Birren B. (2002). The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Research,* 12(4), 532-542.

Galperin MY. (2006). Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *Journal of Bacteriology,* 188(12), 4169-4182.

Galperin MY. (2010). Diversity of structure and function of response regulator output domains. *Current Opinion in Microbiology,* 13(2), 150-159.

Galperin MY, Higdon R, & Kolker E. (2010). Interplay of heritage and habitat in the distribution of bacterial signal transduction systems. *Molecular Biosystems,* 6(4), 721-728.

Galperin MY, Makarova KS, Wolf YI, & Koonin EV. (2018). Phyletic distribution and lineage-specific domain architectures of archaeal two-component signal transduction systems. *Journal of Bacteriology,* 200(7).

Gao R, Bouillet S, & Stock AM. (2019). Structural basis of response regulator function. *Annual Review of Microbiology,* 73, 175-197.

Gao R, Mack TR, & Stock AM. (2007). Bacterial response regulators: versatile regulatory strategies from common domains. *Trends in Biochemical Science,* 32(5), 225-234.

Gao R, & Stock AM. (2009). Biological insights from structures of two-component proteins. *Annual Review of Microbiology,* 63, 133-154.

Gärtner W. (2020). Unequal twins: Unraveling the reaction mechanism of dimeric histidine kinases. *The Journal of Biological Chemistry,* 295(23), 8118-8119.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, & Smith HO. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods,* 6(5), 343-345.

González-Paredes A, Manconi M, Caddeo C, Ramos-Cormenzana A, Monteoliva-Sánchez M, & Fadda AM. (2010). Archaeosomes as carriers for topical delivery of betamethasone dipropionate: *In vitro* skin permeation study. *Journal of Liposome Research,* 20(4), 269-276.

Grozdanov L, Raasch C, Schulze J, Sonnenborn U, Gottschalk G, Hacker J, & Dobrindt U. (2004). Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917. *Journal of Bacteriology,* 186(16), 5432-5441.

Guarnieri MT, Blagg BS, & Zhao R. (2011). A high-throughput TNP-ATP displacement assay for screening inhibitors of ATP-binding in bacterial histidine kinases. *Assay and Drug Development Technologies,* 9(2), 174-183.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, & Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology,* 59(3), 307-321.

Gumerov VM, Ulrich LE, & Zhulin IB. (2023). MiST 4.0: a new release of the microbial signal transduction database, now with a metagenomic component. *Nucleic Acids Research,* 52(D1), D647-D653.

Guss AM, Rother M, Zhang JK, Kulkarni G, & Metcalf WW. (2008). New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for *Methanosarcina* species. *Archaea,* 2(3), 193-203.

Haldimann A, Fisher SL, Daniels LL, Walsh CT, & Wanner BL. (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. *Journal of Bacteriology,* 179(18), 5903-5913.

Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, & Winther O. (2022). DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv*, 2022.2004.2008.487609.

Hanahan D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology,* 166(4), 557-580.

Hargreaves KR, Kropinski AM, & Clokie MRJ. (2014). What does the talking?: Quorum sensing signalling genes discovered in a bacteriophage genome. *PLoS ONE,* 9(1), e85131.

Hérivaux A, So YS, Gastebois A, Latgé JP, Bouchara JP, Bahn YS, & Papon N. (2016). Major sensing proteins in pathogenic fungi: The hybrid histidine kinase family. *PLoS Pathogens,* 12(7), e1005683.

Hiratsuka T. (2003). Fluorescent and colored trinitrophenylated analogs of ATP and GTP. *European Journal of Biochemistry,* 270(17), 3479-3485.

Hoch JA. (2000). Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology,* 3(2), 165-170.

Hoch JA, & Varughese KI. (2001). Keeping signals straight in phosphorelay signal transduction. *Journal of Bacteriology,* 183(17), 4941-4949.

Huber C. (2024). The bathy phytochrome PaBphP of the human pathogen *Pseudomonas aeruginosa*: Introducing a novel perspective on light sensing in Proteobacteria. doctoral thesis. *Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau*.

Huber C, Strack M, Schultheiß I, Pielage J, Mechler X, Hornbogen J, Diller R, & Frankenberg-Dinkel N. (2024). Darkness inhibits autokinase activity of bacterial bathy phytochromes. *The Journal of Biological Chemistry,* 300(4), 107148.

Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, Martin J, Schultz JE, Lupas AN, & Coles M. (2006). The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell,* 126(5), 929-940.

Hunter T. (2022). A journey from phosphotyrosine to phosphohistidine and beyond. *Molecular Cell,* 82(12), 2190-2200.

Immormino RM, Silversmith RE, & Bourret RB. (2016). A variable active site residue influences the kinetics of response regulator phosphorylation and dephosphorylation. *Biochemistry,* 55(39), 5595-5609.

Janczarek M, Vinardell JM, Lipa P, & Karaś M. (2018). Hanks-Type Serine/Threonine protein kinases and phosphatases in bacteria: Roles in signaling and adaptation to various environments. *International Journal of Molecular Sciences,* 19(10).

Jiang M, Shao W, Perego M, & Hoch JA. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Molecular Microbiology,* 38(3), 535-542.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, &

Hassabis D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature,* 596(7873), 583-589.

Jung K, Tjaden B, & Altendorf K. (1997). Purification, reconstitution, and characterization of KdpD, the turgor sensor of *Escherichia coli Journal of Biological Chemistry,* 272(16), 10847-10852.

Kabbara S, Hérivaux A, Dugé de Bernonville T, Courdavault V, Clastre M, Gastebois A, Osman M, Hamze M, Cock JM, Schaap P, & Papon N. (2019). Diversity and evolution of sensor histidine kinases in eukaryotes. *Genome Biology and Evolution,* 11(1), 86-108.

Kim D-j, & Forst S. (2001). Genomic analysis of the histidine kinase family in bacteria and archaea. *Microbiology,* 147(5), 1197-1212.

Köhler S. (2021). Sensing of redox imbalances and its impact on gene expression in *Methanosarcina acetivorans*. master thesis. *Technische Universität Kaiserslautern*.

Kollmann R, & Altendorf K. (1993). ATP-driven potassium transport in right-side-out membrane vesicles via the Kdp system of *Escherichia coli*. *Biochimica et Biophysica Acta - Bioenergetics,* 1143(1), 62-66.

Koretke KK, Lupas AN, Warren PV, Rosenberg M, & Brown JR. (2000). Evolution of two-component signal transduction. *Molecular Biology and Evolution,* 17(12), 1956-1970.

Koshland DE, Jr. (1977). A response regulator model in a simple sensory system. *Science,* 196(4294), 1055-1063.

Krell T. (2018). Exploring the (almost) unknown: Archaeal two-component systems. *Journal of Bacteriology,* 200(7).

Kwiatkowski K. (2013). Integration of pyrrolysine and characterization of a sensorkinase from *Methanosarcina acetivorans*. master thesis. *Ruhr-Universität Bochum*.

Kwon A, Scott S, Taujale R, Yeung W, Kochut KJ, Eyers PA, & Kannan N. (2019). Tracing the origin and evolution of pseudokinases across the tree of life. *Science Signaling,* 12(578).

Laemmli UK. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227(5259), 680-685.

Laub MT, Biondi EG, & Skerker JM. (2007). Phosphotransfer profiling: systematic mapping of two-component signal transduction pathways and phosphorelays. *Methods in Enzymology,* 423, 531-548.

Laub MT, & Goulian M. (2007). Specificity in two-component signal transduction pathways. *Annual Review of Genetics,* 41, 121-145.

Lemoine F, Domelevo Entfellner JB, Wilkinson E, Correia D, Dávila Felipe M, De Oliveira T, & Gascuel O. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature,* 556(7702), 452-456.

Letunic I, & Bork P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research,* 49(W1), W293-W296.

Letunic I, Khedkar S, & Bork P. (2020). SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research,* 49(D1), D458-D460.

Li J, Zheng X, Guo X, Qi L, & Dong X. (2014). Characterization of an archaeal two-component system that regulates methanogenesis in *Methanosaeta harundinacea*. *PLoS One,* 9(4), e95502.

Lu JM, Deschenes RJ, & Fassler JS. (2003). *Saccharomyces cerevisiae* histidine phosphotransferase Ypd1p shuttles between the nucleus and cytoplasm for SLN1-dependent phosphorylation of Ssk1p and Skn7p. *Eukaryotic Cell,* 2(6), 1304-1314.

Mascher T, Helmann JD, & Unden G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiology and Molecular Biology Reviews,* 70(4), 910-938.

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, & Steinegger M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods,* 19(6), 679-682.

Miroux B, & Walker JE. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology,* 260(3), 289-298.

Molitor B. (2013). A heme-based redox sensor in the methanogenic archaeon *Methanosarcina acetivorans*. doctoral thesis. *Ruhr-Universität Bochum*.

Molitor B, Stassen M, Modi A, El-Mashtoly SF, Laurich C, Lubitz W, Dawson JH, Rother M, & Frankenberg-Dinkel N. (2013). A heme-based redox sensor in the methanogenic archaeon *Methanosarcina acetivorans*. *The Journal of Biological Chemistry,* 288(25), 18458-18472.

Moore JC. (1964). Gel permeation chromatography. I. A new method for molecular weight distribution of high polymers. *Journal of Polymer Science Part A: General Papers,* 2(2), 835-843.

Mukherjee S, Jemielita M, Stergioula V, Tikhonov M, & Bassler BL. (2019). Photosensing and quorum sensing are integrated to control *Pseudomonas aeruginosa* collective behaviors. *PLoS Biol,* 17(12), e3000579.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, & Erlich H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology,* 51 Pt 1, 263-273.

Muok AR, Briegel A, & Crane BR. (2020). Regulation of the chemotaxis histidine kinase CheA: A structural perspective. *Biochimica et Biophysica Acta - Biomembranes,* 1862(1), 183030.

Najnin T, Siddiqui KS, Taha T, Elkaid N, Kornfeld G, Curmi PM, & Cavicchioli R. (2016). Characterization of a temperature-responsive two component regulatory system from the Antarctic archaeon, *Methanococcoides burtonii*. *Scientific Reports,* 6, 24278.

