UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

BIOCHEMICAL CHARACTERIZATION OF SPORULATION-RELATED HISTIDINE KINASES IN *CLOSTRIDIOIDES DIFFICILE*

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfilment of the requirements for the

Degree of

MASTER OF SCIENCE

By

KRITI SHUKLA

Norman, Oklahoma

2022

BIOCHEMICAL CHARACTERIZATION OF SPORULATION-RELATED HISTIDINE KINASES IN *CLOSTRIDIOIDES DIFFICILE*

A THESIS APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

BY THE COMMITTEE CONSISTING OF

Dr. Ann West, Chair

Dr. Christina Bourne

Dr. Elena Zgurskaya

Acknowledgements

I would like to extend my gratitude towards the University of Oklahoma Department of Chemistry and Biochemistry. This thesis would not have been possible without the inspiration and support of several wonderful individuals- my thanks and appreciation to all of them for being part of this journey. I owe my deepest gratitude to my supervisor Dr. Ann West and the members of my graduate committee, for their continuous support and encouragement that helped me arrive at this culminating work.

I have benefited from and thoroughly enjoyed- the camaraderie among the researchers in Dr. West's lab, for which I owe thanks to Dr. Ann West, Dr. Smita Menon, Dr. Fabiola Janiak-Spens, Dr. Skyler Hebdon, Dr. Megan Kempher and Savannah Morris.

I must extend that gratitude to everyone on the second floor, East wing of the Stephenson Life Sciences building- GBRA lab, Bourne lab and Rajan lab. The open lab setup certainly aided in clear and smooth communication. I highly appreciate the support and assistance from Dr. Philip Bourne (Protein Purification Core). All the endless protein purification would not have been possible otherwise.

I want to thank Price Family Foundation (AHW, GBRA and EAK), the Oklahoma Center for the Advancement of Science and Technology (HR18-110, AHW), Grayce B. Kerr Endowment funds (AHW), Centers of Biomedical Research Excellence (COBRE, P20GM103640) and Research Corporation for Science Advancement (27333, Cottrell Seed) for funding this work.

Finally, my sincere gratitude to my family for their unparalleled love and support. I am grateful to my sister and my best friend, Shruti Shukla, for always being there for me. I am forever indebted to my parents for everything. They selflessly encouraged me to explore new directions in life and seek my own destiny. This journey would not have been possible if not for them, and I dedicate this milestone to them.

Table of Contents

Acknowledgements ..... **iv**
List of Tables ..... **vi**
List of Figures ..... **vii**
List of Abbreviations ..... **ix**
Abstract ..... **x**

Chapter 1: General Introduction: *C. difficile* and Histidine Kinases **1**
Research Scope 6

Chapter 2: Biochemical Characterization of HK_1587 ..... **7**
Overview **7**
Results **8**
Overexpression of sensor domain of HK_1587 8
Oligomeric form of cytoplasmic HK_1587 10
Autophosphorylation activity of HK_1587^{CA} 14
Phosphotransfer from HK_1587 to the cognate partner RR_1586 16
Phosphatase activity of HK_1587^{CA} towards RR_1586~P 18
Discussion **21**
Methods **21**
Overexpression and purification of HK_1587 constructs 21
Western Blotting 22
Size-exclusion chromatography-multi-angle light scattering (SEC-MALS) 23
Autophosphorylation, phosphotransfer, and phosphatase assays 23
Quantitation of phosphorylated proteins 24

Chapter 3: Biochemical Characterization of Orphan HKs: CD1492 and CD2492 ..... **25**
Overview **25**
Results **26**
Overexpression of the catalytic and PAS domains of CD1492 and CD2492 26
Functional activities of CD1492^{CA} and CD2492^{CA} towards Spo0A 28
Sensor molecules for CD1492 and CD2492 33
Discussion **36**
Methods **37**
Overexpression and purification of CD1492 and CD2492 constructs 37
Autophosphorylation, phosphotransfer, and phosphatase assays 38
Fluorescent thermal shift assays 40

Chapter 4: Future directions **41**
HK_1587-RR_1586 41
Orphan histidine kinases for Spo0A 41

Supplementary Information **43**

References Cited **50**

List of Tables

Table 2.1 Constructs designed for soluble expression of the putative sensor domain of HK_1587 9

Table 3.1 List of molecules tested against $CD1492^{PAS}$ and $CD2492^{PAS}$ in a thermal shift assay 33

Table S1 List of plasmids used for protein expression in this study..... 43

Table S2 List of primers used in this study 43

List of Figures

Figure 1.1 Domain arrangement in a dimeric histidine kinase 3

Figure 1.2 Structures of histidine kinase domains 4

Figure 2.1 Domain organization of HK_1587 8

Figure 2.2 SDS-PAGE for protein expression and solubility studies of the various HK_1587 N-terminal sensor domain constructs, HK_1587^{SD} 10

Figure 2.3 Cytoplasmic HK_1587 constructs designed to study its functional activity 11

Figure 2.4 Expression and purification profile for NΔ1-171 12

Figure 2.5 GST-column affinity chromatography for all three cytoplasmic constructs of HK_1587 13

Figure 2.6 Oligomeric state of HK_1587^{CA} analyzed by SEC-MALS 14

Figure 2.7 Autophosphorylation activity of HK_1587^{CA} 15

Figure 2.8 Autophosphorylation of HK_1587^{CA} and its phosphotransfer activity towards cognate partner RR_1586 17

Figure 2.9 Phosphostability and phosphatase assay for RR_1586 19

Figure 2.10 Quantification of RR_1586~P phosphostability and phosphatase activity of HK_1587 20

Figure 3.1 Domain organization and different constructs of CD1492 and CD2492 26

Figure 3.2 Elution profiles of $CD1492^{PAS}$ (A) and $CD1492^{CA}$ (B) from S200 size-exclusion column 27

Figure 3.3 Elution profiles of $CD2492^{PAS}$ (A) and $CD2492^{CA}$ (B) from S200 size-exclusion column 28

Figure 3.4 Autophosphorylation activity of $CD1492^{CA}$ and $CD2492^{CA}$ orphan kinases 29

Figure 3.5 Phosphotransfer activity of $CD1492^{CA}$ and $CD2492^{CA}$ towards Spo0A 30

Figure 3.6 Phosphatase assay for $CD1492^{CA}$ and $CD2492^{CA}$ 32

Figure 3.7 Protein melt curves from thermal shift assay for $CD2492^{PAS}$ 34

Figure 3.8 Protein melt curves from thermal shift assay for $CD1492^{PAS}$ with cholic acid 36

Figure S1 Expression study for full-length HK_1587 in different cell lines 44

Figure S2 Expression and affinity column purification of cytoplasmic His-tagged HK_1587 in pET-MHL15 45

Figure S3 Autophosphorylation activity for CD1492^{CA} and CD2492^{CA} tested in different buffer conditions 46

Figure S4 Protein melts from thermal shift assay for CD2492^{PAS} 47

Figure S5 Protein melts from thermal shift assay for CD1492^{PAS} 49

List of Abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
CA	Catalytic ATP-binding
Ci	Curie
COBRE	Center of Biomedical Research Excellence
DHp	Dimerization and histidine phosphotransfer
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FAD	Flavin adenine dinucleotide
GST	Glutathione S-transferase
HK	Histidine kinase
HPt	Histidine phosphotransfer
IPTG	Isopropyl β- d-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kDa	kiloDaltons
LB	Luria broth
MALS	Multiple angle light scattering
NAD	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NP-40	Nonyl phenoxypolyethoxylethanol-40
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PMSF	Phenylmethylsulphonyl fluoride
RNA	Ribonucleic acid
RR	Response regulator
SD	Sensor domain
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TCS	Two-component system
T_m	Melting temperature

Abstract

Histidine kinases are one of the two proteins that constitute a two-component signal transduction system. These systems are used by bacteria, fungi, and several plants to sense the conditions of their environment and alter their behavior to ensure survival. In bacteria, two-component systems control basic cellular processes like motility, virulence, cell division, sporulation etc. While histidine kinases are responsible for signal perception in the host, the other protein of this system, called response regulator, conducts the response output. These response outputs can be downstream gene regulation via DNA binding, RNA binding, enzymatic reactions etc. Although a significant number of histidine kinases have been identified for *Clostridioides difficile* till date, it is not yet clear how these histidine kinases affect the pathogenicity of this bacterium, especially sporulation, a process where bacterial cell goes into a dormant form until placed in a favorable environment. The precise prediction of progression mechanism of sporulation adopted by this bacterium could enable researchers to develop techniques to mitigate this pathogenic behavior. In general, these two-component systems (TCS) are absent in mammals and hence can be a good target for therapeutics for several bacterial and fungal infections.

This thesis presents a detailed script of studies done towards unraveling the mechanism behind sporulation of *Clostridioides difficile.* This includes using γ-32P ATP functional assay to provide quantitative and qualitative evidence of functional activities of sporulation-related histidine kinases as well as initial binding screens to examine the impetus behind these phosphorelay systems. Histidine kinases covered in this study are HK_1587 from the hypervirulent R20291 strain and CD1492 and CD2492 from the historic CD630 strain of *C. difficile*.

Chapter 1: General Introduction: *Clostridioides difficile* and Histidine Kinases

Clostridioides difficile is an obligate anaerobic, spore-forming, gastrointestinal pathogen. It colonizes the lower intestinal tract and causes symptoms that can range from mild diarrhea to fatal pseudomembranous colitis.[1] It is responsible for over 12% of hospital-acquired infections, mainly in developed countries.[2] These infections often require disruption of the healthy intestinal microbiome, for example through antibiotic treatment, colonization with *C. difficile*, and release of toxins that cause mucosal inflammation and damage. The bacterial colonization occurs by the fecal-oral route. When anaerobic requirements are not met, the bacterium goes into a dormant spore and uses this morphotype to transmit itself. Since spores are metabolically dormant, they are intrinsically resistant to antibiotics[3] and attacks from our immune systems.[4] Once shed into the environment, spores are also resistant to regular disinfectants used in healthcare settings.[5] Sporulation is crucial to *C. difficil*e pathogenesis and, unlike *Bacillus subtilis*, the sporulation mechanism in *C. difficile* is not very well defined in the literature. Its spore proteome is also poorly conserved as compared to members of the *Bacillus* genus.[6] Currently used treatments rely on only a few broad-spectrum antibiotics, which aggravate the microbial imbalance in the gastrointestinal tract leaving the patient prone to re-infection.[7] *C. difficile* has become a serious problem due to the emergence of the hypervirulent strain- PCR ribotype 027 R20291[8] (CDR20291) and the ever-increasing incidences of antibiotic resistance.[9,10] Thus, there is a critical need to develop novel therapeutic means that will require a detailed knowledge of crucial cellular processes in the bacterium, of which signaling and communication with the outside environment are of prime importance. Along the same lines, two-component systems (TCS) present suitable targets for new therapeutics. Especially the TCS involved in bacterial sporulation[11], which is an absolute requirement for *C. difficile* pathogenesis.

Two-component systems (TCS) are protein systems commonly found in bacteria and are also reported to be present in fungi and few higher plants.[12] They help these organisms in carrying out cellular functions through adaptation in response to environmental changes. They

present a good target for the development of novel antibiotics and/or other therapeutic means because of their conspicuous absence in humans and other metazoans.[13,14] TCS consists of two proteins, namely a histidine kinase (HK) and a response regulator (RR). HKs are a class of typically transmembrane proteins that allow the bacteria to appraise environmental stimuli such as nutrients, quorum signals, antibiotics, osmotic stress along with others. These signals bring a conformational change in the protein's structure stabilizing either the active or inactive form of the protein. TCS signaling has a general mechanism, whereupon sensing certain environmental cue(s), an activated HK autophosphorylates on a conserved histidine residue. Subsequent phosphoryl-group transfer occurs from this histidine residue of the HK to a conserved aspartate residue in the active site of the RR. Once this negatively charged group gets incorporated in the RR, it triggers a conformational change in the structure of the protein, which in turn modifies its biological activity. Hence, the chemical information in the form of a phosphoryl group gets translated into a biological response in the cell that helps the microbe to respond to environmental changes. This biological response occurs mainly by the RR acting as a transcription factor and altering the expression of target genes.[15]

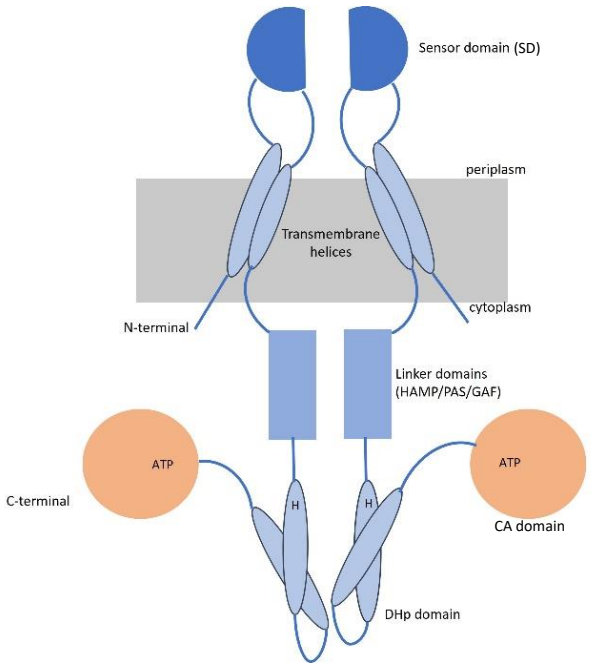


Figure 1.1. Domain arrangement in a dimeric histidine kinase.

Histidine kinases are multidomain, dimeric proteins, which, depending upon the specific function, can be either completely cytoplasmic or membrane bound (Figure 1.1). A typical dimeric HK consists of an N-terminal sensor domain (SD) with α-helices embedded in the membrane. These α-helices are missing in the cytoplasmic HKs. This N-terminal domain arrangement is followed by a linker region which consists of HAMP/PAS/GAF domains. The C-terminal HK module includes the dimerization and histidine phosphotransfer (DHp), and the catalytic adenosine triphosphate (ATP) binding (CA) domains (Figure.1.1). Based on the domain architectures of the histidine kinases, TCSs can be classified into three classes.[16] Class I HKs consist of an N-terminal extracellular stimulus-specific sensor domain and a C-terminal HK module that includes the DHp and catalytic ATP binding domains (Ex: WalK, Figure 1.2B).[17] Class II HKs contain an N-terminal histidine-containing phosphotransfer (HPt) domain and a C-terminal HK module and are specific for chemotaxis[17,18] (Ex: CheA). Class III HKs have features of both class I and class II HKs, with the class I like sensor and HPt domain and class II like C-

terminal HK module.[19] Class III is the most unexplored family of signal transduction histidine kinases and, until to date, are only identified in a relatively short list of bacterial species. The key residues necessary for activity and the 3D-structures of the different domains of histidine kinases are largely conserved. However, recent studies have revealed considerable diversity in the molecular mechanisms used by these proteins to communicate (mode of autophosphorylation, phosphotransfer, signal transduction mechanism etc.).[20]

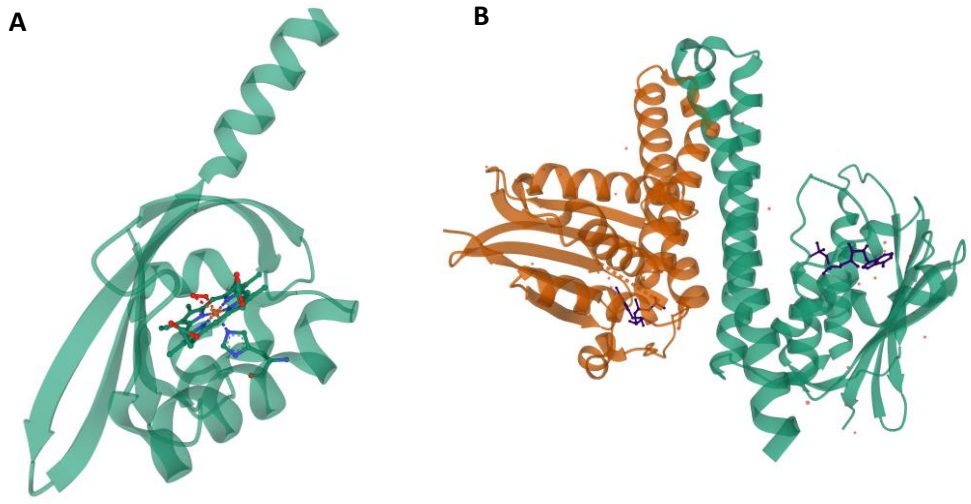


Figure 1.2. Structures of histidine kinase domains. A. Monomeric sensor domain: FixL heme domain is shown bound with oxygen (red) (PDB: 1DP6) and **B. Dimeric C-terminal DHp and kinase domain:** active histidine kinase showing interaction of ATP molecule (blue) with the key residues of C-terminal catalytic domain of WalK (PDB: 4U7O).

Histidine kinases depend on their sensor domains to perceive the environment around them. These sensor domains are not very well conserved across these proteins, which is justified by the diverse signals that they are meant to detect. For example, a cytoplasmic histidine kinase in *Bradyrhizobium japonicum,* FixL has a sensor domain in which oxygen binds directly to a heme moiety that is coordinated to a histidine residue within the PAS domain

(Figure 1.2A).[21] On the other hand, the sporulation kinase KinA in *Bacillus subtilis* has three PAS domains in the N-terminal, but the stimuli sensed by the protein is still unknown.[22] Hence these sensor domains which have a very low sequence conservation and high input specificity may be exploited to develop inhibitors of specific TCS signaling pathways. This makes them potential drug targets.

