DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

MA rer. nat. Hubert Mayerhofer

born in

Hallein

Oral examination

7. Dezember 2011

Structural studies of the *Arabidopsis thaliana* ethylene signal transduction pathway

Referees: Dr. Paul Tucker

Prof. Dr. Irmgard Sinning

Table of contents

Abstract

Since ethylene was first recognized as a phytohormone, the scope of its profound and multi-faceted impact on plant growth and development has been continuously growing. In *Arabidopsis thaliana* the response to ethylene is regulated by a group of five receptors (ETR1, ETR2, ERS1, ERS2 and EIN4), which are located in the endoplasmic reticulum membrane. Ethylene is bound by the membrane-embedded N-terminal part, which is followed by a GAF domain. The remaining C-terminal, cytosolic domains resemble the classical bacterial two-component system consisting of a histidine kinase (HK) and in some receptors a receiver domain. Constitutive triple response 1 (CTR1), whose kinase domain bears most resemblance to the RAF family of Ser/Thr protein kinases, directly interacts with the ethylene receptors and thus links signal reception to the intracellular signalling pathway. Therefore, this signalling pathway presents the interesting case, wherein a two-component signalling system manipulates a MAPKKK and possibly a MAPKKK signalling cascade. Still, the question of receptor deactivation by ethylene and the ensuing signal transduction to CTR1 remain unanswered.

A number of constructs comprising domains of the receptors as well as the kinase domain of CTR1 were expressed and purified. The dimerization domain (DHp) of ERS1 could be crystallized and solved by MAD to 1.9 Å resolution. The domain is structurally similar to other domains of this family, normally found in bacteria, consisting of a homodimer forming a coiled-coil and a four-helix bundle. Different to the previously available structures a larger portion of the N-terminal coiled-coil could be determined. When the DHp domain structure is compared with other HKs it is most similar to the phosphatase-competent state, one of the three activities attributed to HKs. A *trans*-phosphorylation mechanism in the dimer is predicted by topological arrangement of the structure. In addition SAXS data of the cytoplasmic part of ETR1 was collected and a model of the domain architecture is presented, showing their relative location with respect to each other. A different location of the receiver domain compared to a recent bacterial structure is suggested.

In addition the three-dimensional structures of the active, tri-phosphorylated and the unphosphorylated, inactive kinase domain of CTR1 in complex with staurosporine were determined at 3.0 Å and 2.5 Å resolution, respectively. They illustrate the

conformational rearrangements that form the basis of activity regulation. The active kinase domain forms back-to-back dimers in solution, while the unphosphorylated kinase is a monomer. The back-to-back dimer interface is virtually identical to the one found in B-RAF and a number of mutants were identified interfering with the dimerization and also affecting the kinase activity. Furthermore the effects of activation loop phosphorylation on the activity were explored. The results strongly suggest another layer of activity regulation of CTR1 through dimerization *in vivo*. Steric restraints further indicate regulation of kinase activity across dimers with a 'front-to-front' activation interface, which points to CTR1 mediated ethylene receptor crosstalk generating a continuous head-to-tail oligomer of kinase domains.

Zusammenfassung

Seit der Klassifizierung von Ethylen als Phytohormon ist das Wissen um seine weitreichende und facettenreiche Rolle und seinen Einfluss auf eine Vielzahl von Entwicklungs- und Wachstumsprozessen von Pflanzen stetig gewachsen. In *Arabidopsis thaliana* wird die Reaktion auf Ethylen durch eine Gruppe von fünf Rezeptoren reguliert (ETR1, ETR2, ERS1, ERS2 und EIN4), die im endoplasmatischen Retikulum verankert sind. Ethylen wird von der N-terminalen, in die Membran eingebetteten Domäne gebunden gefolgt von einer GAF Domäne. Die restlichen C-terminalen Domänen ähneln dem klassischen Zweikomponentensystem bestehend aus einer Histidinkinase (HK) und einer so genannten Empfängerdomäne die letztere jedoch in nur drei der Rezeptoren. Constitutive triple response 1 (CTR1), dessen Kinasedomäne jener der RAF Familie der Ser/Thr Kinasen ähnelt, interagiert direkt mit den Ethylenrezeptoren und stellt damit eine Verbindung zwischen der Signalerkennung und dem intrazellulären Signalübertragungsweg her. Dieser Signalübertragungsweg stellt einen interessanten Fall dar, in dem das Zweikomponentensystem eine MAPKKK und möglicherweise eine MAPKKK Signalübertragungskaskade kontrolliert. Trotz aller Fortschritte ist der genaue Ablauf der Rezeptordeaktivierung und der Übertragung des Signals auf CTR1 nicht bekannt.

Eine Anzahl von Konstrukten, bestehend aus Domänen der Rezeptoren als auch der Kinasedomäne von CTR1 wurde exprimiert und gereinigt. Die Dimerisierungsdomäne (DHp) von ERS1 konnte kristallisiert und durch MAD, bis zu einer Auflösung von 1.9 Å, gelöst werden. Die Struktur ähnelt bakteriellen Strukturen dieser Domäne mit einem Homodimer bestehend aus einem coiled-coil und einem Vier-Helixbündel. Konträr zu anderen Strukturen konnte ein größerer Teil der Struktur des N-terminalen coiled-coil bestimmt werden. Der Vergleich der ERS1 DHp Domäne mit anderen Histidinkinasen legt den Schluss nahe, dass die beobachtete Struktur dem Phosphatase-kompetenten Status entspricht, einem von drei Aktivitätszuständen, die im Allgemeinen mit HK verbunden werden. Ein *trans*-Phosphorylierungsmechanismus des Dimers wird durch die Topologie der Struktur vorhergesagt. Zusätzlich wurden SAXS Daten der zytoplasmatischen Domänen von ETR1 gesammelt. Ein Modell der zytoplasmatischen Domänen wird präsentiert, mit deren relativen Positionen zueinander. Im Gegensatz zu einer bakteriellen

Struktur befindet sich die Empfängerdomäne an einer anderen, weniger zentralen Position.

Zusätzlich wurden die 3-D Strukturen der aktiven, dreifach phosphorylierten und der inaktiven, unphosphorylierten Kinasedomäne von CTR1 jeweils im Komplex mit Staurosporine bis zu einer Auflösung von 3.0 Å beziehungsweise 2.5 Å bestimmt. Sie veranschaulichen die strukturellen Änderungen, welche die Basis der Aktivitätsregulierung bilden. Die aktive Kinase liegt als ´Rücken-an-Rücken´-Dimer in Lösung vor, während die unphosphorylierte, inaktive Kinase als Monomer vorliegt. Die Kontaktfläche des ´Rücken-an-Rücken´-Dimers ist jener von B-Raf sehr ähnlich und mehrere Mutanten wurden untersucht, welche die Dimerisierung und zum Teil auch die Aktivität beeinflussen. Außerdem wurde der Einfluss der Aktivierungsschleifenphosphorylierung auf die Aktivität analysiert. Die Resultate deuten auf ein weiteres Level der Aktivierungsregulation durch Dimerisierung *in vivo* hin. Sterische Einschränkungen deuten auf eine Regulierung zwischen den Kinasedimeren über eine aktivierende ´Kopf-an-Kopf´ Interaktionsfläche hin, mit deutlichen Hinweisen auf eine, durch CTR1, vermittelte Kommunikation zwischen den Ethylenrezeptoren was zur Generierung eines kontinuierlichen Oligomers aus Kinasedomänen in einer Kopf-Schwanz-Anordnung führt.

Introduction

Ethylene Signalling

Ethylene and its effects

Ethylene is a very simple, organic, gaseous molecule with the formula C_2H_4, belonging to the unsaturated hydrocarbons due to its double bond (Figure 1). The molecule is planar with the angles of the sp^2 hybridized carbon close to the ideal angles.

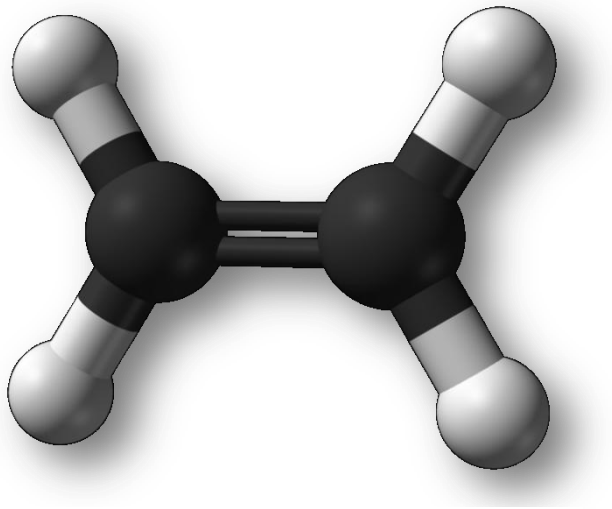


Figure 1: The plant phytohormone ethylene

Abnormalities in the growth of plants in the vicinity of gas lamps were observed in the middle of the 19th century, but the causative agent was unknown. Leakage of gas was shown to be responsible and hydrocarbons were detected, but the exact nature of the molecule was not discovered until 1886, when Neljubov found that ethylene promotes the horizontal growth of etiolated pea seedlings (Neljubov, 1901). Unlike other small volatile molecules produced and emitted in small amounts by plants, ethylene was subsequently suggested to be a plant phytohormone (Crocker et al., 1935) after it was shown previously, that ethylene is indeed also produced by the plants themselves (Gane, 1934). Ethylene forms a group together with other phyto-hormones like auxin, gibberelline, cytokinine and abscisic acid. Plant or phyto-hormones act already at very low concentrations and have a relatively low specificity when compared with hormones found in animals. Phyto-hormones regulate growth, cell division or differentiation processes with the effect depending on the differentiation state of the cell. Furthermore, the signalling of the different hormones in plants is closely connected leading to complex phenotypes.

The effects of ethylene include promotion of fruit ripening in climacteric fruits, the senescence of plants, leaf abscission, influence on the seed germination, flower senescence and it plays a role in response to biotic and abiotic stresses to name some

(Abeles et al., 1992). The triple response phenotype of etiolated seedlings treated with ethylene - inhibition of the growth of stem and roots, curvature of the apical hook and radial swelling of hypocotyl - has been used to screen for defects in the ethylene response. These effects support the seedling to push through the soil while protecting its sensitive tip.

The effects of ethylene have been utilized for thousands of years but the underlying molecular mechanisms have only now begun to get unraveled. Understanding of ethylene biosynthesis, its signal transduction pathway and of the multitude of associated regulatory mechanisms has therefore big implications for basic plant sciences as well as agriculture and economy. The role of ethylene in the ripening of a number of commercially important plants like apples or tomatoes and its role in the senescence and abscission of flowers makes it an appealing research target. Ethylene or its precursors are used to induce fruit ripening, flowering or to de-green citrus fruits. In other cases ethylene production and its signal transduction are blocked by a number of compounds like silver thiosulfate, extending the life of flowers, or 1-methylcyclopropane, a competitive ethylene inhibitor, slowing down the ripening process and extending the shelf life of fruits. Also genetically modified organisms (GMO) have been produced exhibiting strong ethylene insensitivity, for example preventing ripening in tomatoes or delaying senescence of flowers in petunia.

Ethylene synthesis in plants

The ethylene biosynthesis pathway, also termed the Yang cycle (Yang and Hoffman, 1984) (Figure 2), is tightly controlled and starts with the precursor S-adenosyl-L-methionine (SAM) which is produced by SAM synthase utilizing methionine (Figure 2(1)). In a second step SAM is converted into 1-aminocyclopropane-1-carboxylic acid (ACC) and 5'-methylthioadeosine (MTA) catalysed by ACC synthase (ACS) (Figure 2(2)). This is also the rate limiting step in the ethylene biosynthesis. In *Arabidopsis thaliana* nine functional *ACS* genes can be found with distinct spatial and temporal expression patterns during development (Tsuchisaka and Theologis, 2004) and translation can be induced by a number of factors. Furthermore ACS proteins are phosphorylated by MPK6 (Liu and Zhang, 2004) and probably also by calcium-dependent protein kinases (CDPKs), leading to increased stability (Hernández Sebastià et al., 2004). Specific regulation of the ethylene biosynthesis genes exists transcriptionally in response to internal and external stimuli (Tsuchisaka and Theologis, 2004) including ethylene itself (Liu and Zhang, 2004). Also

protein turnover serves as a key regulator of ethylene production in plants (Wang et al., 2004). Methionine is recycled from MTA allowing for the continuous production of ethylene without an increasing need of methionine. In a third step ACC is converted to ethylene (Figure 2(3)) and the by-products CO_2 and HCN by ACC oxidase in an oxygen dependent reaction requiring Fe^{2+} and ascorbate as cofactor and co-substrate, respectively. ACC is transcriptionally regulated. Production of ethylene can go up to 500nl g^{-1} h^{-1} (Burg, 1962) but ethylene was found to be active at concentrations as low as 0.2 nl $l^{-}1$ (Binder et al., 2004a).

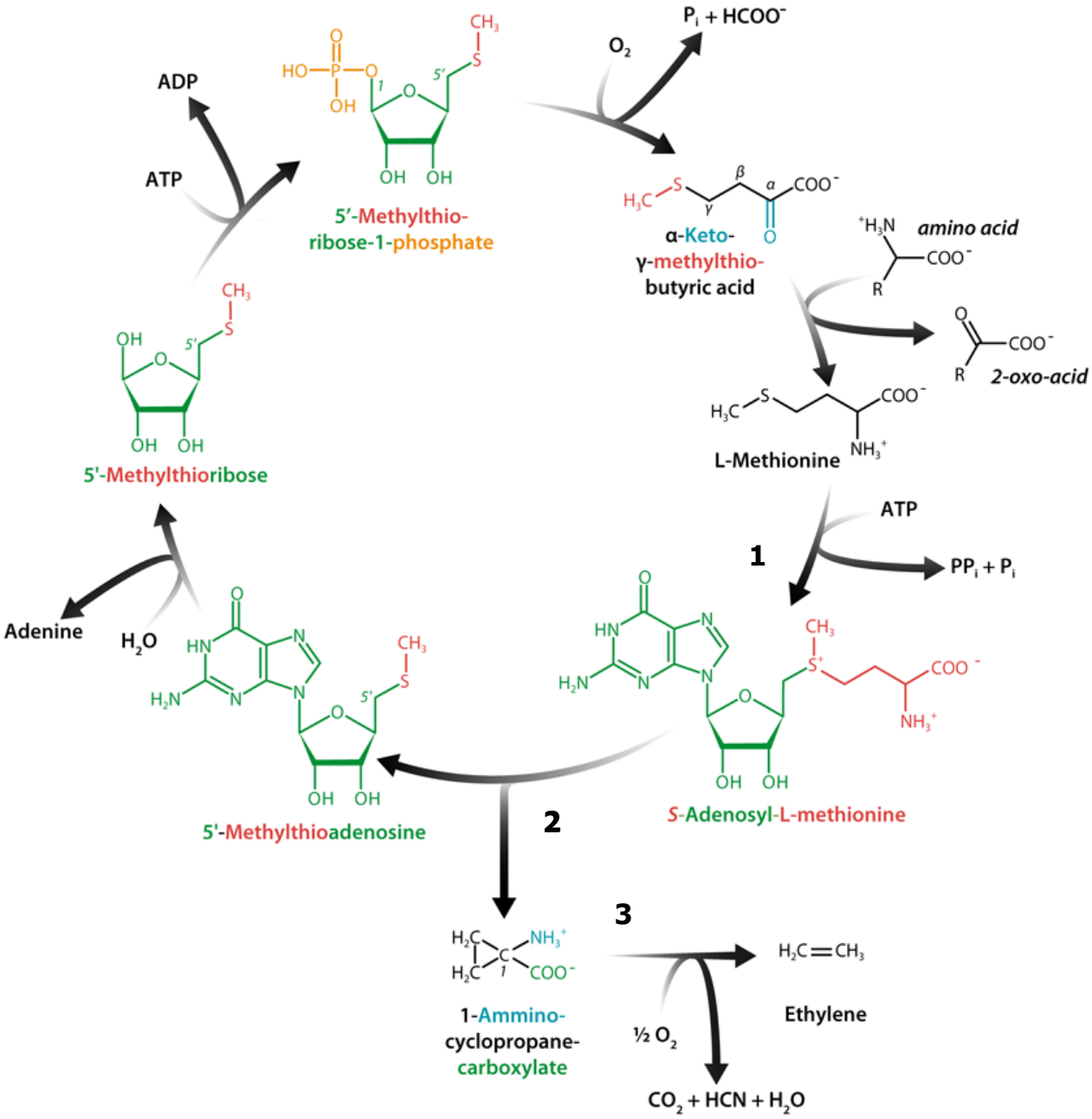


Figure 2: Yang cycle and ethylene biosynthesis pathway with the three steps described marked 1-3. Modified from commons.wikimedia.org/wiki/File:Yang-cycle.png

Ethylene perception and signal transduction

The ethylene signal transduction pathway has been elucidated mainly by using a combination of genetic and molecular analyses to place the members in the pathway. In *Arabidopsis thaliana* ethylene is detected by five integral membrane receptors (Figure 3) termed ETR1, ETR2, ERS1, ERS2 and EIN4 all of which form dimers (Schaller et al., 1995). All receptors are able to bind ethylene with high-affinity and show similar slow-release kinetics for ethylene (O'Malley et al., 2005). Initially the search for LOF mutations of the receptors itself remained unsuccessful, which is explained by their redundant nature. While gain of function mutations lead to ethylene insensitivity, LOF mutations only display an ethylene response phenotype when two or more of the five ethylene receptors are mutated simultaneously.

The receptors mainly insert in the endoplasmic reticulum (ER) membrane (Chen et al., 2002). This is possible because ethylene is soluble in both aqueous and lipid environments. The ethylene binding occurs in the N-terminal membrane domain for which a Cu^{+}-cofactor is needed (Rodríguez et al., 1999), which had been predicted based on the metal affinity of compounds that have ethylene-like or ethylene antagonistic activities. The binding pocket formed by the N-terminal membrane domains of an ethylene receptor dimer accommodates a single Cu^{+}-ion. In ETR1 Cys^{65} and His^{69} were shown to be crucial for copper association (Rodríguez et al., 1999). The copper ion is provided by RAN1, a copper ion transporter probably localized in the same compartment as the receptors. LOF mutants of RAN1 also show a strong constitutive ethylene response (Woeste and Kieber, 2000). While other transition metal ions, such as silver, enable the binding of ethylene (Beyer, 1976), they do not support the necessary structural changes and therefore act as inhibitors. Other inhibitors commonly used are 1-methylcyclopropene (1-MCP) - a gaseous binding site competitor - or aminoethoxyvinylglycine (AVG), which blocks ethylene biosynthesis.

In the absence of ethylene the receptors are in an active signalling state and keep the interaction partner CTR1 (Clark et al., 1998) in an active state, leading to the down regulation of the ethylene signalling pathway (Figure 3). Binding of ethylene causes deactivation of the receptors and subsequently of CTR1. The receptors form disulphide linked homodimers, but also larger oligomers through non-covalent interactions between different receptors have been reported (Gao et al., 2008) comparable to those found in bacterial two-component signalling receptors. This is probably the key to respond

adequately to the broad range of ethylene concentrations ranging from 0.2nl l^{-1} to 1ml l^{-1} (Binder et al., 2004a), (Chen and Bleecker, 1995). In these higher order complexes the receptors show different affinities for each other and the relative ratio of receptor interactions changes upon ethylene treatment. Additionally big complexes (up to 900kDa) consisting of individual ethylene receptor dimers and a number of unidentified, non-receptor proteins were observed *in vivo*. Again the composition of the multi component ERS1 complexes changes in response to ethylene. All domains of the ethylene receptors were shown to contribute to this big complex, most of whose members remain to be defined (Chen et al., 2010).

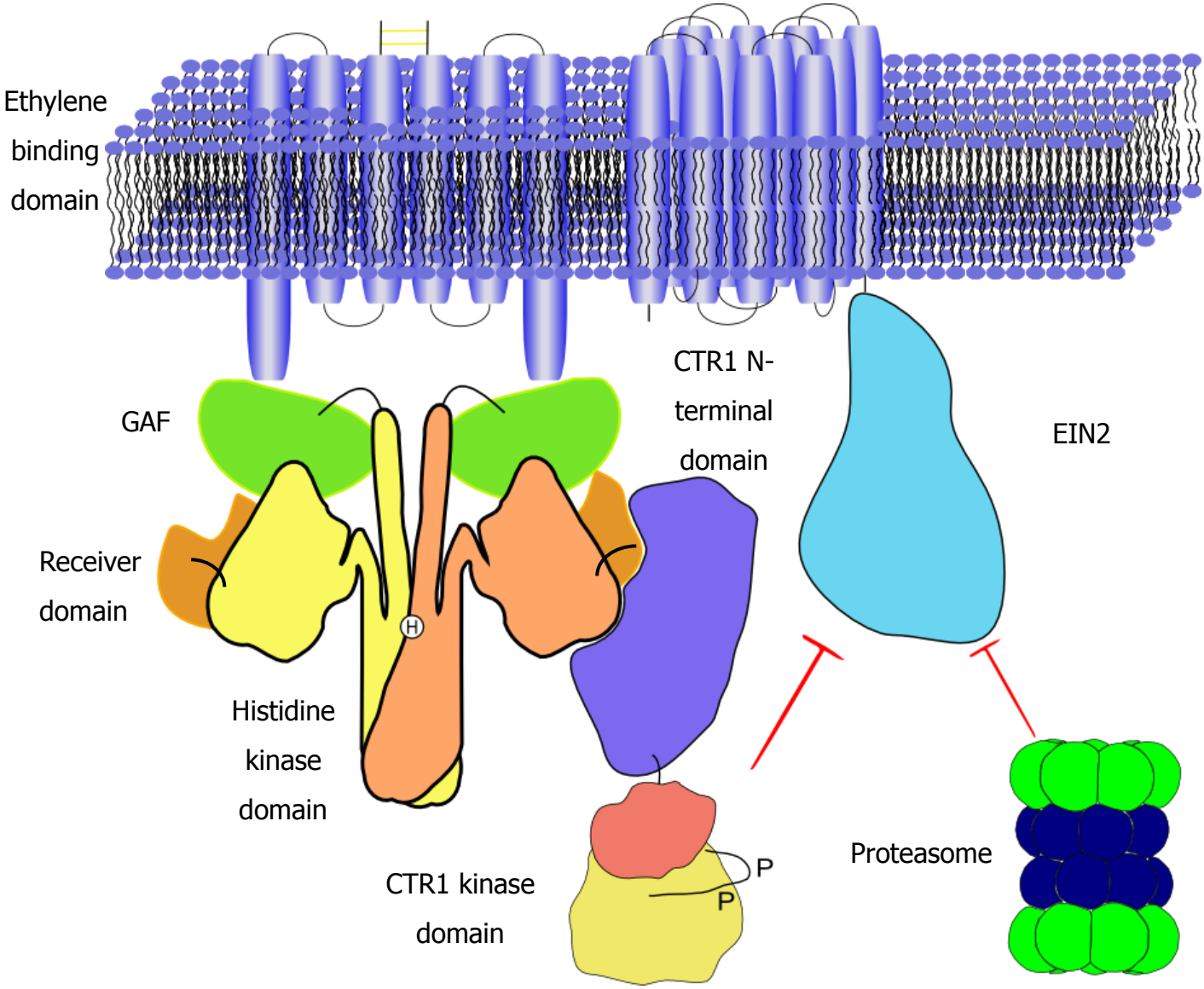


Figure 3: Schematic representation of the first part of the ethylene signalling pathway. The ethylene receptors are bound to the membrane at their N-terminus through 3-4 transmembrane helices, followed by the GAF domain, the histidine kinase domain and the receiver domain. In the absence of ethylene the N-terminal domain of CTR1, shown in

dark blue, binds to the HKD and the receiver domain. Active ethylene receptors do not interact with EIN2 also being membrane bound, which is prone to proteasome degradation (shown in red).

The next member of the signal transduction pathway is constitutive triple response 1 (CTR1), the prime downstream interaction partner of the ethylene receptors, which also localizes to the ER through interaction with the receptors (Gao et al., 2003) (Figure 3). Mutation of CTR1 proved to be particular valuable, as it was the first constitutive ethylene response phenotype and also allowed other members of the pathway to be placed through double mutants. CTR1 is active in the absence of ethylene and is, like the ethylene receptors, deactivated in its presence. The localization of CTR1 to the ER through the interaction with the receptors and its active Ser/Thr kinase domain are crucial for its function in the pathway. Ethylene binding to the receptor leads to the deactivation of CTR1, possibly through a conformational rearrangement or its dissociation from the receptors. The interaction occurs between the C-terminal cytosolic part of the receptors and the N-terminal part of CTR1, which exhibits no similarity to other domains and whose only known function is to bind to ethylene receptors. The influence of the N-terminal domain on the activity of CTR1, as found in RAF, is unclear as both the absence (Huang et al., 2003) and the presence of regulation (Yanagisawa et al., 2003) by the N-terminal domain are reported. Interaction between the two domains of CTR1 has been reported (Larsen and Cancel, 2003). The C-terminal domain of CTR1 is most similar to the kinase domain of B-RAF. Unlike canonical MAPK pathways, active CTR1 acts as a negative regulator of its pathway. MKK9 and MPK6/3 were suggested as the associated MAPKK and MAPK of CTR1, respectively (Yoo et al., 2008). Constitutively active MKK9 leads to EIN3 (discussed later) mediated activation of transcription and was found to be localized to the nucleus as is MPK6, which can also be found in the cytoplasm (Yoo et al, 2008).

CTR1 is also regulated by the phosphatase PP2A-C and shows auto phosphorylation activity although the exact sites have not been described (Huang et al., 2003). CTR1 has also been connected with other plant hormone pathways as auxin, gibberellin or glucose response (Yanagisawa et al., 2003).

For the B-RAF kinase, with 35% sequence similarity to CTR1, catalytic function and activation are dependent on a specific mode of dimerization of kinase domains (Rajakulendran et al., 2009). Homodimerization of B-RAF and heterodimers with the

pseudo kinase Ksr were observed and even kinase dead B-RAF was able to activate the other functional B-RAF molecules through dimerization. Mutations promoting dimerization act as activators (Rajakulendran et al., 2009). The Ser/Thr kinase PknB from *Mycobacterium tuberculosis* exhibits a dimerization mechanism mediated by the kinase domain, which is able to activate the kinase through an allosteric mechanism even in its unphosphorylated state (Lombana et al., 2010). In both cases the interactions are mediated by a number of conserved residues in the interface and mutations of these residues prevent dimerization. The monomeric mutants of PknB show changes in the position of helix C, breaking of the conserved Lys-Glu ion pair, a change of the DFG motif position and alterations of the metal coordination found in the active conformation. In total the monomer seems to be structurally more flexible than the dimer (Lombana et al., 2010). So far no similar mechanism has been described for CTR1.

CTR1, the first downstream target of the ethylene receptors, binds phosphatidic acid (PA), an intermediate in lipid biosynthesis, through its kinase domain, which completely inhibits kinase activity (Testerink et al., 2007). Additionally the interaction between the N and C-terminal domain of CTR1 and the interaction with the ethylene receptors are affected by PA (Testerink et al., 2007). In plants PA is produced in response to osmotic stress or wounding and could induce an early ethylene pathway response providing an explanation how the ethylene pathway is initially activated.

EIN2 is an integral membrane protein with 12 predicted transmembrane helices in the N-terminal domain, showing similarity to the Nramp family of cation transports (about 30% identity in the N-terminus). EIN2 was shown to locate to the ER through the N-terminal domain and to interact specifically with ETR1 through its C-terminus (Bisson et al., 2009), (Bisson and Groth, 2010) which protects EIN2 from degradation by the proteasome (Figure 3). Overexpression of the C-terminal domain in light grown plants leads to an ethylene response. The precise role of the N-terminal domain has not been defined so far. EIN2 seems to play a critical role in the signalling, as the LOF mutations show the strongest constitutive ethylene response observed. EIN2 function is also connected with a number of other pathways, such as ABA and auxin, possibly helping to integrate a number of signals. Despite its key role in ethylene signalling the knowledge of this part of the pathway and the mode of signalling on a molecular level is still limited.

EIN3 acts downstream of EIN2 and is the founding member of a family of six proteins, the EIN3 like (EIL) proteins. Little is known about the role of EIL2 to EIL5 yet the various EIL proteins seem to have at least partially overlapping functions (Chao et al., 1997). These proteins are transcription factors and are found in the nucleus (Chao et al., 1997), acting as a positive regulator of the pathway and start a transcriptional cascade. Signals from the CTR1 initiated MAPK pathway and also from EIN2 converge on the EIN3 and EIL1 transcription factors, which seem to mediate most of the responses. LOF mutants of EIL1 and EIN3 nearly completely abolish the ethylene response in etiolated seedlings (Alonso et al., 2003). Overexpression of EIN3 leads to a constitutive ethylene response (Solano et al, 1998). The EILs bind as homodimers to the primary ethylene response element (PERE) promoter.

The PERE is found in the promoter region of genes encoding ethylene response element binding proteins (EREBPs), which are transcription factors themselves. EREBPs can bind to the specific promoters of ethylene inducible genes through the GCC box, an element which can be found regulating several defence related genes. The best studied member is ethylene response element binding factor 1 (ERF1) (Ohme-Takagi and Shinshi, 1995). The expression of the ERF1 gene is rapidly induced by treatment with ethylene (Solano et al., 1998).

ERF1 overexpression not only affects pathogen associated genes but leads also to a partial triple response (Solano et al., 1998). Other factors seem to play a role in the activation of the ethylene response and while some members of the ERF family were found to act as activators others function as repressors. Furthermore, some contain GCC boxes in their promoter and thereby are regulated by other members of the family. Eventually this leads to the transcriptional regulation of hundreds of genes.

The growth related responses to ethylene can be split into two phases, with a first rapid but transient phase, which is EIN3 and EIL1 independent, reaching its plateau of growth inhibition in etiolated seedlings after 30 min. The second phase leads to a sustained and stronger growth inhibition and depends on the two transcription factors and is less sensitive to ethylene than the first phase. Here the new, low steady-state growth rate is reached after another 30 min. EIN2 is required for both phases, but no information on the cellular process controlling the first phase is available (Binder et al., 2004a). The first phase responds to ethylene levels as low as 0.2nl L^{-1}. This is much lower than the K_d

found for ETR1 expressed in yeast, suggesting a signal amplification process comparable to the one found in the two-component system in bacteria (Thomason et al., 2002). The second phase needs at least 0.5 µl L^{-1} of ethylene for a long term effect. This can be interpreted as an adaptive system as also found in bacteria. An initial response can only be sustained by a higher ethylene concentration or the system will return to the pre-treatment growth rate.

An additional way of signal transduction is starting to get dissected in detail, exhibiting a mode of signal transduction more similar to the one found in the bacterial two-component system (See page 21 for more detailed description). The Arabidopsis response regulator 2 (ARR2) was already shown to contribute to cytokinin signalling (Hwang and Sheen, 2001), acting as a transcription factor. ARR2 has an N-terminal receiver domain and was found to act as a transcription factor for ERF1 (Hass et al., 2004). This activity and the protein stability depend on phosphorylation of Asp^{80} in the receiver domain. The receiver domain interacts with the histidine phosphotransfer protein AHP1 and AHP2 (Hass et al., 2004). They themselves were shown to interact with ETR1 (Scharein et al., 2008), which possess histidine kinase activity, probably acting as shuttling proteins for the phosphoryl group (histidine containing phosphotransfer protein). Mimicking phosphorylation of Asp^{80} in ARR2 causes a triple response phenotype in plants.

Effects of ethylene on gene expression and protein stability

Ethylene promotes its own biosynthesis through MPK6, which was shown to phosphorylate and stabilize ACC2/6 (Liu and Zhang, 2004). MPK6 itself is part of the CTR1 initiated MAPK cascade and furthermore gets activated by a number of stress stimuli like wounding, drought or extreme temperature providing a link to ethylene signalling.

Ethylene seems to have different effects on the receptor protein levels themselves. Some ethylene receptors in plants were shown to be degraded in response to ethylene (Kevany et al., 2007), which could be a way to regulate the response or might also be necessary to supply unbound receptors due to the long half-life time (12hr) of ethylene binding to the receptors (Schaller et al., 1995). Two receptors from *A. thaliana*, ERS1 and ERS2, show elevated expression when ethylene is present, while ETR2 is degraded if ethylene levels are too high (Gao et al., 2008). The receptor protein level correlates with the rate of transcription and the up-regulation of the remaining receptors in LOF mutants is not able to functionally compensate the missing receptor (O'Malley et al., 2005).

Activated MPK3 or MPK6 can phosphorylate EIN3 *in vitro* at position Thr^{174} which stabilizes EIN3. A second phosphorylation site, Thr^{592} has a diametric effect promoting the degradation of EIN3 (Yoo et al., 2008). CTR1 is not found in the nucleus; thereby the phosphorylation of EIN3 at Thr^{592} cannot be carried out by CTR1 and must function via an unknown kinase (Yoo et al., 2008).

EIN3 is regulated on the protein level and gets ubiquitinated in the absence of ethylene by SCF E3 complexes containing the two F-box proteins, EBF1 and EBF2, triggering degradation by the 26S proteasome (Guo and Ecker, 2003) (Qiao et al., 2009) (Potuschak et al., 2003). EIN3 activates the EBF2 promoter (Konishi and Yanagisawa, 2008). EBF1 and EBF2 influence the ethylene response at all stages from repression in the absence and attenuating the signal in the presence of ethylene to a fast recovery following the exposure having temporally distinct roles (Binder et al., 2007).

As EIN3, EIN2 is degraded by the proteasome in the absence of ethylene. Again two F-box proteins, ETP1 and ETP2 were shown to interact with the C-terminus of EIN2 and target it for degradation. The levels of ETP1 and ETP2 are down regulated in response to ethylene (Qiao et al., 2009). The interaction of ETR1 with EIN2 is dependent on the phosphorylation state of the receptor. A reduction of autophosphorylation activity, as mediated through ethylene, increases the affinity for EIN2, allowing the HK domain to bind to EIN2 (Bisson and Groth, 2010). This interaction protects EIN2 from the degradation by the proteasome.

2 component system

The ethylene receptors are similar in sequence to the histidine kinases and their receiver domain of the bacterial two-component system. Eukaryotic proteins with homology to the bacterial histidine kinases have been identified in fungi and plants (Swanson et al., 1994). Two-component systems comprise of a sensor histidine kinase (HK) and its cognate response regulator. They are utilized by bacteria to couple environmental stimuli with adaptive responses regulating a wide variety of behaviours. Up to one hundred TCSs can exist in parallel in bacteria being very specific for the cognate partner. Sensor HKs consist in most cases of an extracellular, N-terminal input domain, tailored at the specific signal, followed by a transmembrane domain and a cytoplasmic kinase domain. The actual HK domain can be subdivided into a DHp domain or dimerization domain, containing the conserved histidine residue for phosphorylation followed by the catalytic and ATP binding (CA) domain (Figure 5A). The CA domain is connected with the DHp domain by a flexible linker and does not contribute to the dimer interface.

Two component system HKs autophosphorylate a conserved histidine residue in the dimerization domain often in response to an environmental stimulus. The phosphoramidate bond present at the histidine is much less stable than the phosphoester bond of hydroxyl kinases. This high-energy phosphoryl group is subsequently transferred to a conserved aspartic acid in a receiver domain, which is part of the response regulator (RR) (Stock et al., 2000). This leads to a conformational change, which activates the output domain of the response regulator, in most instances a transcription factor. The phosphotransfer mechanism is sometimes expanded to by phosphorelay mechanism involving three steps of transfer of the phosphate moiety, involving multiple proteins (Figure 5B). Such a mechanism is found in approximately 25% of all prokaryotic HKs. There the HK domain is fused to a receiver domain to form a hybrid kinase. The receiver domain accepts the phosphoryl group from the histidine in the DHp domain and transfers it to histidine containing phosphotransfer protein (HPt), which eventually transfers it to a second receiver domain fused to the output domain. On top of the autophosphorylation the HK often possess RR phosphatase activity controlling the level of phosphorylation in the receiver domain and thereby the length and strength of the response. The phosphatase activity is located in the DHp domain (Zhu et al., 2000). It is influenced in its activity through the interaction with the CA domain of the kinase. While a covalently linked

CA domain was shown to boost the phosphatase activity, the isolated DHp domain alone shows reduced activity in EnvZ (Zhu et al., 2000).

The receiver domain is also characterized by a number of conserved residues. Two acidic residues (Asp^{616}, Glu^{617} numbering for ETR1) and the phosphoryl accepting Asp^{659} form an acidic pocket, binding together with a backbone carbonyl group and two water molecules, a divalent cation (Bourret, 2010), often being Mg^{2+}. Lys^{714} points to this pocket helping to coordinate the phosphoryl during phosphorylation and is important for phosphorylation mediated conformational changes in the β5α5 loop, forming a salt bridge with one of the phosphoryl group oxygens. The cation also interacts with the phosphoryl group as well as a highly conserved Ser/Thr (Ser^{692}), which causes the movement of the β4α4 loop. Again this aspartyl-phosphate is relatively unstable, especially at a low pH. The receiver domain has been proposed to catalyse its own de-phosphorylation often stimulated by other proteins (Bourret, 2010), which is thought to be a similar mechanism to phosphorylation involving the same residues and an additional water. Another protein, often the cognate HK, stimulates the intrinsic autophosphorylation activity. Residues C-terminal to the Asp phosphorylation site and the conserved Thr/Ser were shown to influence the dephosphorylation rate (Thomas et al., 2008).

Three functional states can be observed within many HKDs: a phosphatase competent state able to remove the phosphate from the cognate receiver domain, a kinase-competent state able to phosphorylate the conserved histidine residue in the dimerization domain and a phospho-transferase state able to interact with the receiver domain and able to transfer the phosphate between the two proteins. These different activities were shown to be accompanied by structural changes including also asymmetric conformations between the two molecules in the dimer (Albanesi et al., 2009). The following states are described based on structures obtained for the cold sensor DesK.

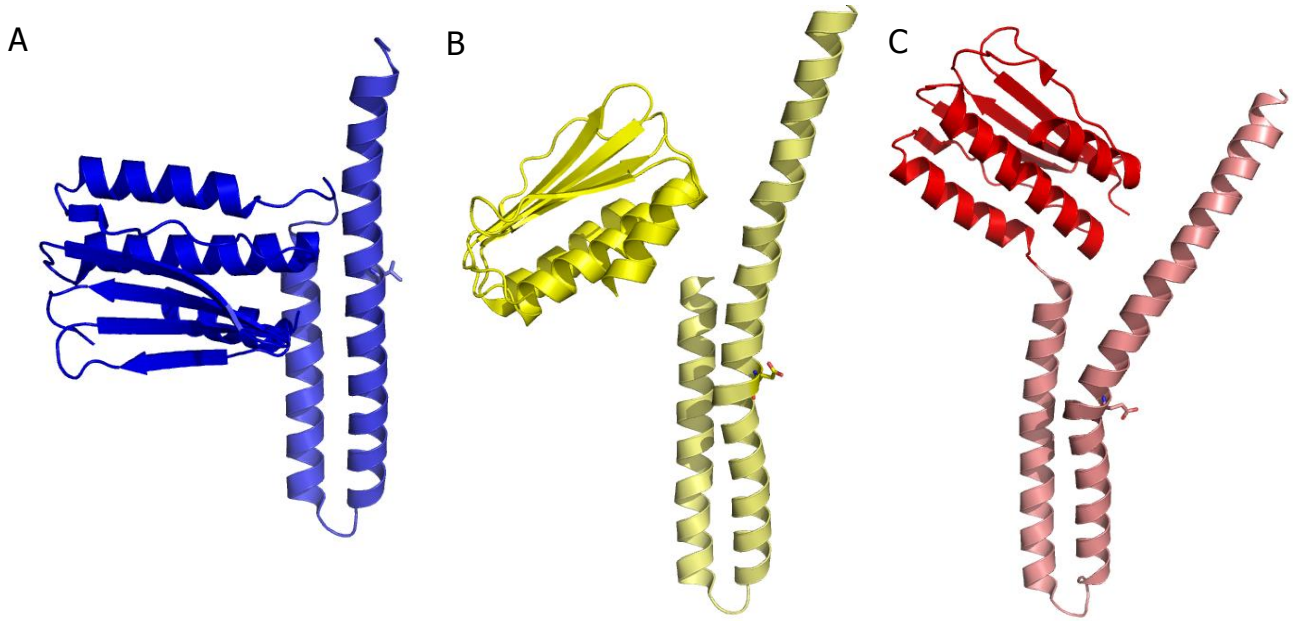


Figure 4: Structures of the DesK histidine kinase domain in different activation states. The dimerization domains were superimposed. The catalytic domain is shown in all pictures darker than the dimerization domain. **A** H188V mutant in the phosphatase-competent state **B** H188E mutant in the autokinase-competent state **C** H188E mutant in the phosphotransferase-competent state. The residue at the position of the conserved histidine is shown in sticks.