Ortet P, Whitworth DE, Santaella C, Achouak W, & Barakat M. (2015). P2CS: updates of the prokaryotic two-component systems database. *Nucleic Acids Research,* 43(Database issue), D536-541.

Osakabe Y, Yamaguchi-Shinozaki K, Shinozaki K, & Tran LS. (2013). Sensing the environment: key roles of membrane-localized kinases in plant perception and response to abiotic stress. *Journal of Experimental Botany,* 64(2), 445-458.

Padilla-Vaca F, de la Mora J, García-Contreras R, Ramírez-Prado JH, Alva-Murillo N, Fonseca-Yepez S, Serna-Gutiérrez l, Moreno-Galván CL, Montufar-Rodríguez JM, Vicente-Gómez M, Rangel-Serrano Á, Vargas-Maya Nl, & Franco B. (2023). Two-component system sensor kinases from asgardian archaea may be witnesses to eukaryotic cell evolution. *Molecules,* 28(13).

Parkinson JS. (1993). Signal transduction schemes of bacteria. *Cell,* 73(5), 857-871.

Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L, Gough J, Haft DH, Letunić l, Marchler-Bauer A, Mi H, Natale DA, Orengo CA, Pandurangan AP, Rivoire C, Sigrist CJA, Sillitoe l, Thanki N, Thomas PD, Tosatto SCE, Wu CH, & Bateman A. (2023). InterPro in 2022. *Nucleic Acids Research,* 51(D1), D418-d427.

Pearson G, Robinson F, Beers Gibson T, Xu B-e, Karandikar M, Berman K, & Cobb MH. (2001). Mitogen-Activated Protein (MAP) kinase pathways: Regulation and physiological functions. *Endocrine Reviews,* 22(2), 153-183.

Perry J, Koteva K, & Wright G. (2011). Receptor domains of two-component signal transduction systems. *Molecular Biosystems,* 7(5), 1388-1398.

Plesniak L, Horiuchi Y, Sem D, Meinenger D, Stiles L, Shaffer J, Jennings PA, & Adams JA. (2002). Probing the nucleotide binding domain of the osmoregulator EnvZ using fluorescent nucleotide derivatives. *Biochemistry,* 41(47), 13876-13882.

Rivera-Cancel G, Ko WH, Tomchick DR, Correa F, & Gardner KH. (2014). Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. *Proceedings of the National Academy of Sciences of the United States of America,* 111(50), 17839-17844.

Rudolph J, & Oesterhelt D. (1995). Chemotaxis and phototaxis require a CheA histidine kinase in the archaeon *Halobacterium salinarium*. *The EMBO Journal,* 14(4), 667-673.

Sanger F, Nicklen S, & Coulson AR. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America,* 74(12), 5463-5467.

Santhosh PB, & Genova J. (2023). Archaeosomes: new generation of liposomes based on archaeal lipids for drug delivery and biomedical applications. *American Chemical Society Omega,* 8(1), 1-9.

Schaap P, Barrantes I, Minx P, Sasaki N, Anderson RW, Bénard M, Biggar KK, Buchler NE, Bundschuh R, Chen X, Fronick C, Fulton L, Golderer G, Jahn N, Knoop V, Landweber LF, Maric C, Miller D, Noegel AA, Peace R, Pierron G, Sasaki T, Schallenberg-Rüdinger M, Schleicher M, Singh R, Spaller T, Storey KB, Suzuki T, Tomlinson C, Tyson JJ, Warren WC, Werner ER, Werner-Felmayer G, Wilson RK, Winckler T, Gott JM, Glöckner G, & Marwan W. (2016). The *Physarum polycephalum* genome reveals extensive use of prokaryotic two-component and metazoan-type tyrosine kinase signaling. *Genome Biology and Evolution,* 8(1), 109-125.

Schultz JE, Kanchan K, & Ziegler M. (2015). Intraprotein signal transduction by HAMP domains: A balancing act. *International Journal of Medical Microbiology,* 305(2), 243-251.

Sexauer A. (2021). Characterization of a bacterial-like signal transduction phosphorelay in *Methanosarcina acetivorans*. doctoral thesis. *Technische Universität Kaiserslautern*.

Shrivastava R, Ghosh AK, & Das AK. (2007). Probing the nucleotide binding and phosphorylation by the histidine kinase of a novel three-protein two-component system from *Mycobacterium tuberculosis*. *FEBS Lett,* 581(9), 1903-1909.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, & Higgins DG. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology,* 7(1), 539.

Singh D, Gupta P, Singla-Pareek SL, Siddique KHM, & Pareek A. (2021). The journey from two-step to multi-step phosphorelay signaling Systems. *Current Genomics,* 22(1), 59-74.

Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, & Laub MT. (2008). Rewiring the specificity of two-component signal transduction systems. *Cell,* 133(6), 1043-1054.

Skerker JM, Prasol MS, Perchuk BS, Biondi EG, & Laub MT. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biology,* 3(10), e334.

Sourjik V, & Schmitt R. (1998). Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. *Biochemistry,* 37(8), 2327-2335.

Sousa EH, Gonzalez G, & Gilles-Gonzalez MA. (2005). Oxygen blocks the reaction of the FixL-FixJ complex with ATP but does not influence binding of FixJ or ATP to FixL. *Biochemistry,* 44(46), 15359-15365.

Sowers KR, Baron SF, & Ferry JG. (1984). *Methanosarcina acetivorans* sp. nov., an acetotrophic methane-producing bacterium isolated from marine sediments. *Applied and Environmental Microbiology,* 47(5), 971-978.

Stancik IA, Šestak MS, Ji B, Axelson-Fisk M, Franjevic D, Jers C, Domazet-Lošo T, & Mijakovic I. (2018). Serine/threonine protein kinases from bacteria, archaea and eukarya share a common evolutionary origin deeply rooted in the tree of life. *Journal of Molecular Biology,* 430(1), 27-32.

Stephenson K, & Hoch JA. (2002). Evolution of signalling in the sporulation phosphorelay. *Molecular Microbiology,* 46(2), 297-304.

Stewart RC, Jahreis K, & Parkinson JS. (2000). Rapid phosphotransfer to CheY from a CheA protein lacking the CheY-binding domain. *Biochemistry,* 39(43), 13157-13165.

Stewart RC, VanBruggen R, Ellefson DD, & Wolfe AJ. (1998). TNP-ATP and TNP-ADP as probes of the nucleotide binding site of CheA, the histidine protein kinase in the chemotaxis signal transduction pathway of *Escherichia coli*. *Biochemistry,* 37(35), 12269-12279.

Stock A, Chen T, Welsh D, & Stock J. (1988). CheA protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that control gene expression in response to changing environmental conditions. *Proceedings of the National Academy of Sciences of the United States of America,* 85(5), 1403-1407.

Stock A, Robinson VL, & Goudreau PN. (2000). Two-component signal transduction. *Annual Review of Biochemistry,* 69, 183-215.

Studier FW, & Moffatt BA. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology,* 189(1), 113-130.

Swain KE, & Falke JJ. (2007). Structure of the conserved HAMP domain in an intact, membrane-bound chemoreceptor: A disulfide mapping study. *Biochemistry,* 46(48), 13684-13695.

Swain KE, Gonzalez MA, & Falke JJ. (2009). Engineered socket study of signaling through a four-helix bundle: Evidence for a yin−yang mechanism in the kinase control module of the aspartate receptor. *Biochemistry,* 48(39), 9266-9277.

Szurmant H, & Ordal GW. (2004). Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiology and Molecular Biology Reviews,* 68(2), 301-319.

Thoma J, & Burmann B. (2020). Fake it 'till you make it—The pursuit of suitable membrane mimetics for membrane protein biophysics. *International Journal of Molecular Sciences,* 22, 50.

Tomomori C, Tanaka T, Dutta R, Park H, Saha SK, Zhu Y, Ishima R, Liu D, Tong KI, Kurokawa H, Qian H, Inouye M, & Ikura M. (1999). Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. *Nature Structural Biology,* 6(8), 729-734.

Ulrich LE, Koonin EV, & Zhulin IB. (2005). One-component systems dominate signal transduction in prokaryotes. *Trends in Microbiology,* 13(2), 52-56.

Valentine DL. (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology,* 5(4), 316-323.

Varughese KI, Tsigelny I, & Zhao H. (2006). The crystal structure of beryllofluoride Spo0F in complex with the phosphotransferase Spo0B represents a phosphotransfer pretransition state. *Journal of Bacteriology,* 188(13), 4970-4977.

Vega-Baray B, Domenzain C, Poggio S, Dreyfus G, & Camarena L. (2022). The histidine kinase CckA is directly inhibited by a response regulator-like protein in a negative feedback loop. *mBio Journal,* 13(4), e01481-01422.

Vidakovic I, Kornmueller K, Fiedler D, Khinast J, Fröhlich E, Leitinger G, Horn C, Quehenberger J, Spadiut O, & Prassl R. (2024). Archaeosomes for oral drug delivery: From continuous microfluidics production to powdered formulations. *Pharmaceutics,* 16(6).

Wang FF, Cheng ST, Wu Y, Ren BZ, & Qian W. (2017). A bacterial receptor PcrK senses the plant hormone cytokinin to promote adaptation to oxidative stress. *Cell Reports,* 21(10), 2940-2951.

Wegener-Feldbrügge S, & Søgaard-Andersen L. (2009). The atypical hybrid histidine protein kinase RodK in *Myxococcus xanthus*: spatial proximity supersedes kinetic preference in phosphotransfer reactions. *Journal of Bacteriology,* 191(6), 1765-1776.

Whelan S, & Goldman N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution,* 18(5), 691-699.

Whitworth DE, & Cock PJ. (2009). Evolution of prokaryotic two-component systems: insights from comparative genomics. *Amino Acids,* 37(3), 459-466.

Wojnowska M, Yan J, Sivalingam GN, Cryar A, Gor J, Thalassinos K, & Djordjevic S. (2013). Autophosphorylation activity of a soluble hexameric histidine kinase correlates with the shift in protein conformational equilibrium. *Chemistry and Biology,* 20(11), 1411-1420.