The decision to enter sporulation in several *Bacilli* and *Clostridia* species is regulated by histidine kinases that have the ability to phosphorylate the master transcriptional regulator Spo0A.[23] Once Spo0A is activated via phosphorylation, it triggers the transcription of a cascade of genes required to go from vegetative state to mother cell and then to endospore. The mother cell is then lysed to release the spore, which can persist indefinitely in unfavorable environments. The *C. difficile* strain 630 genome encodes five orphan histidine kinases, namely CD1352, CD1492, CD1579, CD1949, and CD2492.[24] These histidine kinases share sequence similarity to the kinases in *Bacillus subtilis* which are responsible for initiating sporulation by activating Spo0A through a multistep phosphorelay. To date, very little is known about their exact role in sporulation in *C. difficile*. Childress et al. showed that deletion of CD1492 and CD2492 resulted in an increase in sporulation frequency[25] indicating that these two kinases might possess phosphatase activity towards Spo0A. Underwood et al. showed that inactivation of the kinase CD2492 reduced *C. difficile* spore formation 3.5-fold, suggesting its crucial role in sporulation mechanism.[24] When comparing the expression of genes of *C. difficile* recovered from monoxenic mouse ceca, Jnoir et al. found that these putative orphan histidine kinases were not modulated 4 hours post-infection or later, suggesting that the genes upregulate at an earlier time point or remain constant during the course of infection.[26] *In vitro,* CD1579 has been shown to be the only orphan histidine kinase that is able to autophosphorylate and transfer a phosphoryl group directly to Spo0A.[24] Other than these sporulation related histidine kinases, *C. difficile* possesses some uncharacterized TCSs which may have a link to its pathogenicity through sporulation. There have been studies showing that RR_1586, a response regulator

(RR) in a TCS of *C. difficile* strain R20291 is involved in sporulation and germination.[7,27] Recently, the West lab published an *in vitro* study in which the DNA-binding recognition motif of RR_1586 was identified upstream of putative target genes, notably genes encoding for ion transporters, regulators of sporulation and spore structural proteins.[28] However, very little is known about the role and environmental stimuli of its cognate histidine kinase, HK_1587.

Research Scope

This study is focused on the histidine kinases of *C. difficile* (strains R20291 and CD630), which directly or indirectly regulate sporulation in this bacterium. The overall goal of this study was to provide insights into the gaps in the sporulation pathway of *C. difficile,* particularly with respect to sensor histidine kinases that are involved in regulation of sporulation. Chapter 2 provides biochemical characterization of the autophosphorylation, phosphotransfer and phosphatase activities of the histidine kinase, HK_1587, and its cognate partner, response regulator RR_1586. An insight into the functional activities of HK_1587 will help us discern the way it regulates the activity of RR_1586. Chapter 3 presents an account of the *in vitro* quantification of functional activities of two of the five orphan histidine kinases in *C. difficile*, CD1492 and CD2492, towards their predicted cognate RR, Spo0A. This chapter also proposes plausible binding molecules for the PAS (Per-Arnt-Sim) molecular sensor domains of these proteins. These sensor molecules are hypothesized to regulate sporulation in the bacterium, either independently or in conjunction with each other. Finally, Chapter 4 presents the future directions that should be taken to understand in more detail sporulation regulatory mechanisms adopted by this intricate pathogen.

Chapter 2: Biochemical Characterization of HK_1587

Overview

The intricate balance of autophosphorylation, phosphotransfer and phosphatase activity of a histidine kinase is necessary to modulate the levels of activated response regulator available in the system at any given time[29] which, in turn, regulates the physiological output in the bacterial cell. Histidine kinases can sense a range of specific environmental cues, which include but are not limited to- nutrient limitation, pH change, osmotic stress etc. Once sensed by the HK, these cues trigger autophosphorylation and subsequent phosphotransfer to the cognate RR, which results in a specific cellular output. The environmental signals that initiate major physiological changes, like sporulation, are currently unknown in *C. difficile*. A previous study by Hebdon et al. has predicted that RR_1586 might be regulating various sporulation and germination-associated genes in *C. difficile*.[28] Moreover, a recent study from our lab shows that a *RR_1586* gene knockout in *C. difficile* results in 75% sporulation frequency as compared to 20% in wild-type (data unpublished). In a transcriptome analysis of non-epidemic *C. difficile* strain CD630, HK_1587 has been shown to be down-regulated after 8 hours of infection along with other virulence-associated genes.[30] Hence, HK_1587-RR_1586 presents a pathogenically important TCS in *C. difficile* and exploring the unknowns related to its functioning can prove advantageous to understand this complex pathogen.

In this chapter, I present my findings regarding functional activities of HK1587. I expressed and purified the cytoplasmic (HK_1587^{CA}) and sensor domain (HK_1587^{SD}) of this histidine kinase since soluble expression of this full-length transmembrane protein was arduous and inefficient.

I employed a γ-32P ATP functional assay[31] for studying autophosphorylation, phosphotransfer and phosphatase activity of HK_1587. This chapter provides quantitative *in vitro* evidence of mechanistic activities of the HK_1587-RR_1586 TCS. This study was carried out with an objective to understand the molecular level functioning of both kinase-activated and

kinase-inactivated protein to gain insights into the signaling mechanism. The chapter also reports the experimental progress made towards identifying the sensor molecule/input in the extracellular environment of HK_1587 that enables the protein to regulate its cognate response regulator, RR_1586, through transmembrane signaling.

Results

Overexpression of sensor domain of HK_1587

To identify the stimuli that activate the RR_1586-HK_1587 TCS, I undertook the recombinant protein expression of the N-terminal sensor domain (SD) of HK_1587 (HK_1587^{SD}) in *E. coli*. Several constructs of the predicted sensor domain (residues 38-147, Figure 2.1) with different affinity tags were constructed (Table 2.1).

Figure 2.1. Domain organization of HK_1587. HK_1587 is a 410-amino acid long protein: predicted transmembrane domains are in orange, the extracellular sensor domain is in green, and the cytoplasmic domain(s) are in blue. The sensor domain corresponds to the extracellular region from Arg38 to Gln147.

Table 2.1. Constructs designed for soluble expression of the putative sensor domain of HK_1587.

Construct No.	HK_1587 Amino acids	Construct Name	Affinity tag	Expression level	Solubility
1	38-147	HK_1587^{SD}-His	C-terminal His tag	No expression	Insoluble
2	40-145	HK_1587$^{SD40-145}$-His	C-terminal His tag; N-terminal truncation was done at Leu40	Good	Insoluble
3	38-147	HK_1587^{SD}-GSlinker-GST	A flexible GS linker inserted between sensor domain and GST-tag	Very good	Insoluble
4	38-147	HK_1587^{SD}-GST	C-terminal GST tag	Very good	Insoluble

Results from the expression study are shown in Figure 2.2. Fused with a His-tag, protein constructs 1 and 2 (Table 2.1) had very low expression and/or were insoluble in nature and hence were not pursued further. The sensor domain was fused with a GST-tag at the C-terminus (construct 4, Table 2.1) and although the level of protein expression was high, solubility for these proteins was very low to almost negligible (Figure 2.2A, insoluble pellet). To improve solubility, another construct was made (construct 3, Table 2.1) with a short GS-linker region, A$(GGGGS)_3$A, inserted between the affinity-tag and the predicted sensor domain. However, this approach like previous trials was also unsuccessful (Figure 2.2B, insoluble pellet). Given the high protein expression of the sensor domain constructs, the possibility of protein aggregation and formation of inclusion bodies was explored. I employed several ionic detergents and different concentrations of urea to solubilize these inclusion bodies (data not shown) and was ultimately able to successfully solubilize the protein (construct 4, Table 2.1) from inclusion bodies using 0.2% N-Lauroylsarcosine (Figure 2.2C). Using dialysis, buffer

exchange was performed later to remove the detergent. Dialyzed protein was subjected to a Superdex 200 column for further purification. HK_1587^{SD}-GST will be referred to as HK_1587^{SD} in the subsequent text.

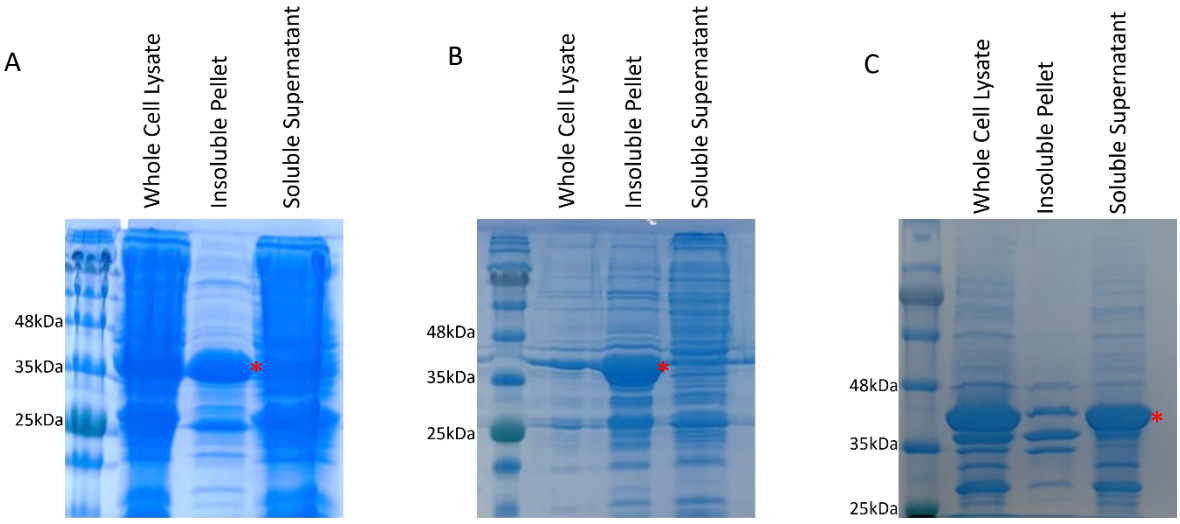


Figure 2.2. SDS-PAGE for protein expression and solubility studies of the various HK_1587 N-terminal sensor domain constructs, HK_1587^{SD}. A. A 12% acrylamide gel showing expression of HK_1587^{SD} (residues 38-147, construct 4, Table 2.1). **B.** HK_1587^{SD}-GSlinker-GST (construct 3, Table 2.1). **C.** SDS-PAGE showing solubilization of HK_1587^{SD} (construct 4, Table 2.1) inclusion body using 0.2% N-Lauroylsarcosine. Expected molecular weight of protein is 39 kDa (panels A and C) and 41 kDa (panel B). Red asterisks indicate band of predicted molecular weight that corresponds to the GST-sensor domain fusion proteins.

Oligomeric form of cytoplasmic HK_1587

A full-length HK_1587 construct with N-terminal His tag (in pET28a) was designed in the earlier stage of this study. The protein was induced using 0.5 mM IPTG in different cell lines, but no expression was observed (Figure S1). Therefore, I designed several constructs of cytoplasmic

HK_1587 by removing the N-terminus (Figure 2.3) and fused them with an N-terminal His-tag in pET-MHL15 and an N-terminal GST-tag in pGEX-KG (to improve solubility). Three different truncations, very close to each other, were made right after the membrane-bound alpha helix to see which one gives the highest expression and/or solubility. The pET-MHL15 constructs expressed well but as shown by the western blot, the His-tagged protein was present in inclusion bodies (Figure S2). Hence, further studies were done only with the GST-tagged constructs in pGEX-KG.

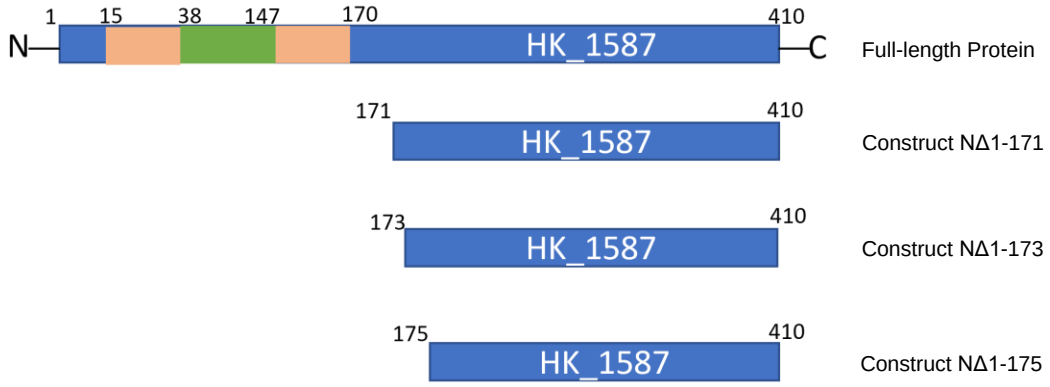


Figure 2.3. Cytoplasmic HK_1587 constructs designed to study its functional activity. Three different truncations, very close to each other, were made right after the membrane-bound alpha helix to see which one provides highest expression and/or solubility. All constructs were fused with GST-tag at the N-terminal. Orange: transmembrane regions, green: extracellular sensor domain, blue: cytoplasmic domain.

All these three constructs (NΔ1-171, NΔ1-173 and NΔ1-175) expressed very well and were also soluble. As an example, expression and purification is shown for NΔ1-171 (Figure 2.4). NΔ1-171 was chosen for further studies since it was the longest construct with slightly higher purification yields as compared to other two (Figure 2.5, yield reported is average of

three experimental replicates). The construct NΔ1-171 will be referred to as HK_1587^{CA} in subsequent text in this document.

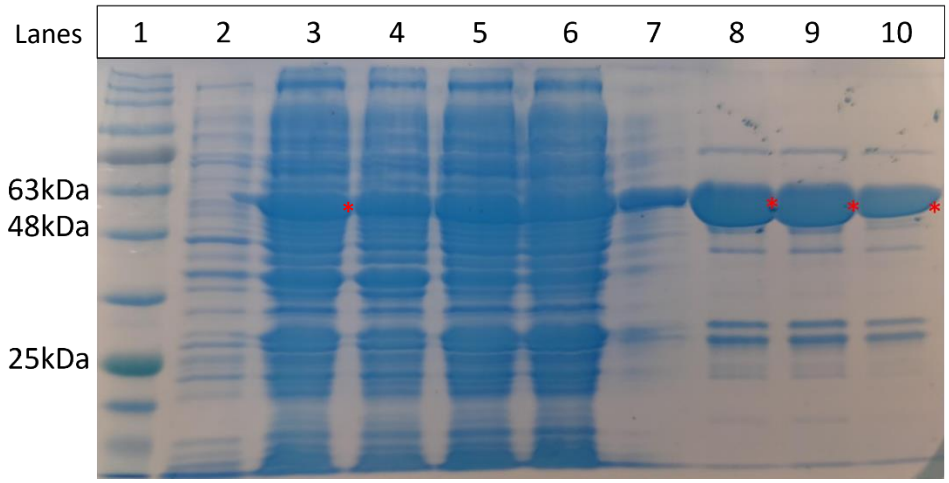


Figure 2.4. Expression and purification profile for NΔ1-171. Lane 1 shows the protein ladder. Lane 2: uninduced cells; Lane 3: IPTG induced whole cell lysate after sonication; Lane 4: insoluble cell pellet after sonication; Lane 5: total soluble protein; Lane 6: flow-through from GST-column; Lane 7: wash flow-through; Lane 8-10: elutions from GST-column using 20 mM reduced glutathione. Expected size of the construct NΔ1-171 is 57 kDa (Red asterisks).

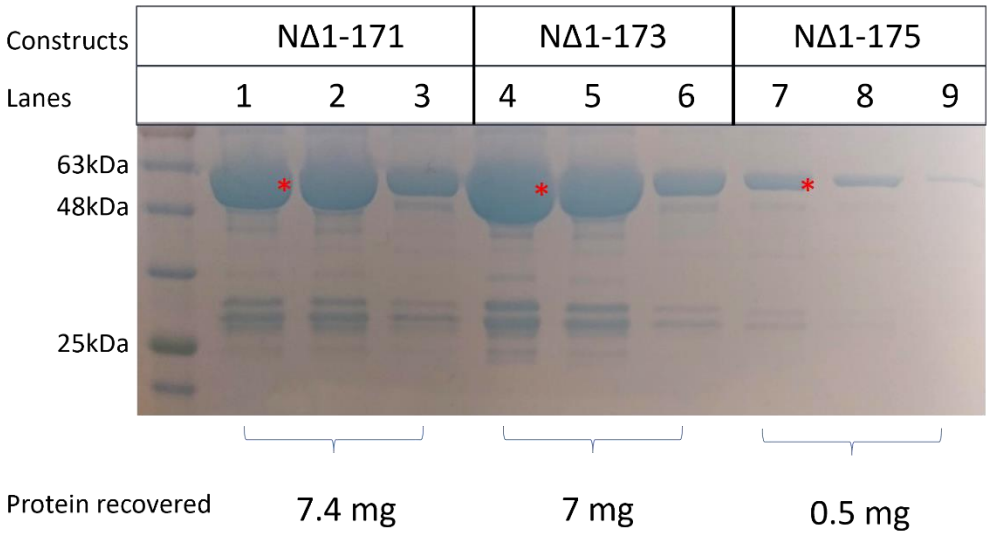


Figure 2.5. GST-column affinity chromatography for all three cytoplasmic constructs of HK_1587. The three grouped lanes (Lanes 1-3: NΔ1-171, Lanes 4-6: NΔ1-173, Lanes 7-9: NΔ1-175) show three subsequent elutions of proteins from the column using 20 mM reduced glutathione. NΔ1-171 had the highest yield (7.4 mg per g of pellet). Red asterisks indicate band of predicted molecular weight that corresponds to the GST-fused cytoplasmic constructs of HK_1587.

The oligomeric state of cytoplasmic HK_1587, HK_1587^{CA}, was explored by size-exclusion chromatography-multi-angle light scattering (SEC-MALS). A typical histidine kinase is dimeric in nature and the HK_1587^{CA} that we expressed for this study was also found to exist as a dimer in solution state, using SEC-MALS (Figure 2.6). The apparent molecular mass of this recombinant protein was found to be 107.7 kDa whereas the monomeric protein is expected to have a molecular mass of 54 kDa. This confirms the structural stability of the truncation we created in the protein sequence to express the cytoplasmic domain and provides confidence in the results from subsequent activity assays.