For the autokinase-competent state (Figure 4 B) a high internal flexibility was described between the two domains in the HKD (Albanesi et al., 2009). A simple rotation around a glycine acting as a hinge residue could bring the ATP close to the conserved histidine residue. The surfaces of the CA and DHp domains would be in good complementarity after the rotation. In the phospho-transferase competent state (Figure 4 C) the two CA domains in DesK show different conformations, with only one interacting with the first helix A1 of the dimerization domain (Albanesi et al., 2009). This CA domain of chain A now also interacts with a basic patch in the A1 helix of chain B. Chain B is also phosphorylated while chain A is not which is also lacking contacts with the second CA domain in DesK. The kink in this helix A1 is increased compared to the autokinase-competent state, also changing the hydrogen bonds around the conserved histidine residue. The phosphatase-competent state (Figure 4 A) has a more rigid conformation with increased interactions between the DHp and CA domains. This conformation has been observed in a number HKDs like TM0853, KinB or DesK. A longer part of the N-terminal coiled-coil is visible, increasing the dimerization interface which also attenuates the

observed kink in the A1 helix (Albanesi et al., 2009). In this state the kinase activity is reduced.

Transition between the states is companied by a rotation and shearing of the helices in the DHp domain, which abandons the coiled-coil and affects the interaction with the CA domain. Therefore a model was suggested with the membrane domain of the protein stabilizing the coiled-coil, inhibiting the kinase and activating the phosphatase activity. In the case of a signalling event the structural changes would free the CA domain thereby activating the kinase. Auto-phosphorylation is expected to cause another structural change, allowing for the interaction with the receiver domain.

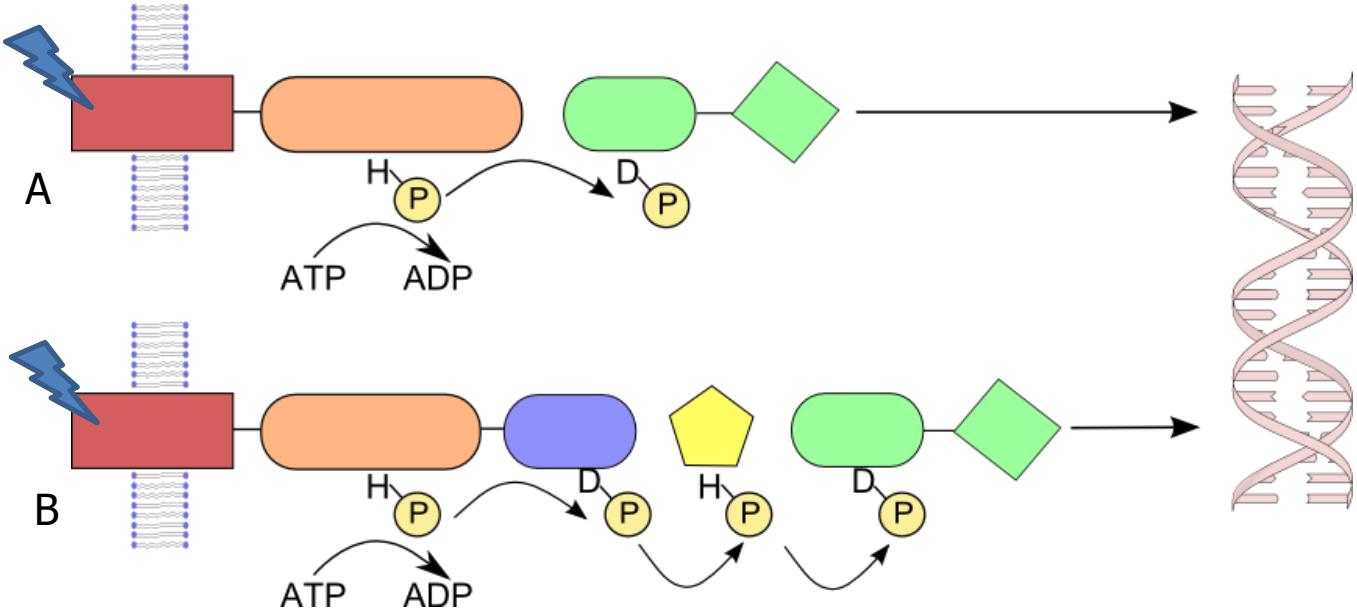


Figure 5: Schematic layout of the 2-component system. **A** The sensor domain receives a signal and transmits it to the histidine kinase domain, thereby activating it. A conserved histidine residue is phosphorylated and the phosphoryl group is subsequently transferred to the receiver domain of its cognate response regulator. This activates the output domain, often a transcription factor. **B** The hybrid kinase with a fused receiver domain shows a more complex phosphoryl relay. After autophosphorylation on the conserved histidine the phosphoryl group is transferred to the aspartate in the receiver domain, which is covalently linked to the HK. From there it is moved to a phosphotransfer protein (HPt), which finally transfers it to the response regulator carrying out the response.

Structures of prokaryotic two-component system elements

At present more than 160 structures of two-component proteins in prokaryotes are available with about 60 of the HK domain or parts of it and about 100 of the receiver domain, allowing predictions of the expected structures in eukaryotes. Most of the structures represent isolated domains but recently a few complete histidine kinase domains (HKD) (Marina et al., 2005) (Figure 6) and also the complex of a HK with its cognate receiver domain became available (Casino et al., 2009). The dimerization domain consists of a helical hairpin domain packing with a second DHp domain into a four-helix bundle as observed for CheA (Bilwes et al., 1999) or EnvZ (Tomomori et al., 1999). They contain a hydrophobic core aligned along the long helix axes. Interestingly depending on the kinase the relative angle of the helices to each other in the four-helix bundle can vary. The conserved histidine residue is found in the middle of first helix, most of the time exposed to the solvent and thereby accessible by the catalytic domain. The DHp stem region, including the linker connecting the two helices of the dimerization domain, was found to be important for the interaction with the cognate response regulator. When it was swapped in EnvZ with the corresponding region of another HK the specificity for the response regulator was changed (Skerker et al., 2008). Helix 1 of the DHp domain is bound by helix 1 and the β5-α5 loop of the phosphorylated receiver domain (Casino et al., 2009). Prior to the four-helix bundle the N-terminal part of the first helix of each subunit form a coiled-coil motif making either a direct connection to the transmembrane section of the HK (Singh et al., 1998) or through an additional interposed domain. Structural changes in the sensor domain are thought to transmit through the helices, affecting the activity with a piston like translation of one helix with respect to the other (Falke and Hazelbauer, 2001), a rotation movements of the helices or a combination of them. The signal probably reaches the HK domain in the form of a helical rotation and tilting. This affects the interaction of the helices in the dimerization domain and the contact with the CA domain leading to a change of the catalytic domain position (Albanesi et al., 2009), (Casino et al., 2009).

The CA domain exhibits a mixed α/β sandwich fold consisting of five β strands and a single layer three α helices (Figure 6). The CA domain is part of the ATPase/kinase GHKL superfamily (Dutta and Inouye, 2000). The CA domain binds ATP between the ATP lid and

the central helix with the substrate binding pocket of the CA domain facing the α helix of the DHp domain such that it points to the conserved histidine residue.

The DHp and the CA domains are connected by a hinge region playing a role in regulating the phosphatase activity and mutations in the C-terminal part of helix 2 of the DHp domain and the linker only affected the phosphatase but not the kinase activity (Hsing et al., 1998). This region shows a number of inter-domain contacts and interacts also with the response regulator, especially with two loops changing most upon phosphorylation (Casino et al., 2009). Mutations might affect the recognition of the phosphorylated response regulator thereby affecting the phosphatase activity.

So far *cis* (Casino et al., 2009) and *trans* (Yang and Inouye, 1991) phosphorylation mechanisms were shown for HKs, which seems to be a consequence of the different connections between the helices in the DHp domain leading to different connections to the CA domain. Also the length of the DHp-CA linker has been proposed to play a role (Casino et al., 2010).

HKs contain a number of conserved motifs termed the H, N, G1, F, G2 and G3 boxes. They are all needed to form a functional histidine kinase (Parkinson and Kofoid, 1992). The H box contains the site of phosphorylation in the DHp domain, while the other motifs are located in the CA domain forming the conserved ATP binding site. The G2 box facilitates the flexibility of the ATP lid, being located at its N-terminus after the conserved phenylalanine or F box (Marina et al., 2001), the only box which is unique for the HKs in the GHKL superfamily. The HKs ATP lid is unusually long compared with other ATP lids in the superfamily. It shows variability in conformation and composition as it can contain α helices and thereby the structures also differ in the amount of solvent exposure of the bound ATP (Dutta and Inouye, 2000). The lid is highly mobile in the absence of ATP as was observed in the apo structures of CheA and NtrB. The G1 and G3 box contact the adenosine moiety, while the N box coordinates the Mg^{2+} bridging the three phosphates (Marina et al., 2001). Although general features can be assigned to the motifs the molecular details of the nucleotide binding vary for different HKs.

A number of prokaryotic receiver domains have been described as CheY or PhoP. The receiver domain changes its structure upon phosphorylation leading to different contacts with the HK domain and therefore affecting the reaction being carried out. Structural details are discussed for the case of the ETR1 receiver domain.

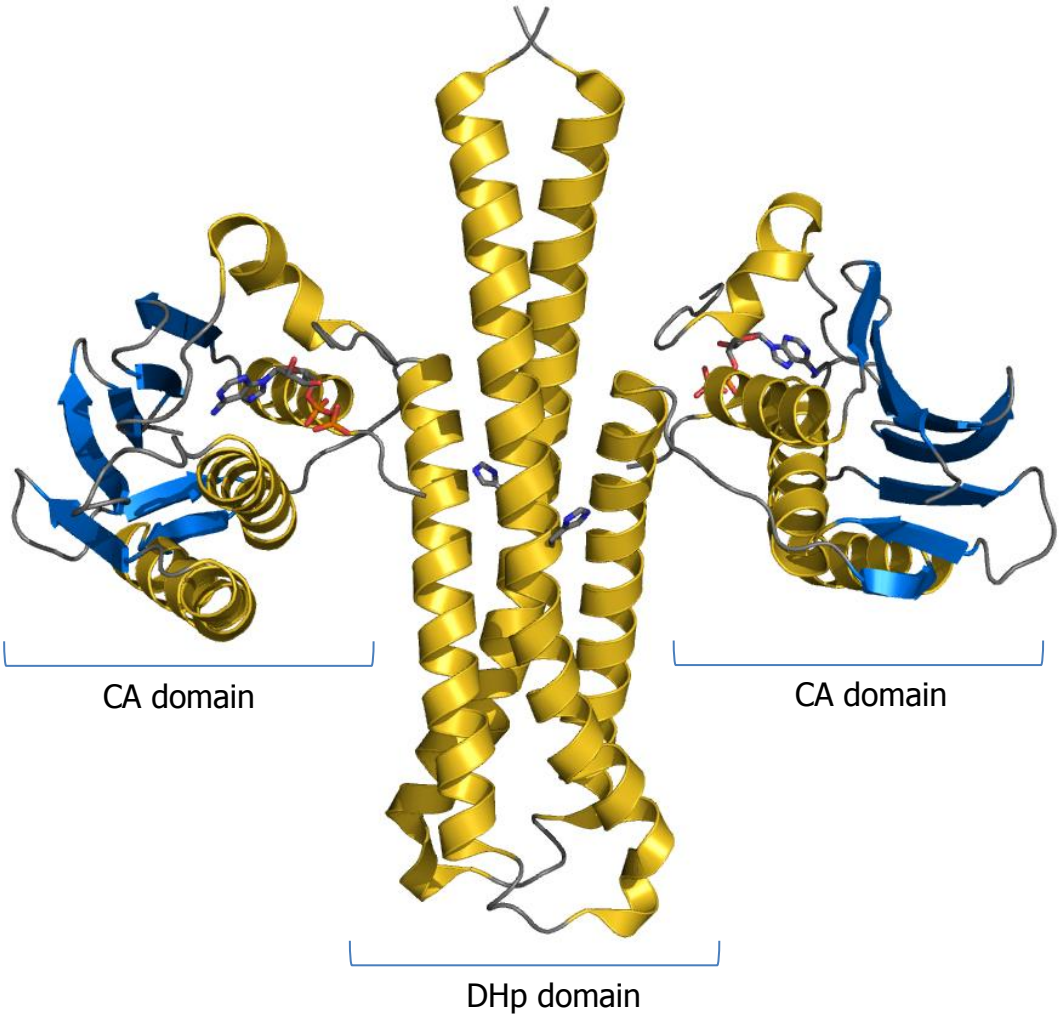


Figure 6: Crystal structure of the histidine kinase domain of *Thermotoga maritima* TM0853 (PDB ID: 2C2A). The conserved histidine and the ADP are shown in sticks. In the centre the dimerization domain with the histidine is forming a four-helix bundle with the two catalytic domains sitting on the side.

Structures of the ETR1 response regulator

Structural information from eukaryotic HKs and the associated molecules is scarce. The structure of the ETR1 response regulator (Müller-Dieckmann et al., 1999) resembles the structures observed in prokaryotes (Bourret, 2010). It forms a globular $(\beta/\alpha)_5$ domain with a central β-sheet packed between two and three helices, respectively (Figure 7). The main differences in ETR1 to other receiver domains are found in the loop regions. The biggest differences to CheY and NarL are being found in loop 5 ($\beta3$-$\alpha3$), located right after the phosphorylation site Asp^{57}. It makes a 180° turn initiated by a conserved proline. In ETR1 this loop is flipped to the other side compared with other response regulators, making a rearrangement prior to phosphorylation necessary. This loop is thought to play

an important role during intermolecular recognition with the HKD. This loop in the structure of the TM0853 kinase with its cognate, phosphorylated receiver domain is essential for the contact with the CA domain (Casino et al., 2009). In the structure the ETR1 response regulator forms a dimer, an oligomeric state, which was also observed in solution (Müller-Dieckmann et al., 1999). The contact is mediated by its C-terminus being responsible for the majority of the dimer interface, consisting of $\alpha 4 \beta 5 \alpha 5$. The interface used for dimerization in ETR1 is used in other receiver domains for the interaction with the DNA-binding domain, as the biggest structural changes upon phosphorylation of the receiver domain occur here. The $\beta 5$-$\alpha 5$ loop was shown in bacteria to bind to helix 1 of the DHp domain and using bacterial structures as models the two receiver domains in ETR1 are predicted to be on opposite sides of the dimers, so the biological relevance of this dimer is unclear. In NtrC and PhoB dimerization was reported to be phosphorylation dependent.

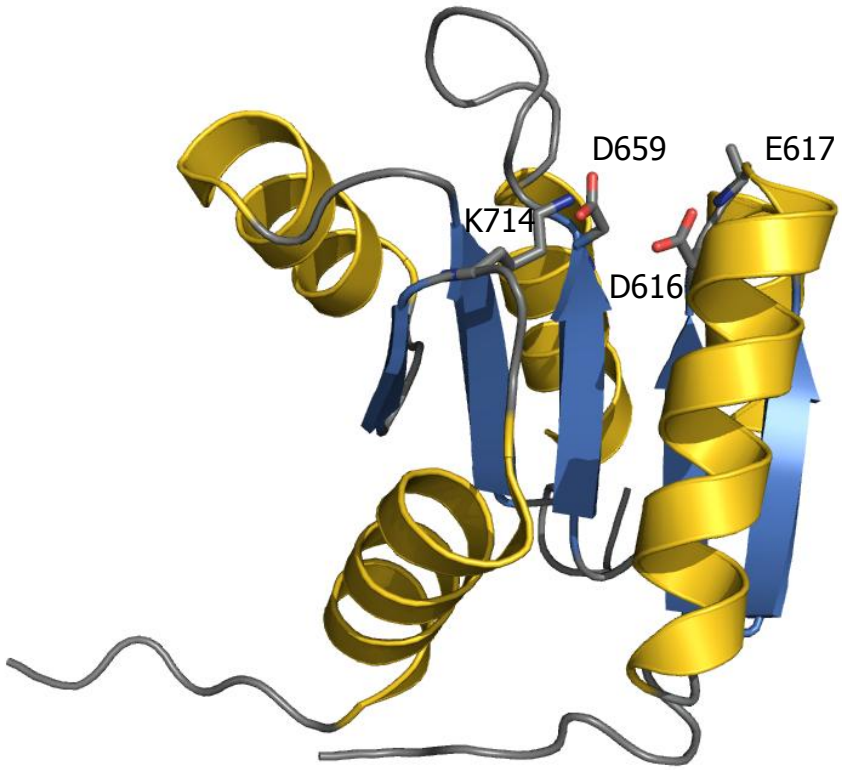


Figure 7: Cartoon representation of the ETR1 response regulator (pdb ID 1DCF). Asp^{616}, Glu^{617}, Asp^{659} and Lys^{714} shown as sticks

Domain organization of the ethylene receptor

Sequence analysis of the ethylene receptors identifies four to five domains depending on the receptor. The first 150 residues comprise an N-terminal membrane domain. Here the only two disulphide bridges present in the receptor are located right at the N-terminus (residue 4 and 6 for ETR1). The membrane domain is crucial for the localization of the ethylene receptors to the ER (Chen et al, 2010). Ethylene binding is known to occur in the hydrophobic N-terminal part of the receptor dimer. The required copper ion is coordinated by the conserved Cys^{65} and His^{69} residues (numbering for ETR1) (Rodríguez et al., 1999). The first cytoplasmic domain, a GAF domain, was found to play a role in the interaction and communication between different receptors (Gao et al., 2008) yet the interaction with other receptors seemed to be reduced in a truncated version lacking the HK and receiver domain highlighting their contribution. Apart from this its role is unknown. GAF domains, named after some of the proteins it is found in (c**G**MP-specific phosphodiesterases, **a**denylyl cyclases and **F**hlA), are normally found in phytochromes and cGMP-specific phosphodiesterases. The GAF domain belongs to a big family of small-molecule binding proteins and exhibits a central anti-parallel β-sheet which is sandwiched by two α-helices. The helices pack together in the homodimer to form a four-helix bundle (Levdikov et al., 2006).

The GAF domain is followed by a HK domain and in three out of the five receptors (ETR1, ETR2 and EIN4) a C-terminal receiver domain. Both domains were shown to be needed for the interaction with CTR1 (Clark et al., 1998).

The receptors can be split into two subfamilies with subfamily 1 composed of ETR1 and ERS1 possessing three transmembrane helices and all conserved residues connected with histidine kinase activity. For subfamily 2 four N-terminal transmembrane helices are predicted and only some of the conserved residues needed for HK activity are present (Hua et al., 1998).

The central histidine kinase domain can, as found in prokaryotes, be split into the N-terminal, H box containing DHp domain followed by the catalytic CA domain. The CA domain has in case of ETR1 and ERS1 all conserved residues known from prokaryotes to be crucial for histidine kinase activity *in vitro* histidine kinase activity has been demonstrated (Moussatche and Klee, 2004). The subfamily 2 members have only some of these motifs consistent with the absence of histidine kinase activity in these proteins.

Subfamily 2 members and also, under certain conditions, ERS1 display Ser/Thr kinase activity despite the sequence similarity to histidine kinases (Moussatche and Klee, 2004).

Functional roles of the receptors and the domains

The receptors seem to work synergistically and are not able to replace each other functionally. Subfamily 1 plays a predominant role as the double mutant *etr1* and *ers1* shows an extremely strong constitutive ethylene response including growth inhibition. The gain of function allele *etr1-1,* representing a constitutively active receptor, leads to a dominant ethylene insensitive phenotype (Schaller et al., 1995). Single LOF mutants do not show an obvious phenotype and double or triple LOF mutants are needed to yield an ethylene response phenotype. Phenotypes of the mutants are dependent on EIN2 (Hall and Bleecker, 2003). A limited ethylene response could be detected in the subfamily 1 double mutant, (Qu et al., 2007) and a triple LOF of the subfamily 2 leads to ethylene hypersensitivity highlighting the fact that subfamily 2 has a role in the signalling (Cancel and Larsen, 2002).

The growth inhibition of an *etr1-7 ers1-2* double mutant, which still had some ERS1 transcript detectable could be alleviated by a truncated, dominant mutant *etr1-1* (1-349) converting other receptors into an ethylene-insensitive conformation, but not in the *etr1-7 ers1-3* mutant lacking any subfamily 1 receptors. The truncated receptor might not be able to signal itself but could convert other wild-type subfamily 1, but not subfamily 2, receptors to an ethylene-insensitive signalling state (Qu and Schaller, 2004). This strongly supports communication between the different receptors, at least within a subfamily.

HK activity was initially demonstrated not to be important for the function as a kinase-dead *etr1* mutant was able to restore the growth inhibition exhibited by a subfamily 1 double mutant (Wang et al., 2003) while receptors from subfamily 2 failed to do so. In the *etr1 etr2 ein4* mutant a truncated version of ETR1 lacking the receiver domain (1-603) needed the histidine kinase domain but not its activity to rescue the triple response phenotype, yet it showed hypersensitivity to ethylene, suggesting a role of the receiver domain. Furthermore when measuring the growth recovery upon ethylene removal, mutants of receptors possessing a receiver domain showed slower recovery while the *ers1 ers2* mutant had no effect on recovery (Binder et al., 2004b) supporting a role of the receiver domain. More recently it was shown that a full length ETR1 kinase dead mutant could not completely rescue the triple response phenotype (Qu and Schaller, 2004), a

result in contrast to other studies. In the *etr1 ers1* double mutant the slow recovery was rescued by full length ETR1 but not by kinase dead ETR1. Ethylene was shown to control the autophosphorylation activity of ETR1 which supports a phosphorelay mechanism controlling aspects of ethylene signalling (Voet-van-Vormizeele and Groth, 2008). In addition autophosphorylation of ETR1 seems to control the interaction with EIN2. This interaction is promoted in the absence of the phosphoryl group. All these observations contribute to a complex picture of the ethylene receptors and the different roles the domains play, with some having a pivotal role while others are engaged in the control of finer details.

Heterodimers and higher order interactions between different receptor homodimer were reported. ETR1 was shown to copurify with all other receptors and these interactions are recruiting a substantial part of the available ETR1 molecules (Gao et al., 2008). ETR1 seems to interact preferentially with subfamily 2 receptors. The percentage of interaction with the other receptors changed upon ethylene treatment. These interactions could allow an ethylene insensitive mutant to amplify its signal keeping other receptors in the active state.

Serine/threonine protein kinases

Ser/Thr kinases play a role in virtually all signalling pathways in eukaryotes, for example in response to extra- and intracellular messengers, metabolism, immune defence or growth regulation, all requiring a tight and multi-layered control (Schenk and Snaar-Jagalska, 1999). In their activated state, protein kinases transfer the gamma phosphoryl of ATP onto the hydroxyl moiety of their target protein leading to conformational changes or change of affinities, which affect their activity. Defining the phosphorylation target through the surrounding sequence, Ser/Thr kinases often exhibit a relatively low specificity enabling them to phosphorylate whole protein families. Specificity therefore also depends on docking or scaffolding interactions (Bhattacharyya et al., 2006).

Ser/Thr kinases often function in a network of phosphorylation events as found in the MAPK signalling pathway with one phosphorylation triggering another one. MAPKKK phosphorylates MAPKK which in turn phosphorylates MAPK thereby transmitting the activation like in the classical mammalian RAF1 – MEK11/2 – ERK1/2 network.

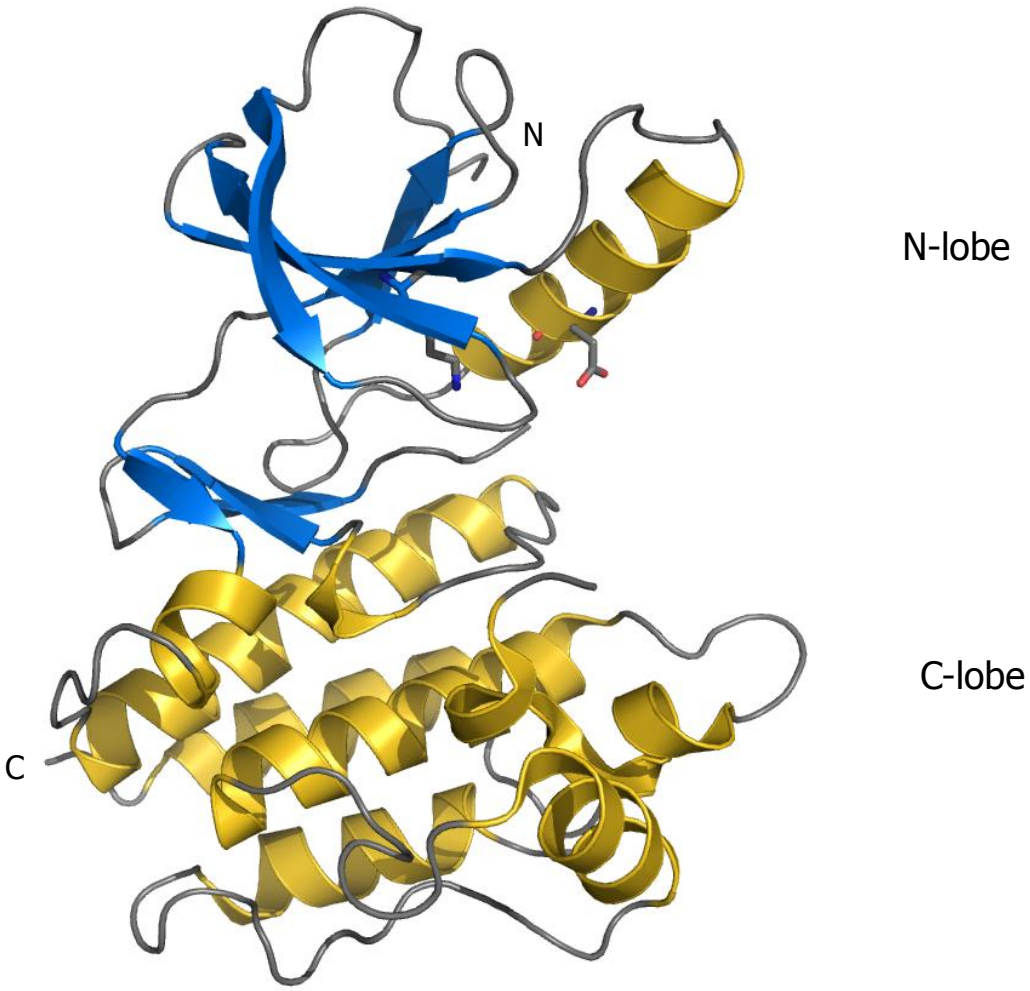


Figure 8: Inactive structure of the RAF-1 kinase domain (pdb ID: 3C4C) bound to the inhibitor PLX4720 (omitted). The conserved Lys and Glu in the N-lobe are shown in sticks, β-strands shown in blue and α helices in gold. The activation loop is disordered.

The protein kinase domains are typically 250-300 amino acids in size adopting a common core structure, which can be subdivided into a smaller N-terminal lobe and a bigger C-terminal lobe (Figure 8). A cleft between the two lobes contains the binding site for ATP utilizing a hydrophobic pocket formed by the N-lobe which helps orienting the nucleotide. The C-lobe is responsible for the substrate binding and the initiation of the phosphotransfer. The N-lobe consists of an antiparallel β-sheet made of five β-strands and one to three helices, including on prominent termed helix C, while the C-lobe consists mainly of α helices. The lobes are connected by a loop forming the back side of the ATP binding pocket.

Numerous kinase structures are available in their active and inactive conformation, with or without nucleotide or inhibitor bound. They have revealed a multitude of ways in which kinase activity is controlled. There are, however, several hallmark features, which can unambiguously be associated with the active and inactive states of the kinase. Those features should be regarded as necessary, but not always sufficient conditions of activity, because many kinases can display a continuum of activity in between their inactive and most active state. Going from a completely inactive to a more active state requires structural rearrangements which normally include changes of the position of helix C, a change of the DFG position and the structure of the activation loop (Nolen et al., 2004). Latter is frequently, but not always, phosphorylated at one or several positions in the active state of a kinase leading to a more structured conformation. Once activated the kinase toggles between an open and closed conformation while moving through the catalytic cycle (Taylor and Kornev, 2011). The hallmark features, conserved among the kinase family, can help defining the state of the kinase. One of the most important features of an active kinase is the highly conserved Lys^{72}-Glu^{91} pair found in the N-lobe (PKA numbering used in this section, pdb ID: 2CPK) which helps to orient the Lys^{72}. It can be observed forming a salt bridge. In active kinases the lysine is positioned to interact with the α and β phosphates of the ATP, thereby orientating it. Helix C in the N-lobe, including the conserved Glu^{91} residue, changes its position upon activation, helping the Lys-Glu salt bridge to form. A number of different arrangements of the helix C have been observed. While the C-terminus of the helix does not change too much in inactive kinases the N-terminus moves away from the C-lobe a movement which can also include a rotation (De Bondt et al., 1993). The interaction with the activation loop is often linked with the out position of the conserved DFG motif which couples, together with the Lys-Glu interaction, the conformation of the helix to nucleotide binding (Huse and Kuriyan, 2002).

The highly conserved DFG motif is located at the beginning of the activation loop in the C-lobe. The aspartate residue in the motif interacts with the three phosphates of the nucleotide, utilizing magnesium. The correct orientation of Phe^{185} is thought to position the aspartate residue in the DFG motif and to support the inward movement of the helix C. In some kinases the phenylalanine of the same motif changes its position between the DFG-in (inactive) and –out (active) position. The phenylalanine is part of a hydrophobic spine in the active state (also including Tyr^{164}, Leu^{106} and Leu^{95} in PKA). These residues form a hydrophobic stack connecting the core of the C-lobe to the N-lobe (Kornev et al.,

2006). It was suggested that this stack helps coordinate the movement of the lobes when going through the catalytic cycle. Yet the movement of the phenylalanine is not a precondition of the activation, as in Cdk2 or Src inactivation is achieved by a single twist of the DFG glycine (Jeffrey et al., 1995). The glycine residue performs a twist upon activation, changing its φ angle by 140° and forms a hydrogen bond to Asp^{184} (PKA), which turns towards the catalytic site (Kornev et al., 2006). This swing was also suggested as a primary event in the deactivation, loosening the surrounding and allowing for other changes as the flip of the DFG Phe^{186}. In order to reach the active conformation a number of events, caused by the phosphorylation, have been described switching from the inactive state, being finally stabilized by a number of hydrogen bonds (Kornev et al., 2006).

Positioned in the C-lobe, between the DFG and APE motifs and 20-30 residues in length (Figure 9) is the activation loop, playing the pivotal role during activation. The activation loop shows great conformational diversity between different kinases. Helix C and the activation loop seem to be structurally coupled as a change in one calls for the other to change as well.

The APE motif, which is part of the EF helix, is connected by a loop to the helix F, which has been suggested to form a platform on which the activation loop could fold. The loop shows no sequence conservation but extensive interactions with the activation loop. It changes its position as does the activation loop and may have coevolved with the activation loop facilitating its stabilization (Nolen et al., 2004). The P+1 loop is found before the APE motif and is named for its contacts with the P+1 residue of the substrate just after the phosphorylation site (Knighton et al., 1991). The P+1 loop has the capacity to undergo big conformational changes (Pike et al., 2008). Specific for active Ser/Thr kinases is a Thr/Ser residue in the P+1 loop, making an interaction with the Lys^{168} residue in the catalytic loop (Figure 9). This conserved threonine also interacts with the aspartate in the HRD motif. This interaction has been suggested to help locking the aspartate in the conformation needed for the interaction with the hydroxyl of the substrate.

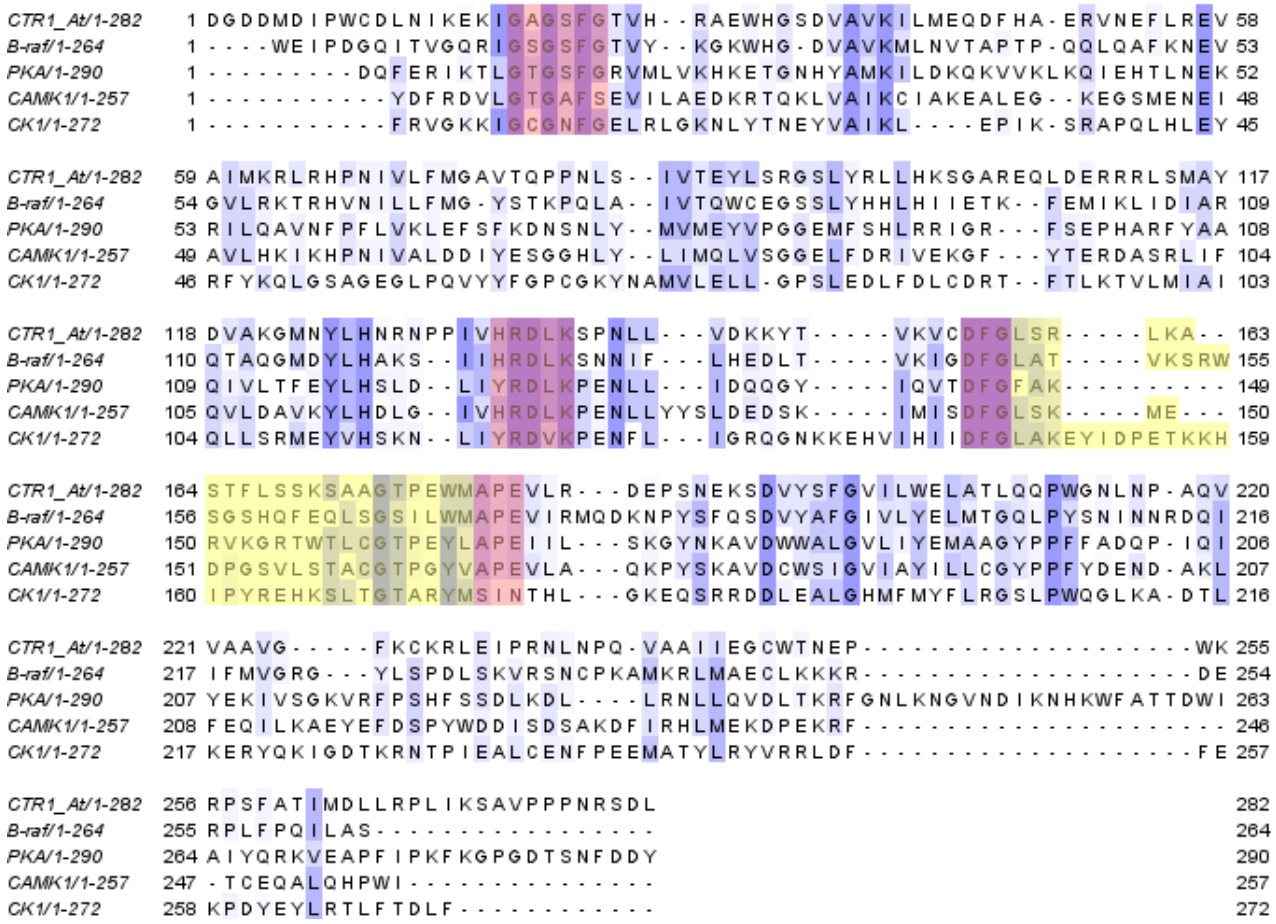


Figure 9: Alignment of the CTR1 kinase domain with four human Ser/Thr kinases. The activation loop is highlighted in yellow. Highly conserved motifs, crucial for the kinase action are shown in red. Residues are coloured in blue according to conservation.

In the kinase inactive state the activation loop is often found to be disordered or to bind to the active site, thereby blocking it as in case of the insulin receptor (Hubbard et al., 1994). The activation loop often gets phosphorylated at one or more positions, leading to an ordered activation loop and converting the enzyme from the inactive to the active state through conformational changes. Upon phosphorylation it adopts a conformation allowing for the substrate binding. Kinases requiring phosphorylation for activation have an arginine before the conserved catalytic aspartate residue and are therefore termed RD kinases (Figure 9). Here a primary phosphorylation site can be defined (Barrett and Noble, 2005) whose position is conserved between kinases. This phosphorylation site acts by neutralizing a number of positively charged residues, which repel each other in the absence of the phosphate group. Normally this basic pocket consists of a basic residue in the catalytic loop, one after the DFG motif and an amino acid in the helix C (basic pocket His^{87}, Arg^{165} and Lys^{189} in PKA, see also Figure 34). The arginine in the HRD motif of

Ser/Thr kinases interacts with the phosphoryl group of the primary phosphorylation site, supporting the conformation of the activation loop, while in tyrosine kinases the arginine forms a cation-π pair with a conserved phenylalanine (Krupa et al., 2004). At the beginning of the activation loop sits the $\beta 9$ strand forming an antiparallel sheet with the $\beta 6$ strand, being connect by three hydrogen bonds, and this sheet is disordered in inactive kinases. This sheet is important for the DFG orientation and positioning of helix C (Nolen et al., 2004). The attraction of the conserved basic residue after the DFG motif (Lys^{189} in PKA) to the primary phosphorylation site seems to be important for the formation of this sheet and to lock the N-terminal anchor of the activation loop. Kinases with a basic residue at this position normally require primary site phosphorylation. A positively charged residue in the helix C is often found in RD kinases interacting with the primary phosphorylation site in the activation loop (Protein kinase B His^{196}, PKA His^{87}).

Found in all kinases, located in the N-lobe in a turn between β-strand 1 and 2, is the glycine rich or P loop (Gly-x-Gly-x-φ-Gly, φ being usually a tyrosine or phenylalanine, Figure 9). It covers and anchors the α and β phosphates of ATP through backbone interactions of the glycines, while the aromatic side chain caps the site of phosphate transfer. In the C-lobe the catalytic loop with the Ser/Thr kinases consensus motif HRDxKxxN is found (Figure 9). The catalytic motif does not seem to be the recipient of activating signals as indicated by its relative static position upon activation. The highly conserved HRD aspartate Asp^{166} is responsible for the correct orientation of the substrate's phosphorylated hydroxyl side chain and acts as a catalytic base. The histidine moiety (tyrosine in some cases), binds to the carbonyl of the DFG aspartate and interacts with the DFG phenylalanine being part of the hydrophobic spine linking it to the N-lobe.

While phosphorylation plays an important role in the activation/deactivation process regulatory domains and other domains play also an important role. In RAF the N-terminal domain acts as a negative regulator of the C-terminal kinase domain (Stanton et al., 1989). Src kinase gets inactivated through intramolecular interactions of the kinase with the N-terminal SH2 and SH3 domains upon phosphorylation of a tyrosine residue, stabilizing the helix C in an inactive conformation (Sicheri et al., 1997). In the cyclin-dependent kinase Cdk2 helix C helix gets oriented in the active conformation upon binding of cyclin (Jeffrey et al., 1995). In CK1 the activation loop is extended allowing for additional contacts with the C-lobe stabilizing the loop and making phosphorylation unnecessary.