Wolanin PM, Thomason PA, & Stock JB. (2002). Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biology,* 3(10), Reviews3013.

Wolfe AJ. (2010). Physiologically relevant small phosphodonors link metabolism to signal transduction. *Current Opinion in Microbiology,* 13(2), 204-209.

Wuichet K, Cantwell BJ, & Zhulin IB. (2010). Evolution and phyletic distribution of two-component signal transduction systems. *Current Opinion in Microbiology,* 13(2), 219-225.

Xu Q, Porter SW, & West AH. (2003). The yeast YPD1/SLN1 complex: Insights into molecular recognition in two-component signaling systems. *Structure,* 11(12), 1569-1581.

Zhulin IB, Nikolskaya AN, & Galperin MY. (2003). Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea. *Journal of Bacteriology,* 185(1), 285-294.

Zschiedrich CP, Keidel V, & Szurmant H. (2016). Molecular mechanisms of two-component signal transduction. *Journal of Molecular Biology,* 428(19), 3752-3775.

Appendix

Accession numbers of employed sequences:

Table S1: Sequences employed for the generation of alignments and the construction of the phylogenetic trees (Figure 3.3, Figure 3.4, Figure 3.6, Figure S1, Figure S2). Only the HK domain sequence or the REC domain sequence was considered.

Histidine kinases of *M. acetivorans*

Protein	Locus tag	Organism	Accession number	Database
MA0014	MA_RS00065	*M. acetivorans*	Q8TUQ1	UniProt
MA0203	MA_RS01090	*M. acetivorans*	Q8TU70	UniProt
MA0490	MA_RS02555	*M. acetivorans*	Q8TTE7	UniProt
MA0551	MA_RS25000	*M. acetivorans*	Q8TT86	UniProt
MA0552	MA_RS02905	*M. acetivorans*	Q8TT85	UniProt
MA0619	MA_RS03260	*M. acetivorans*	Q8TT21	UniProt
MA0620	MA_RS03265	*M. acetivorans*	Q8TT20	UniProt
MA0758	MA_RS03960	*M. acetivorans*	Q8TSN7	UniProt
MA0759	MA_RS03965	*M. acetivorans*	Q8TSN6	UniProt
MA0777	MA_RS04070	*M. acetivorans*	Q8TSL9	UniProt
MA0863 (RdmS)	MA_RS04495	*M. acetivorans*	Q8TSD5	UniProt
MA0970	MA_RS24390	*M. acetivorans*	Q8TS36	UniProt
MA1149	MA_RS05985	*M. acetivorans*	Q8TRM6	UniProt
MA1267	MA_RS06580	*M. acetivorans*	Q8TRB3	UniProt
MA1270	MA_RS06595	*M. acetivorans*	Q8TRB0	UniProt
MA1274	MA_RS06615	*M. acetivorans*	Q8TRA6	UniProt
MA1322	MA_RS06850	*M. acetivorans*	Q8TR62	UniProt
MA1463	MA_RS24430	*M. acetivorans*	Q8TQS7	UniProt
MA1470	MA_RS07625	*M. acetivorans*	Q8TQS0	UniProt
MA1627	MA_RS08450	*M. acetivorans*	Q8TQC1	UniProt
MA1628	MA_RS08455	*M. acetivorans*	Q8TQC0	UniProt
MA1630	MA_RS08465	*M. acetivorans*	Q8TQB8	UniProt
MA1645	MA_RS08530	*M. acetivorans*	Q8TQA5	UniProt
MA1646	MA_RS08535	*M. acetivorans*	Q8TQA4	UniProt
MA1704	MA_RS08850	*M. acetivorans*	Q8TQ48	UniProt
MA1739	MA_RS09030	*M. acetivorans*	Q8TQ13	UniProt
MA1844	MA_RS24495	*M. acetivorans*	Q8TPR2	UniProt
MA1878	MA_RS09795	*M. acetivorans*	Q8TPM8	UniProt
MA1957	MA_RS10200	*M. acetivorans*	Q8TPF6	UniProt
MA1991	MA_RS24525	*M. acetivorans*	Q8TPC2	UniProt
MA2013	MA_RS10480	*M. acetivorans*	Q8TPA1	UniProt
MA2082	MA_RS10815	*M. acetivorans*	Q8TP40	UniProt
MA2256	MA_RS11710	*M. acetivorans*	Q8TNM8	UniProt
MA2266	MA_RS11765	*M. acetivorans*	Q8TNL9	UniProt
MA2294	MA_RS11915	*M. acetivorans*	Q8TNJ1	UniProt
MA2348	MA_RS12170	*M. acetivorans*	Q8TNE0	UniProt
MA2553	MA_RS13275	*M. acetivorans*	Q8TMU8	UniProt
MA2555	MA_RS13285	*M. acetivorans*	Q8TMU6	UniProt
MA2732	MA_RS14295	*M. acetivorans*	Q8TMC7	UniProt
MA2757	MA_RS28770	*M. acetivorans*	Q8TMA7	UniProt
MA2784	MA_RS28790	*M. acetivorans*	Q8TM82	UniProt
MA2890	MA_RS15160	*M. acetivorans*	Q8TLY2	UniProt

MA3066	MA_RS16035	*M. acetivorans*	Q8TLH0	UniProt
MA3346	MA_RS17465	*M. acetivorans*	Q8TKQ3	UniProt
MA3368	MA_RS25835	*M. acetivorans*	Q8TKN3	UniProt
MA3370	MA_RS24765	*M. acetivorans*	Q8TKN1	UniProt
MA3405	MA_RS17780	*M. acetivorans*	Q8TKK0	UniProt
MA3481	MA_RS18195	*M. acetivorans*	Q8TKC7	UniProt
MA3543	MA_RS24775	*M. acetivorans*	Q8TK73	UniProt
MA3962	MA_RS20675	*M. acetivorans*	Q8TJ26	UniProt
MA4026	MA_RS21010	*M. acetivorans*	Q8TIW4	UniProt
MA4377	MA_RS22890	*M. acetivorans*	Q8THY1	UniProt
MA4561 (MsmS)	MA_RS23780	*M. acetivorans*	Q8THF6	UniProt

Response regulators of *M. acetivorans*

Protein	Locus tag	Organism	Accession number	Database
MA_0016	MA_RS00075	*M. acetivorans*	Q8TUP9	UniProt
MA_0018	MA_RS00090	*M. acetivorans*	Q8TUP7	UniProt
MA_1268	MA_RS06585	*M. acetivorans*	Q8TRB2	UniProt
MA_1269	MA_RS06590	*M. acetivorans*	Q8TRB1	UniProt
MA_1366	MA_RS07090	*M. acetivorans*	Q8TR18	UniProt
MA_1468	MA_RS07615	*M. acetivorans*	Q8TQS2	UniProt
MA_1469	MA_RS07620	*M. acetivorans*	Q8TQS1	UniProt
MA_2012	MA_RS10475	*M. acetivorans*	Q8TPA2	UniProt
MA_2861	MA_RS15020	*M. acetivorans*	Q8TM09	UniProt
MA_3068	MA_RS16045	*M. acetivorans*	Q8TLG8	UniProt
MA_4376	MA_RS22885	*M. acetivorans*	Q8THY2	UniProt
MA_4671	MA_RS11705	*M. acetivorans*	Q8TNM9	UniProt
MA_2445	MA_RS12700	*M. acetivorans*	Q8TN48	UniProt
MA_0015	MA_RS00070	*M. acetivorans*	Q8TUQ0	UniProt
MA_3057	MA_RS15990	*M. acetivorans*	Q8TLH9	UniProt

Histidine kinases of bacterial, archaeal, and eukaryotic organisms

Protein	Locus tag	Organism	Accession number	Database
LtrK	Mbur_0694	*Methanococcoides burtonii*	Q12Y17	UniProt
FilI	Mhar_0446	*Methanothrix harundinacea*	E5KK10	UniProt
PcrK	XC_1756	*Xanthomonas campestris*	A0A0H2X6K9	UniProt
ArcB	arcB	*Escherichia coli*	P0AEC3	UniProt
EnvZ	envZ	*Escherichia coli*	P0AEJ4	UniProt
FixL	fixL	*Bradyrhizobium diazoefficiens*	P23222	UniProt
SinK	sinK	*Myxococcus xanthus*	A0A7Y7C172	UniProt
AHK1	AHK1	*Arabidopsis thaliana*	Q9SXL4	UniProt
SLN1	SLN1	*Saccharomyces cerevisiae*	P39928	UniProt
NarX	narX	*Escherichia coli*	P0AFA2	UniProt
PhoQ	phoQ	*Escherichia coli*	P23837	UniProt
CpxA	cpxA	*Escherichia coli*	P0AE82	UniProt
DesK	desK	*Bacillus subtilis*	O34757	UniProt
DHKA	dhkA	*Dictyostelium discoideum*	Q54U87	UniProt
CheA	cheA	*Escherichia coli*	P07363	UniProt
CheA	cheA	*Halobacterium salinarum* R1	P07363	UniProt
KdpD	kdpD	*Escherichia coli*	P21865	UniProt
TorS	torS	*Escherichia coli*	P39453	UniProt
MoSln1p	PoMZ_02373	*Pyricularia oryzae*	A0A4P7N4M5	UniProt
BarA	barA	*Escherichia coli*	P0AEC5	UniProt