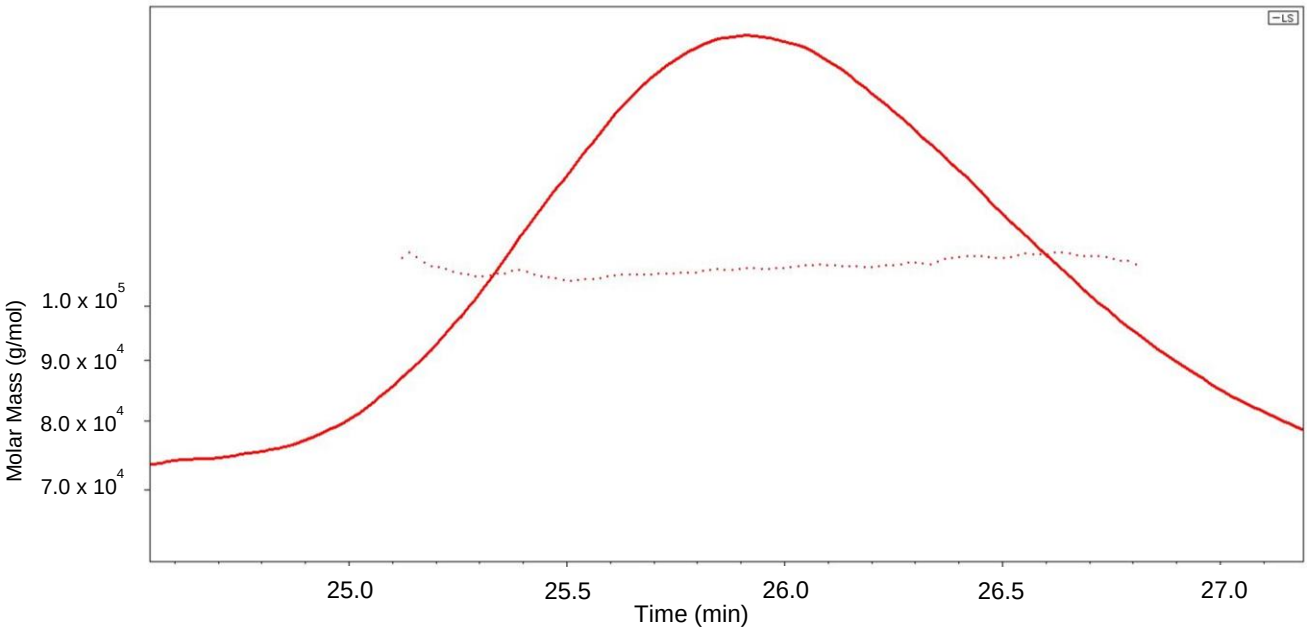


Figure. 2.6. Oligomeric state of HK_1587^{CA} analyzed by SEC-MALS. SEC elution curve (A_{280} curve, solid line) and light-scattering (dotted line) are shown. The experiment was run in triplicate and average molecular mass was found to be 107.7 (±1.8%) kDa. Monomeric HK_1587^{CA} is expected to be 54 kDa.

Autophosphorylation activity of HK_1587^{CA}

To be activated, a histidine kinase must undergo stimulus-induced autophosphorylation where the histidine residue gets phosphorylated in presence of ATP. An activated HK regulates its cognate RR by transferring the phosphoryl group from its histidine residue to the aspartate residue of the RR, hence showing the phosphotransfer towards the cognate RR.[32,33] When autophosphorylation activity is diminished, some histidine kinases can act as a phosphatase towards their cognate response regulators, thus contributing to shutting down the regulated pathway.[21] Theoretically, HK_1587 is expected to have these three functional activities *in vivo* and *in vitro* but this remains to be established. A biological response from the response regulator (typically, regulation of gene expression by DNA-binding) is an outcome of a balance between the phosphotransfer and phosphatase activity of the histidine kinase. Characterizing

these functional activities for HK_1587 would provide insights into regulation of this TCS, which seems to be involved in sporulation, germination and virulence of the bacterium.[7,27,30] Expression of soluble, active transmembrane proteins is highly challenging, so instead I prepared recombinant GST-tagged cytoplasmic *C. difficile* HK_1587 with very high purity (Figure 2.5). Kinase activities were found using these cytoplasmic constructs, thus indicating that the N-terminal transmembrane domains are not essential for it to function as a kinase. Autophosphorylation was assayed by incubating the purified HK_1587^{CA} with radiolabeled ATP mixture and phosphorylated protein could be detected as early as 5 min post-ATP addition which subsequently increased reaching a maximum at 45 min (Figure 2.7). This time-dependent autophosphorylation activity was observed in presence of Mn^{2+} ions. Other ions, including Mg^{2+} which has been found to aid kinase activity[34,35] in other histidine kinases, hindered the autophosphorylation in HK_1587^{CA} (data not shown). Using bioinformatics, the residue that gets phosphorylated in HK_1587^{CA} was found to be H194.

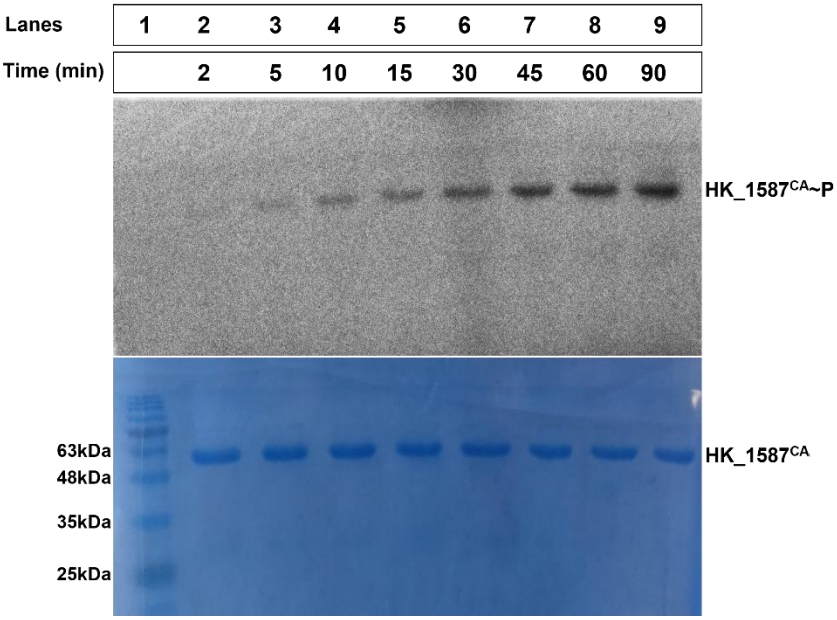


Figure 2.7. Autophosphorylation activity of HK_1587^{CA}. Top panel: Radiograph showing phosphorylation signals saturating at 45 min after HK_1587^{CA} was incubated with γ-32P ATP at

23 °C. **Bottom panel:** Coomassie blue protein stain to indicate equal amounts of sample loads. Lane 1 shows the protein marker for protein sizes.

Phosphotransfer from HK_1587 to the cognate partner RR_1586

Similar to autophosphorylation, the biochemical assays designed to examine phosphotransfer activity were carried out using the truncated HK_1587^{CA}. Phosphotransfer was assessed by incubating RR_1586 together with HK_1587^{CA}~P (Figure 2.8), where HK_1587^{CA} was phosphorylated using radiolabeled ATP mixture as described above. Expected results for phosphotransfer activity were to see a decrease in the HK_1587$^{\mathbf{CA}}$~P band signal and a simultaneous increase in RR_1586~P signal over time. At 60 min post-incubation, the phosphorylation signal for HK_1587$^{\mathbf{CA}}$ was decreased to 40% while RR_1586 saturated with the phosphorylation signal. This residual signal in HK_1587$^{\mathbf{CA}}$ (40% phosphorylation) at the end of phosphotransfer reaction is anticipated to be the result of continued autophosphorylation due to the presence of excess radiolabeled ATP in the reaction. A constant ratio of phosphorylated HK to RR after 45 min of incubation indicates a successful time-dependent phosphotransfer between these two cognate partners.

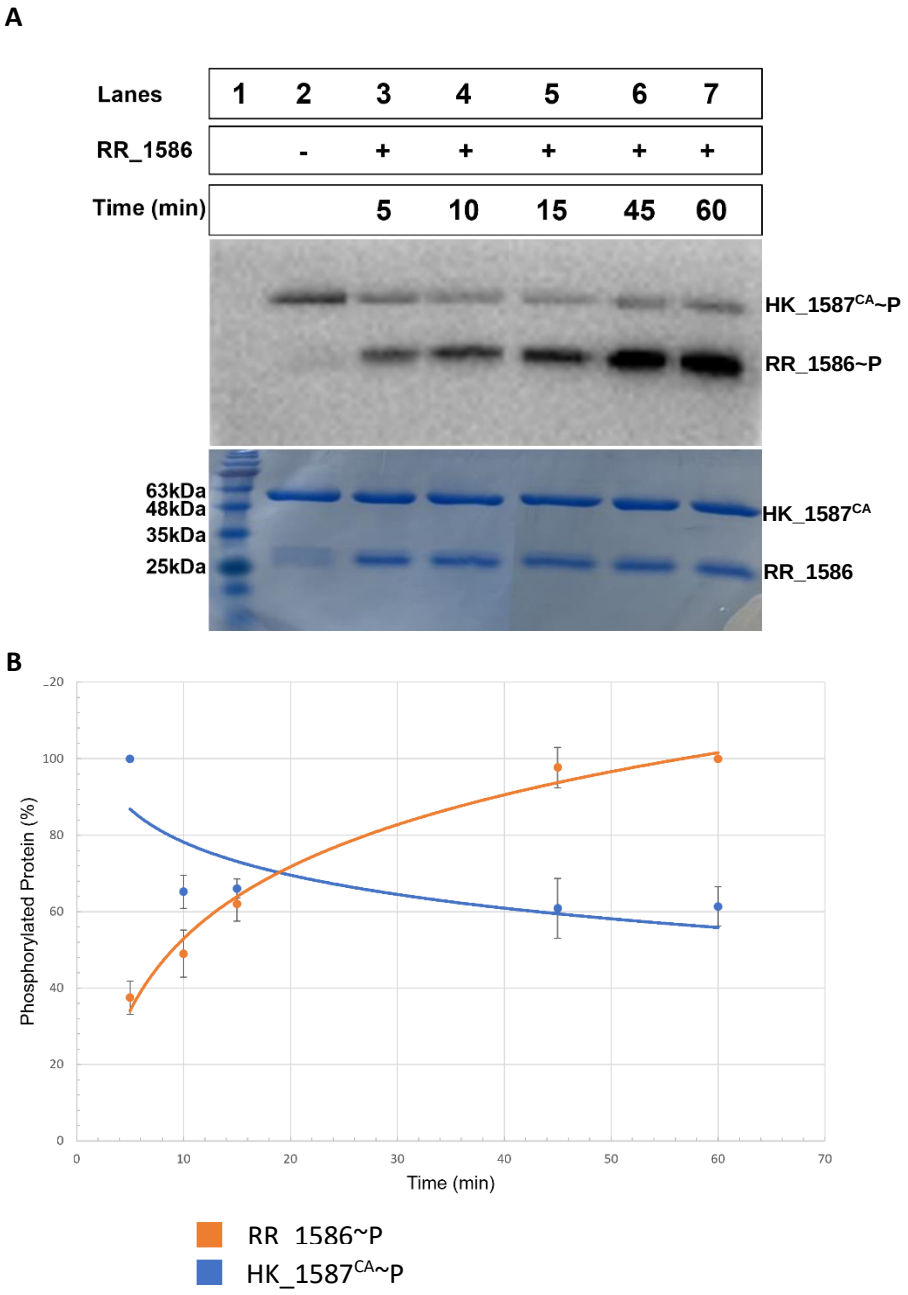


Figure 2.8. A. Autophosphorylation of HK_1587^{CA} and its phosphotransfer activity towards cognate partner RR_1586. Top panel: Radiograph showing phosphorylation signals. Lane 2 shows autophosphorylation of HK_1587^{CA} in absence of RR_1586. Phosphoryl transfer occurs from HK_1587^{CA} to RR_1586 as early as 5 min and saturates around 45 min. **Bottom panel:** Coomassie blue protein stain to indicate equal amounts of sample loads. Expected sizes are: HK_1587^{CA} – 57 kDa and RR_1586- 28 kDa. Lane 1 shows the marker for protein sizes. **B. Quantification of phoshorylated proteins during *in vitro* phosphotransfer reaction.** The

plot shows increasing amounts of phosphorylated RR_1586 (RR_1586~P, orange) when incubated with HK_1587^{CA}~P and a simultaneous decrease in the amounts of HK_1587^{CA}~P (blue) with increasing time. Lane 2 (Figure 2.8A) depicts the saturated HK_1587^{CA}~P signal which was set as 100% and Lane 7 (Figure 2.8A) depicts the saturated RR_1586~P signal which was set as 100% for the quantitation using ImageJ.

Phosphatase activity of HK_1587^{CA} towards RR_1586~P

Many of the sensor histidine kinases display both kinase and phosphatase activities toward their cognate response regulators. While kinase activity leads to transfer of phosphoryl groups from histidine residue of HK to the aspartate residue of RR, phosphatase activity simply removes the phosphoryl group from the aspartate without transferring it back to the histidine. Amino acid sequence analysis has revealed a conserved putative phosphatase motif E/DxxT/N in these kinases where the threonine residue is crucial for the phosphatase activity.[36] Although the motif is known, the molecular mechanisms regulating these opposing activities of these enzymes are not well understood. In HK_1587, the motif 'ELRT' is present adjacent to the catalytic histidine, H194, suggesting that HK_1587 may possess phosphatase activity. To investigate phosphatase activity of HK_1587^{CA} towards RR_1586~P, I used γ-32P ATP assay where RR_1586~P was incubated in the presence and absence of HK_1587^{CA}. If HK_1587^{CA} acts as a phosphatase toward RR_1586~P, I expected the phosphorylation signal in RR to decrease in presence of HK whereas it would remain constant (or decrease very slowly) in the absence of HK, over time.

Hence, as a control, a phosphostability experiment was carried out to measure the half-life of phosphorylated RR_1586 in the absence of HK_1587^{CA} (Figure 2.9A). For the phosphatase assay, HK_1587^{CA} was added to purified, ^{32}P-labelled RR_1586~P (Figure 2.9B) and the phosphatase activity was measured by the percentage of RR_1586~P remaining in the reaction mixture as a function of reaction time. The half-life of RR_1586~P in the presence of

dimeric HK_1587^{CA} was approximately 17 min as compared to half-life of 145 min in the absence of cognate histidine kinase (Figure 2.10).

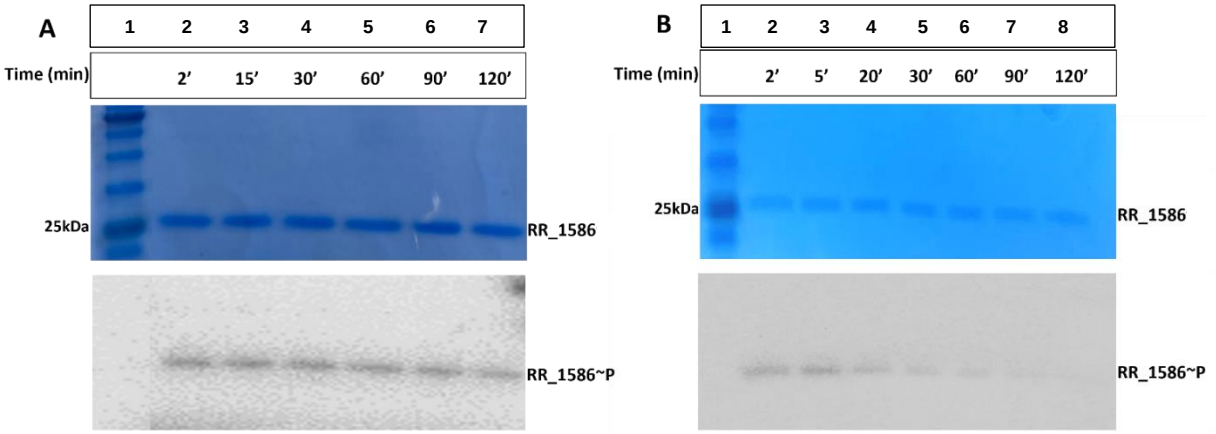


Figure 2.9. Phosphostability and phosphatase assay for RR_1586. A. Stable signal for phosphorylated RR_1586 (RR_1586~P, bottom panel) in the absence of HK monitored over a period of 2 hours (top panel is Coomassie stained gel and shows even load). **B.** RR_1586~P stability in presence of HK (bottom panel); majority of the signal disappears within 20 min in the presence of HK_1587^{CA} (top panel is Coomassie stained gel and shows even load). Lane 2 (Figure 2.9A and B) depicts the saturated RR_1586~P signal which was set as 100% for the quantitation using ImageJ.