Aim of the work

The comparative simplicity of ethylene receptors, which lack extracellular signal binding domains, make them an interesting target for the structural study of signal transduction of membrane bound receptors. At the time of the start of this work, the structural information of the ethylene signalling pathway was still limited to the response regulator of ETR1. While numerous genetic and biochemical results became available over the last 10 years, no new structural information was added. A detailed functional description of this pathway, made even more interesting through the apparent link of a two component system with a MAPK pathway, would significantly contribute to the understanding of the initiation of transmembrane signalling and its primary intracellular event.

Therefore, the aim of the project was to gain insights in the cytoplasmic part of the receptors, focusing on subfamily 1, and the prime interaction partner CTR1. In the project the conditions for the heterologous expression in *E. coli* and purification of mg amounts of the domains and combinations thereof should be developed and based on these results structural and where applicable functional studies should be carried out.

The method of choice is X-ray crystallography and SAXS supported by the local access to synchrotron radiation. SAXS should be used to build a low resolution model which can be used as a scaffold to place the individual high-resolution domains, if the cytoplasmic full length protein fails to crystallize. This could allow assembling a complete model of the cytoplasmic ethylene receptor. While structural predictions of domains are able to give a first impression of the protein they still miss a lot of details and are in a number of cases misleading. Therefore high-resolution structural information is needed to verify models obtained from prokaryotic structures in the eukaryotic context and to add new, detailed information.

Materials and Methods

Bioinformatics

Sequences of the proteins (Uniprot identifiers: ETR1: P49333, ETR2: Q0WPQ2, ERS1: Q38846, ERS2: P93825, EIN4: Q9ZTP3, CTR1: Q05609, MKK9: Q9FX43) were submitted to the Pfam (http://pfam.sanger.ac.uk/) and together with information from the literature the domains were assigned. Jpred (www.compbio.dundee.ac.uk/www-jpred/) and PSIpred (bioinf.cs.ucl.ac.uk/psipred/) were used for secondary structure prediction. Additionally Globplot (www.globplot.embl.de), Disembl (dis.embl.de/) and IUPred (iupred.enzim.hu/) were used for the prediction of disordered regions. This information together with prior knowledge, comparison with structures available and manual inspection of the sequences was used for the design of the constructs.

The Surface Entropy reduction server (services.mbi.ucla.edu/SER/) (Goldschmidt et al., 2007) was used to select residues to be mutated in an attempt to improve the crystallizability.

Construct List

Table 1: List of all designed constructs and their final status. (**GAF** cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA domain, **DHp** dimerization and histidine phosphotransfer domain, **CA** catalytic and ATP-binding domain and **RD** receiver domain)

Name	Protein	Start	End	Domains	Status
H1	ETR1	140	588	GAF and HKD	Low expression
H2	ETR1	140	321	GAF	Purified
H3	ETR1	327	407	DHp	Purified
H4	ETR1	418	588	CA short	Low expression
H5	ETR1	411	588	CA long	Purified
H6	ETR1	327	588	Histidine kinase domain	Purified
H22	ETR1	142	321	GAF domain	Purified
H30	ETR1	327	411	DHp	Purified
H33	ETR1	158	588	GAF + Histidine kinase domain	No Expression
H36	ETR1	308	407	DHp	Crystal
H37	ETR1	158	309	GAF	Expression

H44	ETR1	308	588	Histidine kinase domain	Purified
H46	ETR1	158	738	GAF + HK + RD	Solubility problems
H47	ETR1	308	738	Histidine kinase domain + RD	Purified
H68	ETR1	308	582	Histidine kinase domain	Purified
H69	ETR1	308	592	Histidine kinase domain	Purified
H70	ETR1	308	596	Histidine kinase domain	Purified
H71	ETR1	308	600	Histidine kinase domain	Purified
H72	ETR1	308	604	Histidine kinase domain	Purified
H78	ETR1	158	407	GAF+ DHp	Purified
H94	ETR1	165	407	GAF + DHp	Purified
H13	ETR2	640	771	Receiver domain	Low expression
H17	ETR2	367	616	Histidine kinase domain	No Expression
H29	ETR2	332	615	Histidine kinase domain	No Expression
H27	ETR2	350	615	Histidine kinase domain	No Expression
H80	ETR2	338	773	Histidine kinase domain + RD	No Expression
H83	ETR2	338	427	DHp	Low expression
H7	ERS1	132	589	GAF and HKD	Low expression
H8	ERS1	132	314	GAF	No expression
H9	ERS1	326	407	DHp	Crystal
H10	ERS1	420	589	CA short	Expression
H11	ERS1	414	589	CA long	Expression
H12	ERS1	326	589	Histidine kinase domain	Purified
H24	ERS1	329	589	Histidine kinase domain	Purified
H34	ERS1	158	589	GAF + Histidine kinase domain	No expression
H35	ERS1	308	407	DHp	Structure
H38	ERS1	158	309	GAF	No expression
H45	ERS1	308	589	Histidine kinase domain	Purified
H73	ERS1	308	586	Histidine kinase domain	Purified
H74	ERS1	308	594	Histidine kinase domain	Purified
H75	ERS1	308	599	Histidine kinase domain	Purified
H76	ERS1	308	604	Histidine kinase domain	Purified
H77	ERS1	308	613	Histidine kinase domain	Purified
H79	ERS1	158	407	GAF + Dhp	No Expression
H93	ERS1	308	407	DHp	Purified
H18	ERS2	382	645	Histidine kinase domain	No Expression

H28	ERS2	369	579	Histidine kinase domain	No Expression
H14	EIN4	636	766	Receiver domain	Purified
H16	EIN4	367	613	Histidine kinase domain	No Expression
H19	EIN4	642	766	Receiver domain	Expression
H20	EIN4	644	766	Receiver domain	Expression
H21	EIN4	644	761	Receiver domain	Expression
H26	EIN4	350	611	Histidine kinase domain	No Expression
H81	EIN4	339	766	Histidine kinase domain + RD	No Expression
H84	EIN4	339	430	DHp	Low expression
H15	CTR1	540	821	Kinase domain	Structure
H39	CTR1	540	811	Kinase domain	Structure
H41	CTR1	156	408	N terminal domain	No Expression
H43	CTR1	156	811	Full length∆155	No Expression
H64	CTR1	1	530	N-terminal domain	No Expression
H65	CTR1	54	530	N-terminal domain	No Expression
H66	CTR1	54	408	N-terminal domain	No Expression
H67	CTR1	1	408	N-terminal domain	No Expression
H82	CTR1	1	821	Full length	No Expression
H88	CTR1	204	409	N-terminal domain	No Expression
H89	CTR1	236	409	N-terminal domain	No Expression
H87	EIN2	478	1270	C-terminal domain	Low expression
H90	MKK9	9	310	Full length	Purified

Table 2: List of all prepared mutants and the obtained result

Name	Protein	Based on	Mutation(s)	Status
M10	ETR1	H78	EE177AA	No Expression
M11	ETR1	H78	K271A	No mutant
H52	ETR1	H44	E373T	Purified
H53	ETR1	H44	KK437TT	No Expression
H54	ETR1	H44	KQ476TT	Purified
M1	ERS1	H74	GCG519ACA	Purified
M2	ERS1	H74	H353A	Purified
M3	CTR1	H39	SFGT561AFGA	Purified
M4	CTR1	H39	TFLS704AFLA	Structure
M5	CTR1	H39	T800A	Purified

M6	CTR1	M4	S710A	Purified
M12	CTR1	M6	T714A	Cloned
M14	CTR1	M6	S703A	Purified
M15	CTR1	H39	S703E	Purified
M16	CTR1	H39	T704E	Purified
M17	CTR1	H39	S707E	Purified
M18	CTR1	H39	S710E	Purified
M19	CTR1	H39	T714E	Purified
M20	CTR1	H39	R604H	Purified
M21	CTR1	H39	F611A	Purified
M22	CTR1	H39	M612W	Purified
M23	CTR1	H39	R604A	Purified
M24	CTR1	M23	R602D	Purified
M25	CTR1	H39	FM611AW	Purified
H40	CTR1	H39	D676N	Structure
H42	CTR1	H15	S561E, T564T, S710E, T800E	Structure
M27	MKK9	H90	D167N	Purified

DNA electrophoresis and purification

DNA fragments were analysed on an agarose gel. The agarose gel was made by dissolving 0.8-1% w/v agarose in 1x TBE buffer and boiling. After cooling to 55°C ethidium bromide was added (final concentration of 0.5 µg/ml). The solution was poured into a suitable chamber and, after placing a comb, was allowed to cool and set at RT. Samples were mixed with a 6x loading buffer (glycerol (30%), 0.25% bromophenol blue) and loaded into the slots of the gel. 1x TBE buffer was used as a running buffer. The 100 bp or 1 kbp Gene Ruler DNA Ladder (Fermentas, St. Leon-Rot) were used for size determination. Gels were run at 100 V for 30 min – 1 h depending on the expected size. Bands were cut out of the gel and subsequently purified using the QIAquick Gel Extraction Kit (Qiagen, Hilden) according to the protocol provided by the manufacturer.

TBE Buffer

106 g/l Tris

55 g/l Boric acid

9.3 g/l EDTA pH 8.0

SDS PAGE

In order to separate the proteins by their size SDS-PAGE was used. Normally a 5% stacking gel combined with a 12% separating gel was employed. Gels with 10 to 15 slots for samples were run in at maximum 120 mA for 1h. Prior to loading the sample was mixed with a 4x sample buffer and heated to 95°C for 1min. As a molecular marker the ColourPlus Protein marker (NEB, Frankfurt), the ROTI-Mark (Carl Roth Gmbh, Karlsruhe) or the PAGE Ruler Unstained protein ladder (Fermentas, St. Leon-Rot) were used. As a running TGS buffer (12 g Tris, 57.6 g glycine, 2 g SDS/l) was used. After the run the stacking gel was discarded and the running gel was either stained with Coomassie Brilliant Blue (0.1% Coomassie R250, 10% glacial acetic acid 400 ml methanol 500 ml ddH2O) or the Coomassie blue safe stain. The Coomassie blue safe stain requires the gel to be three times rinsed and boiled in water prior to staining, while Coomassie Brilliant Blue needs to be destained using 20% methanol, 10% glacial acetic acid in ddH2O. The Coomassie blue safe stain is removed after staining and the gel is left in water and does not need to be destained.

PCR

Primers with the appropriate overhangs for restriction enzyme, LIC (see page 44) or PCR cloning (see page 43) were supplied by Biomers (Germany, Ulm). For the PCRs KOD Hot Start polymerase (Merck, Darmstadt), Phusion Flash High-Fidelity PCR Master Mix (NEB, Frankfurt) or Pfu High-fidelity DNA Polymerase (Promega, Mannheim) were used. For colony PCRs Taq Polymerase (NEB, Frankfurt) was used. The PCR reactions were carried out according to the respective protocol provided with the polymerase. A typical PCR setup and reaction is depicted below. Annealing temperature and the extension temperature were adjusted as needed in the specific reaction. When needed a PCR Touch-Down protocol was employed were the annealing temperature was dropped by 2°C for 5 cycles and was continued from this point as usual, keeping the annealing temperature of the 5th cycle.

PCR setup

0.5 µl	template (typical concentration 50 ng/µl)
5 µl	10x PCR reaction buffer
2 µl	forward primer (10 pmol/µl)
2 µl	reverse primer (10 pmol/µl)

PCR reaction

1.	95°C	2 min
2.	95°C	30 sec
3.	60°C	30 sec
4.	72°C	20 sec/kb

5 µl	dNTP mixture (2 mM)	repeat 2-4 30x		
1 µl	Polymerase	5.	72°C	5min
Add water up to 50 µl		6.	4°C	∞

For colony PCR reactions one primer specific for the vector and the other primer specific for the insert were used. A colony was picked and inoculated in the prepared reaction setup and an annealing temperature of 57°C was used. Colonies containing the insert of interest were identified by gel electrophoresis of the PCR product.

Mutagenesis PCR and PCR cloning

For mutagenesis reactions PfuUltra High-fidelity DNA Polymerase AD (Agilent, Waldbronn) was used. Mutagenesis reactions were carried out according to the Stratagene Quikchange protocol using 20 cycles per reaction. Non-mutated parental plasmid was digested for 90 min at 37°C using DpnI (NEB, Frankfurt) and 1.5 µl were used for the subsequent transformation.

PCR cloning (van den Ent and Löwe, 2006) follows a similar protocol as mutagenesis yet instead of the mutagenesis primers an insert containing suitable overhangs to the vector was used. The insert was produced in a prior PCR step and gel purified. The overhangs anneal with the selected sequences in the vector acting as a mega mutagenesis primer. After the extension of the vector now fused with the insert the template vector was digested for 90 min at 37°C using DpnI (NEB, Frankfurt) and 1.5 µl were used for transformation.

Plasmid purification

The plasmid of interest, typically present in the *E. coli* cloning strains SoloPack® Gold Competent (Agilent, Waldbronn) or DH5α, was amplified by growing the cells in 5ml LB media supplied with the appropriate antibiotic at 37°C with vigorous shaking. After 14-16 h the cells were harvested by centrifugation and the plasmid was extracted using the QIAprep Spin Miniprep Kit (Qiagen, Hilden), as described in the manufacturers protocol.

Restriction enzyme digestion

Inserts for the restriction enzyme cloning method were gel purified prior to digestion as described and subsequently incubated at the same time with XhoI/NcoI (NEB,

Frankfurt) at 37°C for 2 h according to the manufacturer's protocol. They were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden) according to the manufacturer's protocol.

The vectors were digested following the same protocol as for the insert. They were purified by gel electrophoresis as a final step as decribed.

Ligation

Ligations were carried out by mixing 1.5 µL 10X Quick ligase buffer, 5 µl of digested insert, 2 µl of digested vector, 1.5 µl of T4 DNA ligase (NEB, Frankfurt) and H_2O up to 15 µl. The reaction was incubated at RT for 3h and then heated to 70°C for 5min. An aliquot of 3 µl was used for the subsequent transformation step; the rest was stored at -20°C.

Ligation independent cloning (LIC)

LIC depends on the vector and the insert having complementary overhangs, which anneal in the absence of ligase. To make complementary overhangs with the vector the primers needed a 5' extension, for the forward primer CAGGGCGCCATGgeneofinterest and for the reverse primer GACCCGACGCGGTTAgeneofinterest. The vector is cleaved. The purified insert and the vector are treated with the T4 polymerase utilizing its 3'->5' exonuclease activity in order to create specific overhangs. The digestion is carried out in the presence of a single dNTP species and comes to a halt when the digesting polymerase reaches this nucleotide. The polymerase enters there a steady-state equilibrium of adding and removing the last nucleotide. Thereby a complementary overhang with a specific length is created.

As a first step 50 µl of vector (1-5 µg of DNA) are cleaved with BsaI as described for restriction enzyme cloning. The insert is amplified as described. The purified vector and insert are treated with the T4 polymerase

2 µl	10x NEB buffer 2
15.9 µl	BsaI digested LIC vector or insert
0.5 µl	of the appropriate nucleotide (dTTP for the vector, dATP for the insert)
1 µl	100 mM DTT
0.2 µl	100x BSA
0.4 µl	T4 DNA polymerase (3 U/µl)

The solution is mixed and incubated at room temperature (RT) for 30 min followed by 20 min at 75°C to inactivate the T4 polymerase. The vector and the insert can be purified at this step for longer storage. If used soon after the digestion purification is not needed. 1 µl of T4 polymerase treated vector was mixed with 2 µl of T4 polymerase treated insert. The annealing reaction is incubated at RT for 5 min. 1 µl EDTA (25 mM) is added and mixed by stirring gently. Following 5-60 min incubation at RT (exact time does not influence result) the reaction mixture is used for transformation as described.

Transformation

When doing a transformation following restriction enzyme cloning or ligation independent cloning 4 µl of the vector-insert reaction were used. Only 0.5 µl of vector were used in case of transformations into expression strains when the vector had been prepared beforehand and was at a concentration of about 50 ng/µl. Chemically competent cells were thawed on ice and for each reaction 40 µl of cells were used. For cloning either SoloPack® Gold Competent (Agilent, Waldbronn) or DH5α cells were used. For protein expression transformations were done in BL21(DE3), Rosetta(DE3), Rosetta(DE3) pLys, BL21(DE3) RP, BL21(DE3) RIL or BL21(DE3) chaperone complemented. Cells mixed with the vector were incubated on ice for 15 min followed by a heat shock at 42°C for 50 sec. Afterwards the cells were cooled on ice for 2 min and 200 µl SOC media was added. The cells were incubated for 1 h at 37°C with moderate agitation. Subsequently the mixture was plated on pre warmed agar plates containing the appropriate antibiotic and stored at 37°C overnight.

Expression

Freshly transformed cells were used to inoculate 5 ml Luria-Bertani (LB) medium (containing the appropriate antibiotic) and were grown overnight at 37°C. The overnight cultures were used to inoculate LB, TB (Terrific Broth) or autoinduction medium (Studier, 2005). They were grown in 2 L baffled flasks at 37°C and 210 rpm to an OD_{600} between 0.7 and 0.9 in case of LB and autoinduction medium and up to OD_{600} of 2 in case of TB media. Thereafter the temperature was lowered to 20°C and cells grown in LB or TB medium were induced with typically 0.4 mM Isopropyl-β-D-thiogalactopyranosid (IPTG) as soon as the lower temperature had been reached. Cultivation was continued for 18 h and

the cells were harvested by centrifugation at 5.500 rpm in a JLA-8.1000 rotor for 25 min at 4°C. The cell pellets were immediately frozen and stored at -20°C.

- The DHp domain of ERS1 of ETR1 were expressed using *E. coli* BL21(DE3) cells in 2L of LB media. The cells were grown to an OD_{600} of 0.6, cooled to 20°C and induced with 0.4 mM IPTG and harvested after 18 h.
- The various constructs of the ETR1 histidine kinase domain were expressed in *E. coli* BL21(DE3) cells co-expressing chaperones DnaK, DnaJ, GrpE, ClpB, GroEL and GroES (**c**haperone **c**omplement 4, cc4) in 1L of TB media. The cells were grown to an OD_{600} of 1.5, cooled to 20°C and induced with 0.4mM IPTG and harvested after 18 h.
- The various constructs of the ERS1 histidine kinase domain were expressed in *E. coli* BL21 cc4 cells (**c**haperone **c**omplement 4, cc4) in 1 L of TB media. The cells were grown to an OD_{600} of 1.5, cooled to 20°C and induced with 0.4 mM IPTG and harvested after 18 h.
- The full cytoplasmic construct of ETR1 (H46) was expressed in Rosetta cells (containing codons rarely used in *E. coli*) in 2 L of TB media. The cells were grown to an OD_{600} of 1.5, cooled to 20°C and induced with 0.05 mM IPTG and harvested after 18 h.
- The GAF and DHp domains of ETR1 (H78) were expressed in Rosetta cells in of media. The cells were grown to an OD_{600} of 2, cooled to 20°C and induced with 0.4 mM IPTG and harvested after 18 h.
- The kinase domain of CTR1 was expressed in *E. coli* strain BL21 cc4 cells. The overnight cultures were used to inoculate 2 L of autoinduction medium. They were grown to an OD_{600} of between 0.7 and 0.9 after which the temperature was lowered to 20°C. Cultivation was continued for 18 h.

SeMet Expression

The media for the SeMet labelling was prepared according to the following protocol:

- 100 ml M9 x 10 (For 1 L: 80 g Na_2HPO_4, 40 g KH_2PO_4, 5 g NaCl, 5 g NH_4Cl)
- 10 ml Trace elements solution (100x) (for 1 L 5 g EDTA, 0.83 g $FeCl_3$ x 6 H2O, 84 mg $ZnCl_2$, 13 mg $CuCl_2$ x 2 H_2O, 10 mg $CoCl_2$ x 6 H_2O, 10 mg H_3BO_3, 1.6 mg $MnCl_2$ x 6 H_2O)
- 20 ml Glucose 20% (w/v)

- 1 ml $MgSO_4$ 1M
- 300 µl $CaCl_2$ 1M
- 1 ml Biotin (1 mg/ml)
- 1 ml Thiamin (1 mg/ml)
- Amino acids mix (40 mg/ml of 19 amino acids except methionine)

The plasmid was transformed into the methionine auxotroph B834 (DE3) strain. A 50 ml preculture was prepared with appropriate antibiotics added and grown overnight at 37°C. 15 ml of the preculture were centrifuged and resuspended in 5 ml of the growth media and added to the flask containing 500 ml of the growth media. Subsequently seleno-methionine was added under the fume hood to a final concentration of 50 mg/L. The flasks were incubated at 37°C at 210 rpm. As soon as the OD_{600} reached 1.0 the temperature was lowered to 27°C and the cells were induced with 0.4 mM IPTG and incubation was continued overnight and harvested as described before.

Cell lysis and NiNTA purification

The initial lysis buffer and its pH were chosen individually for each protein (Table 3) depending on the pI. Later, initial results were refined also taking into account the outcomes of Thermofluor (see page 50) and solubility screening. After thawing on ice, cell pellets were resuspended in lysis buffer in a final volume of 20 ml per 5 g of cells and lysed by pulsed sonication on ice for 2 × 90 s. The specific buffer used for each construct can be found in Table 3. The lysates were centrifuged at 19.000 rpm using an SS-34 rotor at 4°C for 60 min. The supernatants were filtered through a 0.22 µm membrane. The NiNTA column had been stripped using 0.5 M EDTA, washed with water followed by 2 column volumes (CV) 8 M Urea washed again with water followed by equilibrated against 100 mM $NiSO_4$ and washed again with 5 CV water. For a preparative purification the supernatant was loaded onto a 5 ml Ni-NTA Superflow Cartridge (Qiagen, Hilden) and for expression screening 0.5 ml NiNTA in a gravity flow column was used equilibrated against buffer A (20 mM buffer, 250 mM NaCl, 5% glycerol and 3 mM β-ME). In case of H39 and H46 0.1% (w/v) CHAPS was added in the buffer A. The column was washed with four column volumes (CV) of buffer A followed by 3 CV of buffer A with 10% buffer B (buffer A containing 500 mM imidazole). The proteins were eluted with a gradient of 10-70% buffer B in buffer A (50-350 mM imidazole) within 8 CV. Fractions were collected and analysed on a SDS PAGE gel.

Table 3: Buffer conditions used during purification for different constructs

Construct	Customized lysis buffer for selected constructs
H35	20 mM Tris pH 8.5, 150 mM NaCl, 5% v/v glycerol, 20 mM Imidazole, 1 mM EDTA-free protease inhibitors, 1 mg ml^{-1} DNAse
H36	20 mM Tris pH 8.0, 100 mM NaCl, 5% v/v glycerol, 20 mM Imidazole, 1 mM EDTA-free protease inhibitors, 1 mg ml^{-1} DNAse
H39	20 mM HEPES pH 7.5, 350mM NaCl, 3 mM β-ME, 5% v/v glycerol, 0.1% w/v CHAPS, 20 mM Imidazole, 1 mM EDTA-free protease inhibitors, 1 mg ml^{-1} DNAse
H46	20 mM Tris pH 8.8, 150 mM NaCl, 1 M Urea, 3 mM β-ME, 0.1% w/v CHAPS, 20 mM Imidazole, 1 mM EDTA-free protease inhibitors, 1 mg ml^{-1} DNAse
H47	20 mM Tris pH 8.5, 250 mM NaCl, 0.1% w/v CHAPS, 20 mM Imidazole, 1 mM EDTA-free protease inhibitors, 1 mg ml^{-1} DNAse
H78	20 mM CHES pH 9.0, 250 mM NaCl, 3 mM β-ME, 5% v/v glycerol, 0.1% w/v CHAPS, 20 mM Imidazole, 1 mM EDTA-free protease inhibitors, 1 mg ml^{-1} DNAse

Dialysis and 2nd NiNTA step

Dialysis was carried out at 4°C. The buffer was chosen according to the buffer used for the NiNTA step. 2 L of 20 mM buffer, 150 mM NaCl and 6 mM β-ME were prepared and allowed to cool down. The protein was mixed with TEV in a minimum ratio of 1 mg TEV : 50 mg protein and filled into a rinsed dialysis tube with an appropriate cut-off for the protein. The sealed dialysis bag was put into the buffer solution, which was stirred at 100 rpm and incubated overnight.

The protein solution in the dialysis tube was collected the next day and passed over a NiNTA column equilibrated against the dialysis buffer. The flow-through was collected and used for further steps. The flow-through and the elution were analysed using a SDS-PAGE gel to assess the uncleaved fraction and the remaining contaminations.

Size-exclusion chromatography

As a final step of purification size-exclusion chromatography (SEC) was used utilizing either a HiLoad 26/60 Superdex 75 or 200 column (Amersham Biosciences)), depending on sample volume. The column was pre-equilibrated in a buffer chosen for each individual

construct. The protein was concentrated using a Vivaspin concentrator (Sartorius, Göttingen) to less than 5 ml (if possible down to 2 ml) and loaded onto the column. Fractions were collected and analysed on a SDS PAGE gel. Fractions containing pure protein were pooled and again concentrated using a Vivaspin concentrator.

Coupled assay

Activity of the purified CTR1 kinase domain and various mutants was determined using a coupled *in vitro* kinase assay (Cook et al., 1982). The reaction buffer consisted of 25 mM HEPES pH 7.5, 350 mM NaCl, 1 mM DTT, 10 mM MgCl2, 1% glycerol 1 mM NADH, 2 mM phosphoenolpyruvate, 2.2 U lactate dehydrogenase, 1.5 U pyruvate kinase and 250 nM kinase unless indicated otherwise. The reaction was started by the addition of 1 mM ATP, which was serially diluted in 100 µl reaction volume at 30°C. The assay was recorded over a period of 1 h using the microplate reader Power Wave X Select (BIO-TEK Instruments) at a wavelength of 340 nm. The maximum reaction velocity was calculated from the slope of the resulting time curve as change in absorbance at 340 nm. The data followed simple Michaelis-Menten kinetics, which was applied for the calculation of kinetic constants with a nonlinear least-squares best fit of the data (Graphpad Prism5, Graphpad Software).

Kinase-Glo® Luminescent Kinase Assays

The assay measures the amount of ATP remaining after the kinase reaction. A luminescent signal is generated, which correlates with the amount of ATP still present. The assay used is linear up to 500 µM ATP. The protein (about 1 mg/ml) is mixed with ATP (in the range of 50-300 µM) in a total volume of 20-50 µl in a 96well plate and incubated for typically 30 min at RT. The same amount of stabilized luciferase (Ultra-Glo™ Luciferase, Promega, Mannheim) is added and the signal is allowed to stabilize for 5min. The luminescent signal is measured using an Infinite M1000 (Tecan, Männedorf, CH) and compared to the signal obtained of a sample in the absence of enzyme.

Mass spectroscopy

Samples were diluted with 0.1% TFA to a final concentration of 1-10 pmol/ µL with the optimum being 5pmol. Subsequently 2 µl were mixed with 2 µl of sinapinic acid (10 mg/ml) and 1 µl was spotted and dried on a plate which was injected into an Applied

Biosystems VOYAGER-DE STR workstation MALDI TOF mass spectrometer. 50 spectra were acquired and summed up during one measurement.

Thermofluor assay

Folded and unfolded proteins can be distinguished through exposure to a hydrophobic fluoroprobe. In water and in the absence of a hydrophobic surface the fluorescence probe is quenched. If a hydrophobic residue is exposed during the unfolding of the protein the quenching is reduced, which leads to a sharp increase of detectable fluorescence emission. Unfolding of the protein is induced by a stepwise increase in temperature causing an unfolding process. Normally a relatively sharp transition between the folded and unfolded states can be observed, where the melting temperature Tm is defined as the inflection point of the curve recording the change in fluorescent signal. The thermofluor assay is performed in a commercially available real-time PCR instrument (MyiQ, Biorad). 94 conditions can be screened simultaneously in order to identify stabilizing buffers and ligands (Table 4).

Typically a protein concentration of about 2 mg/ml, depending on the protein, was used in order to achieve sufficient signal strength. If possible 500 µl of protein were mixed with 500 µl of water to reduce the effect of the buffer used for purification. If this was not possible, due to the instability of the protein in the diluted buffer, 1ml of protein at 1 mg/ml was used. 3 µl of 5000x Sypro Orange (Invitrogen, Darmstadt) were added. 10 µl of this mixture were added to 40 µl of buffer solution already present in the 96-well thin-wall PCR plate (Nerbe plus, Winsen). The plate was sealed with an Optic View cover seal (Henze Laborbedarf, Elmshorn) and cooled to 4°C for 2 min followed by heating from 4 to 85°C in increments of 0.5°C for 20sec at each temperature. The wavelengths for excitation and emission were 490 and 575 nm, respectively. The customized screen consisted of a pH screen from pH 4.5 to pH 10, the anion and cation Hofmeister series, a number of nucleotides, different salt and glycerol concentrations and a number of other additives reported to stabilize the protein.

Table 4: List of the 94 buffer conditions used in the thermofluor screen

Well	Buffer	Salt	Well	Buffer	Salt
A1	60 mM Citric acid pH 4.0		**E1**	2 mM NAD	5 mM MgCl2
A2	60 mM Na-acetate pH 4.4		**E2**	2 mM NADH	5 mM MgCl2
A3	60 mM Na-citrate pH 5.0		**E3**	2 mM NADP	5 mM MgCl2
A4	60 mM Na-citrate pH 5.5		**E4**	100 mM NH4Cl	

A5	60 mM Bis-Tris pH 6.0		**E5**	100 mM KCl	
A6	60 mM Sodium phosphate pH 6.0		**E6**	100 mM NaCl	
A7	60 mM MES pH 6.2		**E7**	100 mM LiCl	
A8	60 mM Bis-Tris propane pH 6.5		**E8**	100 mM MgCl2	
A9	60 mM ADA pH 6.5		**E9**	40 mM CaCl2	
A10	60 mM PIPES pH 6.7		**E10**	40 mM MnCl2	
A11	60 mM Sodium phosphate pH 7.0		**E11**	4 mM NiCl2	
A12	60 mM MOPS pH 7.0		**E12**	10 mM FeCl3	
B1	60 mM Ammonium acetate pH 7.3		**F1**	40 mM ZnCl2	
B2	60 mM Hepes pH 7.5		**F2**	100 mM Sodium sulfate	
B3	60 mM Imidazol pH 7.7		**F3**	100 mM Sodium phosphate	
B4	60 mM Tris-HCl pH 8.0		**F4**	100 mM Na2-malonate	
B5	60 mM Bicine pH 8.0		**F5**	100 mM Sodium acetate	
B6	60 mM Tris-HCl pH 8.5		**F6**	100 mM Sodium formate	
B7	60 mM CHES pH 9.0		**F7**	100 mM Sodium citrate	
B8	60 mM CHES 9.5		**F8**	100 mM Na2-tartrate	
B9	60 mM CHES 10.0		**F9**	100 mM Sodium nitrate	
B10	60 mM Citric acid pH 4.0	250 mM NaCl	**F10**	100 mM Sodium bromide	
B11	60 mM Na-acetate pH 4.5	250 mM NaCl	**F11**	100 mM Sodium iodide	
B12	60 mM Na-citrate pH 5.0	250 mM NaCl	**F12**	100 mM Na-thiocyanate	
C1	60 mM Na-citrate pH 5.5	250 mM NaCl	**G1**	2.5% Glycerol	
C2	60 mM Bis-Tris pH 6.0	250 mM NaCl	**G2**	5% Glycerol	
C3	60 mM Na/K phosphate pH 6.0	250 mM NaCl	**G3**	10% Glycerol	
C4	60 mM MES pH 6.2	250 mM NaCl	**G4**	15% Glycerol	
C5	60 mM Bis-Tris propane pH 6.5	250 mM NaCl	**G5**	20% Glycerol	
C6	60 mM ADA pH 6.5	250 mM NaCl	**G6**	2 mM TCEP	
C7	60 mM PIPES pH 6.7	250 mM NaCl	**G7**	10 mM TCEP	
C8	60 mM Na/K phosphate pH 7.0	250 mM NaCl	**G8**	60 mM NaCl	
C9	60 mM MOPS pH 7.0	250 mM NaCl	**G9**	100 mM NaCl	
C10	60 mM Ammonium acetate pH 7.3	250 mM NaCl	**G10**	150 mM NaCl	
C11	60 mM Hepes pH 7.5	250 mM NaCl	**G11**	200 mM NaCl	
C12	60 mM Imidazol pH 7.7	250 mM NaCl	**G12**	400 mM NaCl	
D1	60 mM Tris-HCl pH 8.0	250 mM NaCl	**H1**	600 mM NaCl	
D2	60 mM Bicine pH 8.0	250 mM NaCl	**H2**	800 mM NaCl	
D3	60 mM Tris-HCl pH 8.5	250 mM NaCl	**H3**	1000 mM NaCl	
D4	60 mM CHES pH 9.0	250 mM NaCl	**H4**	2 mM CHAPS	
D5	60 mM CHES 9.5	250 mM NaCl	**H5**	0.5% octyl-b-D-glucoside	
D6	60 mM CHES 10.0	250 mM NaCl	**H6**	3% D-glucose	
D7	Water		**H7**	3% sucrose	
D8	2 mM ADP	5 mM MgCl2	**H8**	30 mM L-Arg	30 mM L-Glu
D9	2 mM AMP	5 mM MgCl2	**H9**	5% PEG 1000	
D10	2 mM AMPPcP	5 mM MgCl2	**H10**	5% PEG 4000	
D11	2 mM dCMP	5 mM MgCl2	**H11**	Free – own choice	
D12	2 mM dGMP	5 mM MgCl2	**H12**	Free – own choice	

Limited proteolysis

Limited proteolysis is used to determine domain boundaries or detect exposed loop structures. Only protein regions exposed on the surface either present in a loop or in unstructured regions are normally subject to proteolytic cleavage. The protein core or stable secondary structure elements are usually not accessible for proteases. The reaction

conditions are chosen in such a way that mild cleavage is achieved. Small samples are taken from the reaction at different time points and analysed via SDS PAGE. The method is used to identify stable, well-behaved protein fragments.

The protein (5 mg/ml) was incubated on ice and mixed with trypsin and chymotrypsin at a concentration of 1:200. Samples were taken after 5, 10, 15, 20, 40 and 60 min and analysed by SDS PAGE. Fragments that were abundant and of reasonable size were cut out of the gel and analysed by peptide fingerprinting (Proteomic core facility, EMBL Heidelberg) in order to identify the domain boundaries.

SAXS

SAXS measurements were carried out at beamline X33 (EMBL/DESY, Hamburg) (Roessle et al., 2007) at the DORIS-III storage ring using a Pilatus 1M detector. The purified cytosolic domains of ETR1 (construct H46, 64 kDa) was measured at three concentrations (3.0, 1.5 and 0.75 mg/ml) in a buffer containing 20 mM Tris pH 8.8, 150 mM NaCl, 1 mM DTT and 250 mM 3-(1-Pyridinio)-1-propanesulfonate (NDSB). Concentrations were determined using either Bradford assay (Biorad, Munich) or by refractive index using a Rudolph Research Analytical J357 refractometer, due to the interference of NDSB in UV absorbance measurements. For each measurement eight 15 sec frames were collected and averaged, using X-rays at λ=1.5 Å and a sample volume of 50 µl. The measurements were carried out at 8°C using a sample-detector distance of 2.7 m, covering the momentum transfer range 0.10 nm-1 < s < 5 nm-1 ($s = 4\pi \sin(\theta)/\lambda$ where 2θ is the scattering angle). Prior to and following each sample exposure the buffer excluding the protein was measured to allow for background subtraction.

The data were processed using standard procedures, corrected for buffer contribution, scaled for solute concentration and extrapolated to infinite dilution using the program PRIMUS (Konarev et al., 2003). The radius of gyration R_g and forward scattering intensity *I(0)* were determined using Guinier analysis (Guinier, 1939), assuming that at very small angles ($s \leq 1.3/R_g$) the intensity is represented as $I(s)=I(0)\exp((sR_g)2/3)$. R_g and forward scattering *I(0)* were also determined, along with the maximum particle dimension D_{max} using the indirect Fourier transformation approach in the program GNOM (Svergun, 1992). The molecular mass (MM_{SAXS}) of the construct was calculated by comparison of the interpolated forward scattering with reference BSA samples. The

scattering patterns from available high-resolution models were calculated by the program CRYSOL (Svergun et al., 1995) in order to obtain the fit of the model to the data.

Ab initio models were reconstructed from the H46 scattering data using the simulated annealing program DAMMIF (Franke and Svergun, 2009), which is independent of the actual protein concentration. The average of 10 different models, obtained from DAMMIF, was used to generate a representative shape envelope, using the program DAMAVER (Volkov and Svergun, 2003). The discrepancy of the experimental data was calculated as:

$$\chi^2 = \frac{1}{N-1}\sum_j \left[\frac{I_{\exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)}\right]^2$$

where N is the number of experimental points, c is a scaling factor, and $I_{calc}(s)$ and $\sigma(s_j)$ are the calculated intensity and the experimental error at the momentum transfer s_j, respectively. In addition, the ensemble of *ab initio* models was used to calculate an excluded volume of the particle, Vol_{DAM}, which is independent of the estimated protein concentrations. Rigid-body models were generated for the all constructs using the program CORAL (Petoukhov and Svergun, 2005), where high-resolution models of domains are defined as rigid bodies and linkers/loops between the individual subunits are represented as random polypeptide chains.

Crystallization

The concentration of the protein was determined from the absorption at 280 nm using a NanoDrop ND-1000 spectrophotometer and a theoretical absorption coefficient. Suitable concentrations for crystallization were estimated using the pre crystallization test (Hampton, Aliso Viejo, CA) following the protocol provided. Initial crystallization trials were carried out with different 96-well screens from Qiagen or Jena Bio Science at the EMBL Hamburg high-throughput crystallization facility (Mueller-Dieckmann, 2006). All initial screens were performed using the sitting-drop vapour-diffusion method at 20°C in 96 well sitting drop low profile plate (Greiner Bio-One, Frickenhausen). Typically 500 nl protein solution was mixed with 500 nl reservoir solution and equilibrated against 50 µl reservoir solution.

Crystals that were suitable for data collection were obtained through customized grid screens using 24well hanging drop plates. Additionally the Qiagen Opti-Salts Suite screen was employed. Here, the hit solution from the initial screening is mixed 9:1 with 96 different salts, buffers and other additives and used to search for improved crystallization conditions.

The native crystals of the ERS1 DHp domain were grown in 0.1M tri-sodium citrate pH 5.6, 18% (v/v) isopropanol, 18% (w/v) PEG 4000 and 0.32M sodium nitrate. The seleno-methionine derivatized crystals were grown in 0.1M tri-sodium citrate pH 5.6, 18% (v/v) isopropanol, 18% (w/v) PEG 4000 and 0.22 M thiocyanate.

Initial crystals of the ETR1 DHp domain were obtained in 0.1M MES pH 6.5 and 15% PEG 20000 (w/v) or 0.1 M HEPES pH 7.5, 10% PEG 4000 (w/v) and 5% Isopropanol.

Prior to crystallization of the CTR1 kinase domain, AMPPNP with $MgSO_4$ (final concentration of 1 mM) or staurosporine (final concentration of 150 mM), a generic protein kinase inhibitor, were added to the protein. Crystals of CTR1-kd were obtained at 20°C in 0.2 M K_2SO_4 and 15% (w/v) PEG 3350. The best crystals of CTR1-D676N grew from 1 M LiCl, 0.1 M citric acid pH 5.0 and 10% (w/v) PEG 6000.

Crystal freezing and data collection

Prior to data collection, a single crystal of CTR1-kd or CTR1-D676N was briefly immersed in mother liquor augmented with 25% glycerol as a cryoprotectant. Prior to data collection, a single crystal of the ERS1 DHp domain was briefly immersed in mother liquor augmented with 20% ethyleneglycol as a cryoprotectant. Crystals were mounted on CryoLoops™ (Hampton, Aliso Viejo, CA) using a range of loop diameters from 0.05 – 0.5 mm and were then flash-cooled to -196°C in liquid nitrogen.