DcuS	dcuS	*Escherichia coli*	P0AEC8	UniProt
NarQ	narQ	*Escherichia coli*	P27896	UniProt
YehU	btsS, yehU	*Escherichia coli*	P0AD14	UniProt
A0A5B9DCW9		Candidatus Prometheoarchaeum syntrophicum	A0A5B9DCW9	UniProt
A0A5B9DBR6		Candidatus Prometheoarchaeum syntrophicum	A0A5B9DBR6	UniProt
A0A0F8YCA1		*Lokiarchaeum sp.* (strain GC14_75)	A0A0F8YCA1	UniProt
A0A0F8VHJ1		*Lokiarchaeum sp.* (strain GC14_75)	A0A0F8VHJ1	UniProt
A0A841HC45		*Halobacterium salinarum* R1	A0A841HC45	UniProt
KinA2	kinA2	*Halobacterium salinarum* NRC-1	Q9HRE4	UniProt
O26226		*Methanothermobacter thermautotrophicus*	O26226	UniProt
O26460		*Methanothermobacter thermautotrophicus*	O26460	UniProt
O30214		*Archaeoglobus fulgidus*	O30214	UniProt
A9A1W4		*Nitrosopumilus maritimus*	A9A1W4	UniProt
A0A060HLN9		*Nitrososphaera viennensis*	A0A060HLN9	UniProt
L0JM57		*Natrinema pellirubrum*	L0JM57	UniProt
L0HHT4		*Methanoregula formicica*	L0HHT4	UniProt
Q2FR64		*Methanospirillum hungatei* JF-1	Q2FR64	UniProt
A0A1Y1I805		*Klebsormidium nitens*	A0A1Y1I805	UniProt
A0A0U2KSP6		*Guillardia theta*	A0A0U2KSP6	UniProt

Partial amino acid sequence alignment of the F-box region of all HK of *M. acetivorans*:

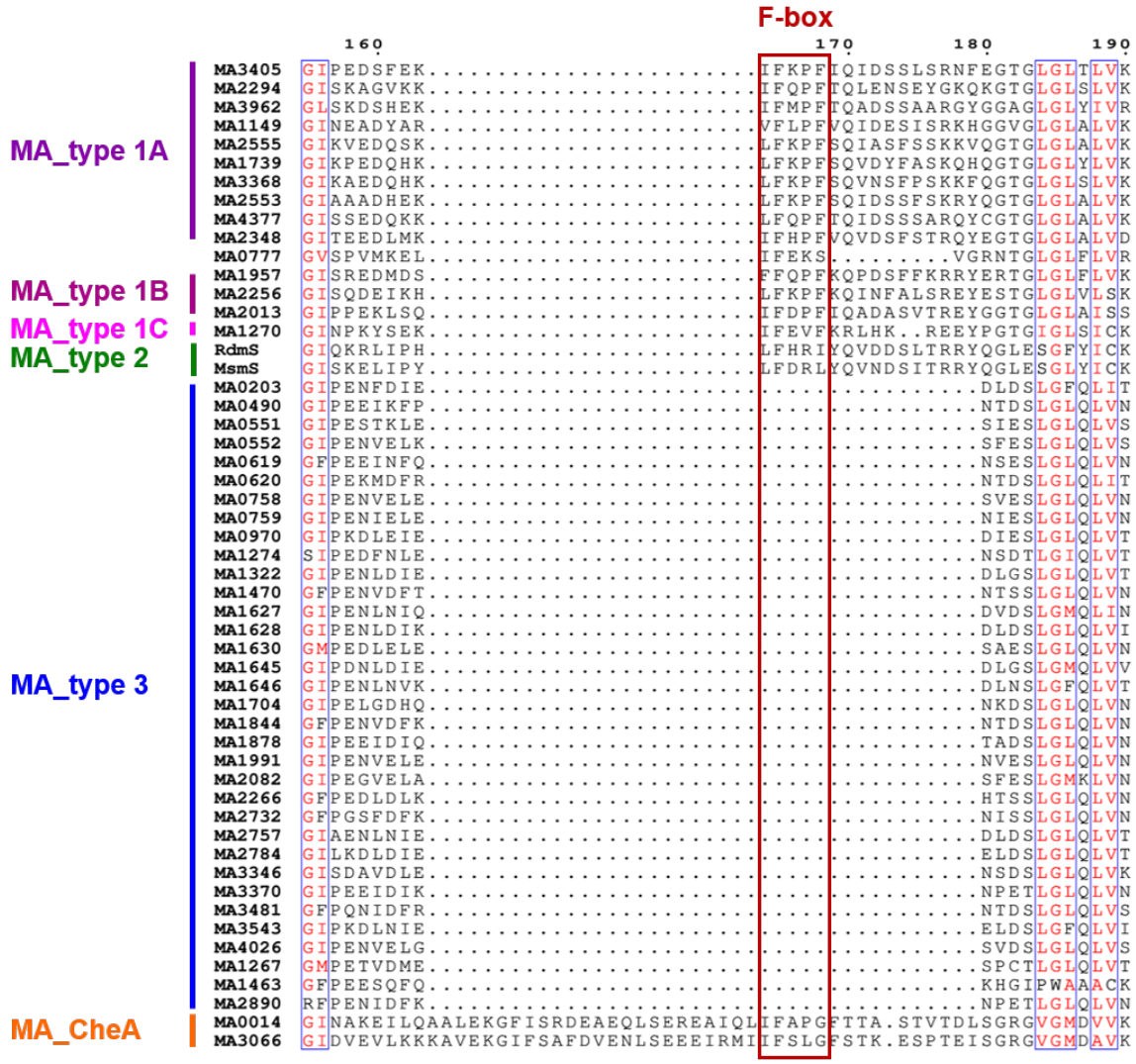

Figure S1: Partial amino acid sequence alignment of the F-box region of all putative HK of *M. acetivorans*. Acession numbers of the employed sequences are listed in Table S1. The alignment was constructed using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) and modified using ESPript v.3.0 (https://espript.ibcp.fr/ESPript/ESPript/). MA_type 3 kinases are characterised by a missing F-box (ATP lid).

Phylogenetic tree of different archaea, bacterial and eukaryotic HK domains:

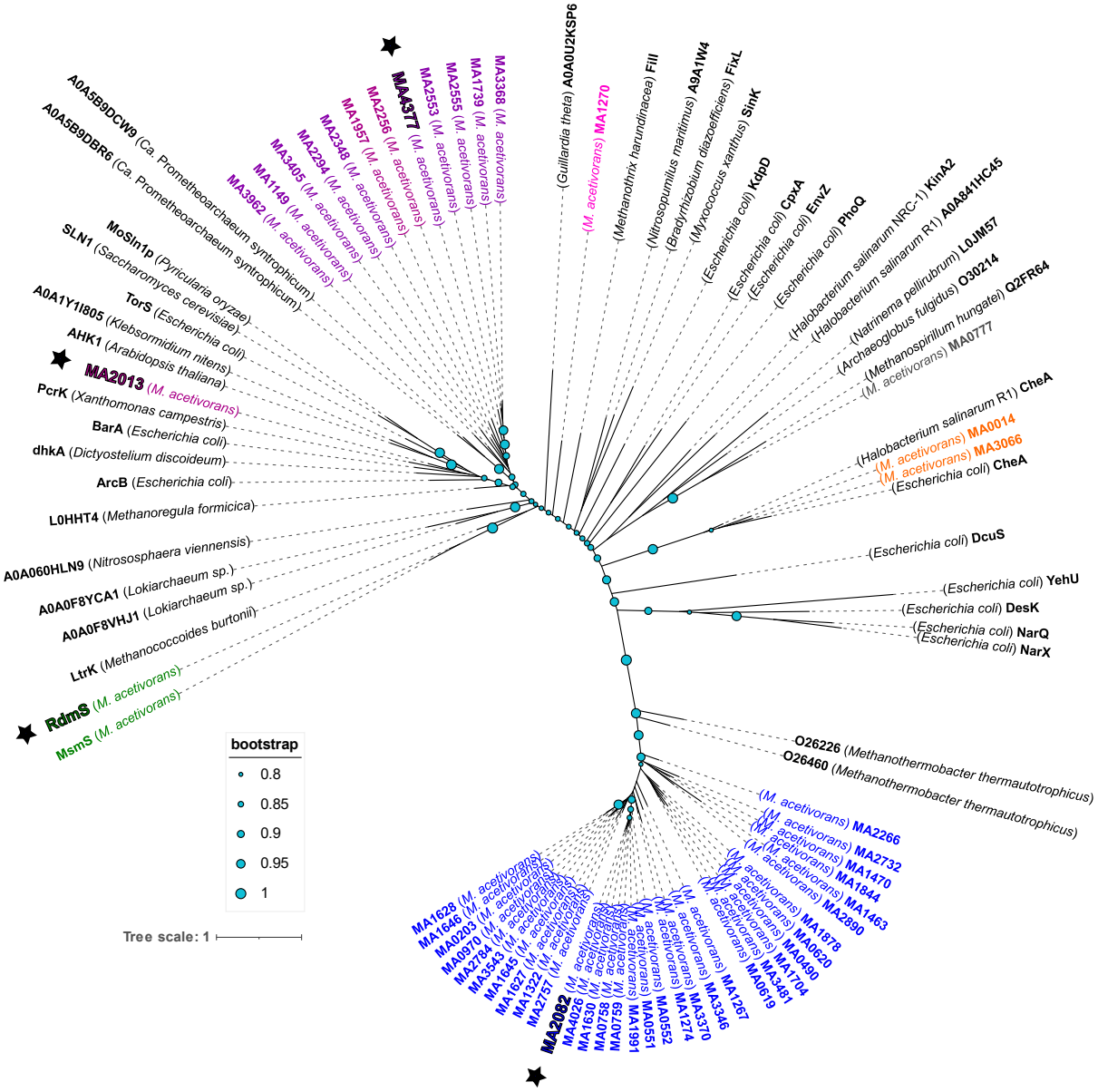

Figure S2: Unrooted phylogenetic tree depicting the HK domain of 53 putative HKs from *M. acetivorans* and additional of putative archaeal, bacterial and eukaryotic HKs. Phylogenetic analysis of the HK domain from *M. acetivorans*. The maximum-likelihood phylogenetic tree was constructed using PhyML 3.0 (Guindon *et al.*, 2010), bootstrap values were calculated using the online tool BOOSTER (Lemoine *et al.*, 2018). Bootstrap values are displayed as light blue circles in different sizes. To generate the tree, only the HK domain of the HKs was considered. Labels at branches are coloured according to the HK type. Accession numbers of the used sequences are found in the Appendix Table S1.