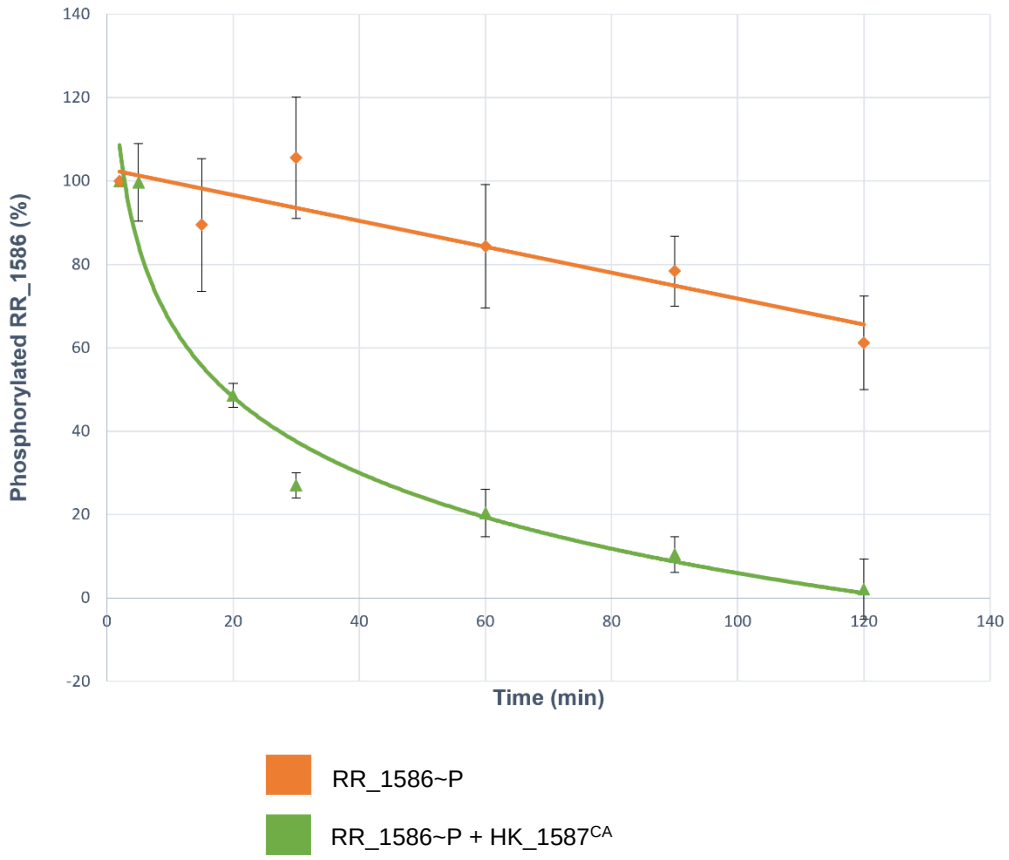


Figure 2.10. Quantification of RR_1586~P phosphostability and phosphatase activity of HK_1587. A typical experimental plot that shows stability of phosphorylated RR_1586 in the absence (orange) and presence (blue) of HK_1587^{CA}. Half-life of RR_1586~P in absence of HK was 145 ± 28 min and in the presence of HK_1587^{CA} was 17.6 ± 3.3 min. Averages are calculated from three replicates of the phosphatase activity assay. The phosphorylation levels for RR_1586 were calculated by setting the band intensity in lanes 2 (saturated RR_1586~P signal, Figure 2.9A and B) as 100% and then comparing the band intensities in lanes 3-7 (Figure 2.9A) or lanes 3-8 (Figure 2.9B) to it.

Discussion

This chapter provided the results of the studies of the functional activities[32] exhibited by the histidine kinase HK_1587 towards its predicted cognate response regulator, RR_1586 *in vitro.* The constructed kinase domain was shown to be responsible for activity. It was well-folded and dimeric in solution, as shown by SEC-MALS which provides more confidence in the results of ^{32}P activity assays. This study provided qualitative and quantitative evidence of functional activities that are important in regulation of a two-component system. HK_1587 was able to autophosphorylate as early as 5 min of incubation with ATP, which is typical for known histidine kinases.[16]. The transfer of phosphoryl groups from HK_1587 to RR_1586 was shown to occur within 5 min which provides evidence for these proteins to be cognate partners.[33] This study also demonstrated that HK_1587 acts as a phosphatase towards RR_1586, since the half-life of phosphorylated RR_1586 decreased from 145 ± 28 min to 17.6 ± 3.3 min in the presence of HK_1587^{CA}. This finding suggests that HK_1587 can modulate the levels of phosphorylated, active form of RR_1586 to regulate gene expression and cellular response.

Methods

Overexpression and purification of HK_1587 constructs

Full-length His-tagged HK_1587 construct in Pet28a was expressed using common methods in cell lines- Rosetta, PlysS, Gold and RIL. Several different constructs of cytoplasmic C-terminal and sensor N-terminal HK_1587 were designed by truncating the protein at its transmembrane region. These truncated genes were cloned using XbaI and XhoI cut sites into a Pgex-KG vector containing a C-terminal GST tag and ampicillin-resistant gene. For soluble expression of sensor domain of HK_1587, a short flexible GS-linker, $A(GGGGS)_3A$ was also used between the protein of interest and GST tag (detailed list of PCR primers and plasmids is provided under supplementary information, Table S1 and Table S2). The plasmid containing the truncated HK_1587 was transformed into *E. coli* DH5α strain and the correct clone was confirmed through

sequencing. Recombinant plasmids were transformed into *E. coli* BL21(DE3) chemically competent cells. Transformed bacterial colonies, grown on LB agar with ampicillin, were used to inoculate 20 mL starter cultures in Luria broth (LB) growth media[37] with 0.1 mg mL^{-1} ampicillin as the selectable marker. Each culture was grown overnight at 37 °C before adding it to a 1 L of LB media with ampicillin which was then grown at 37 °C to an OD600 of ~ 0.55. These cultures were induced for protein expression with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Bacterial cultures were then grown for a further 20 h at 16 °C. Cells were harvested by centrifugation at 6220 x g for 30 min at 4 °C. The cell pellet was resuspended in a lysis buffer containing 150 mM NaCl, 50 mM Tris-HCl pH 8, 0.05% NP-40 and freshly added 0.5 mM PMSF and 25 µg/ml lysozyme. The cells were then lysed with sonication (Sonifier250, Branson). The lysate was separated by centrifugation at 27,000 x g for 50 min at 4 °C. The supernatant was incubated with glutathione agarose beads (Mclab, equilibrated with 50 mM Tris pH 8 and 150 mM NaCl) on a nutator for 30 min at 4 °C and then applied to a gravity flow chromatography column. The flow-through and buffer wash (50 mM Tris pH 8 and 150 mM NaCl) were collected for SDS-PAGE analysis and GST-tagged proteins were eluted with elution buffer (50 mM Tris pH 8, 150 mM NaCl and 10 mM reduced glutathione). The purified proteins were frozen at -80 °C after adding glycerol to a final concentration of 20% (v/v), for subsequent experiments.

Western Blotting

A western blot was performed for the cytoplasmic HK_1587 construct with an N-terminal His-tag in pET-MHL15 to check the expression and solubility. A pre-poured, stain-free 12% SDS-PAGE gel (BioRad) was run at 200 V for 35 min at room temperature with whole cell lysate, insoluble cell pellet, soluble total protein, Ni^{2+}-NTA column elutions and purified RR_1586 (His-tagged) as a control. Protein bands were visualized using stain-free imaging where the gel was illuminated under 300 nm UV transilluminator to produce fluorescent protein bands (red shift). The protein samples were transferred to a nitrocellulose membrane. The membrane was blocked using 1%

milk in TBST (Tris-buffered saline with 0.1% Tween-20) and probed with Alexa Fluor™ 488 anti-His antibody (ThermoFisher) for 30 min at room temperature. The His-tagged proteins were visualized using a ChemiDoc imaging system.

Size-exclusion chromatography-multi-angle light scattering (SEC-MALS)

The oligomeric state of cytoplasmic HK_1587 was determined by subjecting the purified protein to Superdex 200 column chromatography (ÄKTA), equilibrated with 20 mM HEPES pH 7.5 and 150 mM NaCl. The purified protein was subjected to MiniDawn Treos (Wyatt) multiangle light-scattering instrument to measure the molar masses of purified HK_1587.

Autophosphorylation, phosphotransfer, and phosphatase assays

Autophosphorylation of cytoplasmic HK_1587 protein was monitored by radioactivity of phosphorylated HK_1587 using γ-32P ATP (PerkinElmer). Reaction mixtures contained 5 µM of HK_1587^{CA} in autophosphorylation buffer (50 mM Tris-HCl pH 8, 100 mM KCl, 10 mM $MnCl_2$, 5% (v/v) glycerol and 1 mM dithiothreitol). The reactions were started with the addition of the mixture of cold ATP (10 µM) and γ-32P ATP (30 µCi) followed by incubation of the reaction mixture at room temperature. Equal amounts of reaction samples were collected at different time points and the reaction was stopped by adding SDS-PAGE 4X sample buffer (250 mM Tris-HCl pH 6.8, 8% SDS, 20% β-mercaptoethanol, 0.008 % bromophenol blue, and 40% glycerol) containing 40 mM EDTA to a final concentration of 1X and were placed on ice. The protein samples were separated on SDS-PAGE using a 12% acrylamide gel. The ^{32}P-labeled protein bands were visualized by autoradiography (Typhoon FLA 9500) and the amount of radioactive phosphorylated protein was measured by quantifying the bands using ImageJ. The same gel was then stained using Coomassie dye for visualizing total protein present.

For testing the phosphotransfer activity, HK_1587^{CA} was autophosphorylated for 30 min and then equimolar amount of RR_1586 and $MgCl_2$ to a final concentration of 10 mM each were added to the reaction. This phosphotransfer reaction was incubated at room temperature and samples were retrieved at different timepoints. The reaction was stopped with the addition of 4X SDS-PAGE loading buffer containing 40 mM EDTA to a final concentration of 1X. The samples were kept on ice until they were fractionated by SDS-PAGE. The same gel was then stained using Coomassie dye for visualizing total protein present.

Phosphatase activity of HK_1587^{CA} was monitored via the dephosphorylation of RR_1586~P by unphosphorylated HK_1587^{CA}. RR_1586 was phosphorylated using HK_1587^{CA} as described above. Once phosphorylation of RR_1586 was saturated (after 45 min incubation), glutathione agarose beads were added to the reaction to remove HK_1587^{CA}~P from the mixture. The supernatant, containing RR-1586~P, was run through a BioSpin-30 (Biorad) column to remove excess radiolabeled and cold ATP. Unphosphorylated HK_1587 (adsorbed on the glutathione beads) was added to the eluent containing RR-1586~P. Aliquots were removed at specified times and the reaction was stopped by the addition of 4X SDS-PAGE sample loading buffer containing 40 mM EDTA to a final concentration of 1X. The samples were kept on ice until they were fractionated by SDS-PAGE.

Quantitation of phosphorylated proteins

The protein quantitation for phosphostability of RR_1586~P, phosphotransfer activity of HK_1587~P and phosphatase activity of HK_1587~P was done by using ImageJ. 32-bit, high-quality image was obtained for all the concerned autoradiographs. The band intensity of saturated phosphorylation signal was set as 100% and then rest of the bands corresponding to different time points were compared to this saturated band intensity. These comparisons were then reported in % phosphorylated over time.

Chapter 3: Biochemical Characterization of Orphan HKs: CD1492 and CD2492

Overview

The master regulator for sporulation in both *Bacillus subtilis* and *Clostridioides difficile* is the DNA-binding response regulator protein Spo0A.[38] Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay[39] that recruits a series of histidine kinases. KinA-E directly phosphorylate Spo0F which subsequently transfers phosphoryl groups to Spo0A through an additional phosphotransferase, Spo0B.[39] In *C. difficile,* despite the conservation of Spo0A, the absence of Spo0F and Spo0B suggests that there is a direct phosphorylation and dephosphorylation of Spo0A by certain upstream histidine kinase(s) which respond to specific internal or environmental signals.[40] This feature of direct HK-RR communication is representative of the ancestral *Clostridia* before the great oxygen event (2.5 billion years ago), after which additional phosphorelay proteins were evolved in the lineage that led to the *Bacilli*.[41] In the *C. difficile* genome, there are three of the five orphan histidine kinases (CD1492, CD1579 and CD2492) that share some sequence identity with KinA-E.[24] Based on *in vitro* biochemical studies, the cytoplasmic kinase CD1579 is predicted to directly phosphorylate Spo0A.[24] Childress et al. showed that deletion of CD1492 and CD2492 resulted in an increase in sporulation frequency[25] indicating that these two kinases might possess phosphatase activity towards Spo0A. Unlike CD1579, the kinases CD1492 and CD2492 are integral membrane proteins with PAS domains which might be functioning as input modules by sensing oxygen, redox potential, light, or some other stimuli through binding. There are no published studies related to *in vitro* characterization of CD1492 and CD2492. In this chapter, I am providing initial exploration into the functional activities of these proteins towards Spo0A and insights into the sensor molecules that might be the stimuli for these signaling kinases.

Results

Overexpression of the catalytic and PAS domains of CD1492 and CD2492

To establish the functional activities and signaling mechanism of the orphan histidine kinases CD1492 and CD2492 in *C. difficile,* I overexpressed the truncated constructs of these proteins (Figure 3.1) in *E. coli*. Using GST-tags, two constructs were designed for each protein: one containing the PAS domains ($CD1492^{PAS}$ and $CD2492^{PAS}$) and the other including the sequences comprising the catalytic and DHp domains ($CD1492^{CA}$ and $CD2492^{CA}$).

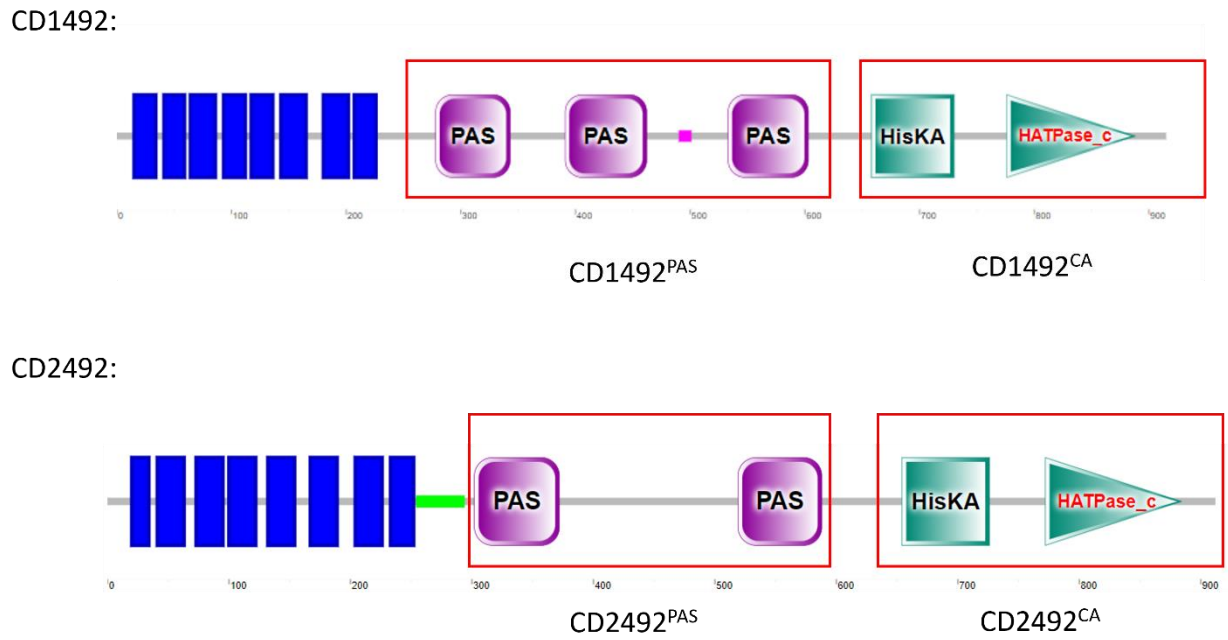


Figure 3.1. Domain organization and different constructs of CD1492 and CD2492. Both proteins contain N-terminal membrane-bound regions shown as blue boxes. CD2492 is predicted to contain a coiled-coil linker region between membrane bound and cytoplasmic part of the protein (light green).

To assess the oligomeric state of the constructs, the purified fusion proteins were loaded onto a calibrated size exclusion column. The PAS domain constructs of CD1492 and CD2492

showed similar elution profiles on a size-exclusion S200 column. The proteins eluted earlier than expected during the run, in fractions 8, 9 and 10 (Figure 3.2 and 3.3) with the peak absorbance at fraction 9, suggesting that the proteins are present in higher oligomeric states; the S200 calibration curve suggests that these proteins are trimeric in solution. SEC-MALS could not be performed with these proteins due to low protein concentrations. The catalytic and dimerization domains of CD1492 and CD2492 also showed elution profiles similar to each other on the size-exclusion S200 column. Unlike the PAS domains, these proteins were spread through multiple fractions (Fractions 9-15) (Figure 3.2 and 3.3) but were much more concentrated and had fewer impurities.

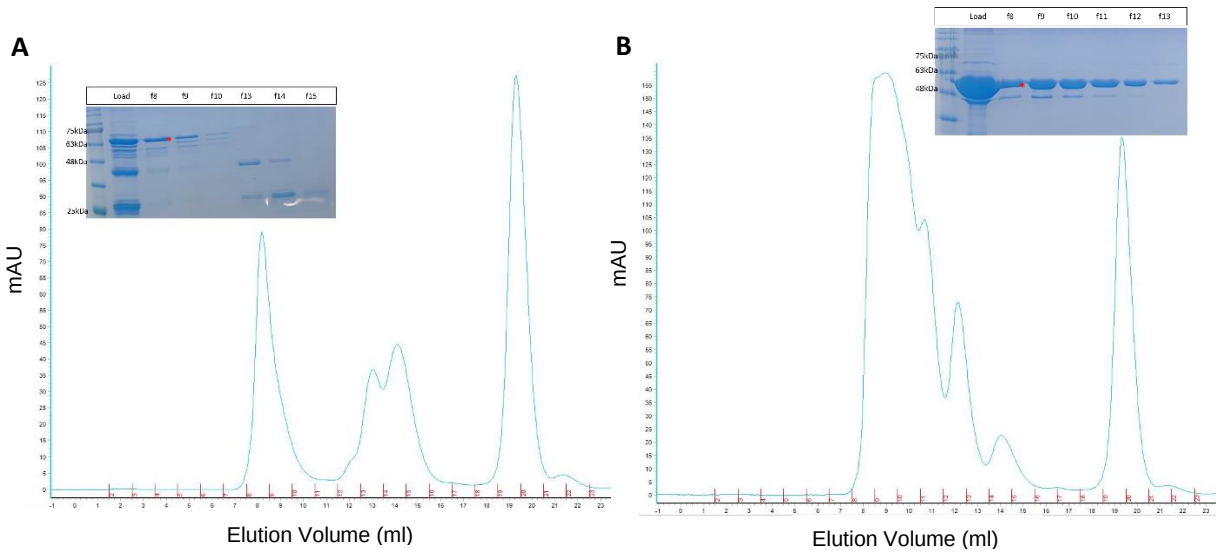


Figure 3.2. Elution profiles of CD1492^{PAS} (A) and CD1492^{CA} (B) from S200 size-exclusion column. The insets show the respective Coomassie protein gels to determine the peaks that contained the protein of interest. Expected sizes of the monomeric fusion proteins are: CD1492^{PAS}- 65 kDa (fractions 8 and 9, panel A) and CD1492^{CA}- 57 kDa (fractions 8-15, panel B). Red asterisks show the bands of correct sizes for these protein constructs.