Data processing

Complete X-ray diffraction data sets of the CTR1 kinase domain crystals were collected on beamlines ID23-2 and ID29 at the ESRF, Grenoble, France using a MAR 225 Mosaic CCD detector and an ADSC Q315R detector, respectively. For CTR1-kd a total of 170 frames were collected with a rotation range of 1.0°, while for CTR1-D676N a total of 220 frames were collected with a rotation range of 0.75°.

Complete X-ray diffraction data sets of the DHp domain crystals were collected on beamlines ID23-1 and ID29 at the ESRF, Grenoble, France and at X12 (DESY, Germany)

using a MAR 225 Mosaic CCD detector (X12), a Pilatus 6M detector(ID29) and an ADSC Q315R detector (ID23-1). For wild type, unlabelled ERS1 DHp domain a total of 1800 frames were collected with a rotation range of 0.1° at ID-29. For the SeMet crystals a total of 140 frames were collected with a rotation range of 1.1° for MAD phasing, while for the same crystals a total of 100 frames with a rotation range of 1.0° were collected for UVRIP phasing.

The data were indexed and integrated using XDS (Kabsch, 2010) and were scaled with SCALA (Collaborative Computational Project Number 4, 1994).

Structure determination

The methodology for the determination of the ERS1 DHp domain SeMet crystals using the UV-induced radiation damage method has already been described (Panjikar et al., 2011).

Prior to the methodological UVRIP phasing, MAD data of the same domain were scaled using *SCALEIT* and subjected to *HKL2MAP* (Pape and Schneider, 2004) using the *SHELX* pipeline running the MAD option. Into the initial map obtained an initial model was built with the help of *Arp/Warp* (Langer et al., 2008). The initial model was refined against the native dataset using *Refmac 5.6* (Murshudov et al., 1997) and *Phenix 1.7* (Adams et al., 2010), followed by iterative cycles of manual model rebuilding in *Coot* (Emsley et al., 2010).

The CTR1 kinase domain was solved by molecular replacement using B-RAF (pdb ID: 3C4C) as a search model utilizing the program *Phaser* (McCoy et al., 2007) as available through the CCP4 package (Collaborative Computational Project Number 4, 1994). The initial model was refined against the native dataset using *Phenix 1.6.4* (Adams et al., 2010), followed by iterative cycles of manual model rebuilding in *Coot* (Emsley et al., 2010).

Figures of the structures were prepared using *Pymol* (The PyMOL Molecular Graphics System, Version 0.99, Schrödinger, LLC). Figures of the SAXS models were prepared using Chimera (Pettersen et al., 2004). Alignments were prepared using *MUSCLE* (Edgar, 2004) and *Jalview2* (Waterhouse et al., 2009).

The atomic coordinates and structure factors for CTR1kd and CTR1D676N were deposited at the Protein Data Bank with accession codes 3PPZ and 3P86, respectively.

Results

Bioinformatics and constructs

The five ethylene receptors were analysed using a number of publicly available servers to analyse and define the domains and their borders, which is needed for the selection of promising constructs. An overview of all the main constructs and domain combinations designed and tested is given in Figure 10, but not all of the individual terminal truncations giving similar constructs are shown. Details are listed in table 1.

Table 1 including the exact start and end position of the constructs.

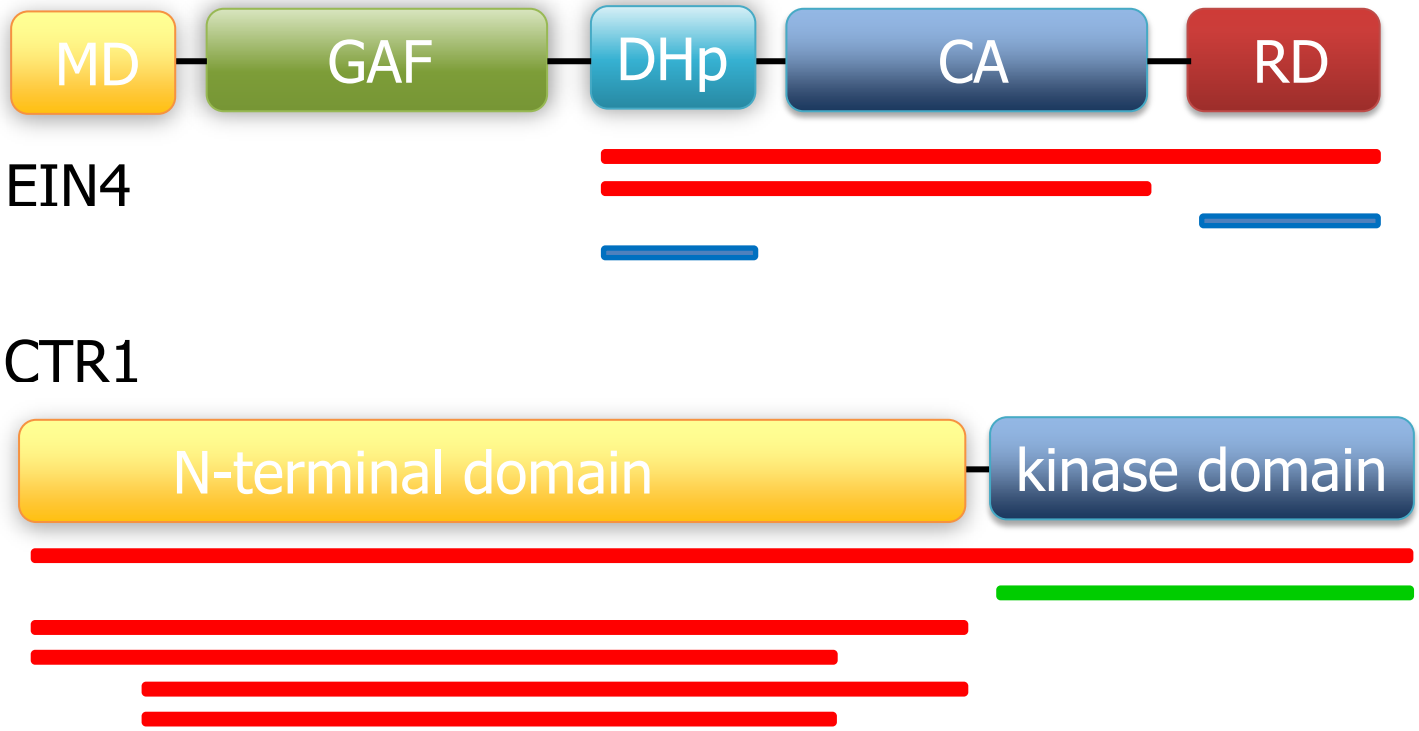


Figure 10: Overview of cloned constructs and their domain organization: **MD** membrane domain, **GAF** cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA domain, **DHp** dimerization and histidine phosphotransfer domain, **CA** catalytic and ATP-binding domain and **RD** receiver domain. All constructs are indicated by bars below the schematic representations of the proteins, indicating which domain combinations were examined therefore not every construct is listed. Red indicates no expression could be obtained, orange a low expression level, blue that the protein was purified, yellow crystals and green a determined structure

ERS1 dimerization domain

Expression and crystallization

Initial expression tests of the DHp domain of ERS1 (construct H9) yielded a good rate of expression (30mg/l). The protein could be purified to homogeneity using NiNTA and SEC, concentrated to more than 20mg/ml and was suitable for crystallization. Initial sitting drop crystallization experiments yielded mainly hexagonal crystals in different conditions. Unfortunately the crystals only diffracted to 7.5 Å. As a consequence a number of different ways were explored to improve the diffraction quality. The His-Tag was removed, customized grid screens using 24well hanging drop plates and an additive screen using the Qiagen Opti-Salts Suite were employed, different cryo conditions were tested, but the diffraction could only be improved to 5 Å using synchrotron radiation. Furthermore, the diffraction pattern exhibited a pseudo-dodecagonal symmetry (Figure 11B). The exact cause of this phenomenon could not be determined, but a possible explanation could be alternating layers of a hexagonal symmetry rotated relative to each other by 30°. This could also give an indication for the poor diffraction properties of the crystals.

As the crystals could not be improved the sequence was reanalysed. A new construct was designed according to the secondary structure prediction by PSIpred, which suggested a different starting point of the first helix. This construct included 18 additional residues to the N-terminus (construct H35). The new construct expressed again well and was purified using a similar protocol as for the H9 construct. Initial crystallization trials yielded again crystals of reasonable size (150 µm x 150 µm x 70 µm). Initial diffraction tests showed the benefit of removing the His-Tag when comparing cleaved and uncleaved protein. Therefore further trials were carried out using the tag-less construct (H35ΔHis). Again the crystals were further optimized using the Opti-Salts Suite (Qiagen) and customized grid screens and eventually diffracted to 1.9 Å at ID29 (Table 5) in space group $C222_1$. The crystals for data collection were grown in 0.1M tri-sodium citrate pH 5.6, 18% (v/v) isopropanol, 18% (w/v) PEG 4000 and 0.32M sodium nitrate.

Despite the availability of dimerization domain structures from bacteria (sequence identity up to 40%), the attempts to solve the structure using molecular replacement failed. This is possibly due to the seemingly continuous stretches of two long α-helices present in the crystal making it hard to place the model.

In general molecular replacement software seems to have problems with this domain and/or crystal packing. After solving the structure later by experimental phasing in space group $C222_1$ the refined model of the DHp domain in space group $C222_1$ had to be modified by removing the first 25 residues for a successful molecular replacement search in space group $P2_12_12$ despite 100% sequence identity and hardly any difference in the structures (R.s.m.d. 1.18 Å).

Phasing

Phasing by soaking sodium bromide, sodium iodide or 5-amino-2,4,6-triiodoisophthalic acid (IC3, (Beck et al., 2008)) into the native crystals was attempted. Unfortunately, the observed resolution was not optimal for phasing (~5 Å) and data were of bad quality due to high mosaicity (>1.3°). Co-crystallization of H35ΔHis with 10mM IC3 yielded crystals, which diffracted to reasonable resolution (3 Å) but no signal of the iodide could be observed, probably because the IC3 did not bind in a specific manner.

As the construct contains seven methionines, labelling with seleno-metheonine was carried out. The protein was purified like the native one and yielded crystals diffracting to 3 Å resolution at X12 (DESY, Hamburg) belonging to space group $C222_1$. The crystals were grown in 0.1M tri-sodium citrate pH 5.6, 18% (v/v) isopropanol, 18% (w/v) PEG 4000 and 0.22M thiocyanate, a similar condition as the native crystal. Following an X-ray fluorescence spectrum over the selenium absorption edge, three datasets at the peak, inflection and high energy remote energies were collected, all to a similar resolution. Data collection statistics for all three datasets are summarized in Table 5.

The structure was solved using the MAD method, by utilizing the *SHELX* pipeline as implemented in *HKL2MAP*. The initial map showed clear electron density for helices. The model showed the expected features of a DHp domain, namely two long helices forming a dimer. The electron density and the data are of good quality (Table 5) and no signs of twinning were detected. The R_{work}/R_{free} of the final model in space group $C222_1$ are 20.4% and 24.6%, respectively.

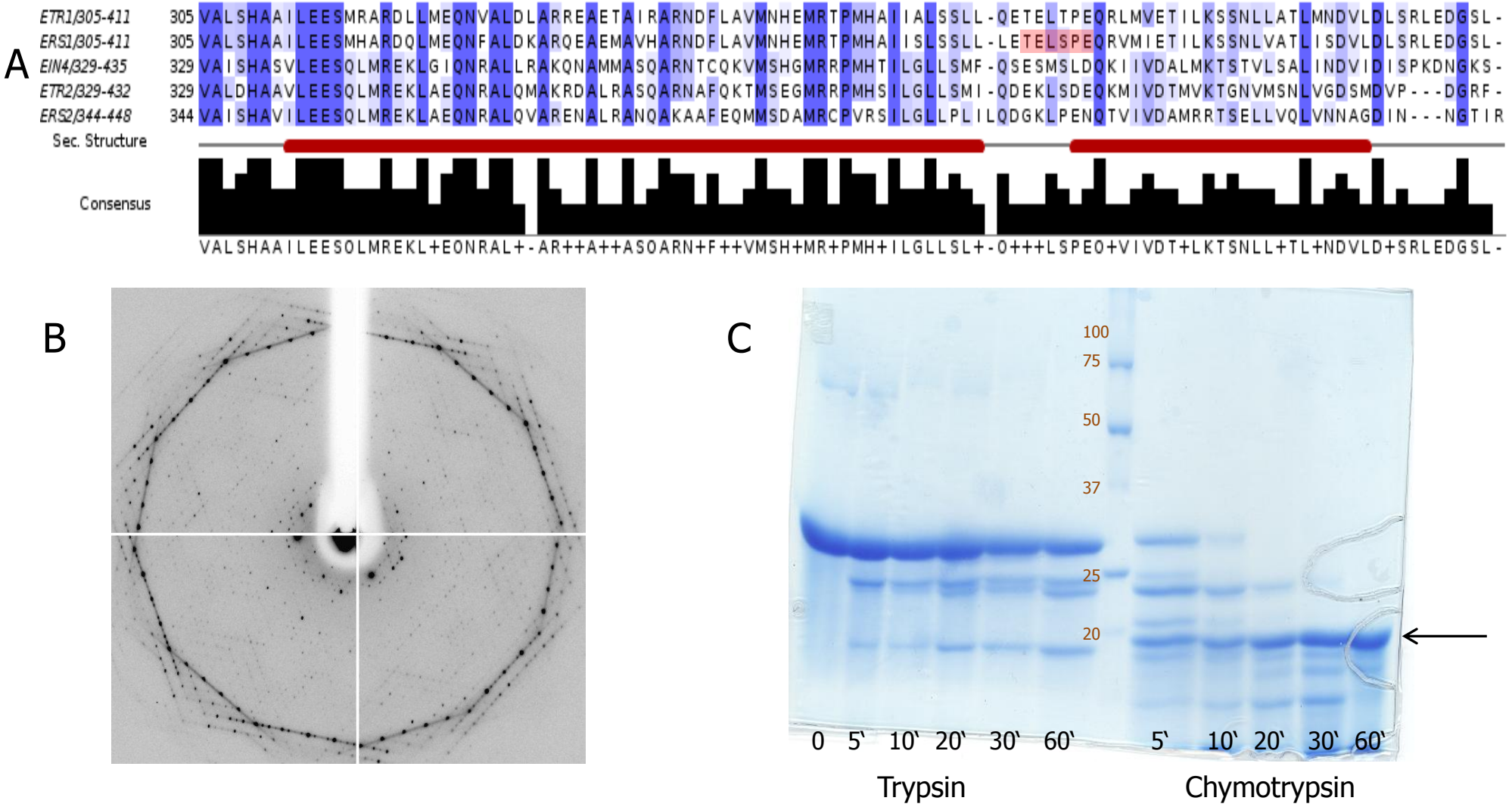


Figure 11: **A** Alignment of the five DHp domains from the *A. thaliana* ethylene receptors using the program *MUSCLE*. The linker between the two helices in the ERS1 DHp structure is indicated in red. **B** Initial diffraction pattern of the ERS1 DHp domain showing the dodecagonal pattern **C** Limited proteolysis of the ETR1 histidine kinase domain (Construct H44)

Table 5: X-ray data-collection and processing statistics of the ERS1 DHp domain. Values in parentheses are for the highest resolution shell.

	Native	SeMet peak	SeMet inflection	SeMet remote	SeMet high res
X-ray source	ID-29, ESRF	X12, DESY	X12, DESY	X12, DESY	ID23-1, ESRF
Wavelength (Å)	0.923	0.9772	0.9777	0.9716	0.9919
Resolution range (Å)	49.3-1.9 (2.0-1.9)	47.3-3.0 (3.15-3.0)	47.3-3.0 (3.15-3.0)	47.3-3.0 (3.15-3.0)	47 – 2.1
Temperature (K)	100	100	100	100	100
Crystal-to-detector distance (mm)	337.13	272	272	272	316.55
Rotation range per image (°)	0.1	1.1	1.1	1.1	1.0
Total rotation range (°)	180	154	154	154	100
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$P2_12_12$
Cell dimensions (Å)	a = 75.03, b = 98.73, c = 77.1	a = 75.20, b = 99.09, c = 77.12	a = 75.16, b = 99.07, c = 77.09	a = 75.65, b = 99.78, c = 77.62	a = 65.92, b = 69.20, c = 108.27
Observed reflections	151061(24280)	36987(5733)	36764(5768)	37543(5837)	109136 (16951)
Unique reflections	22944(3593)	11119(1743)	11090(1755)	11339(1827)	27137 (4328)
Multiplicity	6.58(6.75)	3.3(3.3)	3.3(3.3)	3.3(3.3)	4.0 (3.9)
R_{merge} (%)	5.5(79.9)	7.0(26.0)	7.3(27.5)	5.5(19.6)	5.2 (34.7)
R_{pim} (%)	2.5(35.5)	4.6(17.1)	4.8(18.1)	3.6(12.9)	2.8(19.1)
Completeness (%)	99.4(97.6)	99.8(99.1)	99.9(99.6)	99.7(98.2)	99.2 (99.8)
$\langle I/\sigma(I)\rangle$	14.3(2.5)	17.6(6.1)	16.8(5.4)	22.1(7.9)	17.1 (3.7)
Wilson *B* factor ($Å^2$)	43.7	43	43	43	38.4

Another dataset of the SeMet crystals grown in the same condition as the native crystals also diffracted to 2.1 Å at ID23-1. Unexpectedly they belonged to space group $P2_12_12$, in contrast to all other previously collected datasets. The crystals in space group $C222_1$ contain two molecules in the ASU crystals with 61.1% solvent content, while in space group $P2_12_12$ four molecules are present in the ASU with 54.5% solvent content. In space group $P2_12_12$ the model could be refined to R-values of 21.4% and 25.5% (Table 6). When comparing the models they are similar (for chain A: r.m.s.d = 0.99 Å , r.m.s.d = 0.71 Å when excluding the first 10 residues) with small differences at the N and C-terminus, mainly in the number of residues with clear electron density. For the further analysis of the coiled-coil interface the structure obtained from the crystals in space group $C222_1$ was used due the better resolved N-terminus, especially in chain A.

The structure was also solved with the UVRIP method to demonstrate the general applicability of this approach to phasing (Panjikar et al., 2011). The final model was similar to the one obtained using the classical approach, therefore validating the applicability of this new approach to phasing.

Table 6: Refinement statistics of ERS1 DHp domain in two different space groups

Space group	$C222_1$	$P2_12_12$
Resolution range (Å)	49.3-1.9 (2.0-1.9)	47 – 2.1
*R*work / *R*free %	20.4/24.6	21.4 / 25.5
No. atoms		
Protein	1547	3045
Other	-	8
Water	93	189
B-factors		
Protein	51.2	44.1
Water	50.6	43.4
R.m.s. deviations		
Bond lengths (Å)	0.014	0.016
Bond angles (°)	1.232	1.495

Overall structure

Similar to the DHp domain structure of a sensor histidine kinase from *T. maritima* (pdb ID: 2C2A), the ERS1 DHp domain consists of two anti-parallel helices connected by a short hairpin loop. It forms a dimer, which coils around a central supercoil axis, such that the N-termini are adjacent (Figure 12). The first helix, helix A, in the molecule (307-371, with no observable electron density for the first four residues in chain B) extends for about

90 Å from the N-terminus to the hairpin. Helix A has a pronounced kink, which divides the helix in two parts (A1 and A2). A short hairpin linker (372-376) connects helix A to helix B, which is shorter, with a length of about 40 Å and one to two residues at the C-termini of the different chains not being visible. A proline residue, five residues C-terminal of the phosphorylatable histidine, found in the majority of histidine kinases is located just after the kink in helix A. One helix in each dimer has a stronger kink compared to the other helix in the dimer. When superimposing the structure from the conserved histidine residue to the end (residues 353-404, r.m.s.d.=0.44Å) the bending of helix A1 in chain B is 9° larger than in chain A. Helix A1 contains the conserved phosphoryl-accepting residue His^{353}, which is exposed on the surface and assumes different side chain orientations between the different chains. Helix A2 makes contacts with helices B and B' in the dimer and vice-versa.

Dimer interface

The first 42 residues of the helices A1 and A1' form a parallel coiled-coil (Figure 13) and engage in a left handed twist of about 23°. At the end of the coiled-coil the helices move apart and at the C-terminus helix A2 and helix B together with A2’ and B’ form a

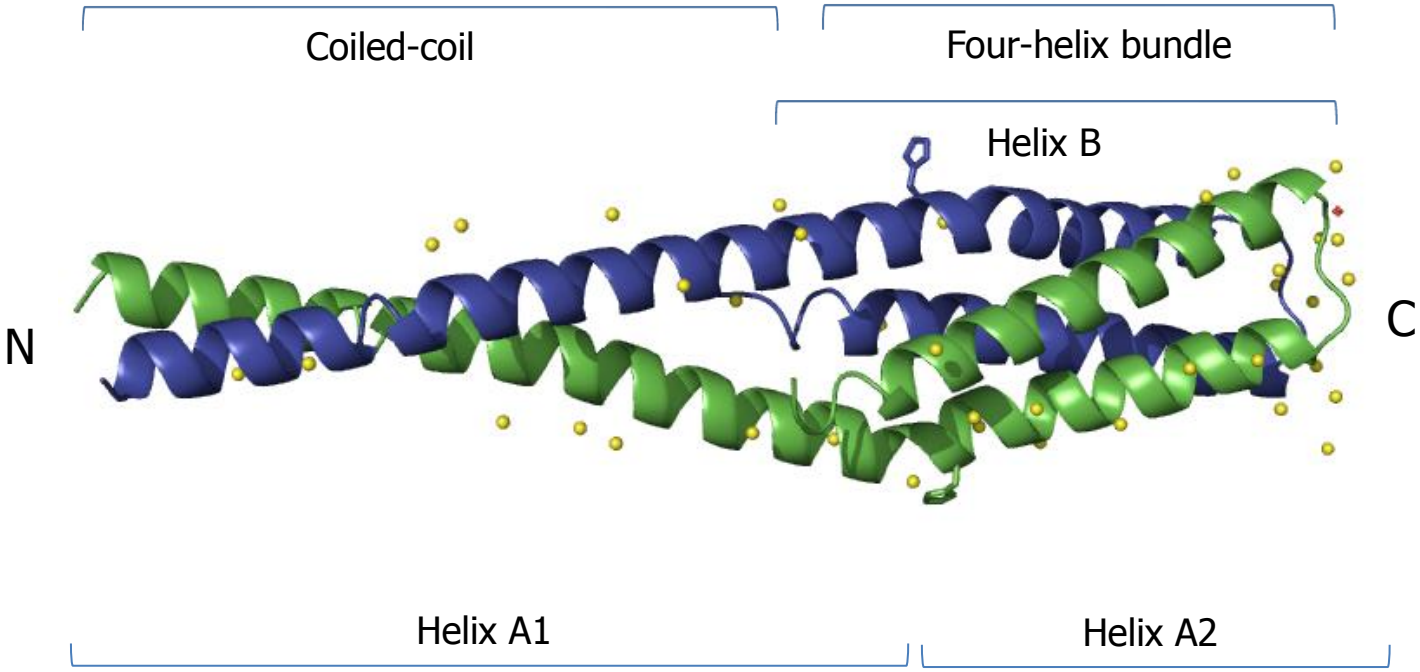


Figure 12: Structure of the ERS1 DHp domain in cartoon representation. The two conserved phospho-receiving histidines are shown as sticks; waters are shown as yellow spheres. Sections of the helices are labeled as referred to in the text.

four-helix bundle with a left-handed twist of about 23°- 28°. This angle and also the anti-parallel topology are favoured by four-helix bundles (Chou et al., 1988). A similar topology is found in the sensor histidine kinase from *T. maritima*, which is different from the unusual, more parallel topology found in EnvZ (pdb ID: 1JOY). In the dimer there is a large interface, burying 2120 Å^2 of solvent accessible surface area from each molecule as calculated by PISA. This creates a large energy barrier for dissociation of the dimers. The majority (62%) of the interface is contributed by the four-helix bundle (1310 Å^2) yet the contribution of the coiled-coil interface (810 Å^2) is more than in other DHp domains. This is due to the fact that a larger portion of the coiled-coil could be observed than in any other structure reported so far.

In general histidine kinases possess coiled-coils with heptads not easily recognized by standard detection algorithms but can be detected by more specialized software (Singh et al., 1998). Compared to the canonical heptade (*a-g*) repeat normally observed, with apolar or hydrophobic residues at position *a* and *d* they frequently have charged residues at heptade positions *a* and *d* as found for example in EnvZ or Sln1 (Tao et al., 2002). They are predicted to destabilize the coil mediated interactions (Tripet et al., 2000). Still in the yeast osmosensor Sln1, which also possess a coiled-coil made up of helix A1, the coiled-coil was found to be important for the dimerization and the kinase activity (Tao et al., 2002), while the four-helix bundle alone could not maintain the dimer.

Only some of the ERS1 heptade agree with the classical model. The coiled-coil is stabilized by three leucine residues pointing to the core (Leu^{313}, Leu^{323} and Leu^{330}), hydrogen bonds formed by the two Ser^{316} side chains and in addition a salt bridge network (Gln^{326}, Asn^{327}) (Figure 13). Similar networks of polar residues within the core of the coiled-coil at heptade positions *a* and *d* can be also be found in other structures for example in HIP1 (pdb ID 3I00), tropomyosin (pdb ID 2EFR) or GIT1 (pdb ID 2W6A). Two arginine residues, Arg^{320} and Arg^{334}, normally not typically found at position *a* or *d*, contribute to the core with the aliphatic straight chain being buried in the centre. The guanidinium groups are pointing away from each other, being exposed to the solvent.

The linker connecting the two helices of the dimerization domain in the close packed hydrophobic core of the four-helix bundle is, with the exception of a few side-chains, relatively rigid, as judged from the low B-factors. Apart from Leu^{374} pointing to the hydrophobic core of the four-helix bundle it consists of polar residues. As shown by the B-factors the coiled-coil seems to be more flexible than the four-helix bundle.

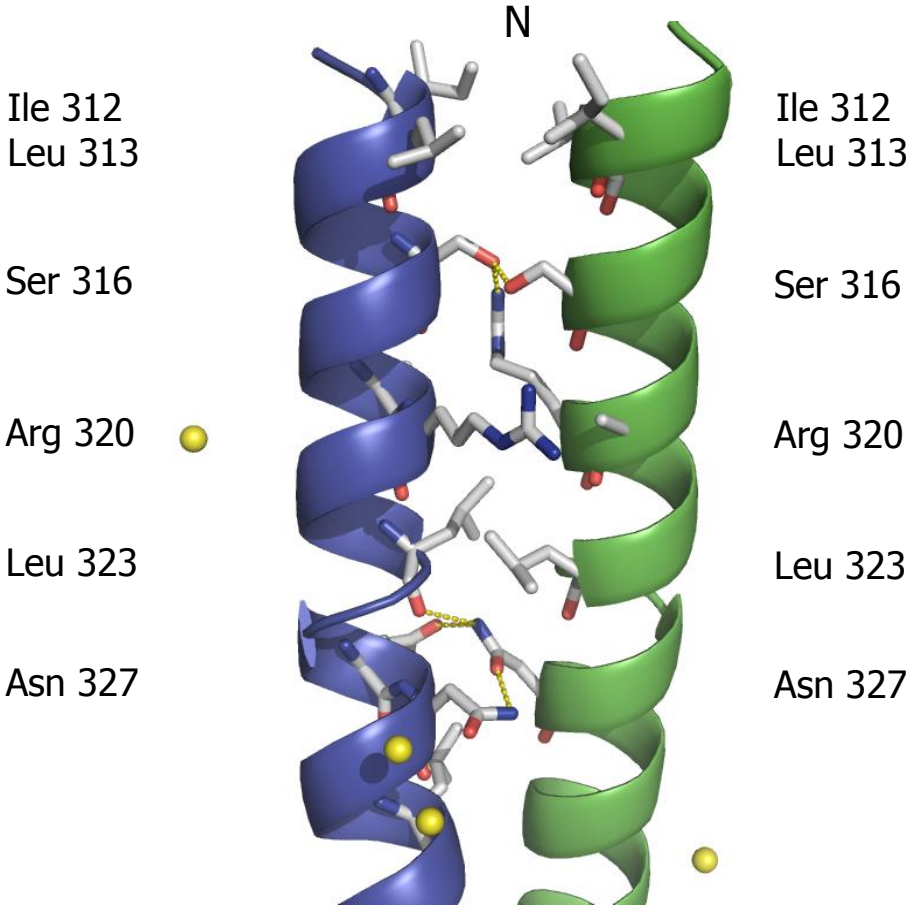


Figure 13: N-terminal coiled-coil of the DHp domain. Side chains involved in the interface are shown as sticks; hydrogen bonds are shown as yellow dashed lines and waters as yellow spheres.

The asymmetric unit of space group $P2_12_12$ contains four molecules forming two dimers. The two dimers are oriented in an anti-parallel orientation, with contacts between the two four-helix bundles. From previous studies, the formation of a dimer was expected. Interestingly the DHp domain eluted from the gel filtration column at about 55 kDa consistent with a tetramer given the tagged monomer of 14 kDa. Analysis using the PISA web service predictes the dimer as well as the tetramer to be stable in solution. Previously also higher-order non-covalent interactions have been observed (Gao et al., 2008). Therefore interactions between the receptors are expected to some degree and dimers could help assemble these clusters. Yet, when expressed together with the catalytic domain only a dimer could be observed. This questions the biological significance of the tetramer observed, also as the concentrations used during gel filtration and crystallization are not the physiological ones and steric clashes with the other domains, omitted in this construct, would occur. In addition the interaction between the two dimers occurs in an anti-parallel fashion, with the N-termini pointing in different directions. This argues against

this interaction to occur *in vivo* as the receptors are inserted into the membrane at the N-terminus, which would require the N-termini to point in the same direction.

ERS1 homodimer

When superimposing the dimerization domains of ERS1 and TM0853 (Casino et al., 2009), different topologies of the helical hairpins become apparent (Figure 15), with a counter-clockwise connectivity in ERS1. TM0853 exhibits the less frequently observed *cis*-phosphorylation mechanism when transferring the γ-phosphate of ATP to the conserved histidine residue. In TM0853 the helical hairpin domains are connected the opposite way compared to DHp from ERS1. The superimposition of the helices in the DHp domain of TM0853 and ERS1 allows for the modelling of a complete ERS1 HK. In this model the catalytic domain from one molecule in the dimer comes close the histidine residue in the other molecule. This indicates a *trans*-phosphorylation mechanism for ERS1 (Yang and Inouye, 1991) as frequently observed in histidine kinases.

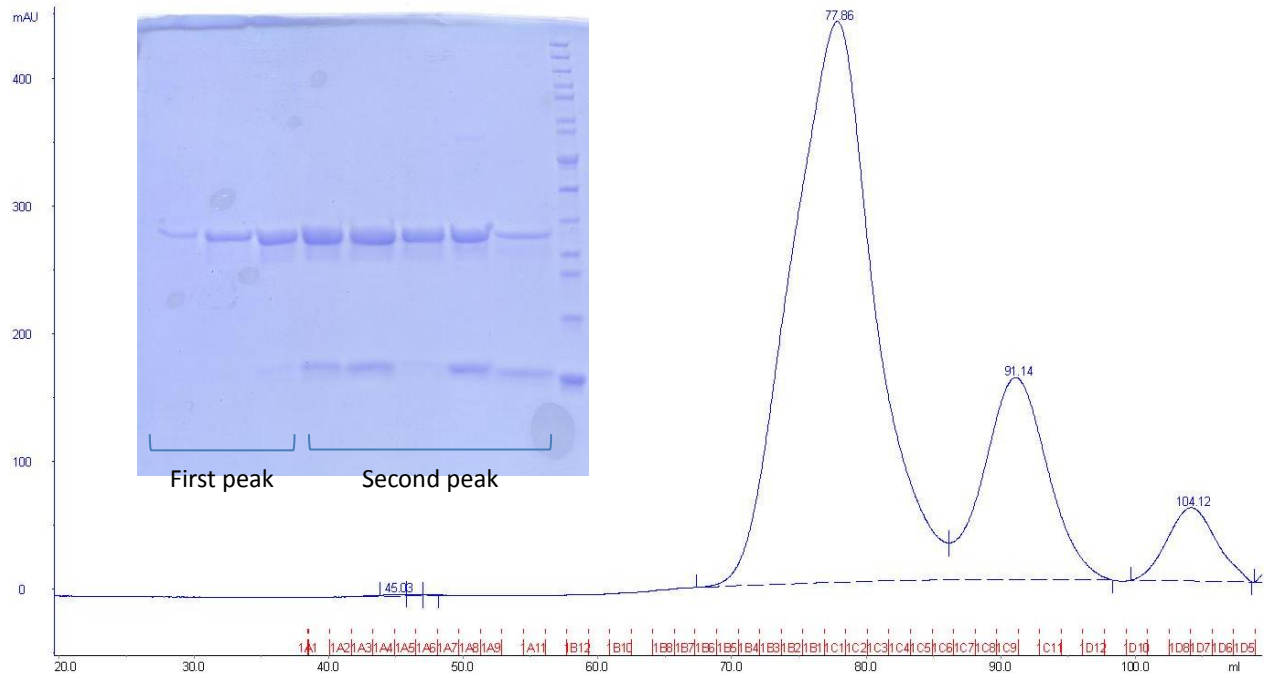


Figure 14: SDS gel and SEC chromatogram of ERS1 HKD co-purified with its DHp domain

In order to examine the inferred mode of action in ERS1, based on the comparison with TM0853, the dimerization domain construct (H35) was modified in order to obtain a completely untagged version. Untagged DHp was then co-expressed with full length histidine kinase domain (H74), which carried a His-tag. Purification via NiNTA, proteolytic removal of the tag and subsequent SEC resulted in the purification of homodimeric full-

length HK and the heterodimer of HK and DHp (Figure 14). Using the size difference between the two constructs in the heterodimer this can be used to discriminate between *cis*- and *trans*-phosphorylation.

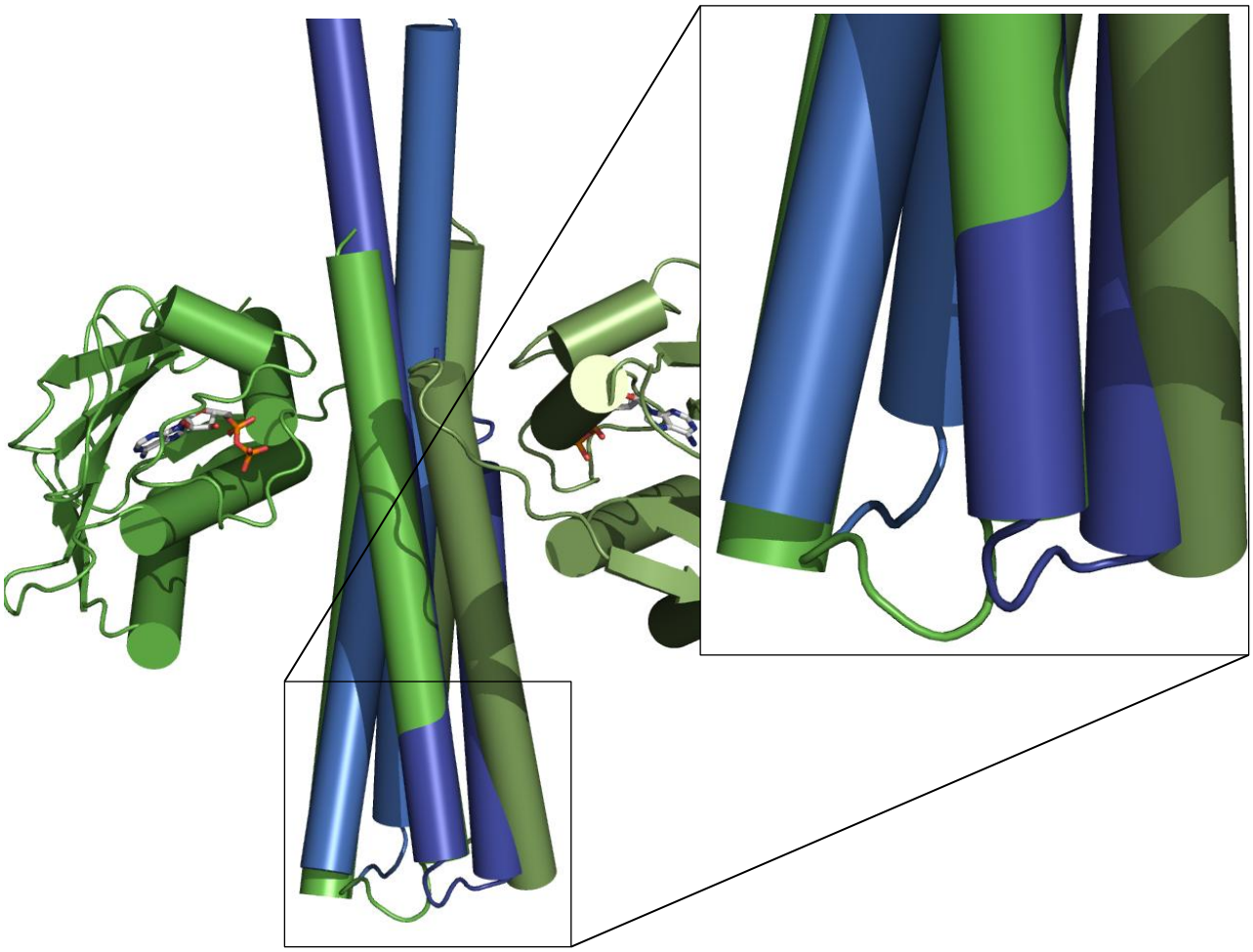


Figure 15: Superimposition of the ERS1 DHp domain (blue) onto the corresponding domain from TM0853 (green) from *T. maritima* with helices shown as cylinders for clarity. The receiver domains in complex with TM0853 were omitted, ADP bound to the TM0853 CA domain is shown as sticks. A zoom on the hairpin connecting the loops is shown on the right.

ETR1 dimerization domain

The initial construct (H3) of the ETR1 DHp domain expressed very well yielding approximately 20mg/L of protein. Yet the construct was unstable and could not be stored even at 4°C having the tendency to precipitate. Different buffer compositions were tested but were not successful in stabilizing the protein.

A new construct, which was extended by four C-terminal residues (H30), was prepared but exhibited the same pathologies. Eventually another construct, with similar start and end points to the one prepared for the ERS1 DHp domain, was prepared extending the N-terminus, which lead to a better behaved protein (H36). It was purified similar to the ERS1 DHp domain by NiNTA and SEC and could be concentrated to 15mg/ml. Initial crystallization trials yielded very fragile and small crystals. The effect of the His-tag removal was explored and yielded crystals of improved mechanical stability. Initial crystals were obtained in 0.1M MES pH 6.5 and 15% PEG 20000 (w/v) or 0.1M HEPES pH 7.5, 10% PEG 4000 (w/v) and 5% Isopropanol. Initial diffraction experiments were unsatisfying due to the poor resolution (7 Å).

Customized grid screens using 24well hanging drop plates and an additive screen using the Qiagen Opti-Salts Suite were employed in order to improve the initial hits. A common scheme of the hits could be observed always involving PEG in the range from 3350 – 20000 Da and HEPES at different pH. The crystals diffracted up to about 3.6 Å but had to be cut to 4.25 Å to obtain a satisfying signal to noise ratio of about 2 in the highest resolution shell. Even though the DHp domain of ERS1 was used as a search model exhibiting 84% sequence identity, no solution could be obtained probably because of the poor data quality. The crystals belonged to space group $P2_12_12$ with a big unit cell (a=59.7 Å b=159.7 Å c=173.6 Å $\alpha=\beta=\gamma=90°$) containing approximately 14 molecules in the ASU (assuming 50% solvent content).