Sequences of CHASE domains employed for the amino acid sequence alignment:

Table S2: Sequences with accession numbers employed for the generation of an amino acid alignment of different CHASE domains (Figure S3). Only the CHASE domain was considered for the alignment.

Proteins for CHASE domain alignment				
Protein	**CHASE domain type**	**Organism**	**Accession number**	**Database**
PA4036	CHASE 2	*Pseudomonas aeruginosa*	Q9HWZ1	UniProt
SPCyaA	CHASE 3	*Spirulina platensis*	O32392	UniProt
XCPcrK		*Xanthomonas campestris*	A0A0H2X6K9	UniProt
ATAHK4		*Arabidopsis thaliana*	Q9C5U0	UniProt
MA2348	CHASE 4	*M. acetivorans*	Q8TNE0	UniProt
MA2553	CHASE 4	*M. acetivorans*	Q8TMU8	UniProt
MA2555	CHASE 4	*M. acetivorans*	Q8TMU6	UniProt
MA1739	CHASE 4	*M. acetivorans*	Q8TQ13	UniProt
MA4377	CHASE 4	*M. acetivorans*	Q8THY1	UniProt
MHFiII	CHASE 4	*Methanothrix harundinacea*	E5KK10	UniProt
AF1721	CHASE 4	*Archaeoglobus fulgidus*	O28553	UniProt

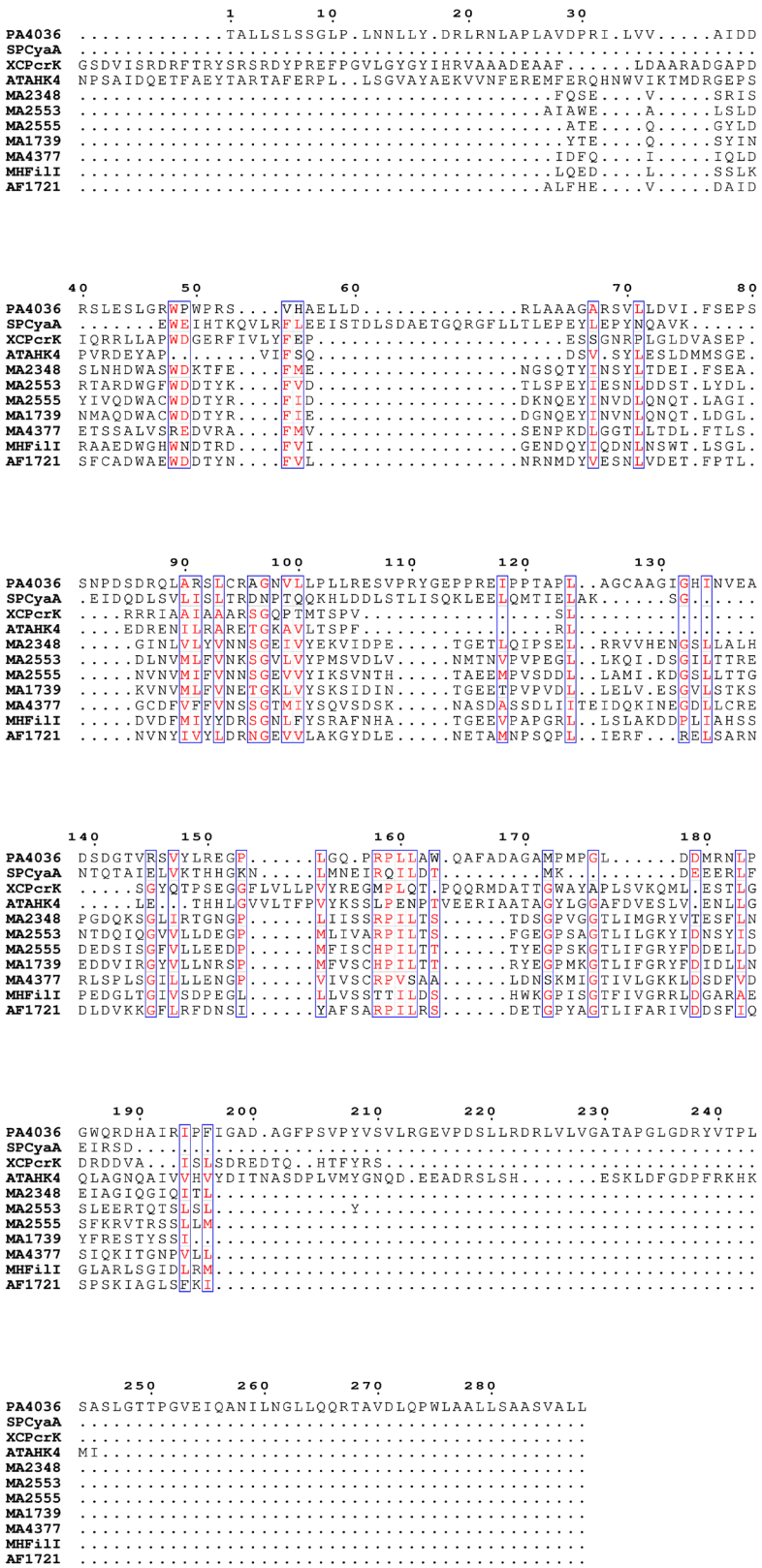

Figure S3: Amino acid sequence alignment of different CHASE domains. Employed sequences are listed in the Appendix Table S2. The alignment was constructed using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) and modified using ESPript v.3.0 (https://espript.ibcp.fr/ESPript/ESPript/).

SEC data and predicted three-dimensional structures of the analysed proteins:

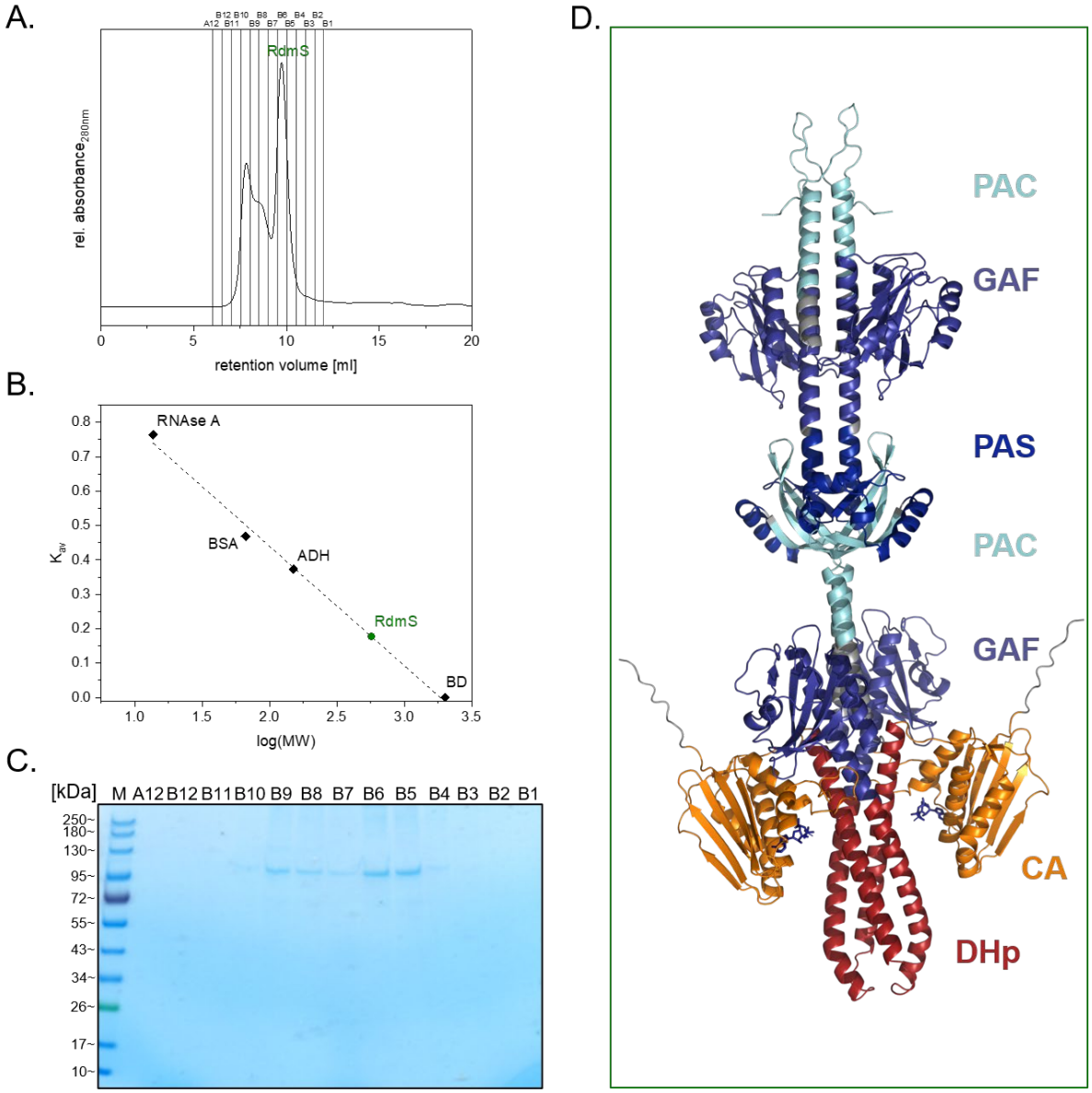

Figure S4: Purification, determination of oligomerisation state and predicted homodimer structure of RdmS. (**A.**) Coomassie stained gel of RdmS after recombinant production and affinity chromatography purification with fractions (E). As reference the colour-stained protein marker (M) from NEB was used. RdmS has a predicted molecular weight of ~82 kDa. (**B.**) Elution profile of RdmS. Size exclusion chromatography was performed with a Superdex® 200 Increase 10/300 GL column equilibrated with kinase buffer, pH 8.0 containing 5 % glycerol. The relative absorbance was detected at 280 nm and plotted against the retention volume. (**C.**) Calibration curve of the Superdex® 200 Increase 10/300 GL column. RNAse A (13.7 kDA), bovine serum albumin (BSA; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), blue dextran (BD; 2000 kDa). The void volume was determined using blue dextran (BD). The actual column volume was determined with Acetone. The calibration curve was obtained by plotting the K_{av} of the standard proteins against the logarithm of the molecular weight (black squares). The fitting was determined using linear regression in OriginLab. RdmS is indicated on the calibration curve (blue dot). (**D.**) AlphaFold3 prediction of RdmS as a homodimer (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