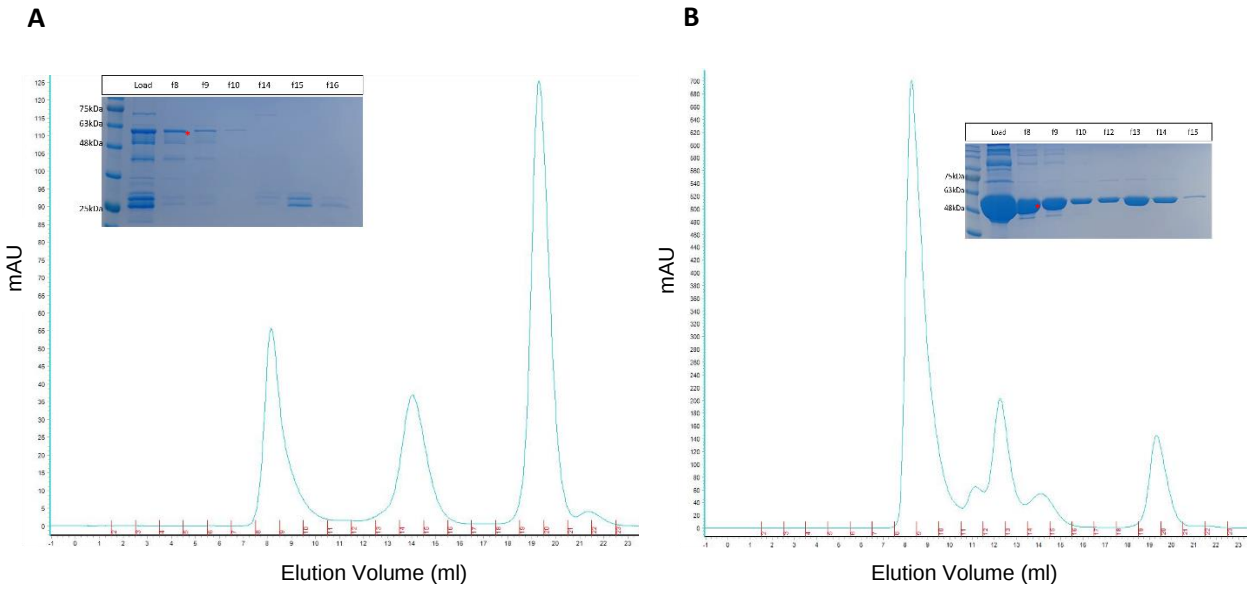


Figure 3.3. Elution profiles of CD2492^{PAS} (A) and CD2492^{CA} (B) from S200 size-exclusion column. The insets show the Coomassie protein gels to determine the peaks that contained the protein of interest. Expected sizes of the monomeric fusion proteins are: CD2492^{PAS}- 60 kDa (fractions 8 and 9, panel A) and CD1492^{CA}- 57 kDa (fractions 8-15, panel B). Red asterisks show the bands of correct sizes for these protein constructs.

Functional activities of CD1492^{CA} and CD2492^{CA} towards Spo0A

I tested the autophosphorylation activity of CD1492^{CA} and CD2492^{CA} using a γ-32P ATP functional assay[31]. Initially, three buffers (50 mM Tris-HCl pH 8, 100 mM KCl, 5% (v/v) glycerol and 1 mM dithiothreitol) were tested containing 10 mM of different metal cations (Buffer 1- K^+, Buffer 2- Mg^{2+}, Buffer 3- Mn^{2+}). The radioactive phosphorylation signal was observed only in the protein sample incubated with Buffer 3 (with Mn^{2+}) (Figure S3) and subsequent experiments were done in Buffer 3. Both CD1492^{CA} and CD2492^{CA} showed a very weak phosphorylation signal starting at 10 min of incubation with the ATP mixture in the presence of Mn^{2+}-containing buffer (Figure 3.4, Lanes 4,5 and 8,9, red asterisks). This time point of *in vitro* autophosphorylation is quite typical of histidine kinases[17], but the signal quality could be

improved upon by optimizing the buffer conditions. This result provides preliminary evidence of autophosphorylation activity of these orphan histidine kinases.

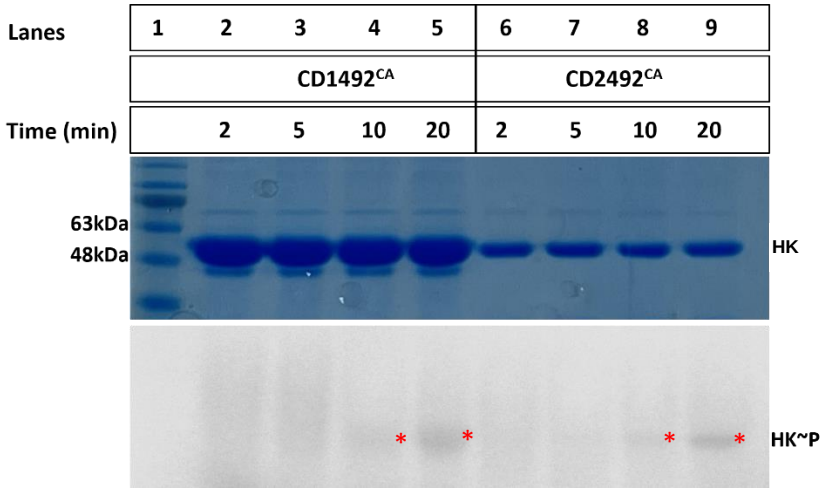


Figure 3.4. Autophosphorylation activity of CD1492^{CA} and CD2492^{CA} orphan kinases. Top panel: Coomassie blue protein stain to indicate equal amounts of sample loads. Lane 1 shows the marker for protein sizes. **Bottom panel:** Radiograph showing phosphorylation signals up to 20 min after CD1492^{CA} and CD2492^{CA} were incubated with γ-32P ATP at room temperature. Red asterisk indicates phosphorylated HKs.

Next, the phosphotransfer activity of these kinases was tested for the predicted cognate response regulator, Spo0A. The catalytic domains were incubated with radioactive and cold ATP for 30 min before adding Spo0A to the reaction mixture. Transfer reactions were stopped after 30 min of incubation and proteins were separated on a 12% SDS-PAGE gel. A very weak autophosphorylation signal was detected for CD1492^{CA} (Figure 3.5, Lane 4, see asterisk). This signal completely disappeared following the 30 min incubation with Spo0A and Spo0A~P could be seen at the same time-point (Figure 3.5, Lane 5). This suggests that, under tested conditions, CD1492 acts as a phosphodonor to Spo0A. Although, as compared to Spo0A, excess amounts of CD1492^{CA} were used and hence this transfer of phosphoryl group might not

occur spontaneously under physiological conditions. A lower size *E. coli* contamination band also appears when CD1492^{CA} was incubated with ATP for 30 min which runs at the same size as Spo0A.

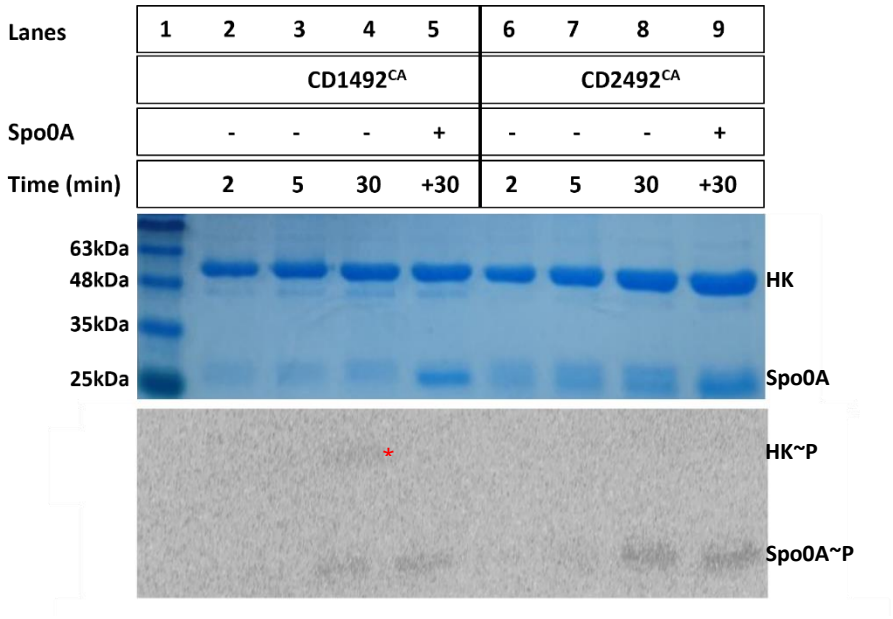


Figure 3.5. Phosphotransfer activity of CD1492^{CA} and CD2492^{CA} towards Spo0A. Top panel: Coomassie blue protein stain to indicate equal amounts of sample loads. Lane 1 shows the marker for protein sizes. **Bottom Panel:** Radiograph showing phosphorylation signals. Autophosphorylation signals for kinases at 2, 5 and 30 min of incubation with ATP, followed by a 30 min incubation with Spo0A.

The results from the corresponding phosphotransfer assay using CD2492^{CA} (expected size- 57kDa) were inconclusive since no signal was detected for autophosphorylation at 30 min (Figure 3.5, Lane 8). Instead, a lower size *E. coli* contamination band appears when CD2492^{CA} was incubated with ATP for 30 min. Hence the band that can be seen in the last lane (Lane 9; where Spo0A was added to the reaction and incubated for extra 30 min) cannot be attributed to Spo0A~P with confidence.

Appearance of these lower size bands in both kinases may indicate proteolytic cleavage of CD1492^{CA} and CD2492^{CA} and this contamination band being the catalytic domain of CD1492 or CD2492 (32kDa) without the GST-tag (25kDa). This might be the other reason of very weak autophosphorylation signals of these proteins as GST-tag can be affecting the proper folding of kinase and/or DHp domains.

Recent experiments done with the histidine kinase CD1579 and Spo0A in our lab showed that CD1579 spontaneously transfers the phosphoryl group from its histidine residue to the aspartate residue of Spo0A. The phosphorylation signal for Spo0A saturates within 15 min of incubation of Spo0A with CD1579 (unpublished data from Dr. Smita Menon, West lab). Taking advantage of this finding, I phosphorylated the Spo0A using CD1579 in appropriate buffer conditions (details in methods section) (Figure 3.6, Lane 2). Once the Spo0A~P signal was saturated, unphosphorylated CD1492^{CA} or CD2492^{CA} were added to the reaction and incubated for up to 45 min, removing aliquots at shown timepoints to test for phosphatase activities that these kinases might possess (Figure 3.6). As shown by the radiograph, under tested conditions, Spo0A~P signal was not decreased with time implying that these specific constructs of kinases do not possess any phosphatase activity. The CD1492^{CA}~P band (Figure 3.6A, Lanes 3-6, red astericks) appears after 5 min of incubation since there is excess ATP present in the reaction mixture, and not because there phosphoryl transfer from Spo0A back to this HK.

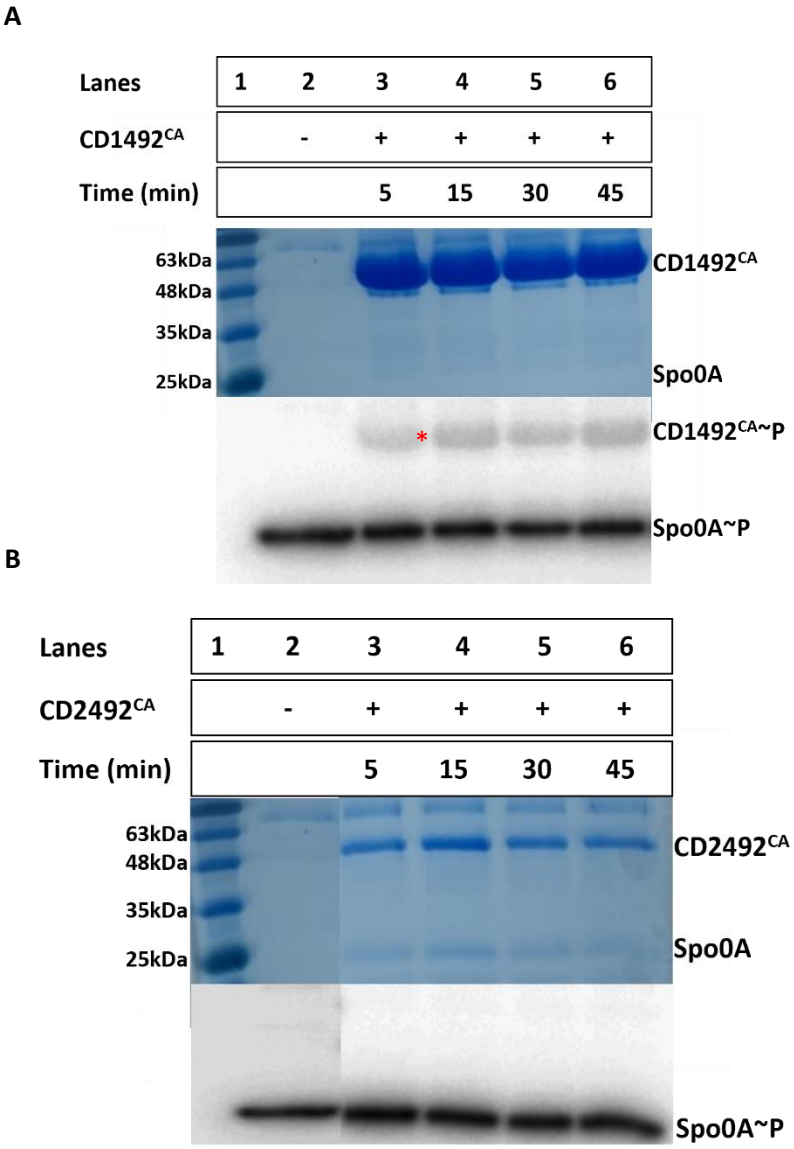


Figure 3.6. Phosphatase assay for CD1492^{CA} (A) and CD2492^{CA} (B). Top Panels: Coomassie blue protein stain to indicate equal amounts of sample loads. Lane 1 shows the marker for protein sizes. Lane 2 has Spo0A phosphorylated using CD1579. **Bottom panels:** Radiograph showing same intensity of Spo0A~P signal both in the absence (Lanes 2) and presence of CD1492^{CA} and CD2492^{CA} (Lanes 3-6).

Sensor molecules for CD1492 and CD2492

Since CD1492 and CD2492 contain multiple cytoplasmic PAS sensor domains, I hypothesized that these proteins are sensing intracellular concentrations of one or more molecules and subsequently regulate the sporulation pathway through Spo0A. To identify the sensor molecule(s)/stimuli for these two orphan histidine kinases, fluorescent thermal shift assays were carried out using SYPRO orange dye. The thermal shift assay measures the thermal stability of a protein in terms of its melting temperature (T_m), which is the temperature at which the protein is 50% denatured. This protein denaturation is monitored via an increase in fluorescence of SYPRO Orange dye which binds to hydrophobic regions that get exposed as the protein unfolds. Any molecule that binds to the target protein, might affect the stability, and hence results in a change in melting temperature. Table 3.1 shows the list of several molecules tested against the PAS domains of CD1492 and CD2492. This list includes sugars, amino acids, and bile salts that have been found to be important for *C. difficile* pathogenesis.[42–44]

Table 3.1. List of molecules tested against CD1492^{PAS} and CD2492^{PAS} in a thermal shift assay.

Saccharides	Amino Acids	Bile Salts/SCFA	Others
• Trehalose • Xylose	• Glycine • Histidine • Cysteine • Proline	• Sodium deoxycholate • Deoxycholic acid • Cholic acid	• ATP • ADP • FAD • NAD • Heme

Out of these 14 ligands, 4 were found to significantly increase the melting temperature[45] (Figure 3.7) of CD2492^{PAS} indicating a tight binding between the protein and the ligand. These three molecules were: ATP, xylose, proline, and glycine. Melt curves with other ligands can be found under supplementary files (Figure S4). Having glycine as one of the potential hits in this

binding screen agrees with the known literature about *C. difficile* pathogenesis. It has been well-established that *C. difficile* spores require an amino acid signal *(*primarily glycine) to trigger the germination process.[46,47] This indicates that CD2492 might be the initiator of this trigger in the bacterium. Further studies to quantify the interaction between these molecules with the protein will be pursued as next steps.