No further attempts to improve were made, also as the structure of the ERS1 DHp domain was solved, with the sequence being very similar to the ETR1 dimerization domain (Figure 11A). No changes in the overall structure are therefore expected and also the topology is anticipated to be similar due to the high similarity of the linker sequence between the two helices (Figure 11 A).

ETR1 and ERS1 histidine kinase domain

Expression of the ETR1 and ERS1 full length histidine kinase domains was only possible in the presence of chaperones (strain BL21 (DE3) cc4). Initial crystallization trials of the constructs H6 and H12 in the presence of the ATP analogue AMPPNP did not yield any crystals. Limited proteolysis was tested in order to define a more stable construct using chymotrypsin and trypsin. While the treatment with trypsin did not yield any

obvious predominant band, a new, strong band emerged and remained stable even after prolonged digestion following the digestion with chymotrypsin (Figure 11C indicated with an arrow). Analysis by peptide mass fingerprinting of the isolated band showed that the cleavage occurred within the catalytic domain of the HK, cutting a disordered or exposed loop, yielding thereby an unusable fragment. The cut occurs around residue 510-525 (an exact position could not be obtained due to the limited resolution of the MALDI-TOF). When comparing the sequences of the ethylene receptor histidine kinase domains to bacterial histidine kinases an obvious insertion in the catalytic domain can be observed ranging approximately from 494 to 507 (ETR1 numbering). Mapping this loop onto the structure of a bacterial histidine kinase (pdb ID 2C2A) it appears to be an insertion in the loop connecting βD with βE and as indicated by the limited proteolysis results, probably leads to some flexibility in this region. This loop is pointing away from the active site of the kinase and one could speculate that it plays a role in the interaction with the receiver domain.

With the information of the favourable N-termini of the DHp domains of ERS1 and ETR1 available, a number of variants of the C-terminus were tested yielding constructs 68-72 in case of ETR1 and 73-77 for ERS1 (Table 1). All expressed equally well and could be purified similar to the original constructs. Both domains were found to be active as judged using the Kinase-Glo Luciferase Assay. Still no crystals could be obtained when attempting to crystallize in the presence of 1mM AMPPNP and 1mM Mn^{2+} (ETR1) or 1mM AMPPNP and 1mM Mg^{2+} (ERS1). The effect of mutations of lysine or glutamate sitting on the surface of the domain and not in a secondary structure element (as predicted by the SER server) was tested with three mutants. $E^{373}T$, $KK^{437}TT$ and $KQ^{476}TT$ were prepared and expressed. $KK^{437}TT$ did not yield any expression, while the other two mutants could be expressed and purified but did also not lead to any crystals.

ETR1 cytoplasmic domains

ETR1 without the membrane domain consisting of the GAF, HK and receiver domain (residues 158-738) was initially expressed in BL21 cc4. However, some of the chaperons co-purified with the overexpressed ETR1 protein, indicating that some portion of the protein was not well folded or had some hydrophobic patches exposed, thereby keeping

the chaperone attached. This impurity could not be removed in the course of the purification.

To overcome this problem expression in other strains was tested and reasonable expression levels could be obtained in Rosetta cells in combination with a very low concentration of IPTG giving the protein more time for folding. NiNTA purification yielded relatively pure protein (Figure 16 A), yet it was not very well suited for crystallization, albeit being pure (99% as judged by SDS PAGE), as it precipitated during concentration. In order to improve the solubility beyond approximately 0.4 mg/ml a number of additives were explored. The addition of 0.5 M of 3-(1-Pyridinio)-1-propanesulfonate, belonging to the non-detergent sulfo betaines (NDSB), permitted the concentration to at least 2.5 mg/ml allowing further studies.

Subsequent trials to explore the necessary concentration range of NDSB needed to keep the protein in solution converged at 200-250 mM NDSB. This concentration of NDSB was chosen as a compromise to minimize interference with crystallization and other biochemical and structural characterization methods, and protein solubility, which declined below 150 mM NDSB.

In order to determine a low resolution model of cytoplasmic portion of ETR1, SAXS measurements were carried out at the X33 BioSAXS beamline in Hamburg. Initially the impact of NDSB at various concentrations was tested. A linear increase of the background scattering was detected with increasing NDSB concentration but did not interfere with successful data collection. An upper limit of 500 mM NDSB was estimated as being still compatible with SAXS measurements. 250 mM NDSB were used in the subsequent measurements in order to reduce the background effects. The SAXS data statistics are summarized in table 7.

Table 7: Overall SAXS parameters of the ETR1 cytoplasmic domain

$Rg_{Guinier}$, nm	Rg_{GNOM}, nm	Dmax, nm	I(0)	MM_{SAXS}, kDa	MMcalc, kDa	Vol_{SAXS}, nm^3	Vol_{DAM}, nm^3
4.7 ± 0.1	4.8 ± 0.1	15.8 ±	28.5	45 ± 5	64.8	282 ± 30	165 ± 20

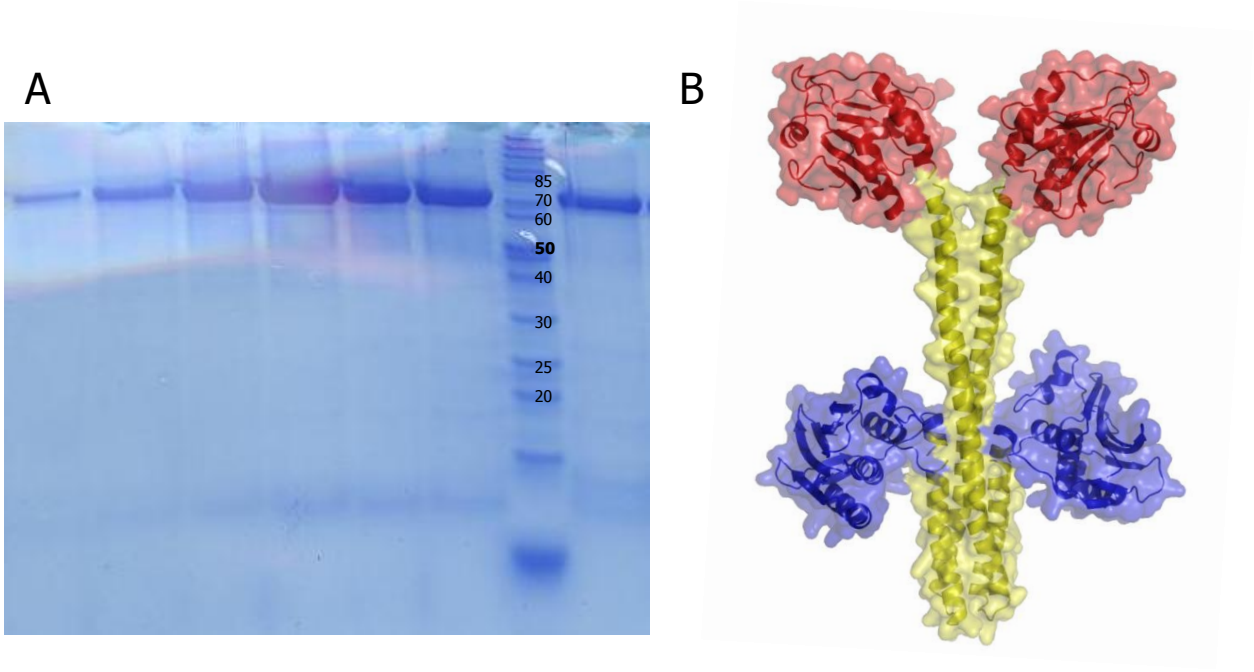


Figure 16: **A** Gel of fractions of the ETR1 full cytoplasmic part after NiNTA purification **B** Model (see text) build of the ETR1 GAF (red), dimerization (yellow) and CA (blue) domains, which was used as a starting point for the SAXS data analysis.

While measurement of the construct at 20°C led to precipitation, measuring at 8°C prevented aggregation. As the strong signal of NDSB at 280 nm interfered with concentration determination using the Nanodrop photometer two different methods very employed namely determination using the Bradford assay and using an analytical refractometer. Unfortunately the results from both methods differed by a factor of 9. While the concentration determined by the Bradford assay was 2.5 mg/ml the refractometer indicated a concentration of about 20 mg/ml, the latter possibly an artefact of the high NDSB concentration. The strength of the SAXS signal of the protein was therefore compared with a similar sized protein (BSA) measured at different concentrations and was estimated to be about 3mg/ml. SAXS curves were determined at 3, 1.5 and 0.75 mg/ml protein concentration (by dilution).

Structure alignment of the ERS1 dimerization domain with the dimerization domain of a bacterial histidine kinase from *T. maritima* in complex with its cognate receiver domain (pdb ID 3DGE) or of the uncomplexed HK (pdb ID 2C2A) placed the catalytic domains in two different conformations reflecting their different orientations in both structures. The residues were mutated in order to match the ETR1 sequence using the SCRWL method as implemented (http://ffas.burnham.org/protmod-cgi/qryByAliForm.pl)

where all template residues are replaced by target residues, thereby giving a model of the full length ETR1 histidine kinase domain. The structure of a cyanobacterial phytochrome GAF domain (pdb ID 2K2N) was chosen to represent the ethylene receptor GAF domain. Small, overlapping helical sections at the C-terminus and N-terminus, respectively, of 2K2N and the dimerization domain were used to link both domains. Since the receiver domain of ETR1 is expected to have a different location from the one observed in a bacterial structure (pdb ID 3DGE) it was omitted from this model (Figure 16 B) and added later to the final SAXS envelope (Figure 17 C). In the bacterial model HKD and receiver domain are two different proteins, while they are linked by a 15 residue linker in ETR1. Yet in the model from *T. maritima* there is a gap of about 50 Å between the C-terminus of the HKD and the N-terminus of the receiver domain. In order to keep the relative position of the two domains as observed in the *T. maritima* structure the 15 residue linker would have to be in a highly extended conformation. For this reason a different location of the receiver domain is assumed.

The model was used for rigid body refinement of SAXS data with short, flexible linkers defined between the GAF and dimerization, the dimerization and CA domain. Between the CA and receiver domain a 15 residue linker was defined as the only restraint for the location of the receiver domain. The model generated fits the data well in the low angle range (<1.5 nm^{-1}) but deviates at higher angles where short range and inter-atomic distances contribute to the scattering (Figure 17 A and B). Yet the excellent fit to the data at low angles allows the description of the global architecture of the dimer with a dumbbell shaped structure (Figure 17 A and B red line). This also fits well in the shape envelop determined by *ab initio* modelling using DAMMIF (Franke and Svergun, 2009) (Figure 17 A and B red dashed line). In this approach the position of the CA domain in the HKD was quite different from those observed in bacterial HKs with a 90-180° rotation relative to the dimerization domain. As there is no other evidence for such a rotation the model was considered to be at least partially artificial and therefore another approach was tested.

The same models were used (Figure 16 B) and defined as a rigid core. The receiver domain was placed relative to this core model by rigid body modelling with a 15 residue flexible linker between the CA and receiver domain. Both models with the flexible receiver domain fit the SAXS data and the *ab initio* model equally well (Figure 17 A and B blue and green line). The difference between the two models is too small to be determined by

SAXS. Again the fit is very good at low angles ((<1.5 nm^{-1}, χ=1.1 for 2C2A and χ=1.13 for 3DGE) and deviates again at higher angles (χ=2.7 and χ=2.9 for the two different models used for rigid body modelling). The data did not show features at higher angles showing mainly background scattering. In these cases the position of the CA domain was fixed to the position found in bacteria and in the absence of other evidence these models are used for the analysis and are more likely to reflect the true architecture.

In both models (Figure 17 C) the position of the receiver domain is different from the one found in the structure of the *T. maritima* HKD in complex with its cognate receiver domain. While in the bacterial structure the receiver domain is placed close to the dimerization domain it is placed further away in ETR1, packing again the CA domain on the side averted from the ATP binding site. In both cases the active site consisting of Asp^{616}, Glu^{617}, Asp^{659} and Lys^{702}, normally invariant in receiver domains, with the Asp^{659} thought to receive the phosphate from the histidine in the dimerization domain, is pointing away from the dimerization domain (Figure 17 C central and right pane). For the phosphotransfer a 180 degree rotation as well as a small translation would be needed as Asp^{659} and His^{353} are, depending on the model, 30 – 40 Å apart.

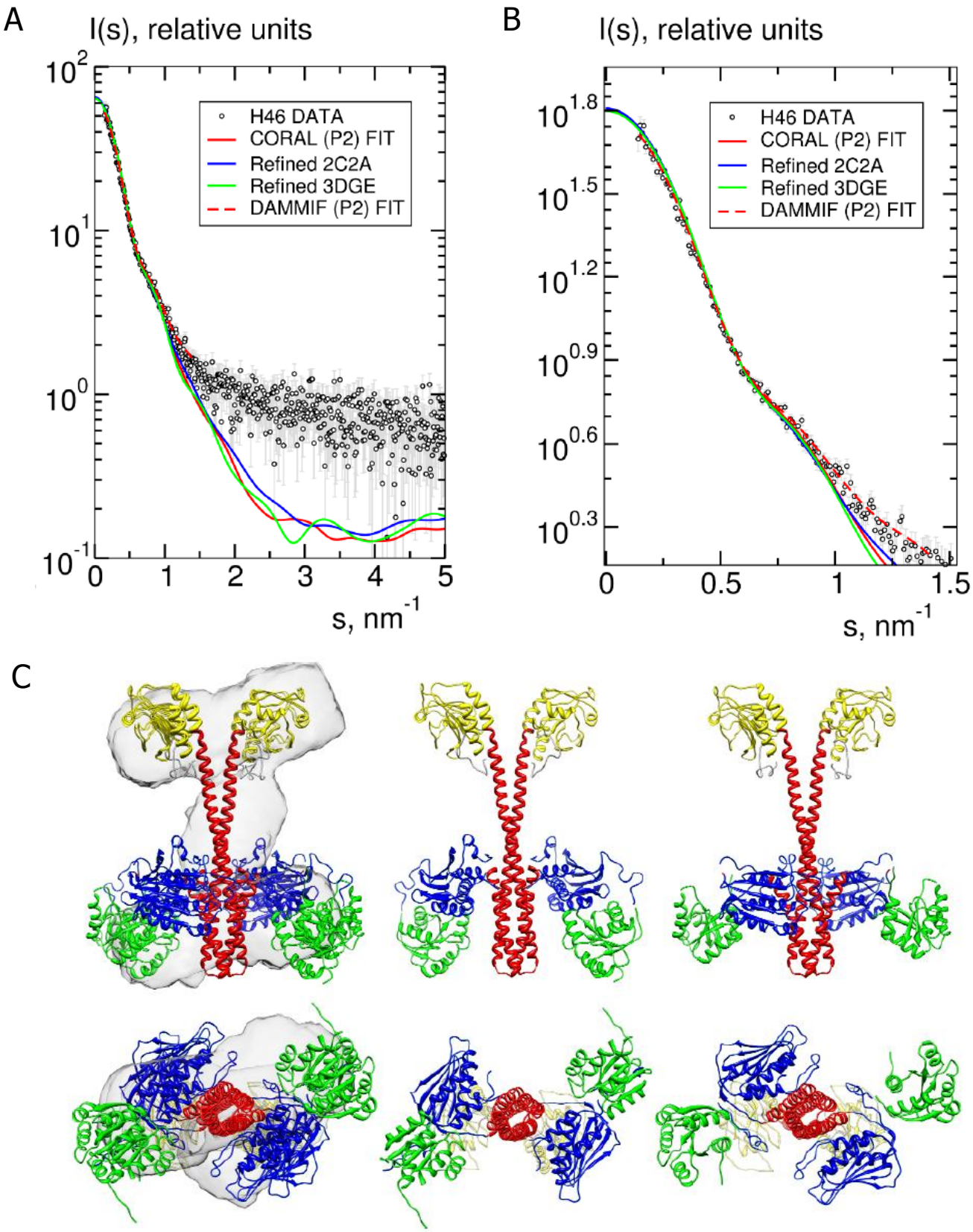


Figure 17: **A** Fit of the different models of the cytoplasmic domains. Dots indicate the measured intensity. The calculated scattering for the model obtained by CORAL is shown in red, in blue and green the scattering for the constructed models when defined as rigid bodies and the red dashed line shows the results obtained from *ab initio* modelling **B**

Zoom of the low angle region of A using the same colours **C** Models of the cytoplasmic domains of ETR1 using a rigid core homology model. On the left an overlay of the DAMMIF shape envelope with the refined models based on the two different structures, each with the receiver domain placed by rigid body modelling with the rest defined as a rigid core. In the centre the model based on 2C2A and on the right the model based on 3DGE. The GAF domain is shown in yellow, the dimerization domain in red, the CA domain in blue and the receiver domain in green.

ETR1 GAF and GAF & DHp domain

The initial expression test yielded soluble protein for the GAF and GAF-DHp domains of ETR1 but never yielded satisfactory results for the same two domains from ERS1. Expression was achieved in the *E. coli* Rosetta strain. Three different constructs were tested comprising the GAF domain and two containing the GAF and DHp domain. Especially the latter construct was prone to precipitation, which could be avoided through the addition of 5 mM DTT and 1 mM EDTA. A thermofluor screen was used to screen for a better buffer and a high pH was found to be beneficial. Subsequent buffers contained 20 mM CHES pH 9.0, 150 mM NaCl, 5 mM DTT and 1 mM EDTA for SEC. Using this buffer the protein was reasonable stable. The oligomeric state was assessed by AUC. As expected the GAF domain together with the DHp domain forms a dimer (Figure 18).

Initial crystallization trials of ETR1 GAF and DHp domains (construct 78) yielded spherically shaped crystals. They seemed to be composed of multiple crystals growing together. A number of optimization screens and the Qiagen Opti-Salts Suite were employed in order to improve the crystals and a few appeared which looked from their shape like single crystals, albeit quite small in size. Unfortunately they did not diffract. A few crystals, apparently consisting of multiple crystals, were tested. They diffracted x-rays to about 5 Å yet it appeared as if multiple lattices were present. Despite several attempts the diffraction quality could not be improved to reasonable levels.

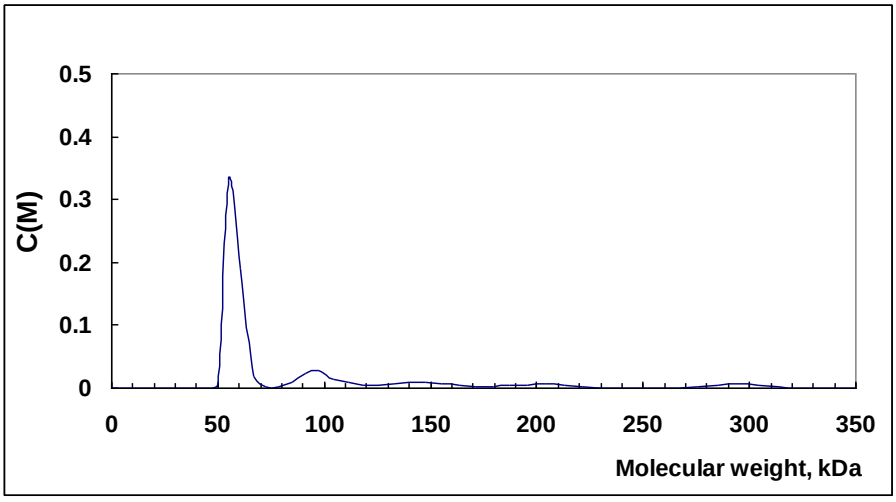


Figure 18: AUC result of the ETR1 GAF and dimerization domain indicates a dimer (calculated size of the monomer 28.4 kDa)

CTR1 kinase domain

The C-terminal and catalytically active protein kinase domain of CTR1 (CTR1-kd) from *A. thaliana* was heterologously expressed in *E. coli* in the presence of chaperones and purified via NiNTA and SEC. In addition to the wild type construct, a kinase dead variant of CTR1-kd was expressed and purified and characterized both kinetically and crystallographically in order to elucidate the mechanisms, which account for the activation of CTR1-kd. Both constructs start 18 residues before the first glycine residue of the conserved P-loop motif (residue 540 of full length CTR1). The wild-type version extends to the natural C-terminus while the inactive kinase lacks the last 10 amino acids, which have been found structurally disordered in the wild-type version.

Heterologously expressed wild type CTR1-kd was purified from *E. coli* complemented with chaperons as a mix of multiply phosphorylated forms. Mass spectrometry indicated 3, 4, 5 and 6 phosphorylation sites (Figure 21A). Those sites, identified by tandem MS (MSMS) peptide fingerprinting, are concentrated within the P-loop motif at Ser^{561} and Thr^{564} and within the activation loop at Ser^{703}, Thr^{704}, Ser^{707} and Ser^{710}. The complete absence of phosphorylation sites in a, likewise heterologously expressed, kinase dead version of CTR1-kd (CTR1-D676N, see below) shows that the identified phosphorylation sites are the result of self-directed activity. Attempts to further fractionate the mixture of differently phosphorylated CTR1-kd using a MonoQ column failed. This finding is consistent with earlier observations, which suggested multiple phosphorylation sites in full length CTR1 as well as in its protein kinase domain and the absence of phosphorylation in a kinase dead version (Huang et al., 2003). There the samples were expressed in insect cells but neither the extent nor the locations of the phosphorylation sites were given.

The initial storage buffer (Tris pH 8.5, 150 mM NaCl, 1 mM DTT, 1 mM $MgCl_2$ and 1% glycerol) was found to be unsuitable for the protein and a thermofluor screen was employed to improve the long-term sample stability. HEPES pH 7.5 and an elevated NaCl concentration (350 mM) were found to be beneficial for protein stability and were subsequently used.

Initial crystallization trials yielded no crystals in the presence of 1mM AMPPNP. The stabilizing effect of a number of ligands was assessed using a thermofluor screen. The addition of staurosporine gave the strongest improvement in terms of thermostability of about 21°C (Figure 19). The addition of the broad range kinase inhibitor staurosporine,

albeit at a very low concentration of 100 µM (approximately 4x the protein concentration), yielded in the initial screen small crystals in a number of conditions.

Supported by kinetic data, the structures therefore represent CTR1-kd in its active and inactive conformation, respectively. Crystallization of CTR1-kd and CTR1-D676N yielded orthorhombic crystals of different subgroups, each with two molecules per asymmetric unit (Mayerhofer et al., 2011). Data collection statistics and refinement for all crystals are summarized in Table 8.

Table 8: X-ray data-collection, processing and refinement statistics of the CTR1 kinase domain. Values in parentheses are for the highest resolution shell

	CTR1-kd	CTR1-D676N
X-ray source	ID23-2, ESRF	ID29, ESRF
Wavelength (Å)	0.8726	0.9763
Resolution range (Å)	46.2-3.01 (3.18-3.01)	47.1-2.50 (2.64-2.50)
Temperature (K)	100	100
Crystal-to-detector distance (mm)	293.3	371.2
Rotation range per image (°)	1.0	0.75
Total rotation range (°)	70	220
Space group	$P4_12_12$	$P4_22_12$
Unit-cell parameters (Å)	a = b = 95.71, c = 179.78	a = b = 122.33, c = 95.01
Observed reflections	93474 (14334)	334176 (52905)
Unique reflections	17053 (2648)	47817 (7636)
Multiplicity	5.5 (5.4)	13.1 (13.2)
R_{merge} (%)	17.6 (84.2)	9.0 (83.3)
R_{pim} (%)	8.5 (40.0)	2.7 (25.3)
Completeness (%)	98.3 (97.4)	99.8 (98.9)
$\langle I/\sigma(I)\rangle$	9.5 (1.8)	16.9 (2.4)
Wilson B factor (Å^2)	73	60
Optical resolution (Å)	2.2	1.9
Refinement		
Rwork / Rfree	20.1 / 25.6	19.5 / 23.2
No. atoms		
Protein	4286	3939
Ligand/ion	34	34
Water	-	61
B-factors		
Protein	60	55

Ligand/ion	73	71
Water	-	46
R.m.s. deviations		
Bond lengths (Å)	0.002	0.01
Bond angles (°)	0.54	0.96

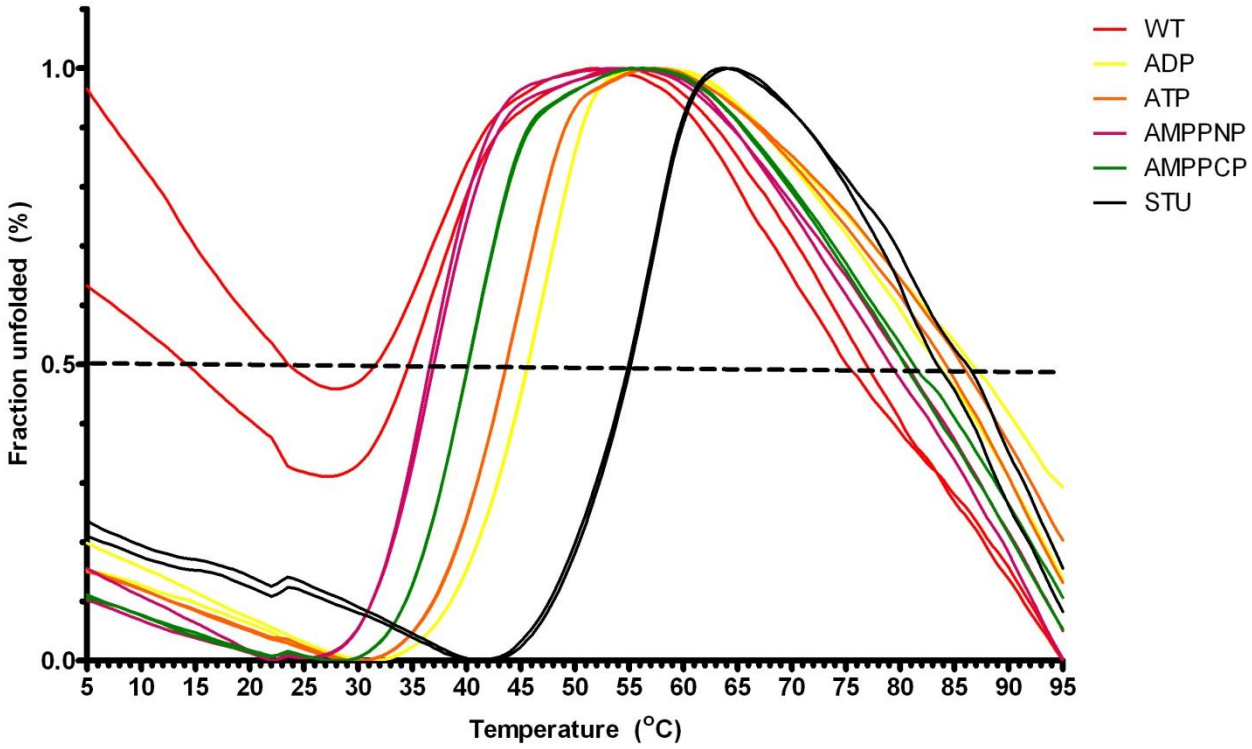


Figure 19: Thermofluor results of the CTR1 kinase domain in the presence of different nucleotides and staurosporine (STU). Every molecule was run as a duplicate.

Structure of active CTR1-kd

Despite the non-homogeneously phosphorylated sample, crystallization succeeded in the presence of staurosporine. Crystals diffracted x-ray radiation up to 3 Å resolution and they contained two independent copies of the kinase in the asymmetric unit. As expected, CTR1-kd adopts the characteristic protein kinase fold with a 5-stranded anti-parallel β-sheet in the N-lobe (res. 540-626), including a characteristic α-helix, and a larger, predominantly α-helical C-lobe (res. 632-821). Both lobes are connected by a short linker (res. 627-631) (Figure 20).

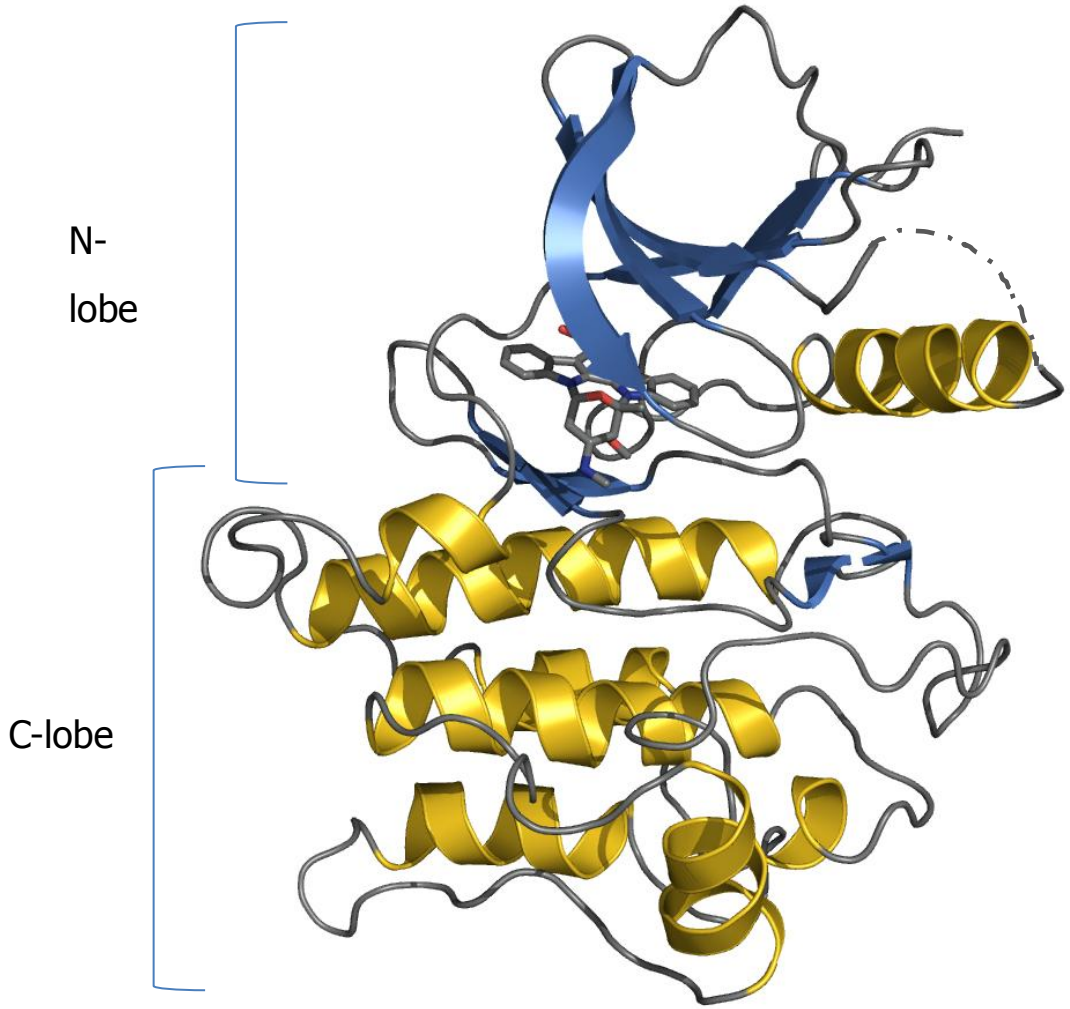
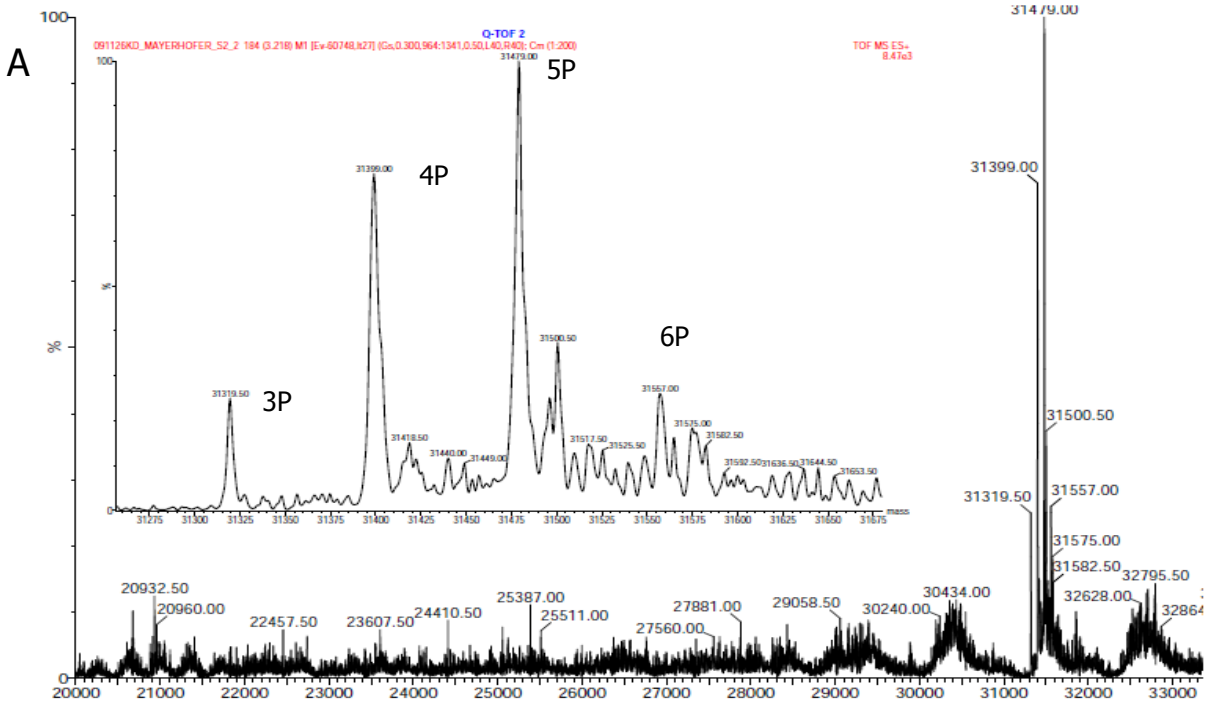


Figure 20: Structure of the wild-type CTR1 kinase domain in cartoon representation. The broad range kinase inhibitor staurosporine is shown in sticks. The missing loop between β3 and helix C is indicated by a dotted line.

Both molecules show well-defined and continuous electron density with the exception of one short stretch in each copy. The P-loop of protomer B shows high B-factors and disconnected density at the tip of the loop from residues Gly^{560} to Gly^{563} in contrast to protomer A where the same stretch is well defined. Conversely, there is no electron density for the loop preceding helix C (res. 583 to 586) in protomer A, while the same region shows well-defined connectivity in protomer B. After several rounds of refinement and manual model building, electron difference maps showed positive density for staurosporine in the active site of protomer A. No ligand was added in the active site of protomer B due to the very weak electron density there. Apparently, the P-loop is very



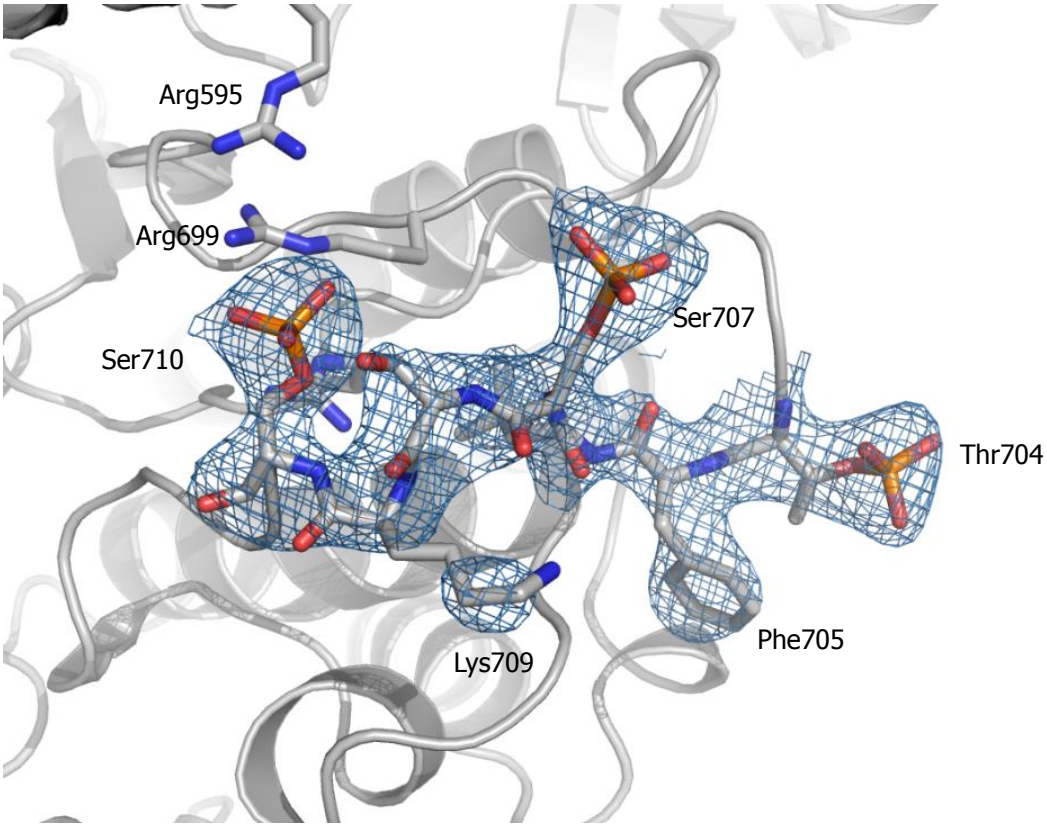


Figure 21: **A** MS result of CTR1 wild-type and a zoom of the region corresponding to the masses of the 3-6 fold phosphorylated CTR1-wt **B** Activation loop of CTR1 wild-type with electron density of residues 704-710 contoured at 1σ, shown in blue. Selected residues are shown as sticks

flexible in the absence of a ligand in the active site. It seems reasonable to assume that the natural interaction with ATP has a similar effect on the P-loop.

In both copies of CTR1-kd the side chains of Thr^{704}, Ser^{707} and Ser^{710}, located in the activation loop, are clearly phosphorylated (Figure 21 B), as indicated by MSMS fingerprinting analyses. Despite the presence of alternatively phosphorylated species in the purified sample, no other phosphorylation sites were found in the structure of CTR1-kd. Apparently, this fraction of the sample is sufficiently abundant and conformationally homogeneous for crystallization.

Structure of inactive CTR1-kd

In order to understand the principles, which govern the activation of CTR1-kd, we produced an enzymatically inactive version of the kinase by replacing a strictly conserved aspartate in the catalytic loop with asparagine (CTR1-D676N). The Asp^{676} serves as a general base in abstracting a proton from the hydroxyl group of the attacking substrate. The crystallographic structure of CTR1-D676N at 2.5 Å resolution reveals that the functionally important activation loop is completely unstructured. Disorder begins at Phe^{695} of the DFG motif, which initiates the activation loop, and continues until about six residues upstream of the loop terminating APE motif (res. 697-712 of protomer A and res. 697-711 of protomer B). Additional disorder exists in the loop of protomer B connecting strand β3 and helix C as well as the N-terminus of helix C. The equivalent region in protomer A however, shows clear connectivity. This could, at least partially, be explained by a different surrounding in the crystal with this region in protomer B having no interaction partner as does protomer A.

CTR1-D676N was also crystallized in the presence of staurosporine. As in CTR1-kd, the ligand is present in the active site of protomer A of CTR1-D676N while it is absent from protomer B. Again the absence of staurosporine is accompanied by significantly higher temperature factors of the P-loop, which is localized further towards the empty active site compared to protomer A whose active site is occupied. This increased flexibility of the P-loop is mirrored at the adjacent strand β3 and the loop connecting it with helix C. In contrast to CTR1-kd, there is no electron density for the entire loop and the amino terminus of helix C between residues 580 and 592 in ligand-free protomer B of the inactive mutant. In the inhibitor bound form of CTR1-D676N, these regions are well defined. This observation underlines the interdependence of structural integrity and active site

composition in CTR1. No phosphorylation sites were found in the structure of CTR1-D676N (Figure 22B). Mass spectrometry corroborates this and further proves that the structurally disordered activation loop of the kinase dead mutant is indeed un-phosphorylated.