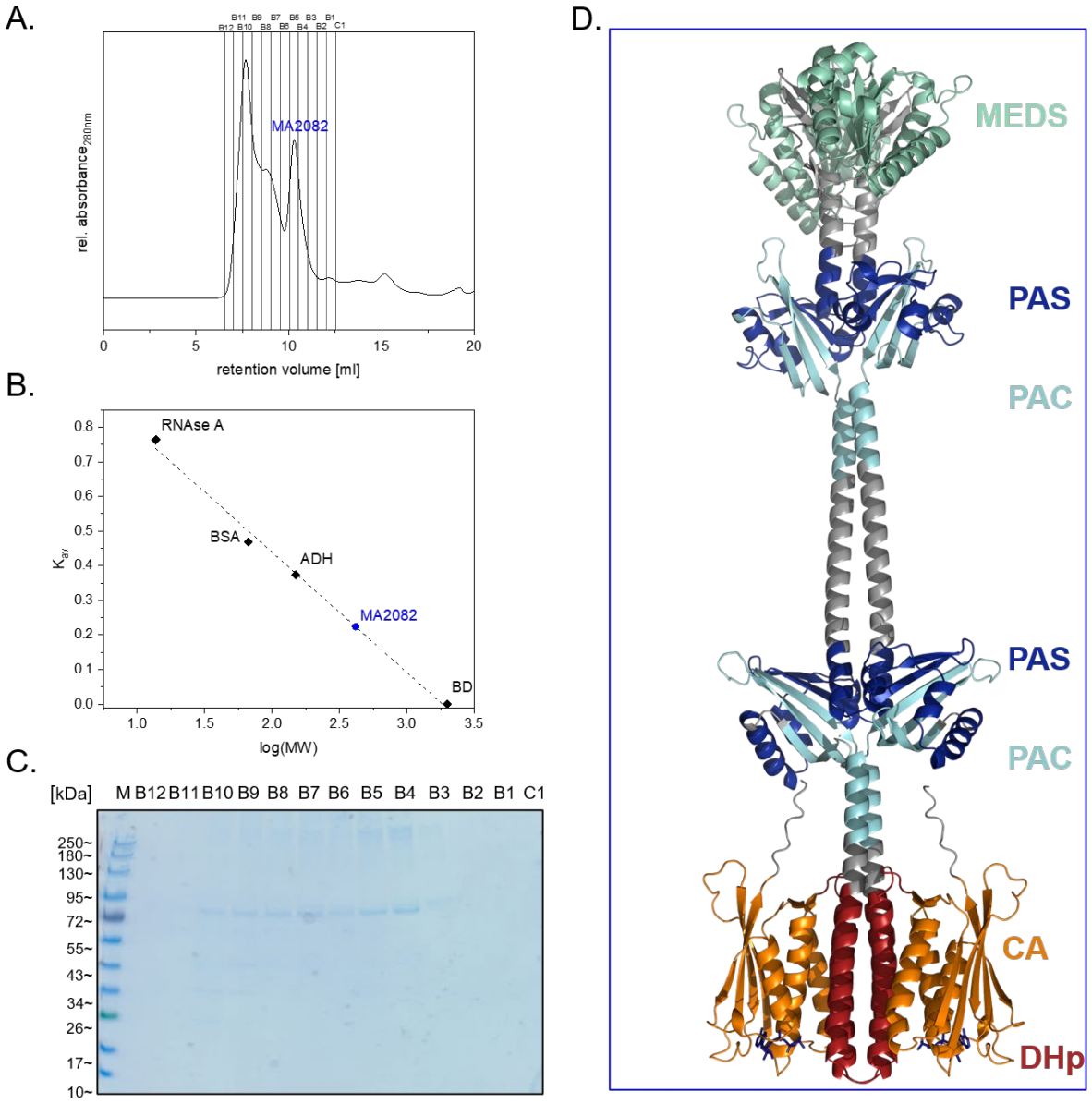

Figure S5: Purification, determination of oligomerisation state and predicted homodimer structure of MA2082. (**A.**) Elution profile of MA2082. Size exclusion chromatography was performed with a Superdex® 200 Increase 10/300 GL column equilibrated with kinase buffer, pH 8.0 containing 5 % glycerol. The relative absorbance was detected at 280 nm and plotted against the retention volume. (**B.**) Calibration curve of the Superdex® 200 Increase 10/300 GL column. RNAse A (13.7 kDA), bovine serum albumin (BSA; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), blue dextran (BD; 2000 kDa). The void volume was determined using blue dextran (BD). The actual column volume was determined with Acetone. The calibration curve was obtained by plotting the K_{av} of the standard proteins against the logarithm of the molecular weight (black squares). The fitting was determined using linear regression in OriginLab. MA2082 is indicated on the calibration curve (blue dot). (**C.**) Coomassie stained gel of the isolated protein fractions to analyses which peak harbours the protein MA2082 (**D.**) AlphaFold3 prediction of MA2082 as a homodimer (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

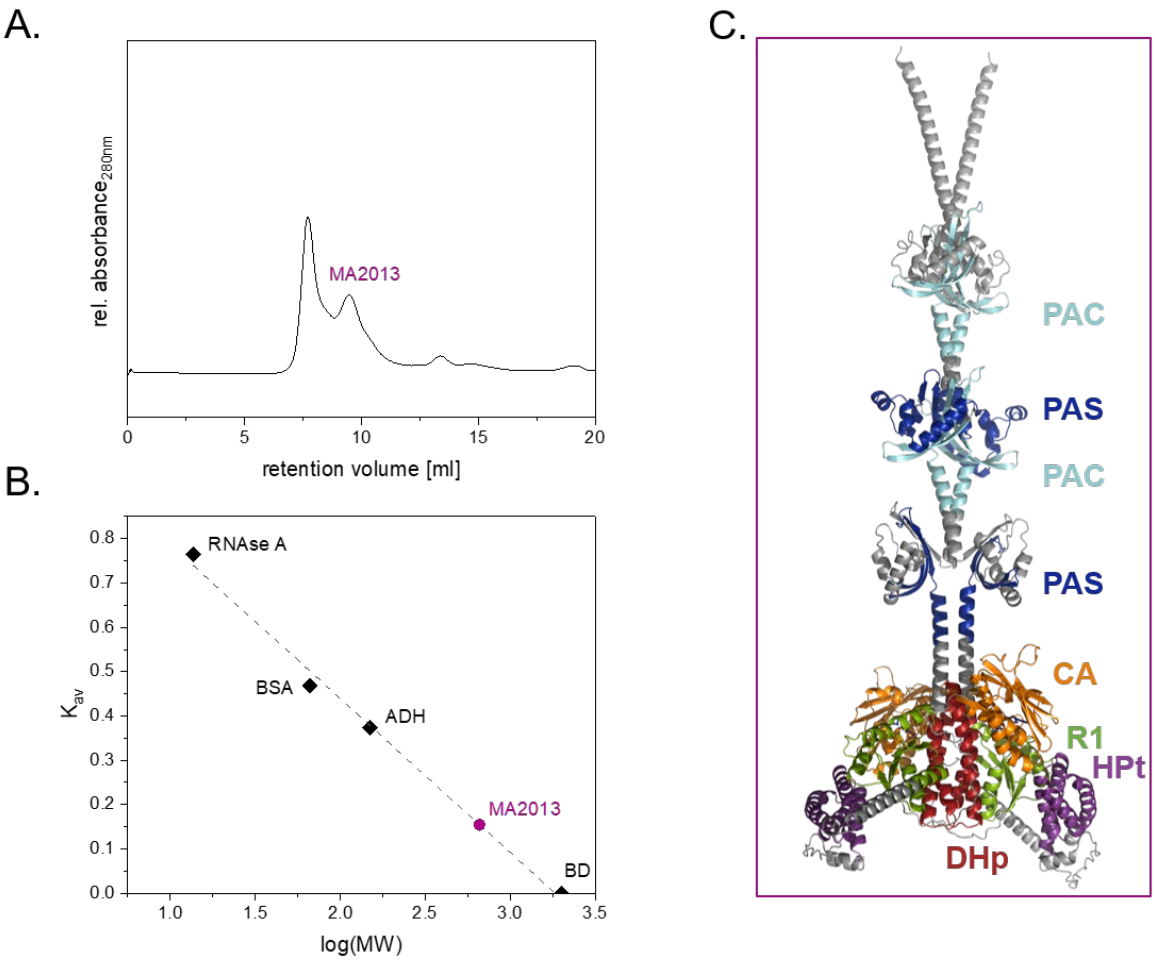

Figure S6: Purification, determination of oligomerisation state and predicted homodimer structure of MA2013. (**A.**) Elution profile of MA2013. Size exclusion chromatography was performed with a Superdex® 200 Increase 10/300 GL column equilibrated with kinase buffer, pH 8.0 containing 5 % glycerol. The relative absorbance was detected at 280 nm and plotted against the retention volume. (**B.**) Calibration curve of the Superdex® 200 Increase 10/300 GL column. RNAse A (13.7 kDA), bovine serum albumin (BSA; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), blue dextran (BD; 2000 kDa). The void volume was determined using blue dextran (BD). The actual column volume was determined with Acetone. The calibration curve was obtained by plotting the K_{av} of the standard proteins against the logarithm of the molecular weight (black squares). The fitting was determined using linear regression in OriginLab. MA2013 is indicated on the calibration curve (pink dot). (**C.**) AlphaFold3 prediction of MA2013 as a homodimer (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