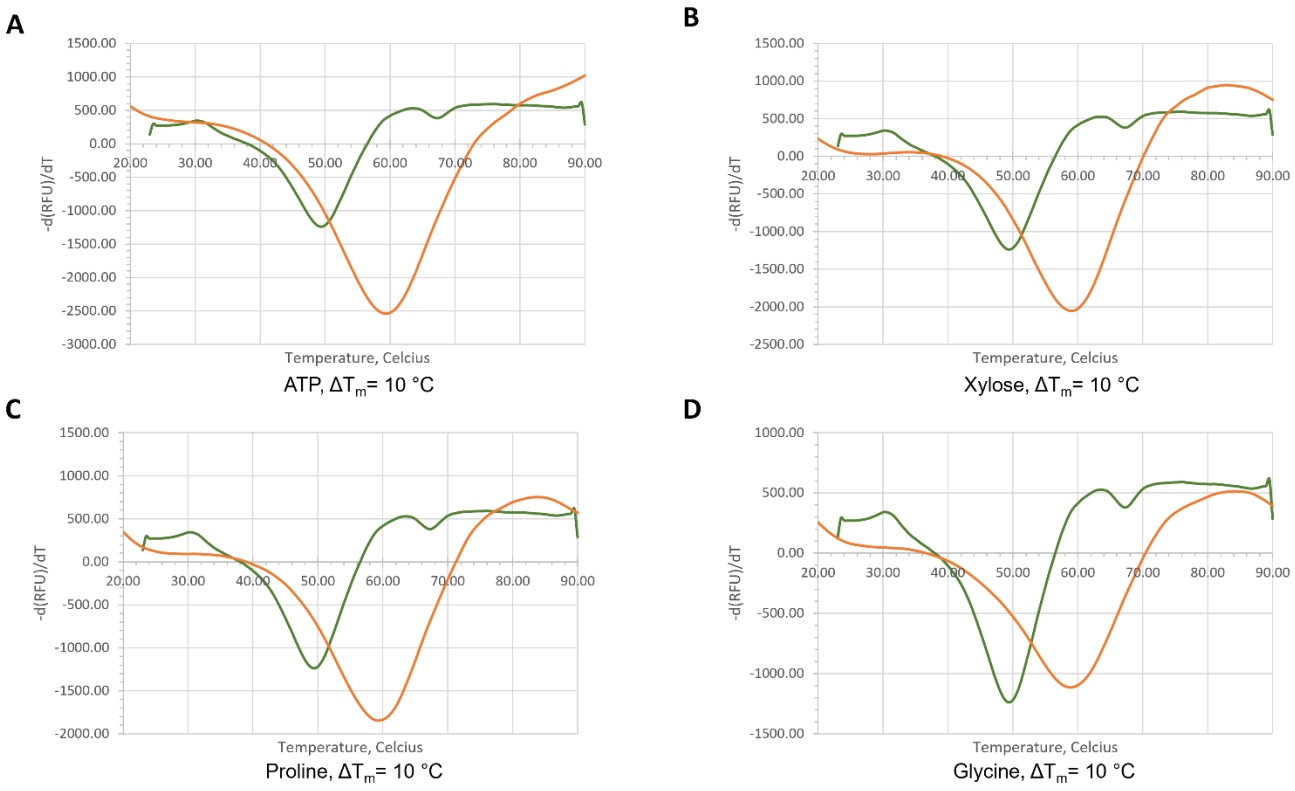


Figure 3.7. Protein melt curves from thermal shift assay for CD2492^{PAS}. The screen showed significant binding between CD2492^{PAS} and **A.** ATP, **B.** xylose, **C.** proline and **D.** glycine. Binding is indicated by a change in the melting temperature of the protein (ΔT_m). Each plots show control melt curve with just protein and SYPRO orange (green) and the melt curve with protein, SYPRO orange and the indicated ligand (orange).

Due to lower purification yields, the thermal shift assays with CD1492^{PAS} were run with only 6 out of 14 ligands listed in the Table 3.1. The molecules included were- xylose, trehalose,

glycine, histidine, proline and ATP. Under the tested concentrations of the protein and these ligands, none of the changes in melting temperature were significant (Figure S5). A second screen was done with using titrated amounts of glycine, xylose, and cholic acid. Cholic acid was included due to the suggested role of primary bile acids in germination/sporulation.[42,47] When different concentrations of these ligands were tested against $CD1492^{PAS}$, only cholic acid showed a significant change in T_m as the ligand concentration was increased from 0.04 mg/ml to 0.32 mg/ml (ΔT_m = 11 °C, Figure 3.8). The critical micelle concentration (CMC) for cholic acid is 12 mM[48] and the highest concentration tested against these proteins was 0.7 mM (0.32 mg/ml) which indicates that there was no denaturation of protein via micelle formation on adding cholic acid. A decrease in T_m was seen which can imply that cholic acid binds preferentially to a less populated conformational state (i.e., a partially unfolded state) of $CD1492^{PAS}$. On the other hand, binding of some small molecules may manifest as decreases in protein stability, especially for very stable proteins.[45] In other cases, a decrease in T_m is observed when binding of a ligand in one region of the protein triggers unfolding in another.[49] The ions and co-solutes used in the buffer conditions also modulate the folding to unfolding equilibria. The precise binding parameters need to be determined for this interaction using a quantitative assay such as isothermal titration calorimetry (ITC).

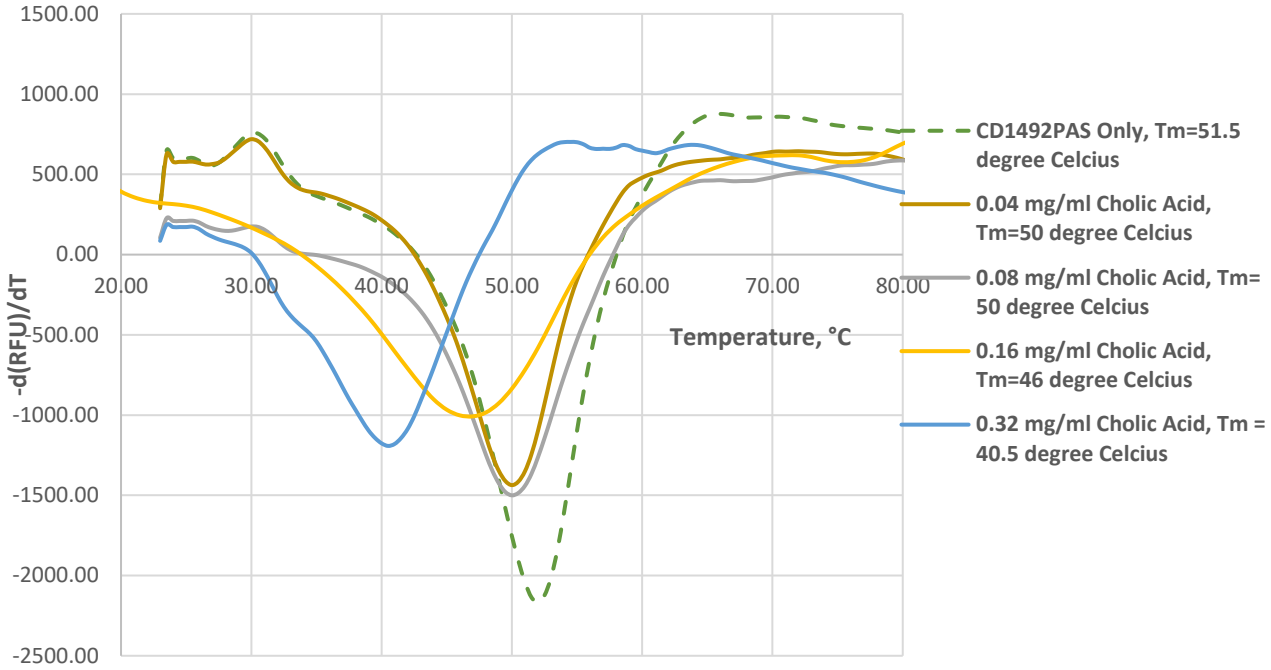


Figure 3.8. Protein melt curves from thermal shift assay for CD1492^{PAS} with cholic acid. The screen shows most significant binding with the highest cholic acid concentration tested (0.32 mg/ml, ΔT_m = 11 °C). Binding is indicated by a change in the melting temperature of the protein (ΔT_m). Control melt curve was run with just protein and SYPRO orange (dotted, green).

Discussion

CD1492 and CD2492 are the sporulation-related orphan histidine kinases in *C. difficile* which are poorly characterized in the literature. Spo0A in *C. difficile* is regulated by one or more of these histidine kinases, either directly or indirectly.[41] This chapter details preliminary studies done in an attempt to characterize these kinases. From the activity assays, these constructs of kinases were shown to autophosphorylate though further experiments may be needed to optimize this activity. Under the conditions given, CD1492 was shown to have phosphodonor ability towards Spo0A and no phosphatase activity (Figures 3.5 and 3.6). This is contradictory to the studies available in the literature[25] and hence more exploration is required to confirm these

results. CD2492 did appear to have phosphotransfer activity, however, low confidence in this conclusion arises from the fact that the signals were quite weak and that only lower molecular weight band showed activity (Figure 3.5). No phosphatase activity was shown by CD2492 under tested conditions (Figure 3.6).

In an attempt to identify the stimuli for these kinases, thermal shift assays were carried out. Out of 6 ligands tested against $CD1492^{PAS}$ in a thermal shift assay, none of the changes in melting temperature were significant (Figure S5). Later, when a titration of ligands was tested against $CD1492^{PAS}$, cholic acid showed a significant change in T_m as the ligand concentration was increased from 0.04 mg/ml to 0.32 mg/ml (ΔT_m = 11 °C, Figure 3.8). This preliminary indication of $CD1492^{PAS}$ binding with cholic acid is interesting since the relationship between *Clostridioides difficile* and primary bile acids is not well understood. *In vitro* studies suggest that the primary bile acids like cholate, taurocholate and glycocholate can stimulate germination of *C. difficile* spores.[50] For $CD2492^{PAS}$, upon screeing the 14 molecules mentioned in Table 3.1, I found ATP, glycine and xylose to affect the melting temperature of the protein significantly. This data is indicative of strong binding between these ligands and $CD2492^{PAS}$ and hence these can be the sensor molecules for this kinase. Subsequent studies are in progress to confirm and quantify the binding shown by the thermal shift assays.

Methods

Overexpression and purification of CD1492 and CD2492 constructs

Truncations were made in the protein sequences of CD1492 and CD2492 to design a PAS domain and a kinase domain construct for both proteins (Figure 3.1). Using XbaI and XhoI sites, the truncated genes were cloned into pGEX-KG vector containing a C-terminal GST tag and ampicillin-resistant gene. The plasmids containing these truncated genes were transformed into *E. coli* DH5α strain and positive constructs were confirmed through sequencing. Recombinant plasmids were transformed into *E. coli* BL21(DE3) cells for protein expression. Transformed

bacterial colonies, grown on LB-agar-ampicillin plates, were used to inoculate 20 mL starter cultures in Luria broth (LB) growth media[37] with 0.1 mg mL^{-1} ampicillin as the selectable marker. Each culture was grown overnight with shaking at 37°C before adding it to a 1 L of LB media with ampicillin which was then grown at 37 °C to an OD600 of ~ 0.55. These cultures were induced for protein expression with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Bacterial cultures were then grown for a further 20 h at 16°C. Cells were harvested by centrifugation at 6220 x g for 30 min at 4 °C. The pellet was then suspended in a buffer containing 150 mM NaCl, 50 mM Tris-HCl pH 8, 0.05% NP-40 and freshly added 0.5 mM PMSF and 25 µg/ml lysozyme. The cells were then lysed by sonication (Sonifier250, Branson). The supernatant was incubated with glutathione agarose beads (MClab, equilibrated with 50 mM Tris pH 8 and 150 mM NaCl) for 30 min at 4 °C and then applied to a gravity flow chromatography column. The flow-through and buffer wash (50 mM Tris pH 8 and 150 mM NaCl) were collected for SDS-PAGE analysis and GST-tagged proteins were eluted with elution buffer (50 mM Tris pH 8, 150 mM NaCl and 10 mM reduced glutathione). These proteins were further purified using a Superdex 200 chromatography column (ÄKTA), equilibrated with 20 mM HEPES pH 7.5 and 150 mM NaCl and purified proteins were stored at 4 °C for subsequent experiments.

Autophosphorylation, phosphotransfer, and phosphatase assays

Autophosphorylation of $CD1492^{CA}$ and $CD2492^{CA}$ was monitored using radioactive ^{32}P-ATP. Reaction mixtures contained 10 µM of $CD1492^{CA}$ or 5 µM of $CD2492^{CA}$ in autophosphorylation Buffer 3 (50 mM Tris-HCl pH 8, 100 mM KCl, 10 mM $MnCl_2$, 5% (v/v) glycerol and 1 mM dithiothreitol). The reactions were started with the addition of the mixture of cold ATP (10 µM) and γ-32P ATP (30 µCi) (PerkinElmer) and the reaction mixture was incubated at room temperature. Equal amounts of reaction samples were collected at different time points and the reaction was stopped by adding SDS-PAGE 4X sample buffer (250 mM Tris-HCl pH 6.8, 8%

SDS, 20% β-mercaptoethanol, 0.008 % bromophenol blue and 40% glycerol) containing 40 mM EDTA to a final concentration of 1X and were placed on ice. The samples were separated on SDS-PAGE using a 12% acrylamide gel. The ^{32}P-labeled protein bands were visualized by autoradiography (Typhoon FLA 9500). The same gel was then stained using Coomassie dye for visualizing total protein present.

For testing the phosphotransfer activity, these catalytic domains were autophosphorylated for 30 min and then equimolar amount of Spo0A was added to the reaction. This phosphotransfer reaction was incubated at room temperature and samples were retrieved at different timepoints by terminating the reaction with the addition of 4X SDS-PAGE loading buffer containing 40 mM EDTA to a final concentration of 1X. The samples were kept on ice until they were fractionated by SDS-PAGE. The ^{32}P-labeled protein bands were visualized by autoradiography (Typhoon FLA 9500). The same gel was then stained using Coomassie dye for visualizing total protein present.

Phosphatase activity of CD1492^{CA} or CD2492^{CA} was monitored via the dephosphorylation of Spo0A~P by unphosphorylated kinases. To phosphorylate Spo0A, 1.4 µM CD1579 was incubated with 20 µM ATP mixture in the phosphorylation buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 15 mM $MgCl_2$, 2 mM dithiothreitol, 5% glycerol) for 30 min at room temperature. Afterwards, 2.5 µM Spo0A was added to this phosphorylated CD1579 and incubated for 30 min. Once phosphorylation of Spo0A was saturated, 2 µM of unphosphorylated CD1492 or 3.5 µM of unphosphorylated CD2492 was added to the reaction. The reactions were stopped using SDS-PAGE sample buffer as described above and the samples were kept on ice. The protein samples were fractionated by SDS-PAGE and the radioactive bands were visualized by autoradiography (Typhoon FLA 9500). The same gel was then stained using Coomassie dye for visualizing total protein present.

Fluorescent thermal shift assays

To measure the melting temperature (T_m) of the proteins with and without the ligands, a semi-skirted 96-well PCR plate (BioRad) was used and 2.5 µM $CD1492^{PAS}$ or 5 µM $CD2492^{PAS}$ was incubated at room temperature with 5X SYPRO orange dye (Invitrogen, 50X stock made in DMSO) in a buffer containing 20 mM HEPES, pH 7.5 and 150 mM NaCl. The ligands mentioned in Table 3.1 were added to a final concentration of 50 mM to the above reaction and the plate was incubated at room temperature for 30 min. The assay plate was placed into the real-time PCR instrument- CFX96 Real Time System (BioRad) and a temperature gradient program (25 °C to 90 °C with an increment of 0.5 °C for 10 sec hold at each temperature) was started for protein thermal denaturation. A no ligand control, protein only, sample was run in one of the wells.

Chapter 4: Future directions

HK_1587-RR_1586

To understand the regulation mechanism of this pathogenically crucial TCS consisting of HK_1587 and RR_1586, an insight into the signal input is necessary[51]. Some of the future directions for this project can involve screening the sensor domain of HK_1587, HK_1587^{SD}, against potential small molecules and carrying out thermal shift binding assays followed by a quantitative analysis of binding with the significant hits using isothermal titration calorimetry (ITC). Structural studies can also be done with the HK_1587^{SD} to elucidate the secondary structure of the protein that will aid in the identification of small molecules that might bind and bring about the conformational change leading to activation of protein. Given their recent reported successes, nanodiscs[52–55] can be employed to purify the full-length transmembrane HK_1587. Having a full-length HK_1587 will enable us to run the HK_1587 activity assays with and without the signal and assess how it affects the regulation of RR_1586. Complete understanding of sporulation mechanism in *C. difficile* requires a detailed and full characterization of sporulation related two-component systems.

Orphan histidine kinases for Spo0A

Since Spo0A is the master sporulation regulator in *C. difficile*, the histidine kinases regulating this transcription factor are of crucial importance when it comes to understanding how this pathogen regulates the sporulation pathway. The caveats presented in the previous chapter can be overcome by either designing better, more active constructs of these orphan histidine kinases or expressing the full-length proteins using nanodiscs. Full-length constructs of CD1492 and CD2492 will enable us to test the functional activities of these proteins in the presence and absence of the ligands/stimulus. Moreover, the autophosphorylation assays should be performed at higher time points (beyond 20 min of incubation) to examine true saturation of phosphorylation signals. To be able to draw conclusions from the phosphotransfer assays,

reaction should be stopped at several time points after adding Spo0A to see if the transfer to Spo0A occurs earlier than 30 min. The ligands which resulted in a significant change in the T_m of the proteins can be tested against the proteins using ITC to quantify the binding which will also provide confidence in the significance of these preliminary findings. Co-crystallizing these PAS domains in the presence of ligands found in this study may also provide unique insight into the stimuli for these proteins.