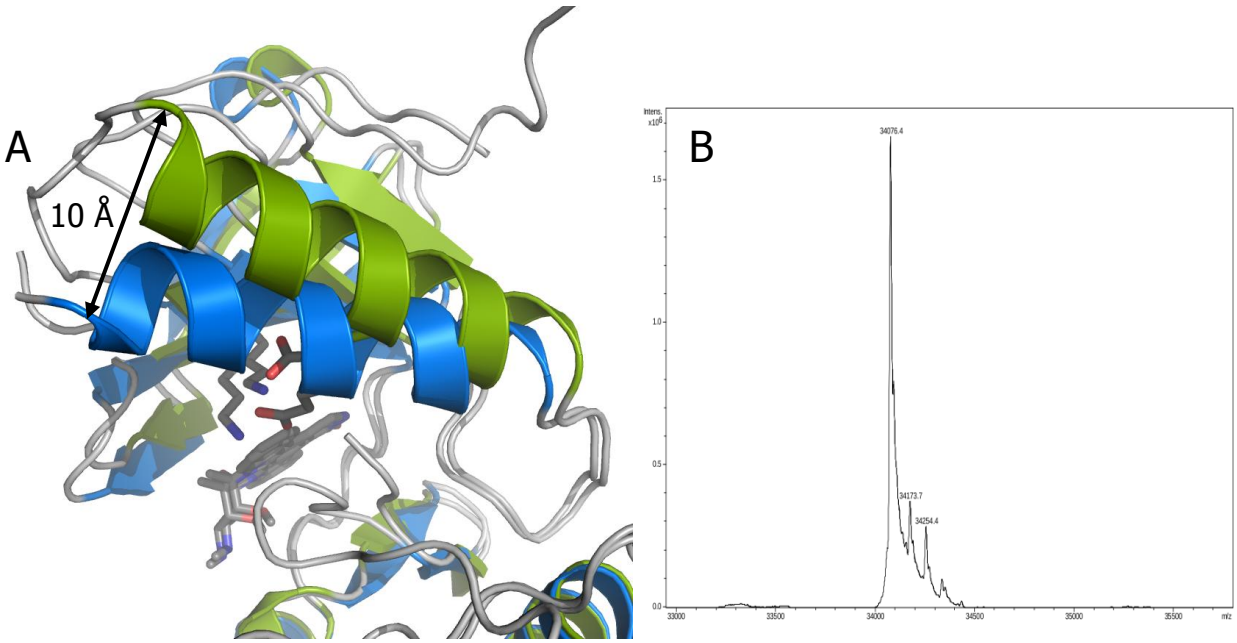


Figure 22: **A** Movement of helix C between the active (blue) and inactive (green) form of the CTR1 kinase. The comparison is based on an LSQ alignment of The C-terminal domains (res. 632 to 810) of CTR1-kd and CTR1-D676N. **B** MS result of inactive CTR1 showing no phosphorylation.

Dimerization

Regulation of kinase activity through the formation of homo- and heterodimers is an established phenomenon. In particular, the ability of B-RAF to homodimerize and its *in vivo* activation through heterodimerization with KSR, a pseudo-kinase, has been demonstrated recently (Rajakulendran et al., 2009). Since active (Figure 25 A) and inactive (Figure 25 B) CTR1-kd form apparent homodimers in the crystal in different space groups across interfaces, which are virtually identical to those found in B-RAF (Figure 25 C) and RAF-1 (Figure 25 D), the possibility of a similar mechanism was explored. In addition all the residues which play a crucial role in the dimerization of B-RAF (Rajakulendran et al., 2009) are conserved in CTR1 (Figure 23). Standard purification via size exclusion chromatography (SEC) indicated a monomer for CTR1-kd and CTR1-D676N rather than a dimer, with a calculated size of 40 kDa compared to the theoretical weight of 35 kDa. Analytical ultra-centrifugation, however, clearly showed a single peak with a maximum at 63kDa for CTR1kd, which correlates well with twice the molecular weight of the monomer

of 35kDa (Figure 26 A). The dimer interfaces (Figure 24 A) observed in the crystal structures when calculated by PISA (Krissinel and Henrick, 2007) cover 1249 $Å^2$ and 1072 $Å^2$ in CTR1-kd and CTR1-D676N, respectively. The interface residues are basically the same in both structures. The variance in surface area arises from small differences in the side chain orientations.

There are several amino acids, which are conserved among the dimer interfaces of CTR1 (Figure 24 B), B-RAF and KSR. Arg^{604}, Phe^{611} and Met^{612} (CTR1 numbering) were shown to be critical to the stability of the dimer interface in B-RAF and KSR by mutation (Rajakulendran et al., 2009). Initially three mutants were made, namely Arg^{604}His, Phe^{611}Ala and Met^{612}Trp (M20 - M22), and purified similar to CTR1-kd. These residues are conserved among CTR1 proteins from various plants (Figure 23 blue stars). Mutations in CTR1-kd, equivalent to those found to monomerize B-RAF, resulted in the formation of predominantly monomers (Figure 26 A) as analysed by AUC. Interestingly when CTR1-D676N was analysed by AUC it is also present mainly as a monomer (Figure 26 B), differing from the dimer observed in the crystal structure. The three interface mutants - despite being mainly monomeric - were found to be phosphorylated by MS implying that they are least partially active (see below).

This supports the importance of the conserved residues on the formation of the interface. On the other hand there exists no simple correlation between the activity of CTR1 and the dimerization, at least when investigating the isolated kinase domain of CTR1 (see next chapter).

Three more mutants were designed and purified namely Arg^{604}Ala (M23), Arg^{602}Asp & Arg^{604}Ala (M24) and FM^{611}AW (M25). The additional Arg^{604}Ala mutant was also designed as the initial Arg^{604}His mutant, when modelled in the structure, could – depending on the actual side chain orientation of the histidine - still contribute a hydrogen bond to the main chain of Met^{600} or Leu^{603} of the dimer partner, which could affect the activity. Arg^{602} forms hydrogen bonds with $Asp^{540,\ 543}$ and Asn^{669}. Arg^{604}Asp was designed to increase the repulsive forces between the two molecules. The FM^{611}AW double mutant was made to test if the reduction in activity of the individual mutations is additive. Again, while all mutants are mainly monomeric, a small fraction was still present as a dimer (Figure 26 C, D and E) and again all of them showed some degree of phosphorylation.

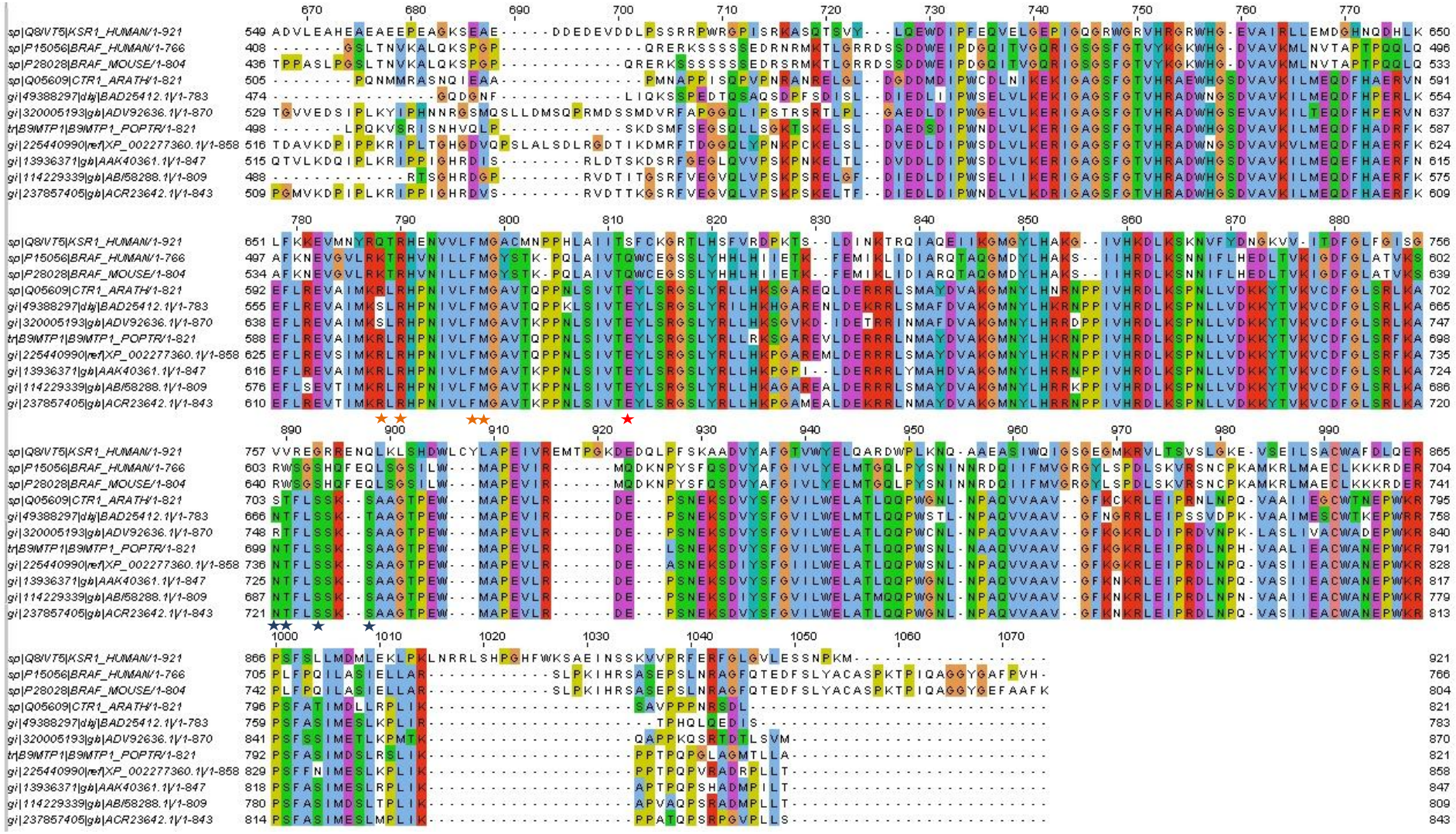


Figure 23: Alignment of the kinase domains of CTR1 from *A. thaliana, Oryza sativa, Cucumis melo, Populus trichocarpa, Vitis vinifera, Rosa hybrid cultivar, Malus x domestica, Prunus persica*, human and mouse B-RAF and human KSR1. Mutated residues are indicated by stars below the amino acid. Red Glue^{626}Lys, orange interface mutants and blue activation loop mutants

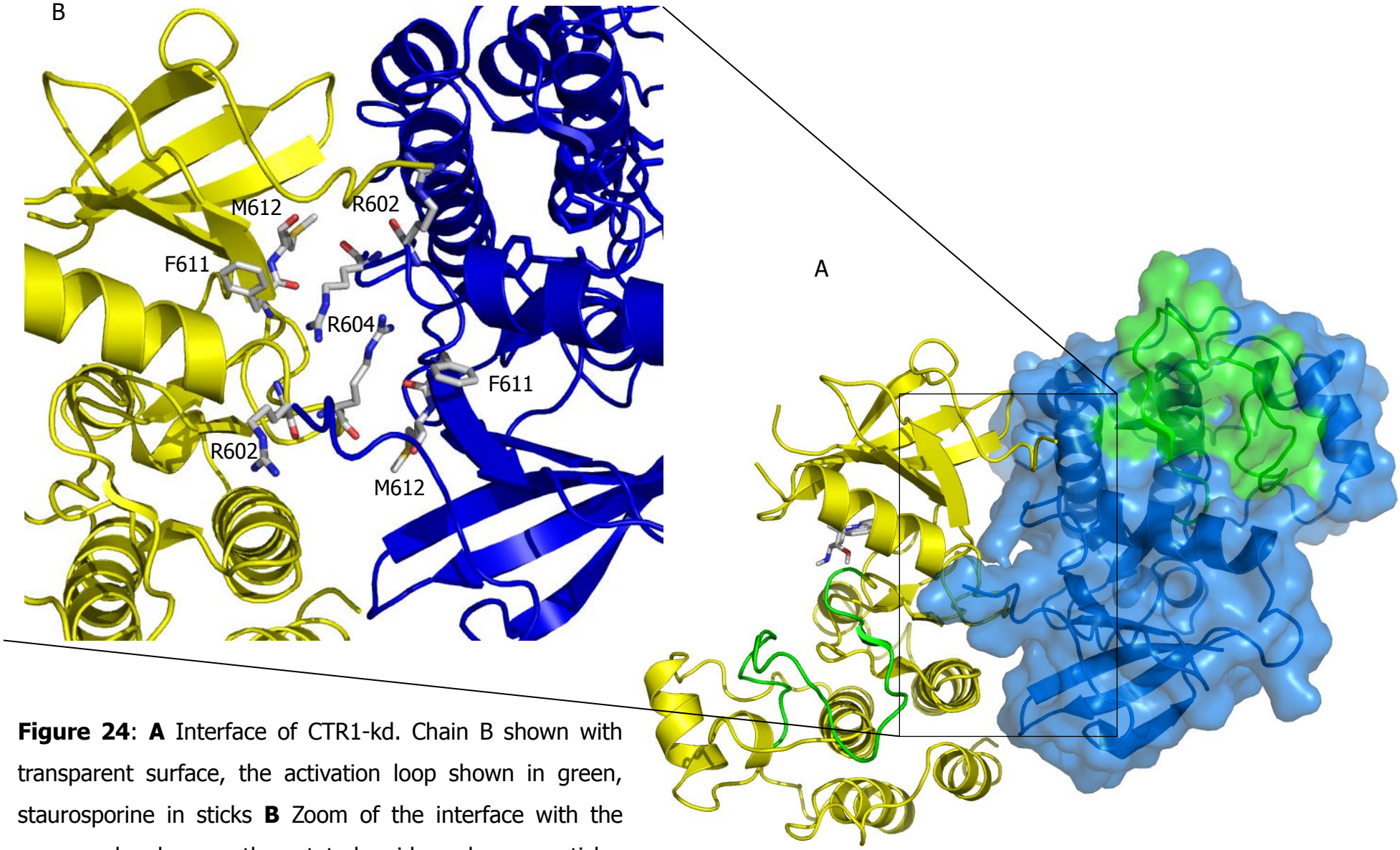


Figure 24: **A** Interface of CTR1-kd. Chain B shown with transparent surface, the activation loop shown in green, staurosporine in sticks **B** Zoom of the interface with the conserved, subsequently mutated residues shown as sticks. Chain A shown in blue, chain B in yellow. Staurosporine is omitted for clarity.

In order to test the effect of ATP onto dimerization the three mutants were run on an analytical S75 SEC column in the presence (1mM) and absence of ATP. While there was a clear difference in elution volume compared to the wild-type, confirming again the monomeric state, ATP had no influence on the oligomerization state. Therefor the observed continuous activity in the coupled assay of M20, M23 and M24 seems to be possible with >95% of the protein present as a monomer.

As the unphosphorylated CTR1-D676N forms a monomer while the phosphorylated CR1-kd forms a dimer the effects of activation loop phosphorylation on the oligomeric state was tested. Two mutants were tested a double mutant of Thr^{704} and Ser^{707} to alanine and triple mutant adding the Ser^{710} to alanine mutant. Both were analysed by AUC and show a mix of monomer and dimers being present (Figure 26 F and G).

The expression of the recently published mutant $Glu^{626}Lys$ (Ikeda et al., 2009) failed unfortunately and the effect on the dimerization could not be assessed.

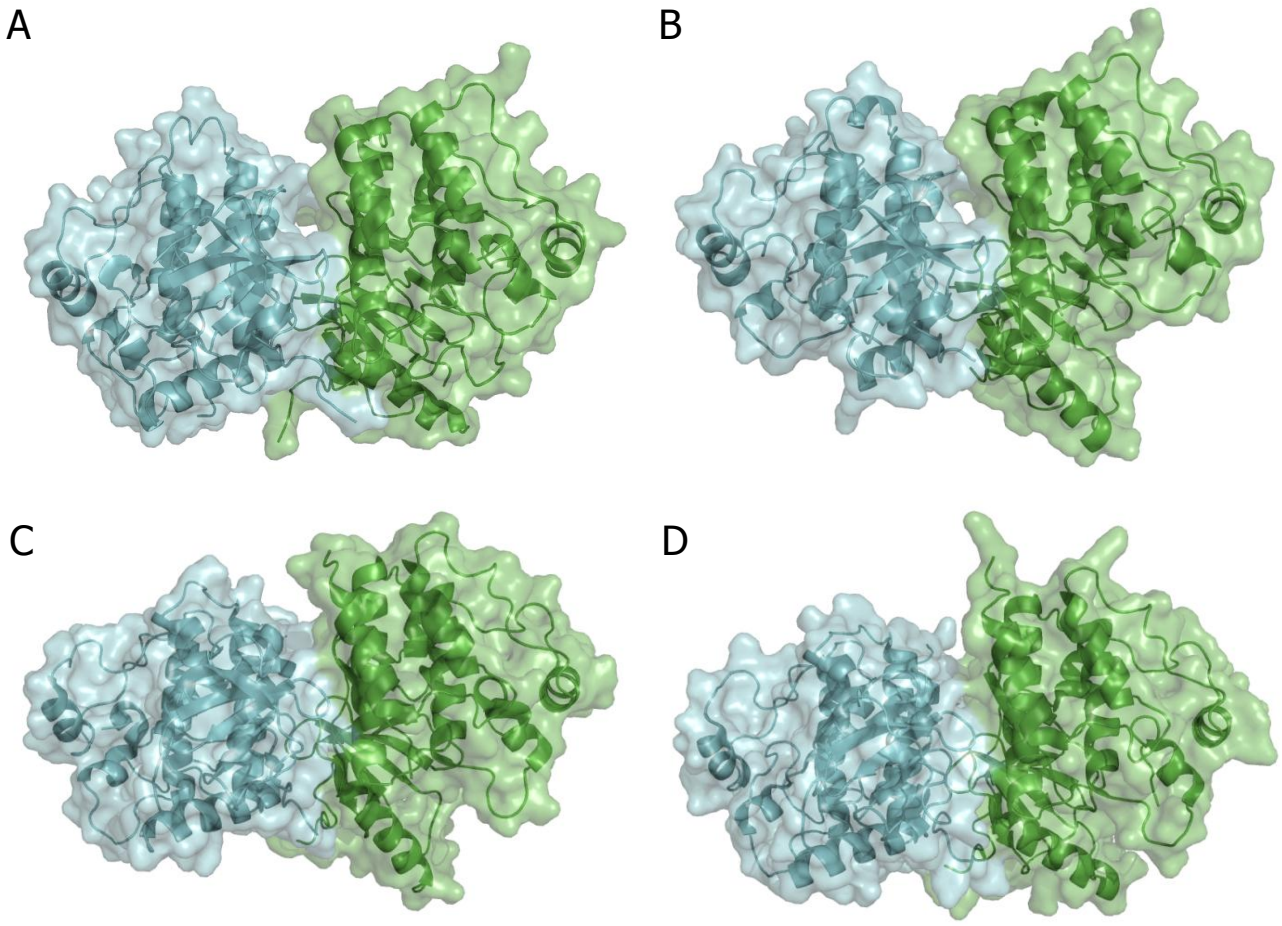


Figure 25: Cartoon and transparent surface representation of the molecules as observed in the crystal structures of **A** CTR1-kd **B** CTR1-D676N **C** B-Raf (pdb ID 3C4C) **D** RAF-1 (pdb ID 3OMV). The individual chains are shown in blue and green

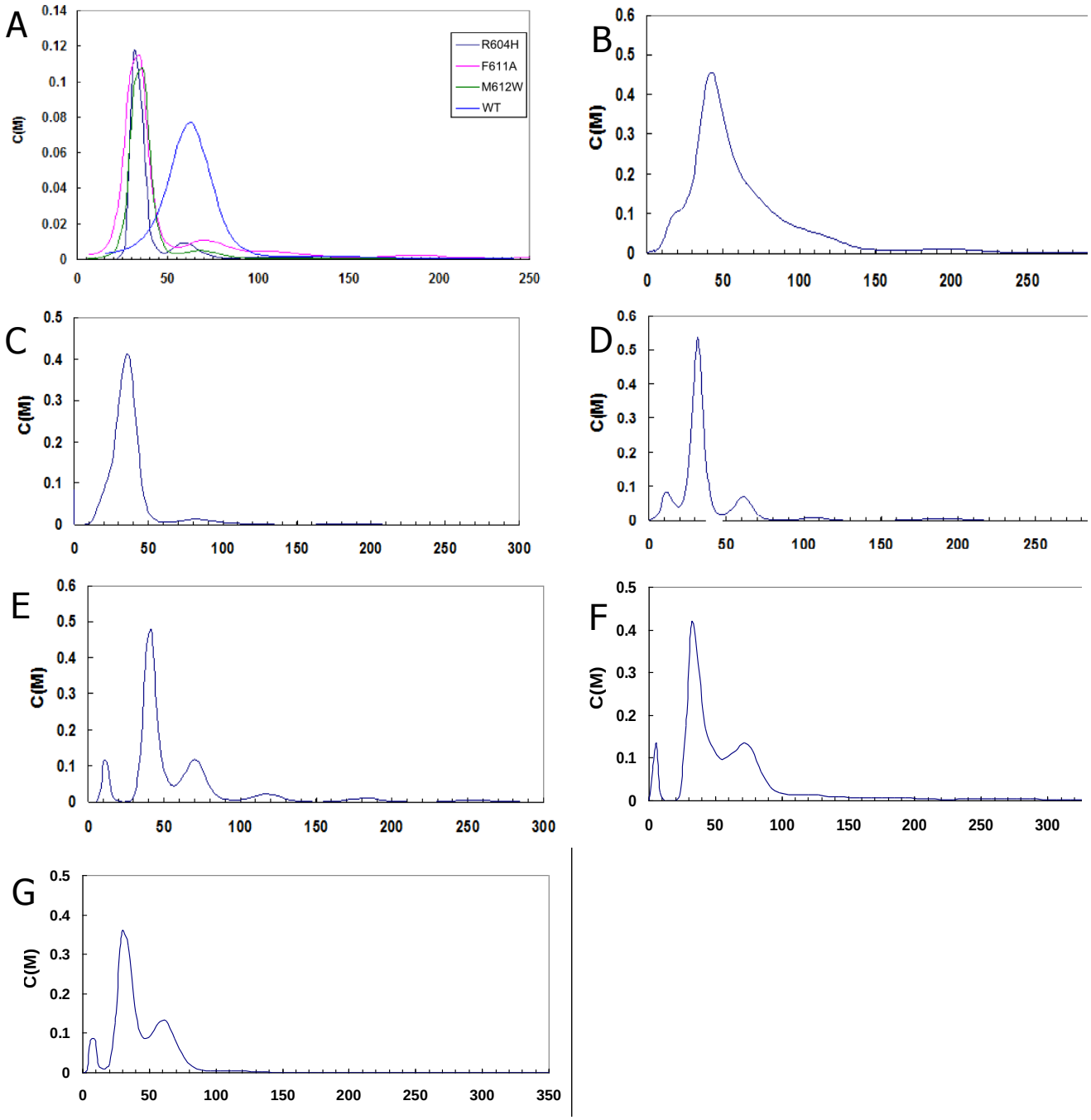


Figure 26: Analytical ultracentrifugation of the **A** CTR1 kinase domain and three interface mutants (Arg^{604}His, Phe^{611}Ala and Met^{612}Trp) affecting dimerization, **B** of the CTR1 kinase dead mutant (CTR1-D676N) and the three additional mutants **C** Arg^{604}Ala, **D** Arg^{602}Asp & Arg^{604}Ala and **E** FM^{611}AW present predominantly in the monomeric state. Mutants of the activation loop **F** Thr^{704}Ala and Ser^{707}Ala and **G** with the additional Ser^{710}Ala mutation show a mix of monomeric and dimeric protein

Activity of CTR1

CTR1-kd is multiply and heterogeneously phosphorylated when expressed in and purified from *E. coli*. The attempts to produce a homogenously phosphorylated form of CTR1-kd *in vitro*, by incubating the protein with an excess of ATP overnight, were unsuccessful as shown by mass spectrometry. To further explore the relationship between structure and activity of CTR1 a coupled assay was employed, which records the consumption of NADH by lactate dehydrogenase (LDH) in the process of reducing pyruvate to lactate. The pyruvate is provided by pyruvate kinase (PK), using phosphoenolpyruvate (PEP) and ADP. The only source of ADP in this assay is CTR1 as long as it consumes ATP.

The assay shows autophosphorylation activity of CTR1-kd. The K_m-value for ATP, calculated by the coupled assay, is 50 µM (Table 9). This value is about 5-times higher than a previously published result of 9.1 µM when using BSA as substrate (Huang et al., 2003). Our MS analyses, however, clearly show that CTR1-kd enters our activity assays in an already multiply phosphorylated state. This suggests that this CTR1-kd is in a dynamic equilibrium where one or several of the phosphorylated side chains of CTR1-kd are labile and undergo rapid hydrolysis or dephosphorylation as in the coupled assay the consumption of ATP by CTR1-kd is continuous. Even at a NADH concentration three orders of magnitude above the initial ATP concentration, the reaction continues until all of the NADH has been consumed. When the assay is resupplied with NADH, the reduction of pyruvate to lactate resumes. A equilibrium is also consistent with the inability to produce a homogeneously modified kinase by prolonged incubation of CTR1-kd with ATP. As expected, CTR1-D676N shows no activity confirming the crucial importance of Asp^{676} for protein kinase activity.

Table 9: K_m and k_{cat} values for autophosphorylation of the CTR1 kinase domain and three interface mutants as determined using the coupled assay.

	WT	R604H	F611A	M612W
K_m (µM)	49,94	51,11	356,1	40,65
k_{cat} (min^{-1})	65,91	63,27	4,15	26,91

As CTR1-kd was shown to be present as a dimer in solution by AUC a number of mutants targeting the interface ($Arg^{604}His$, $Phe^{611}Ala$, $Met^{612}Trp$) were assessed for their effect on the catalytic activity. Surprisingly even though the mutants were clearly predominantly monomeric the influence on the activity was quite divergent (Table 9).

The monomer inducing mutations had different effects on the activity of CTR1-kd *in vitro*. While one mutant ($Arg^{604}His$) had basically no effect on the activity the two other mutants showed a more pronounced influence on the activity. While $Met^{612}Trp$ reduced the activity by approximately 50% the $Phe^{611}Ala$ mutant more or less completely abolished the activity (Figure 27, Table 9). This shows that the dimer interface has a big influence on the catalytic activity as measured in the coupled assay, even though it is located on the opposite site of the activation loop and the catalytic residues. Attempts to obtain the crystal structure of the inactive, monomeric mutants were so far unsuccessful.

All mutants are mainly monomeric but still a small dimer fraction is present (Figure 26). Therefore additional mutations were tested in order to explore if double mutants could completely remove the observed dimer fraction. The three additional mutants $Arg^{604}His$, $RLR^{602}DLA$ and $FM^{611}AW$ (M23-M25) were also assessed regarding their activity. While the $Arg^{604}Ala$ and $RLR^{602}DLA$ mutants did hardly affect the catalytic activity with a comparable activity to CTR1-kd, the M25 mutant was again showing no activity in this assay (Figure 28).

In the structure of CTR1-kd three residues were clearly phosphorylated. One of them, Ser^{710}, is at a highly conserved position, termed the primary phosphorylation site in RD kinases. Interaction between the primary phosphoryl group and the basic pocket in CTR1-kd (Figure 34) structurally resembles the situation in the active forms of other RD kinase representatives from the AGC, CMGC and TK kinase families. An overlay of the C-terminal domains from four AGC (pdb ID 1OL5, 1ATP, 1O6K, 1H1W), three CMGC (pdb ID 1QMZ, 2ERK, 1CM8), and three TK (pdb ID 1IR3, 1K3A, 3LCK) family members and CTR1-kd shows that the respective primary phosphates are on average within 1.8 Å (σ = 0.9 Å), despite dissimilar lengths and conformations of their activation loops.

Therefore the effects of removing the phosphorylation sites in the activation loop were explored. Four residues were chosen, the three observed in the crystal structure (Thr^{704}, Ser^{707} and Ser^{710}) and in addition Ser^{703} which was found by MS but not in the crystal structure. All four residues were mutated to alanine (M14) and, using this mutant, the

effect of individual replacement to glutamate, mimicking phosphorylation, was tested (M15-M18). All mutants abolished the continuous activity observed in the wild-type to the level of the kinase-dead mutant. None of the single glutamate substitutions could restore this activity (Figure 29). Nevertheless when being measured by MS all mutants were found to be phosphorylated.

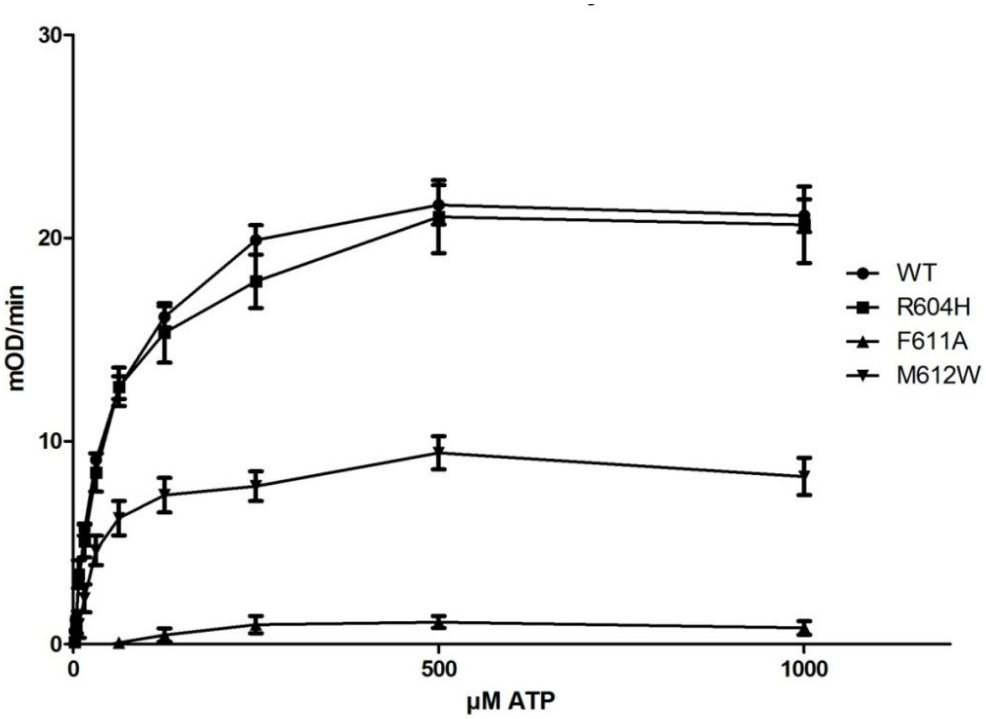


Figure 27: Catalytic activity of CTR1-kd and three single interface mutants

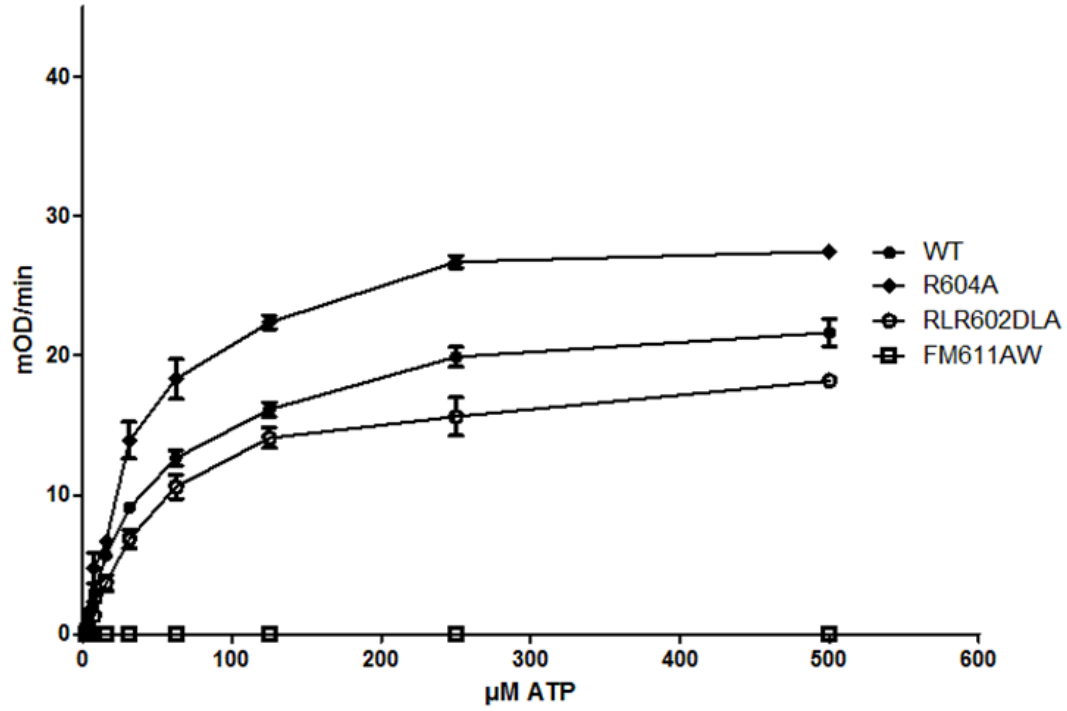


Figure 28: Activity of CTR1-kd and the three additional interface mutants $Arg^{604}Ala$, $RLR^{602}DLA$ and $FM^{611}AW$

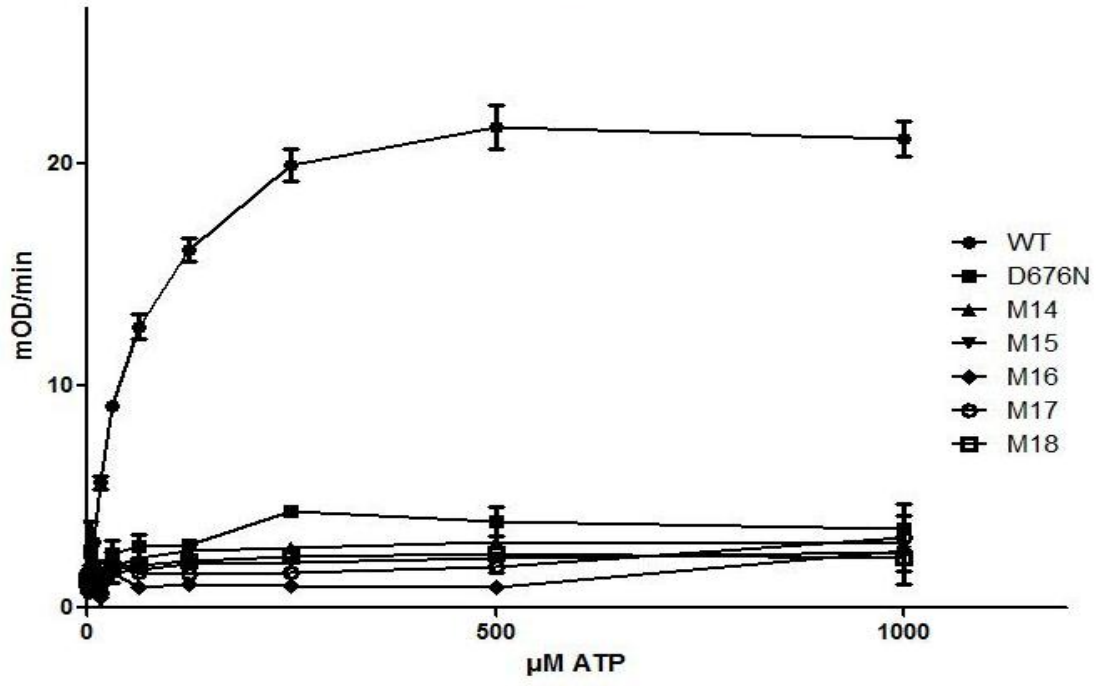


Figure 29: Activity measured by the coupled assay of the activation loop mutants M14-M18 mutating four residues to alanine followed by individual replacements to glutamate, mimicking phosphorylation

Discussion

The effects of ethylene have been utilized and observed for thousands of years starting from burning incense to enhance the ripening of fruits to the observation of early fruit ripening of plants growing close to gas street lamps. For more than hundred years (Neljubov, 1901) ethylene is known to cause the numerous known effects and since the mid-80s, with the first mutants of members involved in ethylene signalling, a complex network has be uncovered. A number of crucial proteins has been discovered enabling the transmission of the ethylene signal from the membrane-bound ethylene receptors and the prime interaction partner CTR1 to the transcription factors EIN3 and EIL1 in the nucleus (Stepanova and Alonso, 2009). Despite decades of research on ethylene signalling using genetic and biochemical techniques leading to well characterized ethylene receptors, structural information of the proteins involved remained incomplete, with only the structure of the ETR1 receiver domain being available (Müller-Dieckmann et al., 1999).

The histidine kinase domain and the receiver domain, the latter not present in all receptors in *A. thaliana*, show similarity to the bacterial two-component system typically found in prokaryotes. On the other hand is the C-terminal kinase domain of CTR1, similar to RAF kinases which are classically the starting point of a MAPK pathway. Structures from prokaryotic HKD containing proteins also in complex with the receiver domain and also of the eucaroytic RAF family are available.

In order to improve the knowledge of this interesting intersection of two different signal networks a number of domains and domain combinations were expressed and the structures of the ERS1 dimerization domain and the CTR1 kinase domain were determined to high resolution as well as the solution structure of the cytoplasmic domains from ETR1 using SAXS.

ERS1 dimerization domain

HKs are involved in a number of reactions namely autokinase, phosphotransferase and phosphatase reactions and the dimerization domain plays a role in all of them. The different activities are often regulated by their N-terminal sensor domain and the signal is thought to be transmitted to the other domains through a rotation and tilting movement of the helices, leading to different conformations of the HKD. The dimerization domain of

ERS1 is so far the first described structure of this domain of eukaryotic origin and extends the observed N-terminal portion compared with prokaryotic structures. It reveals an organization similar to the available results from bacteria, with a dimer consisting of an N-terminal parallel coiled-coil followed by a four-helix bundle. The observed, longer N-terminal part furthermore allowed for a more reliable placement of the GAF domain in the model used for the cytosolic domains of the ethylene receptor.

The original construct, based on bioinformatic domain prediction, needed to be extended by 18 residues at the N-terminus in order to obtain diffraction data to high resolution. The reason for this could be an improved and thereby stronger crystal contacts between the N-terminus of the now extended coiled-coil with the four-helix bundle in symmetry related molecules as exemplified by the different interaction interfaces. While in the determined structure this interface is about 400 Å (as calculated by PISA) it is reduced to 230 Å when artificially reducing the structure to the size of the initial construct H9. The N-terminal extension contributes a number of polar and apolar interactions in the crystal structure between symmetry related elements. Even though the interface is less than 1000 Å in size, too small to be considered stable in solution, this difference can be enough to stabilize the crystal assuming a similar packing between the crystals. This modification also avoided or at least modified the packing observed in the initial construct H9, which possibly consisted of alternating layers in a hexagonal space group rotated by 30° with respect to each other, giving rise to the observed, interesting diffraction pattern (Figure 11 B).