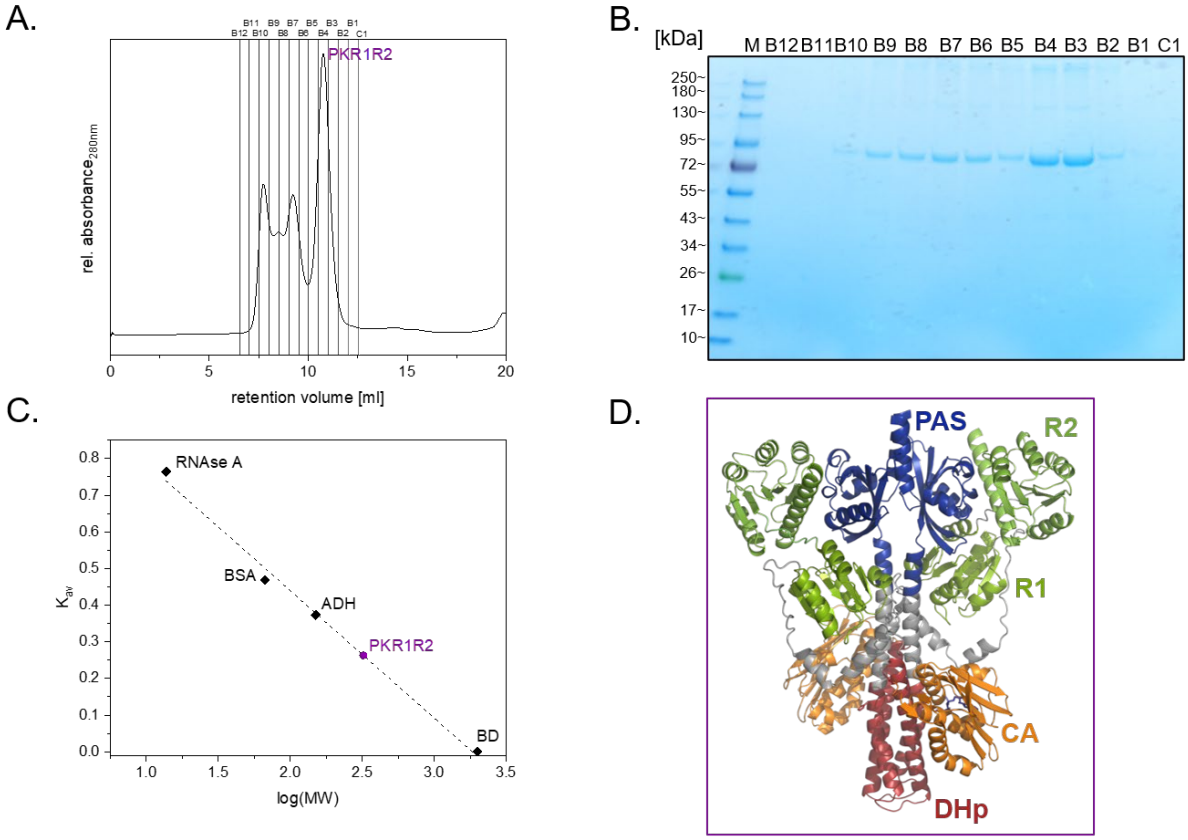

Figure S7: Purification, determination of oligomerisation state and predicted homodimer structure of PKR1R2. (**A.**) Elution profile of PKR1R2. Size exclusion chromatography was performed with a Superdex® 200 Increase 10/300 GL column equilibrated with kinase buffer, pH 8.0 containing 5 % glycerol. The relative absorbance was detected at 280 nm and plotted against the retention volume. (**B.**) Coomassie stained gel of the isolated protein fractions to analyses which peak harbours the protein PKR1R2 (**C.**) Calibration curve of the Superdex® 200 Increase 10/300 GL column. RNAse A (13.7 kDA), bovine serum albumin (BSA; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), blue dextran (BD; 2000 kDa). The void volume was determined using blue dextran (BD). The actual column volume was determined with Acetone. The calibration curve was obtained by plotting the K_{av} of the standard proteins against the logarithm of the molecular weight (black squares). The fitting was determined using linear regression in OriginLab. PKR1R2 is indicated on the calibration curve (purple dot). (**D.**) AlphaFold3 prediction of PKR1R2 as a homodimer (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

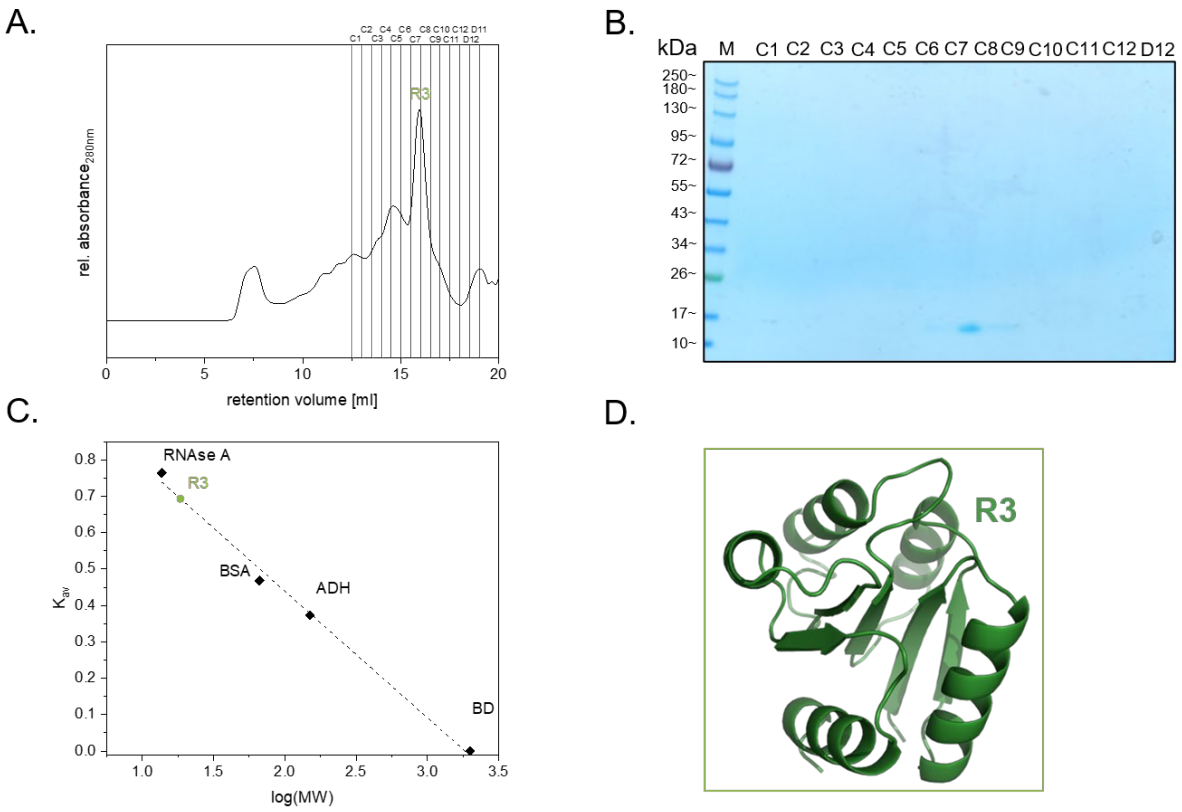

Figure S8: Purification, determination of oligomerisation state and predicted homodimer structure of R3. (**A.**) Elution profile of R3. Size exclusion chromatography was performed with a Superdex® 200 Increase 10/300 GL column equilibrated with kinase buffer, pH 7.0 containing 5 % glycerol. The relative absorbance was detected at 280 nm and plotted against the retention volume. (**B.**) Coomassie stained gel of the isolated protein fractions to analyses which peak harbours the protein R3 (**C.**) Calibration curve of the Superdex® 200 Increase 10/300 GL column. RNAse A (13.7 kDA), bovine serum albumin (BSA; 66 kDa), alcohol dehydrogenase (ADH; 150 kDa), blue dextran (BD; 2000 kDa). The void volume was determined using blue dextran (BD). The actual column volume was determined with Acetone. The calibration curve was obtained by plotting the K_{av} of the standard proteins against the logarithm of the molecular weight (black squares). The fitting was determined using linear regression in OriginLab. R3 is indicated on the calibration curve (green dot). (**D.**) AlphaFold3 prediction of R3 as a monomer (Abramson *et al.*, 2024). Displayed in cartoon representation and coloured regarding the domain structure with PyMOL (DeLano, 2020).

Curriculum vitae

PERSÖNLICHE DATEN

Nora F. K. Georgiev

STUDIUM UND SCHULISCHE LAUFBAHN

seit 10/2020	**Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau** Wissenschaftliche Mitarbeiterin/ Promotion in der Abteilung Mikrobiologie (Prof. Dr. Nicole Frankenberg-Dinkel)
10/2017–06/2020	**Rheinische Friedrich-Wilhelms-Universität, Bonn** Master of Science „Mikrobiologie“
05/2019–07/2019, 10/2019–03/2020	**University of Essex, Colchester (UK)** ERASMUS+ Praktikum
10/2014–09/2017	**Rheinische Friedrich-Wilhelms-Universität, Bonn** Bachelor of Science „Biologie“
09/2012–07/2014	**Naturwissenschaftliches Technikum Dr. Künkele, Landau** Ausbildung: Biologisch-Technische Assistentin (BTA)
08/2003–03/2012	**Nikolaus-von-Weis Gymnasium, Speyer** Allgemeine Hochschulreife

AUSGEWÄHLTE KONFERENZTEILNAHMEN

Vorträge

03/2024	**“A bacterial-type phosphorelay in the methanogenic archaeon *Methanosarcina acetivorans”*** Rheinland-Pfalz Mikrobiologie Symposium; Mainz, Deutschland
09/2023	**“A bacterial-type phosphorelay in the methanogenic archaeon *Methanosarcina acetivorans”*** VAAM Annual Meeting; Göttingen, Deutschland

Poster

07/2022	**“Characterization of signalling kinases in *Methanosarcina acetivorans* - Biochemical analyses give insight into evolution and function”** EMBO Workshop: “Molecular biology of archaea”; Frankfurt, Deutschland
06/2022	**“Signalling kinases of *Methanosarcina acetivorans* - Biochemical analyses give insight into evolution and function”** Halophiles 2022 Conference; Alicante, Spanien
02/2022	**“Characterization of unusual kinases in *Methanosarcina acetivorans* and their role in signal transduction”** VAAM Annual Meeting; online

PUBLIKATIONEN

Nora FK Georgiev, Anne L Andersson, Zoe Ruppe, Loriana Kattwinkel and Nicole Frankenberg-Dinkel. Archaeal signalling networks – new insights into the structure and function of histidine kinases and response regulators of the methanogenic archaeon *Methanosarcina acetivorans.* eingereichtes Manuskript, Environmental Microbiology; BioRXiv: https://doi.org/10.1101/2024.10.04.616672

Eidesstattliche Erklärung

Erklärungen lt. § 6 (4) der Promotionsordnung des Fachbereichs Biologie vom 27.11.2018:

Ich erkläre wahrheitsgemäß, dass ich die eingereichte Dissertation selbst angefertigt und alle für die Arbeit benutzten Hilfsmittel und Hilfestellungen in der Arbeit angegeben habe. Die Dissertation oder Teile hiervon habe ich bisher nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht. Ebenso habe ich die gleiche oder eine andere Abhandlung nicht bei einem anderen Fachbereich oder einer anderen Universität als Dissertation eingereicht.