Supplementary Information

Table S1. List of plasmids used for protein expression in this study

Plasmid	Restriction sites	Protein constructs	N-terminal tag	Source
pET28a	BamHI and XhoI	Full-length HK_1587	His	Novagen # 69864
pET-MHL15	NdeI and BamHI	Cytoplasmic HK_1587	His	Addgene #26092
pGEX	XhoI and XbaI	Cytoplasmic HK_1587	GST	Addgene #77103
pGEX	XhoI and XbaI	All CD1492 and CD2492 constructs	GST	Addgene #77103

Table S2. List of primers used in this study

ID	Sequence	Experiment
pr384_AHW_F	ATGTACGGATCCATGATAAATAAGAATGTATTTACCAGTA	Full-length HK_1587 expression
pr385_AHW_R	TAGCTACTCGAGTTTTAACTTTTTTGGTATTGACAATTCAAATG	Full-length HK_1587 expression
pr386_AHW_F	TGACTGGTGGACAGCAA	pET28a sequencing
pr387_AHW_R	TGTTAGCAGCCGGATCT	pET28a sequencing
pr398_AHW_F	ATGCCGGCCACGATGCGT	pET28a sequencing
pr438_AHW_F	TAGCTTTCTAGA TACTAGAAAAGCTTTAATTC	HK_1587 NΔ1-171
pr407_AHW_R	AAGCTACTCGAG TTATTTTAACTTTTTTG	Cytoplasmic HK1587
pr439_AHW_F	TAGCTTTCTAGA TAAAGCTTTAATTCC	HK_1587 NΔ1-173
pr440_AHW_F	TAGCTTTCTAGA TTTAATTCCAATAGAAAC	HK_1587 NΔ1-175
pr441_AHW_F	TAGCTTTCTAGA TAGGGGACTTACATATAG	HK_1587^{SD}
pr442_AHW_R	AAGCTACTCGAG TTACTGTCTTAAAGAATTC	HK_1587^{SD}

pr487_AHW_F	CGACCATCCTCCAAAATC	pGEX-KG sequencing
pr566_AHW_F	TAGCTTTCTAGA TAAAAACTACATAGACCAGAG	$CD1492^{PAS}$ expression
pr567_AHW_R	AAGCTACTCGAG TTAGTGATTATTTTCATTTACTATATGG	$CD1492^{PAS}$ expression
pr568_AHW_F	TAGCTTTCTAGATTTAAGTGTGCAAAGAGAG	$CD1492^{CA}$ expression
pr569_AHW_R	AAGCTACTCGAG TTATTTATATATATCAAAAAATTCTATCTC	$CD1492^{CA}$ expression
pr570_AHW_F	TAGCTTTCTAGA T AAGGGAATACTTATGAAAAG	$CD2492^{PAS}$ expression
pr571_AHW_R	AAGCTACTCGAG TTACAATTCATATTCCAGTAAAAG	$CD2492^{PAS}$ expression
pr572_AHW_F	TAGCTTTCTAGA T TTAGTACATAAGAAATATTCCG	$CD2492^{CA}$ expression
pr573_AHW_R	AAGCTACTCGAGTTAATAGTATATATCAGA	$CD2492^{CA}$ expression

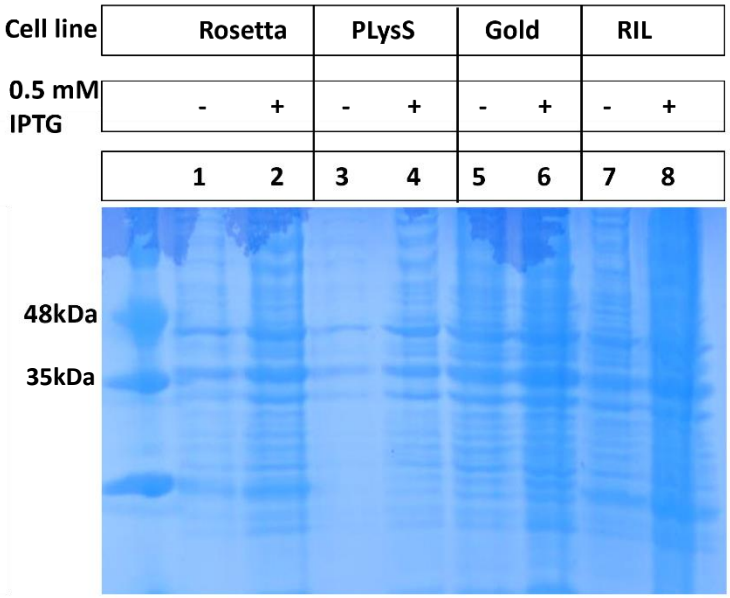


Figure S1. Expression study for full-length HK_1587 in different cell lines. This construct was made in pET28a with an N-terminal His-tag. No expression was observed in tested conditions (0.5 mM IPTG, 16 °C, 20 hours). Expected size is 48 kDa.

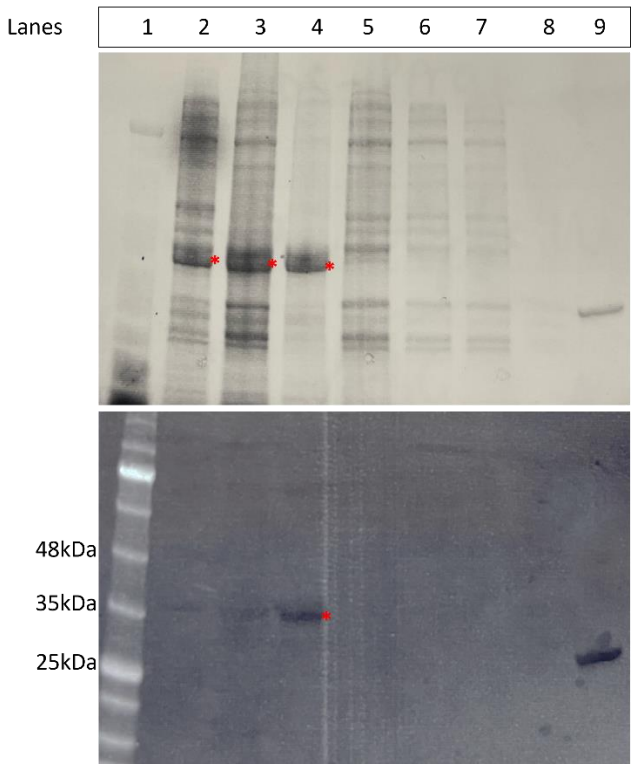


Figure S2. Expression and affinity column purification of cytoplasmic His-tagged HK_1587 in pET-MHL15. Top Panel: stain-free imaging for total protein and **Bottom Panel:** western blot using anti-His alexa488 tagged antibody. Lane 1: Protein ladder; Lane 2: uninduced whole cell; Lane 3: 1 mM IPTG induced whole cell; Lane 4: insoluble cell pellet; Lane 5: soluble total protein; Lane 6-8: three elutions from Ni-NTA column; Lane 9: purified RR_1586 with His-tag (28 kDa) as an anti-His antibody control. Expected size of HK_1587 is 35 kDa (see asterisk).

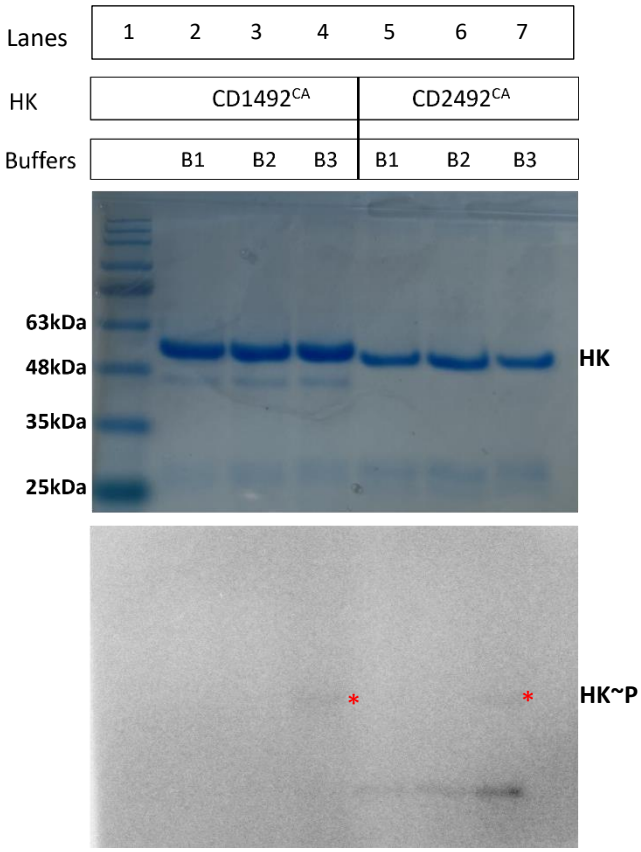
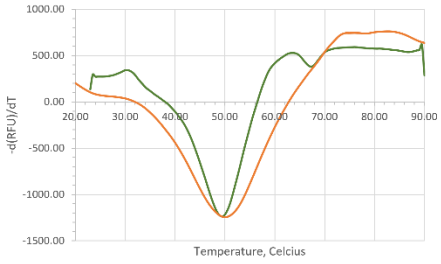
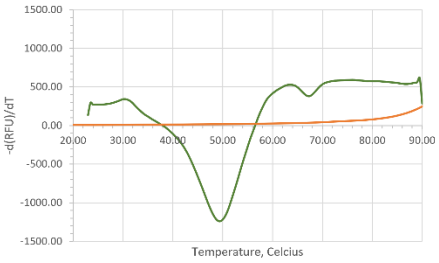


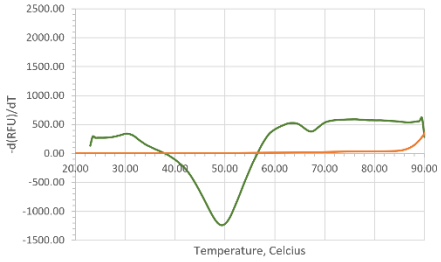
Figure S3. Autophosphorylation activity for CD1492^{CA} and CD2492^{CA} tested in different buffer conditions. Top Panel: Coomassie stain gel showing total protein content. The proteins were incubated with ATP mixture and buffer for 30 min before loading on the gel. Lane 1 shows the protein marker. **Bottom Panel:** Autoradiograph showing the phosphorylated kinase. Reaction buffer was 50 mM Tris pH8, 5% glycerol, 1 mM dithiothreitol containing the following cations B1- 100 mM NaCl, 10 mM $MgCl_2$; B2- 100 mM KCl, 10 mM $MgCl_2$; B3- 100 mM KCl, 10 mM $MnCl_2$.



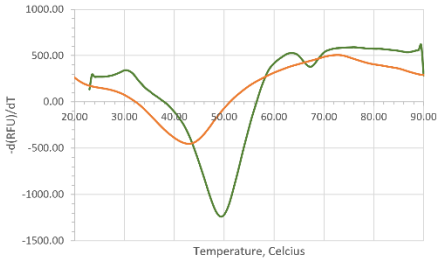
Cholic acid, ΔT_m= 1 °C



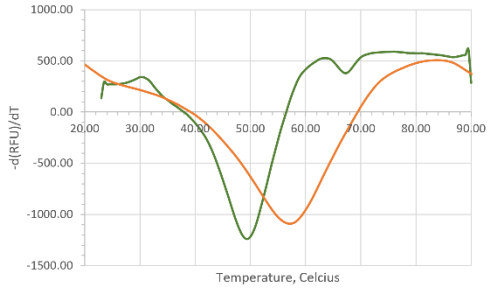
Sodium deoxycholate, ΔT_m= ---°C



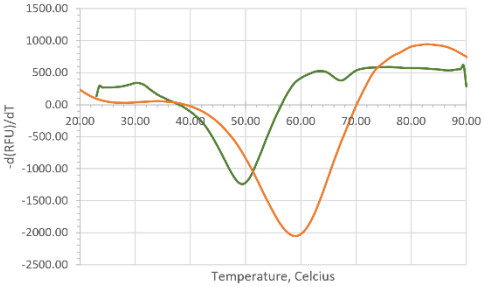
Deoxycholic acid, ΔT_m= ---°C



Heme, ΔT_m= 6 °C



Trehalose, ΔT_m= 8 °C



Xylose, ΔT_m= 10 °C

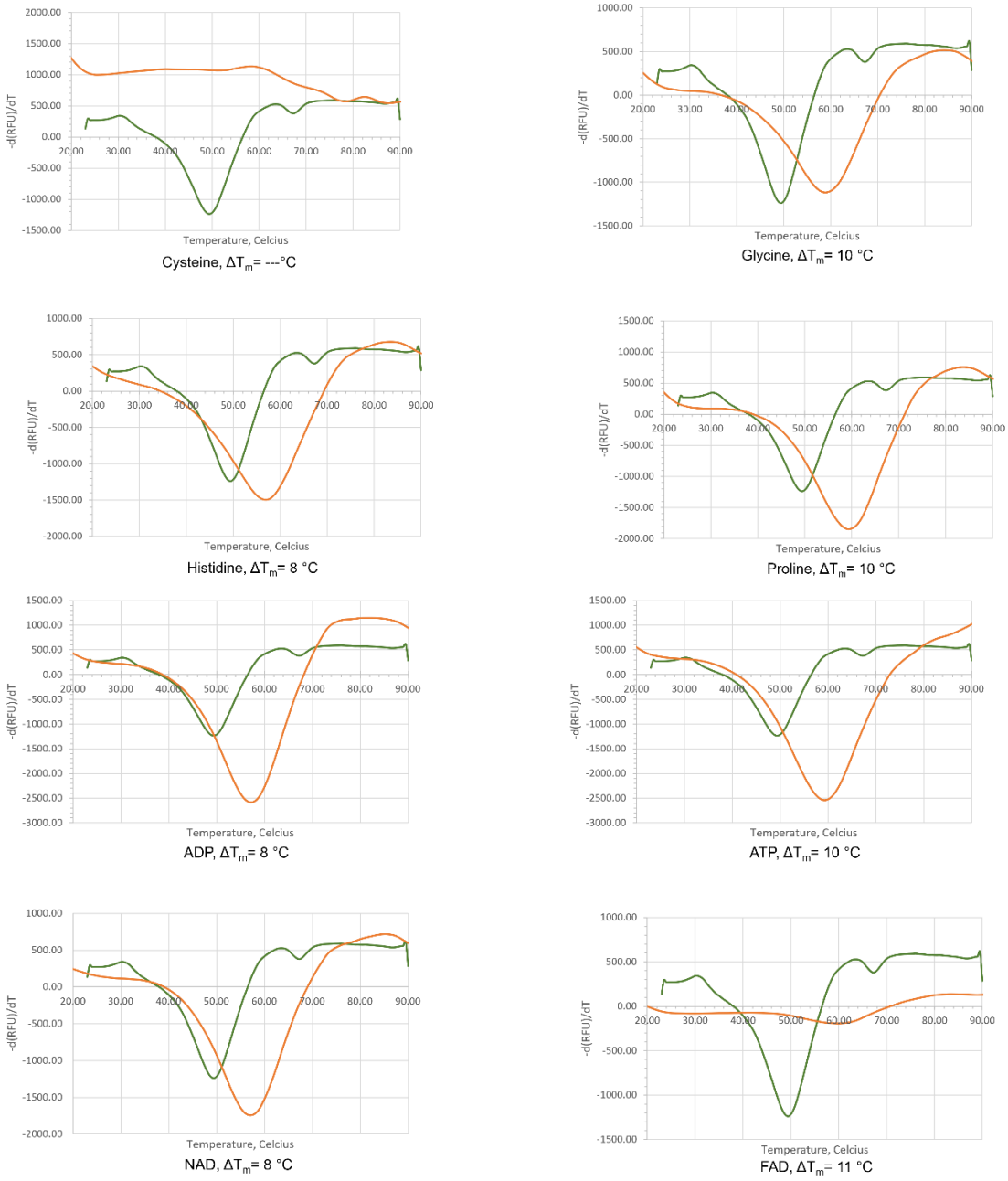


Figure S4. Protein melts from thermal shift assay for CD2492^{PAS}. These 14 ligands were tested against 5 µM protein. Ligand concentrations were: 0.1 mg/ml cholic acid; 2 mg/ml deoxycholic acid; 1 mg/ml heme; 3 mM of proline, ADP, ATP, NAD, or FAD; 50 mM of sodium deoxycholate, trehalose, xylose, cystine, glycine, or histidine.

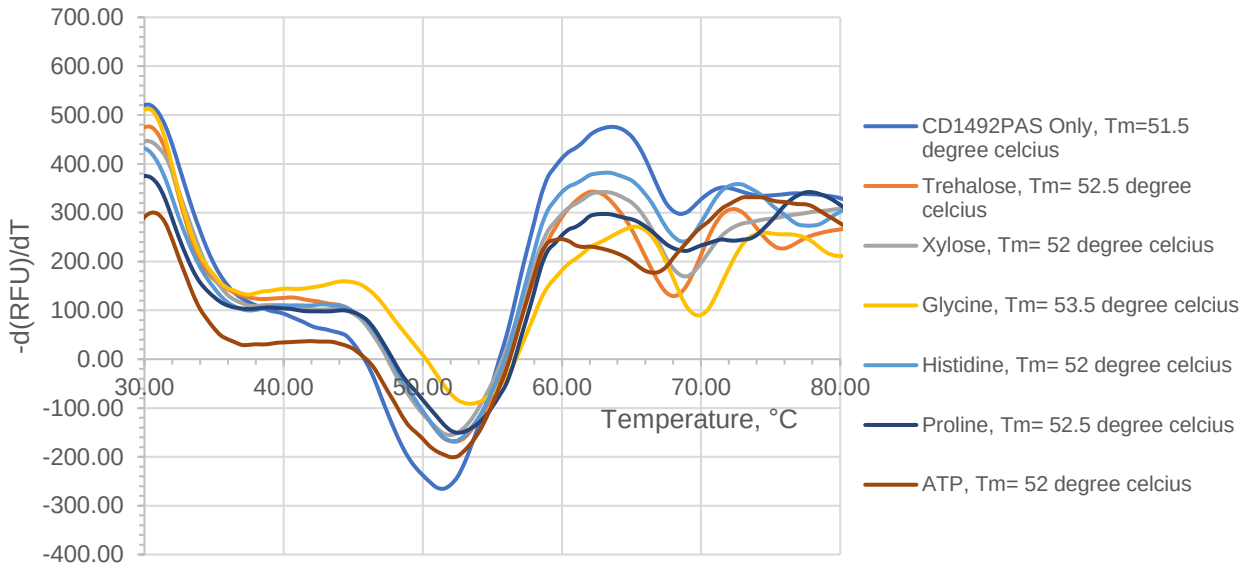


Figure S5. Protein melts from thermal shift assay for CD1492^{PAS}. 6 ligands were tested against 3 µM protein. Ligand concentrations were: 1 mM of proline or ATP; 25 mM of trehalose, xylose, glycine, or histidine. No significant shift in melting temperature was observed.