Interface

Coiled-coils are a relatively simple building block of proteins with the hallmark of a seven-residue heptad repeat with hydrophobic residues at position *a* and *d* forming a seam of knobs-into-holes interactions (Crick, 1953). The interface of the coiled-coil in ERS1 is not always following the classical heptads. There are also polar (Ser^{316}, Gln^{326}, Asn^{327}) and charged residues (Arg^{320}, Arg^{334}) in the interface which are - with the exception of Ser^{316} - located at the *d* or *e* position of the heptad. Charged residues at positions *a* and *d* were also found in the HKs EnvZ or Sln1. The interface residues are conserved among the five ethylene receptors in *A. thaliana*. Out of 11 interface residues between residues 312 and 339 eight are completely conserved and two are conservatively

replaced and one (Ala^{319}) is occupied by Met in subfamily 2 (Figure 11 A). This suggests a very similar interface in all five ethylene receptors.

The coiled-coil in the yeast HK Sln1, was found to be needed for the dimer formation and normal activity (Tao et al., 2002). Recently also the structures of the DesK HK in three different activation states became available. There conformational signals are transmitted through interhelical rearrangements in the DHp domain modulating the positioning of the CA domain and the interaction with the receiver domain (Albanesi et al., 2009). This movement is accompanied by changes in the position and angle of the coiled-coil helices and rearrangements of the four-helix bundle exposing a hydrophobic patch, allowing for interactions with the CA domain leading to its change of position. Similar, in the structure of the isolated TM0853 HKD from *T. maritima*, the N-terminal coiled-coil is visible, while it is more twisted and disordered in the complex with the receiver domain.

When comparing the structure of the ERS1 DHp domain with the three DesK HKDs the highest similarity is present with the structure in the phosphatase-competent state. There, as in ERS1, an N-terminal part is ordered and forms a coiled-coil. Conversely in the other conformations this part is not visible as the helices are disordered or show a very asymmetric conformation.

Mutations in the C-terminus of the DHp domain were found to affect phosphatase activity, while not changing kinase activity (Hsing et al., 1998). These residues, when mapped on the structure of the *T. maritima* HKD, are central in the interface with the CA domain. When comparing this region of the ERS1 and ETR1 DHp domains with bacterial structures, the residues interacting with the CA domain are conserved (ERS1 residues Asp^{401}, Leu^{405} and Glu^{406} and Leu^{450}, Ala^{453}, Leu^{454}, Arg^{459} and Gln^{462} in the CA domain). In the ERS1 DHp domain they are positioned such that they are accessible for an interaction with the CA domain, arguing for a similar positioning of the CA domain as observed in phosphatase-competent bacterial structures. While actual phosphatase activity is not likely in ERS1 due to the lack of a receiver domain this gives some indication that a similar interaction and binding between the DHp and CA domain also exists in eukaryotes.

When the sequences of the linkers between the two helices in the four-helix bundle of the five receptors in *A. thaliana* are compared, they appear to fall into two groups. The linkers of subfamily 1 (ETR1 and ERS1) have nearly identical linker sequence (only ERS1 Ser^{375} exchanged for ETR1 Thr^{375} in the six residues linker), while the linker sequences of

subfamily 2 are diverse (Figure 11 A). The topology in the ETR1 dimerization domain will probably be quite similar to the situation in ERS1, also as the sequence identity between the two dimerization domains in subfamily 1 is very high (84%), while no prediction of the connectivity between the helices can be made for subfamily 2. Given these differences it is possible that the connectivity is different in subfamily 2, which could allow for a *cis*-phosphorylation mechanism.

An interesting feature of the dimerization domain of ERS1 is the different bending angle between helix A1 and A2 observed within the dimer. This difference can be observed in two different space groups, $C222_1$ and $P2_12_12$, and between the two dimers per AU in $P2_12_12$. While the r.m.s.d. between equivalent Cα atoms between the protomers of similar conformation is 1 Å, it increases to 2 Å for those of a different conformation. This is true for the comparison within dimers or across space groups. The observed difference of 9° of the kink within dimers is therefore consistent and unlikely to be accidental. As a consequence, the dimerization domain of ERS1 is asymmetric (Figure 30) and this asymmetry would also be found in the N-terminal GAF domains and possibly all the way into the membrane domain. Because this part of the receptor was also shown to be involved in binding of CTR1, the conformational difference may indicate one way of communication between signal binding domain and the cytosolic output domains.

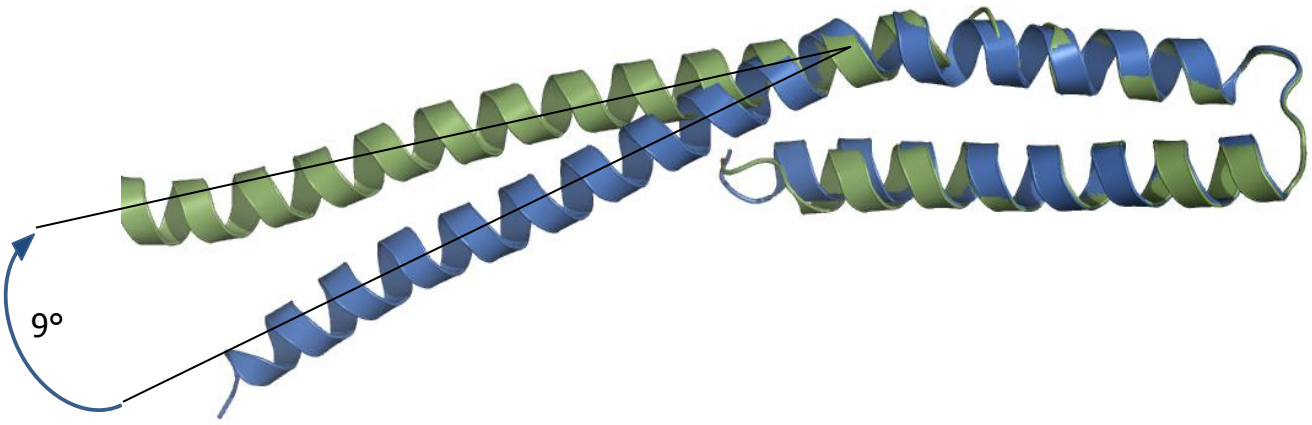


Figure 30: Superimposition of the two helices in the ERS1 dimerization domain from residue 351 to 404. Helix A is shown in green, helix B is shown in blue.

The structure of the entire cytoplasmic portion of the *T. maritima* HK TM0853 is available as well as structural information of the DesK and KinB HKs and of the dimerization domain of EnvZ. Superimposition with the EnvZ DHp domain does not yield a

very good fit caused by the partly disordered nature of EnvZ whose structure was determined in solution by NMR. EnvZ exhibits an unusual orientation angle between neighbouring helices different from those previously determined as being favourable (Chou et al., 1988). The DHp domain of ERS1 exhibits the more stable angle similar to the HKD of TM0853 from *T. maritima*. With the remaining structures the four-helix bundle superimposes well but shows more divergence in the coiled-coil region probably reflecting the different conformation that can be adopted depending on the signal. The hairpin connection topology in the four-helix bundle is different from those found in TM0853, DesK or KinB but similar to the situation in EnvZ, stressing again this surprising variability between different histidine kinases.

When using the structure from *T. maritima* (pdb ID: 3DGE) as model the catalytic domain comes close to the conserved histidine residue of the second protomer in the dimer, which suggests a *trans*-phosphorylation mechanism frequently observed in histidine kinases. A heterodimer of the full HK domain of ERS1 with a construct comprising only of the DHp domain was purified, which can serve as the starting point to elucidate the mode of phosphorylation.

ETR1 full cytoplasmic portion

As no crystals of the cytoplasmic domains of ETR1 could be obtained, a construct, excluding only the membrane-spanning part of ETR1, was analysed by SAXS.

The overall SAXS parameters determined from the scattering data suggest that the cytoplasmic domains of ETR1, construct H46, predominantly exist as dimers (calculated MM 130 kDa) in solution under the conditions used in this study. Yet the obtained value of MM_{SAXS} (Table 7) is low compared with the calculated value of a dimer. Accurate concentration estimation was problematic for this sample (discrepancy between two methods) in the presence of the solubilizing agent NDSB, impacting upon the normalisation of the SAXS intensities and as a result under or over estimating I(0) (and thus MM_{SAXS}). Therefore MM_{SAXS} was considered to be an unreliable parameter (being dependent on accurate sample concentration for normalisation of the scattering intensities). More accurate estimates of the MM were obtained from the parameters describing the hydrated particle volume, Vol_{SAXS} (directly obtained from the SAXS curve) and Vol_{DAM} (determined from the *ab initio* bead models constructed from the SAXS data)

(Table 7). The concentration independent estimates for Vol_{SAXS} and Vol_{DAM} are higher than expected for a monomeric H46, being more in line with a dimer (expected volume for dimeric H46, based upon a molecular weight of ~ 130 kDa, is 208 nm^3 for Vol_{SAXS} and 260 nm^3 for Vol_{DAM}). The hydrated particle volume Vol_{SAXS} (282 nm^3) is consistent with that expected for a dimeric H46, however, Vol_{DAM} (165 nm^3) is significantly lower, suggesting that some monomeric H46 is also present in an equilibrium with dimeric H46. As the solubilizing agent NDSB is likely to interact with the surface of H46 and also contribute to background scattering it is probable that it has had some effect on the estimation of the particle volume. However, taking this into account it remains clear that H46 can form dimers in solution.

To help resolve the ambiguity of the monomer – dimer question for H46 a simple modeling approach was conducted. Using the *ab initio* shape envelope reconstructed from the SAXS data it is possible to manually fit all the high-resolution structures of the domains in this volume as a dimer, supporting that all the domains are indeed present in the measured sample. Rigid body modelling using the high-resolution structures of the domains and a homology model of the GAF domain also provides strong evidence that the sample contains predominantly intact H46 dimers, due to the excellent fit of the resulting H46 models to the data at low angles (< 2-3 nm^{-1}). The quality of the fit deteriorates at higher angles due to the strong background signal from the 250 mM NDSB, however, the SAXS data in this range is essentially featureless and not required for low-resolution modelling of the dimer.

The overall dumbbell shape obtained by *ab initio* modelling correlates well with the constructed dimeric model (Figure 16B). The GAF domain is positioned at the top of a central helical bundle with the catalytic domain of the HK and the receiver domain at the other end. There seems to be some flexibility of the receiver domain position (Figure 17 C) but the position is distinct from the one described for a *T. maritima* receiver domain in complex with its HK (Casino et al., 2009) as expected by the length of the linker between the CA and the receiver domain. In the two obtained models the receiver domain is not in a position compatible with a phosphotransferase-competent state. Unfortunately the SAXS data of the cytoplasmic domains of ETR1 is of too low resolution to assign a functional state. In the two obtained models the active site of the receiver domain is pointing away from the conserved histidine residue in the dimerization domain. Positioning of the aspartate in the receiver domain for phosphoryl-transfer would require a rotation of the

receiver domain relative to the dimerization domain as well as a small translation. This would highly stretch the linker between the domains therefore a movement of the CA domain as observed in the structures of DesK or TM0853 would also be required.

ETR1 HK activity was demonstrated to have a role in growth recovery after removal of ethylene (Binder et al., 2004b) and in regulation of growth (Qu and Schaller, 2004). In addition genomic ETR1 could rescue a slow growth recovery phenotype, while a mutant replacing the conserved Asp^{659} in the receiver domain could not. This underlines the role of the receiver domain in this process and the contribution of a TCS-like signalling pathway in the ethylene response. Given the functional role of the receiver domain changes of its position are likely to be possible and to occur. The interface between the CA domain of the HKD and the receiver domain alone does not seem to be very strong as SEC of the isolated domains did not indicate complex formation and also SAXS analysis of a construct comprising only the catalytic and the receiver domain suggest an extended conformation (results of S. Panneerselvam). Hence, also as shown by the two different positions in the SAXS models of the receiver domain, it seems to occupy a number of positions in solution in the absence of other signals.

The here presented SAXS model gives the first impression of the overall composition of the cytoplasmic domains of an ethylene receptor in one state, yet leaving structural changes upon signals and protein interactions open.

CTR1 kinase domain

Since the first structural description of a kinase, the cyclic adenosine monophosphate-dependent protein kinase (Knighton et al., 1991), many more examples followed. The conserved motifs and numerous regulation mechanisms have been revealed and still every kinase offers a somewhat different picture regarding its regulation, cellular location or substrate recognition, which allows it to distinguish itself from the hundreds of other kinases found in higher eukaryotes reflecting their different tasks.

Structure comparison

The most conspicuous structural difference between active and inactive CTR1-kd is the conformation of the activation loop. This loop is stabilized through phosphoryl group

mediated interaction with the rest of the enzyme. In the absence of these phosphoryl groups the activation loop is disordered.

Superposition of the structures of active and inactive forms of CTR1-kd revealed further differences. Helix C of the N-terminal lobe assumes two distinct orientations (Figure 22A). A phosphorylated and structured activation loop causes the N-terminal part of helix C to tip towards the C-terminal lobe, moving by more than 7 Å. This movement conforms to a rotation of helix C about a pivot perpendicular to the helical axis and through its C-terminus. The amount of rotation corresponds to about 11° based on the alignment of the C-lobes (res. 632 to 810) of CTR1-kd and CTR1-D676N. The immediate consequence of this helix movement is a tighter conformation of the kinase. In CTR1-D676N, the open form of the kinase, the conserved salt bridge between Lys^{578} in β3 and Glu^{596} in helix C is disrupted and 6.1 Å apart in the better defined protomer A. Lys^{578} is fundamental in the productive orientation of the alpha and beta phosphates of ATP. Consequently, the active site of the open conformation of CTR1-kd would be expected to be incompetent to bind ATP productively, based on the relative position of Lys^{578} (Hanks & Hunter, 1995). In the closed form, however, Lys^{578} and Glu^{596} interact at distances of 2.8 Å and 3.0 Å in the two protomers in the AU. Additionally, the DFG motif of the active form assumes an Asp-out conformation with Asp^{694} forming a hydrogen bond with the backbone amide of Gly^{696}. In the inactive, open form the DFG motives in both protomers are very poorly defined and the glycine residues are disordered.

When comparing CTR1-kd with the active conformation of PKA (pdb code 1ATP) it is obvious that the amino acids needed for phosphorylation activity are correctly positioned. The Asp^{676} of the HRD motif interacts with the side chain of Asn^{681} (Figure 31) allowing it to orientate the attacking hydroxyl from the substrate. The P+1 loop is anchored to the hydrophobic core through interactions of Trp^{717} with Val^{739} and Ile^{740} in helix F, which are themselves part of a hydrophobic network, similar to the situation found in PKA (Yang et al., 2004). This anchor is supported by the interaction of Glu^{721} with Arg^{795}, which packs against Tyr^{735} in helix F. This creates a link to the catalytic loop through the hydrogen bond formed by Asp^{733} to the backbone amides of His^{674} and Arg^{675} as well as the hydrogen bonds of Glu^{743} in helix F with Trp^{717} and Ser^{679}.

Activation loop phosphorylation causes the catalytic loop segment to move by up to 2 Å and thereby enables several conserved interactions. The conserved Thr^{714} at the beginning of the P+1 loop interacts with Asp^{676} which has been suggested to help lock the aspartate in the conformation needed for the interaction with the hydroxyl of the substrate (Figure 31). Thr^{714} in CTR1-D676N, located right after the disordered region, moves about 3.5 Å away from Asp^{676}, thereby abandoning the interaction. Thr^{714} in CTR1-kd interacts also with Lys^{678} another highly conserved residue in the catalytic loop (Figure 31). The anchoring of the P+1 loop (with Thr^{714} at its beginning) is conserved in most protein kinases and seems to be important for the substrate recognition and activity (Nolen et al., 2004). The P+1 loop interacts with the substrate but is also directly or indirectly connected to the catalytic loop allowing it to sense the presence of substrate.

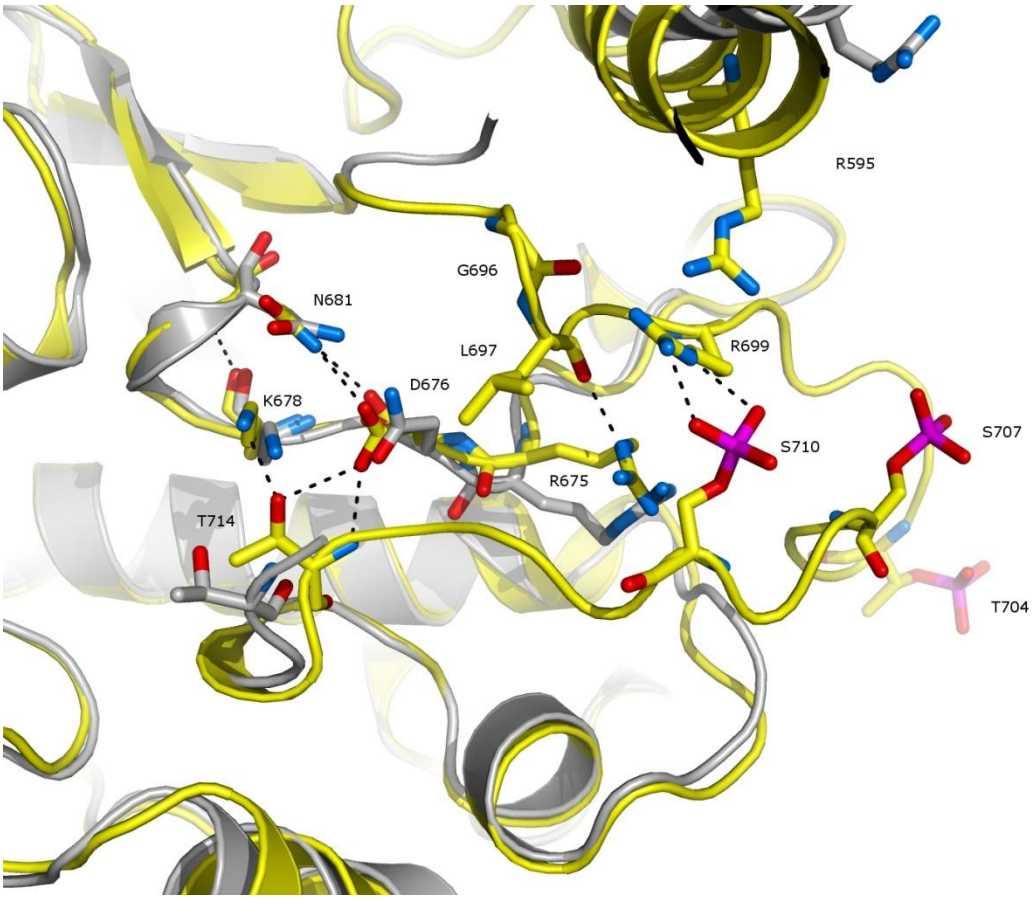


Figure 31: Overlay of the active site of CTR1-kd (yellow) and CTR1-D676N (grey). Selected residues are shown as sticks with the interactions show as black, dashed lines. Only the start and the end of the kinase dead activation loop are shown as it is mainly disordered. Parts of the structures were omitted for clarity.

The crystal structure of active CTR1-kd revealed three phosphorylation sites, all of them located in the activation loop. A triple alanine mutant Thr^{704}Ala, Ser^{707}Ala and Ser^{710}Ala demonstrated the prerequisite of activation loop phosphorylation for the observed CTR1 activity in the coupled assay (Figure 29) with the activity of this mutant being comparable to that of kinase dead CTR1-D676N. Still some phosphorylation was detected by MS showing the phosphorylation is not absolutely crucial for activity (discussed later) yet the biological role of these sites is unclear. Phosphorylation confers structure to the activation loop, which is disordered in inactive and un-phosphorylated CTR1-D676N. Structuring of the activation loop triggers several conformational rearrangements, which are typically found in the active forms of protein kinases. One example is the formation of a short but essential anti-parallel β-sheet between β6 (Pro^{671}/Ile^{672}/Val^{673}), which precedes the catalytic loop, and strand β9 (Arg^{699}/Leu^{700}/Lys^{701}) within the activation loop and. Formation of this β-sheet is initiated through the recruitment of Arg^{675} and Arg^{699} to the so-called basic pocket, which assembles around phosphoserine 710 (SEP^{710}). Arg^{675} is part of the catalytic loop HRD motif, which succeeds strand β6. When compared with the inactive CTR1-D676N the C_α of Arg^{675} moves 1.9 Å towards the activation loop. In this conformation, the side chain of Arg^{675} also forms a hydrogen bond with the backbone carbonyl of Leu^{697}, which is located directly after the loop initiating DFG motif, further stabilizing the activation loop. Recruitment of Arg^{699} triggers the formation of strand β9, which is disordered in the inactive mutant (Figure 31).

The salient feature of the phosphorylated activation loop consists of the electrostatic interface between SEP^{710} and the previously described basic pocket. This interaction likely precedes and therefore triggers the described structural rearrangements. Since neither of the other two activation loop phosphoresidues, phosphothreonine 704 and phosphoserine 707, directly participate in any interactions (cut-off 3.5 Å), SEP^{710} will be referred to as the primary phosphorylation site and CTR1 as an RD kinase as defined in a 2004 review (Nolen et al., 2004). Interaction between the primary phosphoryl group and the basic pocket in CTR1-kd structurally resembles the situation in the active forms of other RD kinase representatives from the AGC, CMGC and TK kinase families. An overlay of the C-terminal domains from showed that the respective primary phosphates are on average within 1.8 Å (σ = 0.9 Å), despite dissimilar lengths and conformations of their activation loops. In addition to the aforementioned canonical Arg^{675} and Arg^{699} a third basic residue,

Arg^{595} from helix C, comes to within 4 Å of SEP^{710} and augments the basic pocket in CTR1. Employment of Arg^{595} probably contributes to the movement and/or stabilization of helix C, which was discussed earlier. Interactions from basic residues located in the helix C have also been described for other Ser/Thr kinases (Krupa et al., 2004).

During the MS analysis a number of additional phosphorylation sites were found not observed in the structure. Yet similar observations have been reported for residues determined to be phosphorylated by MS, which were found unphosphorylated in the structure (Knowles et al., 2006), (McTigue et al., 1999). The most probable explanation for this observation is that only a sub-fraction of molecules being sufficiently abundant is forming the crystal.

Mode of staurosporine binding

Crystallization trials in the presence of AMPPNP resulted in no crystals. In a thermofluor assay staurosporine improved the thermal stability considerably while for all nucleotides tested smaller effects were observed. In both structures staurosporine is bound in the large groove between the lobes by hydrogen bonds (Glu^{626}, Leu^{628} and Pro^{680}) and CH–π interactions, with the plane structure of the staurosporine being sandwiched by hydrophobic residues (Figure 32). The binding of the ligand between the two structures is very similar. When compared with the Lck, Fyn and Lyn kinases in complex with staurosporine (pdb ID: 1QPD, 2DQ7, 3A4O), the position of the staurosporine and its interactions are highly conserved. Even though in both CTR1 structures magnesium was present in the crystallization buffer, no obvious density for an ion could be observed.

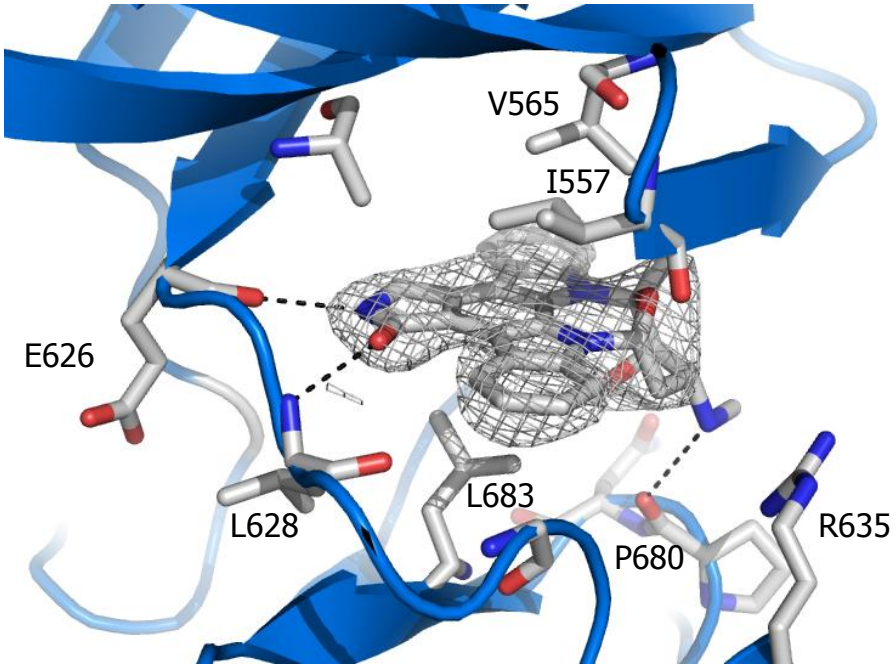


Figure 32: Staurosporine bound to CTR1-kd. The electron density for staurosporine is shown in grey at 1σ. Residues around the ligand are shown in sticks. Polar interactions are indicated as black dashed lines

Activity and Dimerization of the CTR1 kinase domain

Despite the extensive research targeted at B-RAF the presence and importance of dimerization by the enzyme was not recognized until recently (Rajakulendran et al., 2009). When the side-to-side dimer interface was disturbed by specific mutations most of the substrate phosphorylation was lost. Interestingly CTR1 crystallized in the active and inactive state in a very similar dimeric arrangement as exhibited by B-RAF (Figure 25).

The interface residues found to be crucial in B-RAF dimerization are also present in the RAF-related pseudo-kinase KSR, which can interact with and activate B-RAF. These residues are conserved in CTR1 (Figure 23). After the initial verification of CTR1-kd being a dimer in solution, confirming the crystal structure, the role of these conserved residues was investigated. All of them play a role in formation of the *in vitro* dimer in solution as already single mutations at these three positions (Arg^{604}, Phe^{611} and Met^{612}) lead to the formation of predominantly monomers. Yet mutations of the positions $Arg^{604}Ala$, $Arg^{602}Asp$ & $Arg^{604}Ala$ and $FM^{611}AW$ could not completely abolish the observed small dimer fraction. This is probably due to equilibrium between the states and despite disturbing the interface some interaction is still possible. Probably other residues involved provide a still strong enough interface allowing for transient binding. The activity is partially affected by these

mutations with some mutations (Met^{612}Trp, FM^{611}AW) causing the loss of the continuous activity observed in the wild type. Also in B-Raf kinase the mutants had varying effects on the activity (Rajakulendran et al., 2009). All of the CTR1 mutants also retained some of the autophosphorylation activity. Some off the residues found to be involved in the interface (Arg^{602}, Arg^{604}) are part of a basic patch and a similar motif was shown in Raf-1 to be important for binding to the inhibitor phosphatidic acid (PA) (Ghosh et al., 2003). While the same residues don't play a role in PA binding in a mutant a reduction of the kinase activity was also observed (Testerink et al., 2007).

Remarkably in the crystal structure CTR1-D676N is forming the same dimer interface as does the wild-type, yet in solution it was present as a monomer. No obvious structural changes of the interacting residues can be observed when aligning the interfaces of CTR1-kd and CTR1-D676N. Also the mutated residue Asn^{676} and the biggest conformational changes found in N-terminus of helix C and the activation loop are not part of the interface, making it difficult to determine the exact reason for the monomeric form. The high protein concentration in the crystal could allow CTR1-D676N to overcome the weak affinity preventing dimerization in solution and to form a dimer, similar to the wild-type.

When analysing the effects of mutants of the activation loop a mix of dimers and monomers (Figure 26 F and G) was observed highlighting the contribution of activation loop phosphorylation to the dimerization. Phosphorylation is compacting the structure through the rotation of N-lobe, which pivots by 7.4° towards its C-lobe in CTR1-kd, possibly stabilizing the dimer through subtle changes. This also allowed assigning a role to phosphorylated Thr^{704} and Ser^{707}, which in the structure seem to have no role in stabilization compared to Ser^{710}, but are needed to shift the equilibrium to the dimer in solution. A double mutant of Thr^{704} and Ser^{707} to alanine already causes the shift to a dimer/monomer mixture when compared with the wild type. Addition of the Ser^{710}Ala mutant does not change this shift or enhance it. This is in line with preliminary results of a low resolution structure (3.25 Å) of the Thr^{704} and Ser^{707} mutant being more similar to the CTR1-D676N structure than the CTR1-kd one. Therefore even in the absence of a clear role in activation loop stabilization they contribute to the structural integrity and are needed for dimerization.

Mutants of the three observed phosphorylation sites in the activation loop are still phosphorylated to some extent outside the activation loop. This must be due to autophosphorylation as CTR1-D676N was found to be completely unphosphorylated. This

was relatively unexpected, as phosphorylation of the activation loop is often triggering activity in RD kinases. Even the primary phosphorylation site Ser^{710}, which is highly conserved in position when comparing with other kinases where it often plays a pivotal role during activation of the kinase, is not crucial for auto-phosphorylation activity *in vitro*. Moreover CTR1 exhibits a continuous consumption of NADH in the coupled assay in the absence of a substrate. The simple explanation is it being a rapid hydrolysis of one residue, which gets subsequently again phosphorylated therefore leading to an unproductive circle. Yet this explanation clashes to some extend with Ser/Thr phosphorylation normally known to be relatively stable. Mutations in the activation loop and some affecting the dimerization can stop this phenomenon making it unfortunately difficult to assign specific crucial residues.

Also in other Ser/Thr kinases a varying number of phosphorylation sites in the same batch were reported (Young et al., 2003). Yet as in CTR1 only phosphorylation of the activation loop was consistently observed and connected with a clear role. In CTR1 the remaining sites are scattered throughout the molecule and only observed in fractions of the proteins and are possibly an artefact of the high concentration of the protein in the cell and the expression in the absence of the N-terminal domain.

In a recent structure of the *M. tuberculosis* Ser/Thr Kinase PknB bound to an ATP-competitive inhibitor an asymmetric dimer was observed interacting via the helix G (Mieczkowski et al., 2008). One molecule possesses an ordered activation loop (suggested as the active conformation), while the other has a disordered activation loop (inactive conformation) with the ends pointing to the active site of the active molecule. This was suggested to be the activation complex with a front-to-front association allowing for the *trans*-phosphorylation and the activation of the kinase (Mieczkowski et al., 2008). Mutations in the helix G were sufficient to abolish kinase activation. In both CTR1 structures a similar interaction via the helix G of symmetry related molecules can be observed with similar interface sizes. Yet the interaction is different as chain B of the CTR1-D676N front-to-front interface is rotated about 30° and translated slightly relative to chain B of the CTR1-kd front-to-front interface when superimposing the C-lobes of chain A of the two structures. When using this superimposition a hybrid front-to-front model can be created, consisting of the active chain A (CTR1-kd) and the inactive chain B (CTR1-D676N), interacting through helix G. There the C-terminal, structured end of the activation loop on CTR1-D676N points to the active site of the active wild-type structure similar to

the situation in the PknB structure. The first structured residue in the inactive structure Gly^{713} is about 13 Å away from the active site in the wild-type molecule. Therefore in this conformation Ser^{707} and the primary phosphorylation site Ser^{710} of the inactive molecule could be phosphorylated by an active kinase in a different dimer (Figure 33) in a similar fashion as in PknB, leading to its activation. Due to steric restraints phosphorylation of Thr^{704} would probably require a different conformation. This would allow for a cross-receptor signal transmission, allowing one active dimer to phosphorylate a number of neighbouring, inactive dimers (see next section).

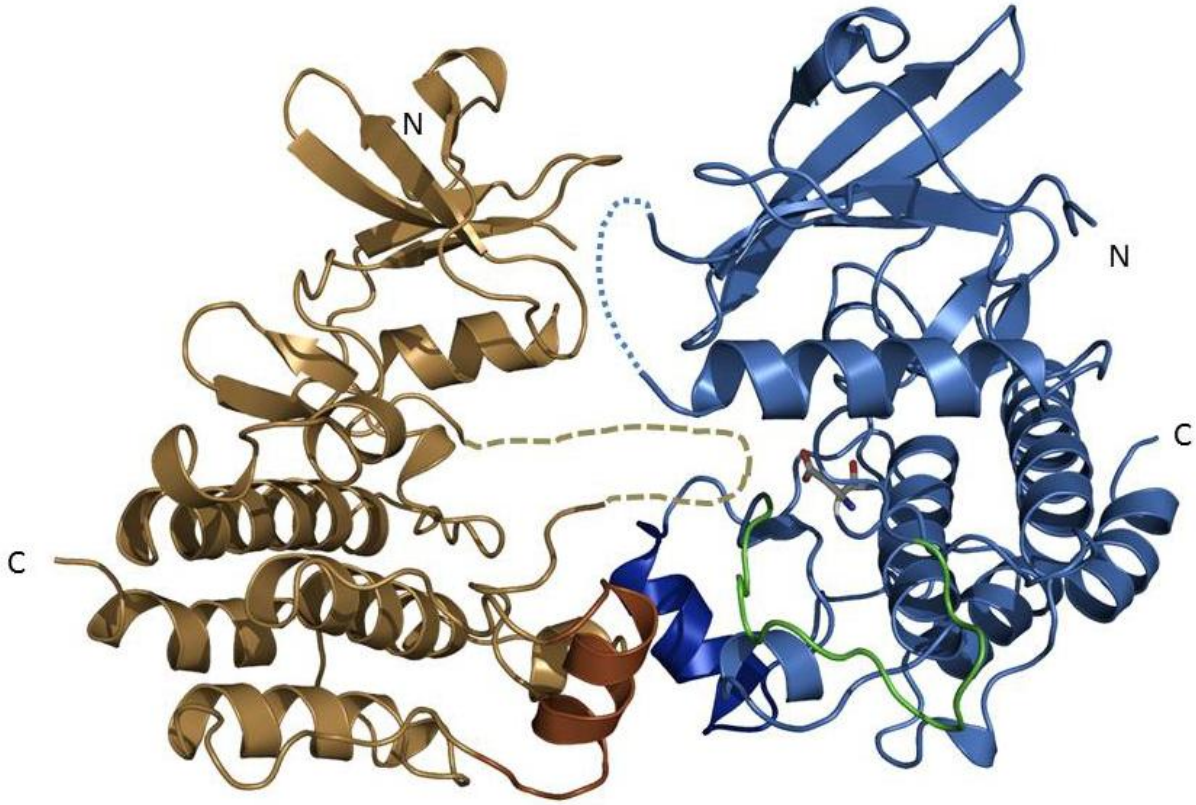


Figure 33: Hybrid model of an inactive CTR1-D676N molecule shown in sand interacting via helix G with an active CTR1-kd protomer shown in blue. The disordered activation loop (dashed brown line) is positioned at the entrance of the active site of the CTR1-kd molecule with Asp^{676} shown in sticks. The ordered activation loop is shown in green; both helixes G are shown in brown and dark blue, respectively.

In general all mutated interface and activation loop residues and the residues in helix G are highly conserved when comparing with other CTR1 molecules from other species (Figure 23). This suggests that the results obtained for CTR1 from *A. thaliana* are very likely to be also valid for other plants and mutants will have similar effects the dimerization

and phosphorylation patterns. The residues involved in dimerization of the kinase domain are also conserved in human and mouse B-RAF and KSR1 underlining the importance of the dimerization for the function, allowing for the evolutionary conservation. In contrast the activation loop is, not unexpectedly, quite diverse when comparing human and plant kinases, reflecting their different cellular roles.

The loss of dimerization shifts the kinase to some extend to the inactive state showing reduced activity. Furthermore the activity is not only depending on activation loop phosphorylation *in vitro* and the presence of the substrate ATP does not affect dimerization. CTR1 is initially present as an unphosphorylated monomer, yet able to autophosphorylate to some extent supporting the formation of a stable dimer. This leads to a gradual shift of the kinase from the inactive to the active state with all parts contributing but none being the only decisive element making the on/off decision.

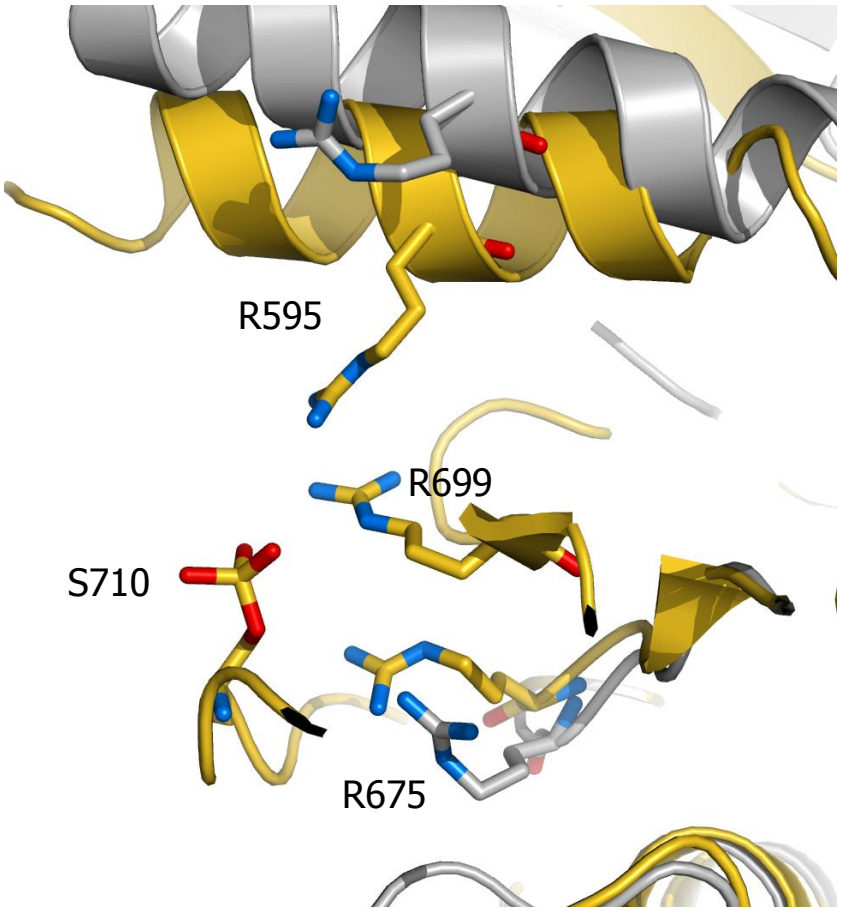


Figure 34: Superimposition of the basic pocket of CTR1-wt (yellow) and CTR1-D676N (grey) with the phosphorylated serine and the three arginines shown as sticks. Ser^{710} and Arg^{699} are disordered in CTR1-D676N

Dimerization has an influence on the activity of CTR1 which is, given the results from the AUC and the highly conserved interface residue, very likely to act as a dimer *in vivo*. A

recently identified mutant of CTR1 was found to affect planar polarity of roothair placement in *A. thaliana,* exhibiting a $Glu^{626}Lys$ amino acid exchange located at the loop connecting the N-lobe with the C-lobe. The highly conserved side chain is pointing towards the dimer interface and the substitution with lysine could lead to repulsion and subsequent destruction of the dimer. A reduction of the *in vitro* kinase activity was observed and an increase of auxin biosynthesis (Ikeda et al., 2009). This gives some indication of the physiological role of dimerization, with the mutant only partially affecting the activity, while still being able to block the cellular role. Unfortunately this mutant could not be purified despite testing different tags and media, preventing the assessment of the mutation on the dimerization.

All this supports CTR1 as a very active kinase which needs to be strongly inhibited in order to allow the ethylene signal to proceed to the nucleus. As CTR1 autophosphorylates in the absence of activating molecules, one can speculate that CTR1 is structurally flexible in the absence of phosphorylation, with a fraction being in a phosphorylation-competent state, being able to interact with other kinases allowing for their phosphorylation.