Kaiserslautern, 15.11.2024

Nora Georgiev

Darlegung des Eigenanteils

Nora Georgiev

Histidine kinases and response regulators in *Methanosarcina acetivorans*: Structural and functional insights into archaeal signalling networks

Die Konzeption, die Planung, das Versuchsdesign und die Methodenentwicklung der vorliegenden Arbeit wurde in Teilen von Prof. Dr. Nicole Frankenberg-Dinkel unterstützt. Die Methodenentwicklung wurde in Teilen von Dr. Eugenio Pérez Patallo unterstützt.

Die Literaturrecherche, die Datenerhebung, die Datenanalyse, die Interpretation und Diskussion der Ergebnisse wurden, sofern im Folgenden nicht anders dargelegt, nicht von Dritten unterstützt.

Bakterienstämme sowie Plasmide, die von dritter Seite generiert wurden, sind, in Tabelle 2.3 und Tabelle 2.4, mit der entsprechenden Quelle, aufgeführt. Die Klonierung von folgendem Plasmid wurde von Zoe Ruppe unterstützt: pASK-IBA3-MA_4377-CHPK

Der *E. coli* TKR 2000 Stamm wurden freundlicherweise von Prof. Dr. Kirsten Jung vom Institut für Mikrobiologie (Ludwig-Maximilians-Universität München (LMU), Deutschland) zur Verfügung gestellt.

Die Datenerhebung für Abbildung (Figure 3.8) wurde von Zoe Ruppe (Bachelorstudentin, RPTU) unterstützt. Die Datenanalyse und -auswertung für diese Abbildung wurde nicht von Dritten unterstützt.

Teile dieser Arbeit liegen zur Veröffentlichung vor: “Archaeal signalling networks - new insights into the structure and function of histidine kinases and response regulators of the methanogenic archaeon *Methanosarcina acetivorans*”, Nora FK Georgiev, Anne L Andersson, Zoe Ruppe, Loriana Kattwinkel and Nicole Frankenberg-Dinkel. Environmental Microbiology, eingereichtes Manuskript, BioRXiv: https://doi.org/10.1101/2024.10.04.616672

Die ersten Entwürfe für diese Veröffentlichung wurden von der Autorin dieser Dissertation angefertigt, die Fertigstellung des Manuskripts bis zur Einreichung wurde von Prof. Dr. Nicole Frankenberg-Dinkel unterstützt.

Die vorliegende Einschätzung über die erbrachte Leistung von Dritten wurde mit den genannten Personen einvernehmlich abgestimmt.

Die folgenden Abschnitte dieser Arbeit enthalten Daten, Beschreibungen und Auswertungen, die, in unterschiedlicher Form, im Rahmen von Georgiev *et al.*, 2024 (eingereichtes Manuskript) bereits veröffentlicht wurden:

Materials and Methods: 2.2, 2.5, 2.6

Results: 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.12,3.13, 3.14, 3.15

Discussion: 4.1, 4.2, 4.4

_____ _____

Datum, Unterschrift Doktorandin Datum, Unterschrift Betreuerin

Darlegung aller benutzten Hilfsmittel und Hilfestellungen

Nora Georgiev

Histidine kinases and response regulators in *Methanosarcina acetivorans*: Structural and functional insights into archaeal signalling networks

Die Analyse und Visualisierung der Daten erfolgte überwiegend mit Hilfe von Microsoft Excel (Microsoft 365 Apps, Version 2105; Redmond (WA), USA) und OriginLab (Origin 2022, 9.9.0.225, Northampton, (MA), USA). Abbildungen wurden mit CorelDraw (CorelDRAW GraphicsSuite 2017, Version 19.1.0.448; Ottawa, Canada), Microsoft PowerPoint (Microsoft 365 Apps, Version 2105; Redmond (WA), USA) und Inkscape (Inkscape 1.3.2, 2023) erstellt und bearbeitet.

In silico Analysen wurden mit folgenden Programmen und Webseiten durchgeführt: SMART (http://smart.embl-heidelberg.de/) (Letunic *et al.*, 2020), InterPro (https://www.ebi.ac.uk/interpro/) (Paysan-Lafosse *et al.*, 2023), DeepTMHMM 1.0 server (https://services.healthtech.dtu.dk/services/DeepTMHMM-1.0/) (Hallgren *et al.*, 2022), Microbial Signal Transduction Database 4.0 (MiST 4.0) (https://mistdb.com/mist/genomes) (Gumerov *et al.*, 2023), Procaryotic 2-Component System database (P2CS) (http://www.p2cs.org/) (Ortet *et al.*, 2015), Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) (Sievers *et al.*, 2011) and edited using ESPript v.3.0 (https://espript.ibcp.fr/ESPript/ESPript/), AlphFold2 mit ColabFold (Jumper *et al.*, 2021; Mirdita *et al.*, 2022), AlphFold3 (https://alphafoldserver.com/about) (Abramson *et al.*, 2024), PyMOL (DeLano, 2020), iTOL (interactive Tree of Life, version 6.8) über den Webserver itol.embl.de (Letunic & Bork, 2021), SnapGene (SnapGene® 4.0.8; San Diego (CA), USA).

Handelt es sich nicht um frei nutzbare Programme oder Webseiten wurden sie von der RPTU oder der Abteilung Mikrobiologie zur Verfügung gestellt.

Die Übersetzungsprogramme DeepL und DeepL Write wurde zur Übersetzung von einzelnen Worten, zum Finden von Synonymen und zur Verbesserung der akademischen Sprachqualität verwendet.

Kaiserslautern, 15.11.2024

Nora Georgiev

Danksagung

Zuerst möchte ich mich bei meiner Doktormutter Prof. Dr. Nicole Frankenberg-Dinkel für die Vergabe dieses spannenden und vielseitigen Themas bedanken, das uns manchmal beide herausgefordert hat. Vielen Dank für die stetige Unterstützung, die experimentellen Freiräume und die fachlichen Diskussionen die mich während der gesamten Zeit motiviert haben.

Prof. Dr. Matthias Hahn danke ich für die freundliche Übernahme des Zweitgutachtens. Weiterhin möchte ich mich bei Prof. Dr. Stefan Kins für die Übernahme des Vorsitzes der Prüfungskommission bedanken.

Ein weiterer Dank geht an Prof. Dr. Kirsten Jung (Ludwig-Maximilians-Universität München) und Dr. Stefan Jacob (Institut für Biotechnologie und Wirkstoff-Forschung Mainz) für hilfreiche fachliche Diskussionen und an Prof. Dr. Kirsten Jung für die Bereitstellung des *E. coli* TKR 2000 Stammes.

Bei Dr. Ilka Haferkamp möchte ich mich für die Unterstützung im Isotopenlabor und die regelmäßigen Bestellungen von radioaktivem ATP bedanken. Weiterhin möchte ich mich auch bei der Arbeitsgruppe Zellbiologie (Prof. Dr. Johannes Herrmann) für die Möglichkeit bedanken das ich dort meine Kinase Assays auswerten konnte.

Einen besonderen Dank möchte ich an die gesamte Arbeitsgruppe Mikrobiologie richten. Danke an die Technischen Assistentinnen, allen voran Christine Förster-Schorr für die Hilfe beim finalen Klonierungs-Marathon. Ein großes Dankeschön geht auch an PD Dr. Susanne Zehner, Dr. Eugenio Pérez Patallo und Dr. Michelle Gehringer für immer neue Ideen, die das Projekt vorangebracht haben. Danke auch an alle Studenten, die das Projekt begleitet haben. Besonders möchte ich mich bei Zoe bedanken, die gerade am Ende nochmal alles gegeben hat und eine riesige Unterstützung war. Vielen Dank an das ehemalige Team Methano, besonders an Loriana und Sandra die mich gerade am Anfang bei der Einarbeitung unterstützt haben. Vielen Dank auch an Christina, Anna, Thomas, Jacki, Nicola, Miguel, Bin, Helen, Steffen, Loriana und Federica für die angenehme Arbeitsatmosphäre und die großartige gemeinsame Zeit auch außerhalb des Labors. Ein besonderes Dankeschön geht an Federica! Thanks for your ongoing support! We shared the good and the difficult times from the beginning to the end. It was a pleasure to share the office with you.

Vielen Dank an alle Freunde die immer ein offenes Ohr hatten. Ganz besonders möchte ich mich bei meiner Familie bedanken. Vielen Dank für eure Unterstützung in allen Lebenslagen, euer Interesse an meinem Projekt und einfach dafür, dass ihr immer für mich da seid. Danke an Philipp für deine emotionale Unterstützung und dein Verständnis, wenn „die Arbeit“ mal wieder Vorrang hatte. Ohne euch wäre dieses Projekt in dieser Form nicht möglich gewesen.