References Cited

(1) George, R. H.; Symonds, J. M.; Dimock, F.; Brown, J. D.; Arabi, Y.; Shinagawa, N.; Keighley, M. R.; Alexander-Williams, J.; Burdon, D. W. Identification of *Clostridium difficile* as a Cause of Pseudomembranous Colitis. *Br Med J* **1978**, *1* (6114), 695. https://doi.org/10.1136/bmj.1.6114.695.

(2) Antibiotic Resistance Threats in the United States, **2019**, 148.

(3) Baines, S. D.; O'Connor, R.; Saxton, K.; Freeman, J.; Wilcox, M. H. Activity of Vancomycin against Epidemic *Clostridium difficile* Strains in a Human Gut Model. *J Antimicrob Chemother* **2009**, *63* (3), 520–525. https://doi.org/10.1093/jac/dkn502.

(4) Paredes-Sabja, D.; Sarker, M. R. Interactions between *Clostridium perfringens* Spores and Raw 264.7 Macrophages. *Anaerobe* **2012**, *18* (1), 148–156. https://doi.org/10.1016/j.anaerobe.2011.12.019.

(5) Ali, S.; Moore, G.; Wilson, A. P. R. Spread and Persistence of *Clostridium difficile* Spores during and after Cleaning with Sporicidal Disinfectants. *J Hosp Infect* **2011**, *79* (1), 97–98. https://doi.org/10.1016/j.jhin.2011.06.010.

(6) Paredes-Sabja, D.; Shen, A.; Sorg, J. A. *Clostridium difficile* Spore Biology: Sporulation, Germination, and Spore Structural Proteins. *Trends Microbiol* **2014**, *22* (7), 406–416. https://doi.org/10.1016/j.tim.2014.04.003.

(7) Dembek, M.; Barquist, L.; Boinett, C. J.; Cain, A. K.; Mayho, M.; Lawley, T. D.; Fairweather, N. F.; Fagan, R. P. High-Throughput Analysis of Gene Essentiality and Sporulation in *Clostridium difficile*. *mBio* **2015**, *6* (2), e02383. https://doi.org/10.1128/mBio.02383-14.

(8) Warny, M.; Pepin, J.; Fang, A.; Killgore, G.; Thompson, A.; Brazier, J.; Frost, E.; McDonald, L. C. Toxin Production by an Emerging Strain of *Clostridium difficile* Associated with Outbreaks of Severe Disease in North America and Europe. *Lancet* **2005**, *366* (9491), 1079–1084. https://doi.org/10.1016/S0140-6736(05)67420-X.

(9) Banawas, S. S. *Clostridium difficile* Infections: A Global Overview of Drug Sensitivity and Resistance Mechanisms. **2018**. *BioMed Res. Int.* https://doi.org/10.1155/2018/8414257.

(10) IV, E. C. O.; III, E. C. O.; Johnson, D. A. Clinical Update for the Diagnosis and Treatment of *Clostridium difficile* Infection. *World J Gastrointest Pharmacol Ther* **2014**, *5* (1), 1–26. https://doi.org/10.4292/wjgpt.v5.i1.1.

(11) Regulatory Targets of the Response Regulator RR_1586 from *Clostridioides difficile* Identified Using a Bacterial One-Hybrid Screen. **2018**. *Journal of Bacteriology* https://jb.asm.org/content/200/23/e00351-18 (accessed 2020 -01 -15).

(12) Urao, T.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Two-Component Systems in Plant Signal Transduction. *Trends Plant Sci.* **2000**, *5* (2), 67–74. https://doi.org/10.1016/s1360-1385(99)01542-3.

(13) Gotoh, Y.; Eguchi, Y.; Watanabe, T.; Okamoto, S.; Doi, A.; Utsumi, R. Two-Component Signal Transduction as Potential Drug Targets in Pathogenic Bacteria. *Curr. Opin. Microbiol.* **2010**, *13* (2), 232–239. https://doi.org/10.1016/j.mib.2010.01.008.

(14) Chauhan, N.; Calderone, R. Two-Component Signal Transduction Proteins as Potential Drug Targets in Medically Important Fungi. *Infect. Immun.* **2008**, *76* (11), 4795–4803. https://doi.org/10.1128/IAI.00834-08.

(15) Gao, R.; Mack, T. R.; Stock, A. M. Bacterial Response Regulators: Versatile Regulatory Strategies from Common Domains. *Trends Biochem Sci* **2007**, *32* (5), 225–234. https://doi.org/10.1016/j.tibs.2007.03.002.

(16) Dutta, R.; Qin, L.; Inouye, M. Histidine Kinases: Diversity of Domain Organization. *Mol. Microbiol.* **1999**, *34* (4), 633–640. https://doi.org/10.1046/j.1365-2958.1999.01646.x.

(17) Qin, L.; Dutta, R.; Kurokawa, H.; Ikura, M.; Inouye, M. A Monomeric Histidine Kinase Derived from EnvZ, an *Escherichia coli* Osmosensor. *Mol. Microbiol.* **2000**, *36* (1), 24–32. https://doi.org/10.1046/j.1365-2958.2000.01837.x.

(18) Wuichet, K.; Zhulin, I. B. Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. *Sci Signal* **2010**, *3* (128), ra50. https://doi.org/10.1126/scisignal.2000724.

(19) Adebali, O.; Petukh, M. G.; Reznik, A. O.; Tishkov, A. V.; Upadhyay, A. A.; Zhulin, I. B. Class III Histidine Kinases: A Recently Accessorized Kinase Domain in Putative Modulators of Type IV Pilus-Based Motility. *J. Bacteriol.* **2017**, *199* (18). https://doi.org/10.1128/JB.00218-17.

(20) Zschiedrich, C. P.; Keidel, V.; Szurmant, H. Molecular Mechanisms of Two-Component Signal Transduction. *J. Mol. Biol.* **2016**, *428* (19), 3752–3775. https://doi.org/10.1016/j.jmb.2016.08.003.

(21) Wright, G., SA. Architecture of the complete oxygen-sensing FixL-FixJ two-component signal transduction system. **2018**. *Science Signaling* https://stke.sciencemag.org/content/11/525/eaaq0825.long (accessed 2020 -01 -15).

(22) Taylor, B. L.; Zhulin, I. B. PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. *Microbiol Mol Biol Rev* **1999**, *63* (2), 479–506.

(23) Higgins, D.; Dworkin, J. Recent Progress in *Bacillus subtilis* Sporulation. *FEMS Microbiol Rev* **2012**, *36* (1), 131–148. https://doi.org/10.1111/j.1574-6976.2011.00310.x.

(24) Underwood, S.; Guan, S.; Vijayasubhash, V.; Baines, S. D.; Graham, L.; Lewis, R. J.; Wilcox, M. H.; Stephenson, K. Characterization of the Sporulation Initiation Pathway of *Clostridium difficile* and Its Role in Toxin Production. *J Bacteriol* **2009**, *191* (23), 7296–7305. https://doi.org/10.1128/JB.00882-09.

(25) Childress, K. O.; Edwards, A. N.; Nawrocki, K. L.; Anderson, S. E.; Woods, E. C.; McBride, S. M. The Phosphotransfer Protein CD1492 Represses Sporulation Initiation in *Clostridium difficile*. *Infect Immun* **2016**, *84* (12), 3434–3444. https://doi.org/10.1128/IAI.00735-16.

(26) Janoir, C.; Denève, C.; Bouttier, S.; Barbut, F.; Hoys, S.; Caleechum, L.; Chapetón-Montes, D.; Pereira, F. C.; Henriques, A. O.; Collignon, A.; Monot, M.; Dupuy, B. Adaptive Strategies and Pathogenesis of *Clostridium difficile* from In Vivo Transcriptomics. *Infect Immun* **2013**, *81* (10), 3757–3769. https://doi.org/10.1128/IAI.00515-13.

(27) Dembek, M.; Stabler, R. A.; Witney, A. A.; Wren, B. W.; Fairweather, N. F. Transcriptional Analysis of Temporal Gene Expression in Germinating *Clostridium difficile* 630 Endospores. *PLoS ONE* **2013**, *8* (5), e64011. https://doi.org/10.1371/journal.pone.0064011.

(28) Hebdon, S. D.; Menon, S. K.; Richter-Addo, G. B.; Karr, E. A.; West, A. H. Regulatory Targets of the Response Regulator RR_1586 from *Clostridioides difficile* Identified Using a Bacterial One-Hybrid Screen. *Journal of Bacteriology* **2018**, *200* (23). https://doi.org/10.1128/JB.00351-18.

(29) Buschiazzo, A.; Trajtenberg, F. Two-Component Sensing and Regulation: How Do Histidine Kinases Talk with Response Regulators at the Molecular Level? *Annual Review of Microbiology* **2019**, *73* (1), 507–528. https://doi.org/10.1146/annurev-micro-091018-054627.

(30) Scaria, J.; Janvilisri, T.; Fubini, S.; Gleed, R. D.; McDonough, S. P.; Chang, Y.-F. *Clostridium difficile* Transcriptome Analysis Using Pig Ligated Loop Model Reveals Modulation of Pathways Not Modulated In Vitro. *J Infect Dis* **2011**, *203* (11), 1613–1620. https://doi.org/10.1093/infdis/jir112.

(31) Dubrac, S.; Msadek, T. Identification of Genes Controlled by the Essential YycG/YycF Two-Component System of *Staphylococcus aureus. J. Bacteriol.* **2004**, *186* (4), 1175–1181. https://doi.org/10.1128/jb.186.4.1175-1181.2004.

(32) Bhate, M. P.; Molnar, K. S.; Goulian, M.; DeGrado, W. F. Signal Transduction in Histidine Kinases: Insights from New Structures. *Structure* **2015**, *23* (6), 981–994. https://doi.org/10.1016/j.str.2015.04.002.

(33) Trajtenberg, F.; Albanesi, D.; Ruétalo, N.; Botti, H.; Mechaly, A. E.; Nieves, M.; Aguilar, P. S.; Cybulski, L.; Larrieux, N.; de Mendoza, D.; Buschiazzo, A. Allosteric Activation of Bacterial Response Regulators: The Role of the Cognate Histidine Kinase Beyond Phosphorylation. *mBio* **2014**, *5* (6). https://doi.org/10.1128/mBio.02105-14.

(34) Chamnongpol, S.; Cromie, M.; Groisman, E. A. Mg^{2+} Sensing by the Mg^{2+} Sensor PhoQ of *Salmonella enterica. Journal of Molecular Biology* **2003**, *325* (4), 795–807. https://doi.org/10.1016/S0022-2836(02)01268-8.

(35) Mechaly, A. E.; Soto Diaz, S.; Sassoon, N.; Buschiazzo, A.; Betton, J.-M.; Alzari, P. M. Structural Coupling between Autokinase and Phosphotransferase Reactions in a Bacterial Histidine Kinase. *Structure* **2017**, *25* (6), 939-944.e3. https://doi.org/10.1016/j.str.2017.04.011.

(36) Huynh, T. N.; Noriega, C. E.; Stewart, V. Conserved Mechanism for Sensor Phosphatase Control of Two-Component Signaling Revealed in the Nitrate Sensor NarX. *Proc Natl Acad Sci U S A* **2010**, *107* (49), 21140–21145. https://doi.org/10.1073/pnas.1013081107.

(37) Wood, E. J. Molecular Cloning. A Laboratory Manual by T Maniatis, E F Fritsch and J Sambrook. Pp 545. Cold Spring Harbor Laboratory, New York. 1982. $48 ISBN 0-87969-136-0. *Biochemical Education* **1983**, *11* (2), 82–82. https://doi.org/10.1016/0307-4412(83)90068-7.

(38) Fujita, M.; Losick, R. Evidence That Entry into Sporulation in *Bacillus subtilis* Is Governed by a Gradual Increase in the Level and Activity of the Master Regulator Spo0A. *Genes Dev.* **2005**, *19* (18), 2236–2244. https://doi.org/10.1101/gad.1335705.

(39) Burbulys, D.; Trach, K. A.; Hoch, J. A. Initiation of Sporulation in *B. subtilis* Is Controlled by a Multicomponent Phosphorelay. *Cell* **1991**, *64* (3), 545–552. https://doi.org/10.1016/0092-8674(91)90238-t.

(40) Wörner et al. Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from *Clostridium botulinum*. **2006** - *Molecular Microbiology* - Wiley Online Library https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2958.2005.04988.x (accessed 2022 -03 -08).

(41) Steiner, E.; Dago, A. E.; Young, D. I.; Heap, J. T.; Minton, N. P.; Hoch, J. A.; Young, M. Multiple Orphan Histidine Kinases Interact Directly with Spo0A to Control the Initiation of Endospore Formation in *Clostridium acetobutylicum*. *Molecular Microbiology* **2011**, *80* (3), 641–654. https://doi.org/10.1111/j.1365-2958.2011.07608.x.

(42) Tam, J.; Icho, S.; Utama, E.; Orrell, K. E.; Gómez-Biagi, R. F.; Theriot, C. M.; Kroh, H. K.; Rutherford, S. A.; Lacy, D. B.; Melnyk, R. A. Intestinal Bile Acids Directly Modulate the Structure and Function of *C. difficile* TcdB Toxin. *PNAS* **2020**, *117* (12), 6792–6800. https://doi.org/10.1073/pnas.1916965117.

(43) Eyre, D. W.; Didelot, X.; Buckley, A. M.; Freeman, J.; Moura, I. B.; Crook, D. W.; Peto, T. E. A.; Walker, A. S.; Wilcox, M. H.; Dingle, K. E. *Clostridium difficile* Trehalose Metabolism Variants Are Common and Not Associated with Adverse Patient Outcomes When Variably Present in the Same Lineage. *eBioMedicine* **2019**, *43*, 347–355. https://doi.org/10.1016/j.ebiom.2019.04.038.

(44) Weingarden et al. Microbiota transplantation restores normal fecal bile acid composition in recurrent *Clostridium difficile* infection | American Journal of Physiology-Gastrointestinal and Liver Physiology https://journals.physiology.org/doi/full/10.1152/ajpgi.00282.**2013** (accessed 2022 -03 -14).

(45) Huynh, K.; Partch, C. L. Current Protocols in Protein Science. *Curr Protoc Protein Sci* **2015**, *79*, 28.9.1-28.9.14. https://doi.org/10.1002/0471140864.ps2809s79.

(46) Shrestha, R.; Sorg, J. A. Hierarchical Recognition of Amino Acid Co-Germinants during *Clostridioides difficile* Spore Germination. *Anaerobe* **2018**, *49*, 41–47. https://doi.org/10.1016/j.anaerobe.2017.12.001.

(47) Sorg, J. A.; Sonenshein, A. L. Bile Salts and Glycine as Cogerminants for *Clostridium difficile* Spores. *J Bacteriol* **2008**, *190* (7), 2505–2512. https://doi.org/10.1128/JB.01765-07.

(48) Pavlović, N.; Golocorbin-Kon, S.; Đanić (Stojančević), M.; Stanimirov, B.; Al-Salami, H.; Stankov, K.; Mikov, M. Bile Acids and Their Derivatives as Potential Modifiers of Drug Release and Pharmacokinetic Profiles. Frontiers in Pharmacology **2018**, *9*. https://doi.org/10.3389/fphar.2018.01283.

(49) Cimmperman, P.; Baranauskienė, L.; Jachimovičiūtė, S.; Jachno, J.; Torresan, J.; Michailovienė, V.; Matulienė, J.; Sereikaitė, J.; Bumelis, V.; Matulis, D. A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. *Biophysical Journal* **2008**, *95* (7), 3222–3231. https://doi.org/10.1529/biophysj.108.134973.

(50) Allegretti, J. R.; Kearney, S.; Li, N.; Bogart, E.; Bullock, K.; Gerber, G. K.; Bry, L.; Clish, C. B.; Alm, E.; Korzenik, J. Recurrent *Clostridium difficile* Infection Associates with Distinct Bile Acid and Microbiome Profiles. *Aliment Pharmacol Ther* **2016**, *43* (11), 1142–1153. https://doi.org/10.1111/apt.13616.

(51) Shinar, G.; Milo, R.; Martínez, M. R.; Alon, U. Input Output Robustness in Simple Bacterial Signaling Systems. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (50), 19931–19935. https://doi.org/10.1073/pnas.0706792104.

(52) Boldog, T.; Grimme, S.; Li, M.; Sligar, S. G.; Hazelbauer, G. L. Nanodiscs Separate Chemoreceptor Oligomeric States and Reveal Their Signaling Properties. *PNAS* **2006**, *103* (31), 11509–11514. https://doi.org/10.1073/pnas.0604988103.

(53) Hörnschemeyer, P.; Liss, V.; Heermann, R.; Jung, K.; Hunke, S. Interaction Analysis of a Two-Component System Using Nanodiscs. *PLoS ONE* **2016**, *11* (2), e0149187. https://doi.org/10.1371/journal.pone.0149187.

(54) Mitra, N. Nanodiscs: Membrane Protein Research in Near-Native Conditions. *Materials and Methods* **2020**.

(55) Boldog, T.; Li, M.; Hazelbauer, G. L. Using Nanodiscs to Create Water-Soluble Transmembrane Chemoreceptors Inserted in Lipid Bilayers. In *Methods in Enzymology.*; **2007**; Vol. 423, pp 317–335. https://doi.org/10.1016/S0076-6879(07)23014-9.