Implications for signal transduction

A consequence of the specific dimer interface of CTR1 is a positioning of the expected substrate binding sites (close to the P+1 loop and Asp^{676}) on opposite sides of the kinase dimer being over 40 Å apart (Figure 24 A). This excludes the possibility of an intra-dimer, *trans*-autophosphorylation mechanism. Also the large distance of more than 24 Å between the hydroxyl side chain of phosphorylated Thr^{704} in the activation loop and the catalytic Asp^{676} in the same molecule makes it difficult to imagine a *cis*-phosphorylation mechanism. In addition *cis*-phosphorylation in Ser/Thr kinases is a rare event, which needs special biochemical properties and often an interaction with another protein in order to overcome the steric problems, keeping it from occurring more frequently (Lochhead, 2009). *Trans*-phosphorylation across CTR1 dimers is further supported by the presence of an activation interface via helix G, which closely resembles the complex recently described for PknB (Mieczkowski et al., 2008) and the orientation of the disordered activation loop (Figure 33). Modelling of a complex of CTR1-kd by substituting an active monomer within the observed packing via helix G of CTR1-kd with inactive CTR1-D676N, positions the open ends of the disordered activation loop pointing into the active site of active CTR1-kd. The distance between the Cα-atom of the first

visible residue after the strand break, G713, and the side chain of catalytic D676 is 12Å, which corresponds to three amino acid residues in an extended polypeptide chain. This would position the active site catalytic base in immediate contact with the hydroxyl group of the primary phosphorylation site of CTR1 (Ser^{710}), as required for deprotonation.

The ethylene receptors and CTR1 are active in the absence of ethylene. In addition CTR1 was now found to be a very active kinase and, when expressed without its N-terminal part, being able to self-activate in the absence of other molecules. Furthermore a number of mutants targeting the activation loop and the dimerization interface could not completely abolish the activity when assessed by MS. This suggests a very active enzyme effectively blocking the pathway in the absence of ethylene. The exact mode of CTR1 deactivation is still unknown. Phosphatidic acid was shown to bind to the kinase domain and deactivate CTR1 as well as the interaction with ETR1 and the N-terminal domain. Various biotic and abiotic stresses, also associated with ethylene signalling, are known to lead to the PA synthesis and PA was suggested to be at the start of ethylene signalling acting when ethylene is still absent (Testerink et al., 2008). Also a change in the interaction with the ethylene receptors caused by structural rearrangements resulting from the ethylene binding is a probable mechanism affecting the kinase activity, the phosphorylation state, the dimerization or interactions with the substrate(s) of CTR1. Regardless of the exact way of regulation it appears that a strong deactivation signal is needed in order to overcome the robust activity of CTR1.

The ethylene receptors are known to form dimers and also associations with other receptor dimers were shown (Gao et al., 2008). In addition the receptors can respond to ethylene at a concentration 300-fold lower than the K_d of the receptors for ethylene. This suggests that the ethylene receptors form clusters similar to the situation found in bacteria (Gestwicki et al., 2000), leading to a local increase of the signalling molecules being able to signal in a cooperative manner. With the here presented evidence for dimerization in CTR1 *in vitro*, similar to the RAF kinases, and the presented model of cross-dimer phosphorylation this adds another level of interaction between molecules. Two possible modes of binding to the receptors can be envisioned: A dimer of CTR1 molecules binds to a homodimer of ethylene receptors (Figure 35) or a dimer of CTR1 molecules binds to two different ethylene receptor dimers. While the first model would support the homodimer stability the second could promote the predicted cluster formation through the linkage of

receptor homodimers. In addition both models allow one active CTR1 molecule to phosphorylate other surrounding molecules (Figure 33) and possibly also other substrates. CTR1 mediated inter-receptor communication may therefore rapidly spread across receptors clusters and multiply the output signal from a single ethylene molecule. The combination of dimer and activation interfaces, as revealed by the crystal packing of CTR1-kd, results in a continuous head-to-tail oligomer of kinase dimers (Figure 35).

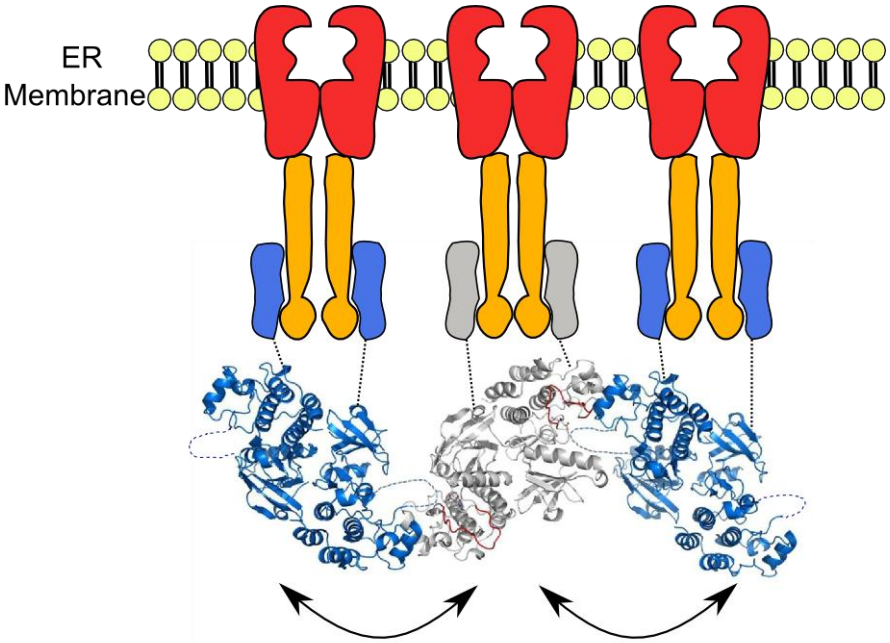


Figure 35: A model of cross-receptor signalling. An active CTR1 dimer (shown in grey) binds to one ethylene receptor dimer and is able to phosphorylate other inactive CTR1 molecules (blue) and allow for signal transduction between different receptors being in different activation states.

CTR1 plays a central role helping to integrate a number of different inputs leading to the choice of the appropriate response. The role of dimerization and phosphorylation with the connected autophosphorylation activity represent one aspect of the CTR1 regulation. Some other factors, as phosphatidic acid, which seems to play a role at the beginning of the ethylene response, the phosphatase PP2A, which was shown to be important for the activity of CTR1 or the $Glu^{626}Lys$ mutant, which was found to influence the auxin biosynthesis, highlight the complexity of the signalling network surrounding CTR1. Therefore further testing, especially *in vivo*, of the here presented insights into the structure, biochemistry and oligomeric state of CTR1 will be needed to gain an in depth understanding of this part of the ethylene signalling network and to be able to understand the wealth of processes regulated.

Acknowledgements

I am thankful to my supervisor, Jochen Müller-Dieckmann, who initiated the thesis project on the ethylene signalling pathway. Without his continuous encouragement, guidance and support my thesis would not have been possible. I especially enjoyed our numerous discussions, patient explanations regarding crystallographic issues and his open-door policy allowing me to discuss any results or problems very quickly.

I want to thank Francesco Fersini for always being a good friend, helpful discussions and advice in the lab and for all the nice days and nights we spent together. Furthermore I want to thank Spyros Chatziefthimiou for numerous help with a lot of issues and questions, always being available for a discussion, all the nice time we spent together, his very generous way and in general for being Spyros. I want to thank Georgios Hatzolopoulos for helping me a lot in order to understand the software needed to process and analyse the data, for help at the beamlines and getting me started in scripting. Most of all I want to thank him for being a good friend who it was always a pleasure to meet and spent the evening with. In addition I would like to thank Xandra Kreplin for help with the setup of a lot of crystallization plates, accepting my last minute bookings and changes and helping me with non-standard setups and making any ordering very straightforward for me. I want to thank Matthew Groves for help and advice in the lab and at the beamline, always being open for a discussion. I want to thank Jacopo Negroni, Marco Salomone Stagni, Matthew Dunne, Georg Eulenburg and Jon Rapley for the nice time we had together. Thanks also to Haydyn Mertens for help with the SAXS measurements, data analysis and interpretation and suggestions. Furthermore I also want to acknowledge Saravanan Panneerselvam for help with the CTR1 project and advice regarding kinases in general. I want to thank Santosh Panjikar, Paul Tucker and Manfred Weiss for discussions and help. I thank as well all the people of the wetlab, present and past members for making the time there enjoyable, giving advice on various issues and discussions regarding scientific and less scientific issues, making it a fun place to work. Special thanks go to the indoor football team at EMBL Hamburg for making the game always a highlight of the week. I would also like to thank the members of my TAC committee Paul Tucker, Darren Hart, Irmgard Sinning and Victor Sourjik for their suggestions during the meetings. I would like to thank Dr. Vladimir Rubin at the EMBL core facility for the execution of the

analytical ultracentrifugation experiments and the proteomics core facility at the EMBL Heidelberg for assistance with the mass spectrometry analysis. I thank the ESRF for provision of beamtime at beamlines ID23-1, ID23-2 and ID29.

I want to thank my flatmate Martin Krings for sharing with me for more than three nice years, making it always a pleasure being in the flat. Furthermore I have to thank far too many people to mention them all at the DAV climbing centre, helping me to nearly forget the flat countryside in Hamburg, for nice days climbing together, a number of outdoor climbing trips and pleasant evenings with special thanks to the Pink Tank.

Finally I want to thank my parents, Ernst and Christiane, my brother Christoph and sister Uli for supporting me, coming numerous times to Hamburg and making me always enjoy my trips back to Austria. Also for all the trips we did together throughout the last years, being always there for me – I'm just happy to have you all. I will also always keep the great memories of my grandmother Friederike Mayerhofer for being a remarkable person and it was always fun to have her around.

Finally I dedicate this thesis to my mother Christiane Mayerhofer who passed away 6 month before the end of my time in Hamburg for all the love and support I got from her over the last 27 years and for encouraging me all the time. I would like to thank you for all the small things, all the care, all the things you taught me and your seemingly endless energy, having always the best for the family in mind, taking care of so many things at school and at home and still at all times having time for others.

I always will miss you.

Abbreviations

ACC	1-aminocyclopropane-1-carboxylic acid
ARR2	Arabidopsis response regulator 2
AU	Asymmetric unit
AUC	Analytical ultra-centrifugation
β-ME	β-Mercaptoethanol
CTR1	Constitutive triple response 1
CA	Catalytic and ATP-binding domain
CTR1-kd	CTR1 kinase domain (wild-type construct)
CTR1-D676N	CTR1 kinase dead mutant
CV	Column volume
DHp	Dimerization and histidine phosphotransfer domain
EIL	EIN3-like
EIN2	Ethylene insensitive2
EIN4	Ethylene insensitive 4
ER	Endoplasmic reticulum
EREBP	Ethylene response element binding protein
ERF1	Ethylene response factor 1
ERS	Ethylene response sensor
ETR	Ethylene receptor
GMO	Genetically modified organism
HK	Histidine kinase
HKD	Histidine kinase domain
IC3	5-amino-2,4,6-triiodoisophthalic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
(k)Da	(kilo)Dalton
LOF	Loss of function
MAPK	Mitogen activated protein kinase
MAPKK	MAPK kinase
MAPKKK	MAPK kinase kinase
MR	Molecular replacement
MS	Mass spectrometry
NDSB	Non-detergent sulfo betaine
NiNTA	Nickel-nitriloacetic acid
PA	Phosphatidic acid
RAN1	Responsive to antagonist 1
R.m.s.	Root-mean-square
SCF	Skp, Cullin, F-box containing complex
SEC	Size exclusion chromatography
SER	Surface entropy reduction
SOC	Super Optimal Broth

Bibliography

Abeles, F. B., Morgan, P. W., and Saltveit, M. E. J. (1992). Ethylene in Plant Biology 2nd ed. (San Diego, CA: Academic Press).

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr *66*, 213-221.

Albanesi, D., Martín, M., Trajtenberg, F., Mansilla, M. C., Haouz, A., Alzari, P. M., de Mendoza, D., and Buschiazzo, A. (2009). Structural plasticity and catalysis regulation of a thermosensor histidine kinase. Proceedings of the National Academy of Sciences *106*, 16185 - 16190.

Alonso, J. M., Stepanova, A. N., Solano, R., Wisman, E., Ferrari, S., Ausubel, F. M., and Ecker, J. R. (2003). Five components of the ethylene-response pathway identified in a screen for weak ethylene-insensitive mutants in Arabidopsis. Proc. Natl. Acad. Sci. U.S.A *100*, 2992-2997.

Barrett, C. P., and Noble, M. E. M. (2005). Molecular Motions of Human Cyclin-dependent Kinase 2. Journal of Biological Chemistry *280*, 13993 -14005.

Beck, T., Krasauskas, A., Gruene, T., and Sheldrick, G. M. (2008). A magic triangle for experimental phasing of macromolecules. Acta Crystallogr D Biol Crystallogr *64*, 1179-1182.

Beyer, E. M. (1976). A Potent Inhibitor of Ethylene Action in Plants. Plant Physiol. *58*, 268-271.

Bhattacharyya, R. P., Reményi, A., Yeh, B. J., and Lim, W. A. (2006). Domains, Motifs, and Scaffolds: The Role of Modular Interactions in the Evolution and Wiring of Cell Signaling Circuits. Annu. Rev. Biochem. *75*, 655-680.

Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. (1999). Structure of CheA, a signal-transducing histidine kinase. Cell *96*, 131-141.

Binder, B. M., Mortimore, L. A., Stepanova, A. N., Ecker, J. R., and Bleecker, A. B. (2004a). Short-term growth responses to ethylene in Arabidopsis seedlings are EIN3/EIL1 independent. Plant Physiol *136*, 2921-2927.

Binder, B. M., O'Malley, R. C., Wang, W., Moore, J. M., Parks, B. M., Spalding, E. P., and Bleecker, A. B. (2004b). Arabidopsis Seedling Growth Response and Recovery to Ethylene. A Kinetic Analysis. Plant Physiol. *136*, 2913-2920.

Binder, B. M., Walker, J. M., Gagne, J. M., Emborg, T. J., Hemmann, G., Bleecker, A. B., and Vierstra, R. D. (2007). The Arabidopsis EIN3 Binding F-Box Proteins EBF1 and EBF2 Have Distinct but Overlapping Roles in Ethylene Signaling. The Plant Cell Online *19*, 509 -523.

Bisson, M. M. A., and Groth, G. (2010). New insight in ethylene signaling: autokinase activity of ETR1 modulates the interaction of receptors and EIN2. Mol Plant *3*, 882-889.

Bisson, M. M. A., Bleckmann, A., Allekotte, S., and Groth, G. (2009). EIN2, the central regulator of ethylene signalling, is localized at the ER membrane where it interacts with the ethylene receptor ETR1. Biochem. J *424*, 1-6.

De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O., and Kim, S. H. (1993). Crystal structure of cyclin-dependent kinase 2. Nature *363*, 595-602.

Bourret, R. B. (2010). Receiver domain structure and function in response regulator proteins. Curr. Opin. Microbiol *13*, 142-149.

Burg, S. P. (1962). The Physiology of Ethylene Formation. Annual Review of Plant Physiology *13*, 265-302.

Cancel, J. D., and Larsen, P. B. (2002). Loss-of-Function Mutations in the Ethylene Receptor ETR1 Cause Enhanced Sensitivity and Exaggerated Response to Ethylene in Arabidopsis. Plant Physiol. *129*, 1557-1567.

Casino, P., Rubio, V., and Marina, A. (2009). Structural insight into partner specificity and

phosphoryl transfer in two-component signal transduction. Cell *139*, 325-336.

Casino, P., Rubio, V., and Marina, A. (2010). The mechanism of signal transduction by two-component systems. Current Opinion in Structural Biology *20*, 763-771.

Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W., and Ecker, J. R. (1997). Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. Cell *89*, 1133-1144.

Chen, Q. G., and Bleecker, A. B. (1995). Analysis of ethylene signal-transduction kinetics associated with seedling-growth response and chitinase induction in wild-type and mutant arabidopsis. Plant Physiol *108*, 597-607.

Chen, Y.-F., Gao, Z., Kerris, R. J., Wang, W., Binder, B. M., and Schaller, G. E. (2010). Ethylene receptors function as components of high-molecular-mass protein complexes in Arabidopsis. PLoS ONE *5*, e8640.

Chen, Y.-F., Randlett, M. D., Findell, J. L., and Schaller, G. E. (2002). Localization of the ethylene receptor ETR1 to the endoplasmic reticulum of Arabidopsis. J. Biol. Chem *277*, 19861-19866.

Chou, K. C., Maggiora, G. M., Némethy, G., and Scheraga, H. A. (1988). Energetics of the structure of the four-alpha-helix bundle in proteins. Proceedings of the National Academy of Sciences of the United States of America *85*, 4295 -4299.

Clark, K. L., Larsen, P. B., Wang, X., and Chang, C. (1998). Association of the Arabidopsis CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. Proceedings of the National Academy of Sciences of the United States of America *95*, 5401-5406.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Cook, P. F., Neville, M. E., Vrana, K. E., Hartl, F. T., and Roskoski, R. (1982). Adenosine cyclic 3',5'-monophosphate dependent protein kinase: kinetic mechanism for the bovine skeletal muscle catalytic subunit. Biochemistry *21*, 5794-5799.

Crick, F. H. C. (1953). The packing of α-helices: simple coiled-coils. Acta Cryst *6*, 689-697.

Crocker, W., Hitchcock, A., and Zimmerman, P. (1935). Similarities in the effects of ethlyene and the plant auxins. Contrib. Boyce Thompson Inst. *7*, 231-48.

Dutta, R., and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. Trends in Biochemical Sciences *25*, 24-28.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

van den Ent, F., and Löwe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J. Biochem. Biophys. Methods *67*, 67-74.

Falke, J. J., and Hazelbauer, G. L. (2001). Transmembrane signaling in bacterial chemoreceptors. Trends Biochem. Sci *26*, 257-265.

Franke, D., and Svergun, D. I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J Appl Crystallogr *42*, 342-346.

Gane, R. (1934). Production of Ethylene by Some Ripening Fruits. Nature *134*, 1008.

Gao, Z., Chen, Y.-F., Randlett, M. D., Zhao, X.-C., Findell, J. L., Kieber, J. J., and Schaller, G. E. (2003). Localization of the Raf-like kinase CTR1 to the endoplasmic reticulum of Arabidopsis through participation in ethylene receptor signaling complexes. J. Biol. Chem *278*, 34725-34732.

Gao, Z., Wen, C.-K., Binder, B. M., Chen, Y.-F., Chang, J., Chiang, Y.-H., Kerris, R. J., Chang, C., and Schaller, G. E. (2008). Heteromeric interactions among ethylene receptors mediate signaling in Arabidopsis. J. Biol. Chem *283*, 23801-23810.

Gestwicki, J. E., Lamanna, A. C., Harshey, R. M., McCarter, L. L., Kiessling, L. L., and Adler, J. (2000). Evolutionary Conservation of Methyl-Accepting Chemotaxis Protein Location in Bacteria and Archaea. J Bacteriol *182*, 6499-6502.

Ghosh, S., Moore, S., Bell, R. M., and Dush, M. (2003). Functional Analysis of a

Phosphatidic Acid Binding Domain in Human Raf-1 Kinase. Journal of Biological Chemistry *278*, 45690 -45696.

Goldschmidt, L., Cooper, D. R., Derewenda, Z. S., and Eisenberg, D. (2007). Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. Protein Sci. *16*, 1569-1576.

Guinier, A. (1939). La diffraction des rayons X aux tres petits angles; applicationa l'etude de phenomenes ultramicroscopiques. Ann. Phys. Paris *12*, 161-237.

Guo, H., and Ecker, J. R. (2003). Plant Responses to Ethylene Gas Are Mediated by SCFEBF1/EBF2-Dependent Proteolysis of EIN3 Transcription Factor. Cell *115*, 667-677.

Hall, A. E., and Bleecker, A. B. (2003). Analysis of combinatorial loss-of-function mutants in the Arabidopsis ethylene receptors reveals that the ers1 etr1 double mutant has severe developmental defects that are EIN2 dependent. Plant Cell *15*, 2032-2041.

Hass, C., Lohrmann, J., Albrecht, V., Sweere, U., Hummel, F., Yoo, S. D., Hwang, I., Zhu, T., Schäfer, E., Kudla, J., et al. (2004). The response regulator 2 mediates ethylene signalling and hormone signal integration in Arabidopsis. EMBO J *23*, 3290-3302.

Hernández Sebastià, C., Hardin, S. C., Clouse, S. D., Kieber, J. J., and Huber, S. C. (2004). Identification of a new motif for CDPK phosphorylation in vitro that suggests ACC synthase may be a CDPK substrate. Arch. Biochem. Biophys *428*, 81-91.

Hsing, W., Russo, F. D., Bernd, K. K., and Silhavy, T. J. (1998). Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ. J. Bacteriol *180*, 4538-4546.

Hua, J., Sakai, H., Nourizadeh, S., Chen, Q. G., Bleecker, A. B., Ecker, J. R., and Meyerowitz, E. M. (1998). EIN4 and ERS2 Are Members of the Putative Ethylene Receptor Gene Family in Arabidopsis. Plant Cell *10*, 1321-1332.

Huang, Y., Li, H., Hutchison, C. E., Laskey, J., and Kieber, J. J. (2003). Biochemical and functional analysis of CTR1, a protein kinase that negatively regulates ethylene signaling in Arabidopsis. Plant J *33*, 221-233.

Hubbard, S. R., Wei, L., and Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature *372*, 746-754.

Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. Cell *109*, 275-282.

Hwang, I., and Sheen, J. (2001). Two-component circuitry in Arabidopsis cytokinin signal transduction. Nature *413*, 383-389.

Ikeda, Y., Men, S., Fischer, U., Stepanova, A. N., Alonso, J. M., Ljung, K., and Grebe, M. (2009). Local auxin biosynthesis modulates gradient-directed planar polarity in Arabidopsis. Nat Cell Biol *11*, 731-738.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature *376*, 313-320.

Kabsch, W. (2010). XDS. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kevany, B. M., Tieman, D. M., Taylor, M. G., Cin, V. D., and Klee, H. J. (2007). Ethylene receptor degradation controls the timing of ripening in tomato fruit. Plant J *51*, 458-467.

Knighton, D., Zheng, J., Ten Eyck, L., Ashford, V., Xuong, N., Taylor, S., and Sowadski, J. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science *253*, 407 -414.

Knowles, P. P., Murray-Rust, J., Kjær, S., Scott, R. P., Hanrahan, S., Santoro, M., Ibáñez, C. F., and McDonald, N. Q. (2006). Structure and Chemical Inhibition of the RET Tyrosine Kinase Domain. Journal of Biological Chemistry *281*, 33577 -33587.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003). PRIMUS : a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr *36*, 1277-1282.

Konishi, M., and Yanagisawa, S. (2008). Ethylene signaling in Arabidopsis involves feedback regulation via the elaborate control of EBF2 expression by EIN3. Plant J *55*, 821-831.

Kornev, A. P., Haste, N. M., Taylor, S. S., and Ten Eyck, L. F. (2006). Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. Proc Natl Acad Sci U S A *103*, 17783-17788.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. Journal of Molecular Biology *372*, 774-797.

Krupa, A., Preethi, G., and Srinivasan, N. (2004). Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. J. Mol. Biol *339*, 1025-1039.

Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc *3*, 1171-1179.

Larsen, P. B., and Cancel, J. D. (2003). Enhanced ethylene responsiveness in the Arabidopsis eer1 mutant results from a loss-of-function mutation in the protein phosphatase 2A A regulatory subunit, RCN1. Plant J *34*, 709-718.

Levdikov, V. M., Blagova, E., Joseph, P., Sonenshein, A. L., and Wilkinson, A. J. (2006). The Structure of CodY, a GTP- and Isoleucine-responsive Regulator of Stationary Phase and Virulence in Gram-positive Bacteria. Journal of Biological Chemistry *281*, 11366 -11373.

Liu, Y., and Zhang, S. (2004). Phosphorylation of 1-Aminocyclopropane-1-Carboxylic Acid Synthase by MPK6, a Stress-Responsive Mitogen-Activated Protein Kinase, Induces Ethylene Biosynthesis in Arabidopsis. Plant Cell *16*, 3386-3399.

Lochhead, P. A. (2009). Protein kinase activation loop autophosphorylation in cis: overcoming a Catch-22 situation. Sci Signal *2*, pe4.

Lombana, T. N., Echols, N., Good, M. C., Thomsen, N. D., Ng, H.-L., Greenstein, A. E., Falick, A. M., King, D. S., and Alber, T. (2010). Allosteric activation mechanism of the Mycobacterium tuberculosis receptor Ser/Thr protein kinase, PknB. Structure *18*, 1667-1677.

Marina, A., Mott, C., Auyzenberg, A., Hendrickson, W. A., and Waldburger, C. D. (2001). Structural and Mutational Analysis of the PhoQ Histidine Kinase Catalytic Domain. Journal of Biological Chemistry *276*, 41182 -41190.

Marina, A., Waldburger, C. D., and Hendrickson, W. A. (2005). Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. EMBO J *24*, 4247-4259.

Mayerhofer, H., Mueller-Dieckmann, C., and Mueller-Dieckmann, J. (2011). Cloning, expression, purification and preliminary X-ray analysis of the protein kinase domain of constitutive triple response 1 (CTR1) from Arabidopsis thaliana. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun *67*, 117-120.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

McTigue, M. A., Wickersham, J. A., Pinko, C., Showalter, R. E., Parast, C. V., Tempczyk-Russell, A., Gehring, M. R., Mroczkowski, B., Kan, C.-C., Villafranca, J. E., et al. (1999). Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: a key enzyme in angiogenesis. Structure *7*, 319-330.

Mieczkowski, C., Iavarone, A. T., and Alber, T. (2008). Auto-activation mechanism of the Mycobacterium tuberculosis PknB receptor Ser/Thr kinase. EMBO J *27*, 3186-3197.

Moussatche, P., and Klee, H. J. (2004). Autophosphorylation Activity of the Arabidopsis Ethylene Receptor Multigene Family. Journal of Biological Chemistry *279*, 48734-48741.

Müller-Dieckmann, H. J., Grantz, A. A., and Kim, S. H. (1999). The structure of the signal receiver domain of the Arabidopsis thaliana ethylene receptor ETR1. Structure *7*, 1547-1556.

Mueller-Dieckmann, J. (2006). The open-access high-throughput crystallization facility at EMBL Hamburg. Acta Crystallogr D Biol Crystallogr *62*, 1446-1452.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Neljubov, D. (1901). Über die horizontale Nutation der Stengel von Pisum sativum und einiger anderer Pflanzen. Beih Bot. Zentralb. *10*, 128-139.

Nolen, B., Taylor, S., and Ghosh, G. (2004). Regulation of Protein KinasesControlling

Activity through Activation Segment Conformation. Molecular Cell *15*, 661-675.

O'Malley, R. C., Rodriguez, F. I., Esch, J. J., Binder, B. M., O'Donnell, P., Klee, H. J., and Bleecker, A. B. (2005). Ethylene-binding activity, gene expression levels, and receptor system output for ethylene receptor family members from Arabidopsis and tomato. Plant J *41*, 651-659.

Ohme-Takagi, M., and Shinshi, H. (1995). Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell *7*, 173-182.

Panjikar, S., Mayerhofer, H., Tucker, P. A., Mueller-Dieckmann, J., and de Sanctis, D. (2011). Single isomorphous replacement phasing of selenomethionine-containing proteins using UV-induced radiation damage. Acta Crystallogr. D Biol. Crystallogr *67*, 32-44.

Pape, T., and Schneider, T. R. (2004). HKL2MAP : a graphical user interface for macromolecular phasing with SHELX programs. J Appl Crystallogr *37*, 843-844.

Parkinson, J. S., and Kofoid, E. C. (1992). Communication modules in bacterial signaling proteins. Annu. Rev. Genet *26*, 71-112.

Petoukhov, M. V., and Svergun, D. I. (2005). Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data. Biophysical Journal *89*, 1237-1250.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Pike, A. C. W., Rellos, P., Niesen, F. H., Turnbull, A., Oliver, A. W., Parker, S. A., Turk, B. E., Pearl, L. H., and Knapp, S. (2008). Activation segment dimerization: a mechanism for kinase autophosphorylation of non-consensus sites. EMBO J *27*, 704-714.

Potuschak, T., Lechner, E., Parmentier, Y., Yanagisawa, S., Grava, S., Koncz, C., and Genschik, P. (2003). EIN3-Dependent Regulation of Plant Ethylene Hormone Signaling by Two Arabidopsis F Box Proteins: EBF1 and EBF2. Cell *115*, 679-689.

Qiao, H., Chang, K. N., Yazaki, J., and Ecker, J. R. (2009). Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in Arabidopsis. Genes Dev *23*, 512-521.

Qu, X., and Schaller, G. E. (2004). Requirement of the histidine kinase domain for signal transduction by the ethylene receptor ETR1. Plant Physiol *136*, 2961-2970.

Qu, X., Hall, B. P., Gao, Z., and Schaller, G. E. (2007). A strong constitutive ethylene-response phenotype conferred on Arabidopsis plants containing null mutations in the ethylene receptors ETR1 and ERS1. BMC Plant Biol *7*, 3.

Rajakulendran, T., Sahmi, M., Lefrançois, M., Sicheri, F., and Therrien, M. (2009). A dimerization-dependent mechanism drives RAF catalytic activation. Nature *461*, 542-545.

Rodríguez, F. I., Esch, J. J., Hall, A. E., Binder, B. M., Schaller, G. E., and Bleecker, A. B. (1999). A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. Science *283*, 996-998.

Roessle, M. W., Klaering, R., Ristau, U., Robrahn, B., Jahn, D., Gehrmann, P., Konarev, P., Round, A., Fiedler, S., Hermes, C., et al. (2007). Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. J. Appl. Cryst. *40*, 190-194.

Schaller, G. E., Ladd, A. N., Lanahan, M. B., Spanbauer, J. M., and Bleecker, A. B. (1995). The ethylene response mediator ETR1 from Arabidopsis forms a disulfide-linked dimer. J. Biol. Chem *270*, 12526-12530.

Scharein, B., Voet-van-Vormizeele, J., Harter, K., and Groth, G. (2008). Ethylene signaling: identification of a putative ETR1-AHP1 phosphorelay complex by fluorescence spectroscopy. Anal. Biochem *377*, 72-76.

Schenk, P. W., and Snaar-Jagalska, B. E. (1999). Signal perception and transduction: the role of protein kinases. Biochim. Biophys. Acta *1449*, 1-24.

Sicheri, F., Moarefi, I., and Kuriyan, J. (1997). Crystal structure of the Src family tyrosine kinase Hck. Nature *385*, 602-609.

Singh, M., Berger, B., Kim, P. S., Berger, J. M., and Cochran, A. G. (1998). Computational

learning reveals coiled coil-like motifs in histidine kinase linker domains. Proc. Natl. Acad. Sci. U.S.A *95*, 2738-2743.

Skerker, J. M., Perchuk, B. S., Siryaporn, A., Lubin, E. A., Ashenberg, O., Goulian, M., and Laub, M. T. (2008). Rewiring the specificity of two-component signal transduction systems. Cell *133*, 1043-1054.

Solano, R., Stepanova, A., Chao, Q., and Ecker, J. R. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. Genes & Development *12*, 3703-3714.

Stanton, V. P., Nichols, D. W., Laudano, A. P., and Cooper, G. M. (1989). Definition of the human raf amino-terminal regulatory region by deletion mutagenesis. Mol. Cell. Biol *9*, 639-647.

Stepanova, A. N., and Alonso, J. M. (2009). Ethylene signaling and response: where different regulatory modules meet. Curr. Opin. Plant Biol *12*, 548-555.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. Annu. Rev. Biochem *69*, 183-215.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification *41*, 207-234.

Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr *25*, 495-503.

Svergun, D., Barberato, C., and Koch, M. H. J. (1995). CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. J Appl Crystallogr *28*, 768-773.

Swanson, R. V., Alex, L. A., and Simon, M. I. (1994). Histidine and aspartate phosphorylation: two-component systems and the limits of homology. Trends in Biochemical Sciences *19*, 485-490.

Tao, W., Malone, C. L., Ault, A. D., Deschenes, R. J., and Fassler, J. S. (2002). A cytoplasmic coiled-coil domain is required for histidine kinase activity of the yeast osmosensor, SLN1. Mol. Microbiol *43*, 459-473.

Taylor, S. S., and Kornev, A. P. (2011). Protein Kinases: Evolution of Dynamic Regulatory Proteins. Trends Biochem Sci *36*, 65-77.

Testerink, C., Larsen, P. B., van der Does, D., van Himbergen, J. A. J., and Munnik, T. (2007). Phosphatidic acid binds to and inhibits the activity of Arabidopsis CTR1. J. Exp. Bot *58*, 3905-3914.

Testerink, C., Larsen, P. B., McLoughlin, F., van der Does, D., van Himbergen, J. A., and Munnik, T. (2008). PA, a stress-induced short cut to switch-on ethylene signalling by switching-off CTR1? Plant Signal Behav *3*, 681-683.

Thomas, S. A., Brewster, J. A., and Bourret, R. B. (2008). Two variable active site residues modulate response regulator phosphoryl group stability. Mol. Microbiol *69*, 453-465.

Thomason, P. A., Wolanin, P. M., and Stock, J. B. (2002). Signal transduction: receptor clusters as information processing arrays. Curr. Biol *12*, R399-401.

Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S. K., Zhu, Y., Ishima, R., Liu, D., Tong, K. I., Kurokawa, H., et al. (1999). Solution structure of the homodimeric core domain of Escherichia coli histidine kinase EnvZ. Nat. Struct. Biol *6*, 729-734.

Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., and Hodges, R. S. (2000). Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d." J. Mol. Biol *300*, 377-402.

Tsuchisaka, A., and Theologis, A. (2004). Unique and overlapping expression patterns among the Arabidopsis 1-amino-cyclopropane-1-carboxylate synthase gene family members. Plant Physiol *136*, 2982-3000.

Voet-van-Vormizeele, J., and Groth, G. (2008). Ethylene Controls Autophosphorylation of the Histidine Kinase Domain in Ethylene Receptor ETR1. Mol Plant *1*, 380-387.

Volkov, V. V., and Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. J Appl Crystallogr *36*, 860-864.

Wang, K. L.-C., Yoshida, H., Lurin, C., and Ecker, J. R. (2004). Regulation of ethylene gas biosynthesis by the Arabidopsis ETO1 protein. Nature *428*, 945-950.

Wang, W., Hall, A. E., O'Malley, R., and Bleecker, A. B. (2003). Canonical histidine kinase activity of the transmitter domain of the ETR1 ethylene receptor from Arabidopsis is not required for signal transmission. Proc. Natl. Acad. Sci. U.S.A *100*, 352-357.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189 -1191.

Woeste, K. E., and Kieber, J. J. (2000). A strong loss-of-function mutation in RAN1 results in constitutive activation of the ethylene response pathway as well as a rosette-lethal phenotype. Plant Cell *12*, 443-455.

Yanagisawa, S., Yoo, S.-D., and Sheen, J. (2003). Differential regulation of EIN3 stability by glucose and ethylene signalling in plants. Nature *425*, 521-525.

Yang, J., Ten Eyck, L. F., Xuong, N.-H., and Taylor, S. S. (2004). Crystal Structure of a cAMP-dependent Protein Kinase Mutant at 1.26 Å: New Insights into the Catalytic Mechanism. Journal of Molecular Biology *336*, 473-487.

Yang, S. F., and Hoffman, N. E. (1984). Ethylene Biosynthesis and its Regulation in Higher Plants. Annu. Rev. Plant. Physiol. *35*, 155-189.

Yang, Y., and Inouye, M. (1991). Intermolecular complementation between two defective mutant signal-transducing receptors of Escherichia coli. Proc. Natl. Acad. Sci. U.S.A *88*, 11057-11061.

Yoo, S.-D., Cho, Y.-H., Tena, G., Xiong, Y., and Sheen, J. (2008). Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signalling. Nature *451*, 789-795.

Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M., and Alber, T. (2003). Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases. Nat Struct Mol Biol *10*, 168-174.

Zhu, Y., Qin, L., Yoshida, T., and Inouye, M. (2000). Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain. Proc. Natl. Acad. Sci. U.S.A *97*, 7808-7813.

List of Figures and Tables

Appendix

The methods for the mass spectrometry analysis carried out by the staff of the EMBL Core facility are summarized below.

In-solution digestion

All reagents used were dissolved in 50 mM ammonium bicarbonate. Approximately 100 µg of protein was taken (50 µL of a 2 µg/ µL solution provided) and diluted to 1 µg/µL in 50 mM ammonium bicarbonate. Proteins were then reduced (2.5 µL of 50 mM DTT, 56 °C, 30 minutes) and alkylated (110 mM iodoacetamide, 2.5 µL, room temperature, in the dark, 20 minutes). Trypsin digestion (1 µL of a 1 µg/µL solution) was carried out overnight at 37 °C.

LC-MS/MS

Peptides were separated using the Proxeon Easynano LC system (Thermo Fisher) fitted with a trapping column (self-packed Hydro-RP C_{18} (Phenomenex), 100 µm x 2.5 cm, 4 µm) and an analytical column (self-packed Reprosil C_{18} (Dr. Maisch) 75 µm x 15 cm, 3 µm, 100 Å). The outlet of the analytical column was coupled directly to an HCT Ultra Ion Trap mass spectrometer (Bruker Daltonics) using the ESI nanoflow source in positive ion mode.

Solvent A was water, 0.1 % formic acid and solvent B was acetonitrile, 0.1 % formic acid. The samples (8 µL) were loaded with a constant pressure (280 bar) of solvent A with a total volume of 15 µL onto the trapping column. Peptides were eluted via the analytical column a constant flow of 0.3 µL/min. During the elution step, the percentage of solvent B increased in a linear fashion from 4 % to 40 % B in 20 minutes.

On the HCT Ultra, a spray voltage of 4.3 kV was applied. The drying temperature was set at 200 °C, with drying gas at 6 L/min and nebulizer gas at 7 psi. Data was acquired in standard enhanced mode. Full scan MS spectra with mass range 350-1500 *m/z* were acquired in profile mode with a maximum fill time of 200 ms or 200000 counts. MS/MS was performed in an automated fashion on the 3 most intense ions from the full scan MS. MS/MS spectra were also acquired in profile mode over the mass range 100 – 2000 *m/z*. Dynamic exclusion was used after 2 MS/MS events and had a maximum period of 30 seconds.

Data analysis

Software scripts within Hystar (Bruker) were used for creating .mgf files, needed for searching in MASCOT version 2.2.03 (Matrix Science). The data were searched against a species specific (arabadopsis thaliana) swissprot database as well as a list of common contaminants. The data were searched with the following modifications: Carbamidomethyl (C) (Fixed) and Oxidation (M) (Variable), Phosphorylation (ST). The mass error tolerance for the full scan MS spectra was set at 0.3 Da and for the MS/MS spectra at 0.3 Da. A maximum of 1 missed cleavage was allowed. Only peptides with Mascot score above 20 were reported.

Intact Protein Sample Preparation

Prior to analysis, protein samples in solution were zip-tipped using C4 ZipTips (Millipore). Conditioning and equilibration of the ZipTip for sample binding was performed (3 aspirations each time) with 50:50 water:acetonitrile, followed by water, 0.1% TFA.

Proteins were bound to the ZipTip by aspiration of the protein solution 10 times. The sample was then washed using water, 0.1% (5 times). Proteins were eluted in 10 µL of 50:50 water:acetonitrile, 1% formic acid. Eluant was aspirated 3 times for good sample recovery.

Intact Protein Sample Analysis

Salt-free samples were loaded into coated (3AP) borosilicate PicoTip Emitters (New Objective) and mounted on the static nanospray source of a Q-Tof II mass spectrometer (Waters/Micromass). After breaking the needle tip, 1.8 kV was applied to the PicoTip. A cone voltage of 40 V was used, with extraction cone at 10 V. A collision energy of 8 eV was used, with Argon in the collision cell. Data were acquired in continuum mode between 500 and 3000 *m/z* with a 1 second scan time and acquired for approximately 1 minute. Data were externally calibrated against a reference standard of intact myoglobin, acquired immediately prior to sample data acquisition. Spectra were then summed and intact mass was calculated using the MaxEnt1 maximum entropy algorithm (Waters/Micromass) to give the zero charge deconvoluted molecular weight.