Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

# The *Halobacterium salinarum* Taxis Signal Transduction Network: a Protein-Protein Interaction Study

Matthias Schlesner

aus Kiel

2008

# Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. Dieter Oesterhelt betreut.

# Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig und ohne unerlaubte Hilfe angefertigt.

Martinsried, am 07.10.2008

Matthias Schlesner

Dissertation eingereicht am: 14.10.2008

1. Gutachter: Prof. Dr. Dieter Oesterhelt
2. Gutachter: Prof. Dr. Wolfgang Marwan

Mündliche Prüfung am: 9.12.2008

Meiner Familie

# Contents

*Contents*

# List of Figures

# List of Tables

# Summary

The archaeon *Halobacterium salinarum* grows optimally in saturated brine environments like salt lakes and solar salterns. To survive in these harsh environments, *H. salinarum* actively seeks the places with the best growth conditions by moving in a biased random walk. This process, called taxis, is driven by two molecular systems: First, the motility apparatus, consisting of the flagellum, the flagellar motor, and its switch, which allows to change the direction of flagellar rotation. Second, the chemotaxis signal transduction system, which targets the flagellar motor switch in order to modulate the switching frequency in response to external stimuli. This modulation is the fundamental basis for the biased random walk. Whereas the signal transduction system is conserved throughout Archaea and Bacteria, the archaeal motility apparatus is unique and only poorly understood on the molecular level. The proteins constituting the flagellar motor and its switch in Archaea have not yet been identified, so that the connection between the bacterial-like chemotaxis signal transduction system and the archaeal motility apparatus is not known.

The aim of this study was to extend the understanding of the taxis signal transduction system of *H. salinarum* using protein-protein interaction (PPI) analysis. By this, the roles of several proteins involved in this signalling system like CheW1/CheW2, CheC1/CheC2/CheC3, or CheD should be enlightened, and previously unrecognised proteins involved in or connected to the system identified. The unknown connection to the archaeal flagellum was a further focus.

In the first step, a method had to be found that allows for the analysis of interactions of halophilic proteins. The yeast two-hybrid system was used to test a set of known interactors from *H. salinarum*, but here it failed in all cases. The main problem was transcriptional activation by the (acidic) halobacterial proteins when fused to a DNA-binding domain, and probably also misfolding of the high-salt adapted proteins when expressed in yeast. Thus an affinity purification method for halobacterial protein complexes was established, which enables the identification of the complex components by mass spectrometry. As affinity tag a cellulose-binding domain from *Clostridium thermocellum* is used, which binds cellulose even in the presence of multi-molar salt concentrations, so that the whole purification can be performed under close to physiological conditions. Labelling with stable isotopes (SILAC) is applied to discriminate specific interaction partners from unspecific contaminants. In test experiments with the established method, it was possible to identify several known and new protein interactions, so it was finally applied to analyse the interactions of all ten halobacterial chemotaxis (Che) proteins. To assist the laboratory work and the handling and

evaluation of results, a bioinformatics environment was created and integrated in the department's database system, HaloLex (`http://www.halolex.mpg.de/`).

The protein interaction study generally confirmed the expected topology of the core of the taxis signalling system: interactions between CheA, CheW1, CheY, and several halobacterial transducers (Htrs) could be detected. With PurH/N and OE4643R two unexpected interactors of the core or CheA, respectively, have been identified, whose role in taxis signalling is completely unclear. For the two CheW proteins, CheW1 and CheW2, different interactions have been detected – CheW1 appeared as the main coupling protein for the formation of stable signalling complexes between the Htrs and CheA. The Htrs as preys revealed nonuniform interaction patterns: some Htrs were associated to CheW1, CheW2, CheA, and CheY, others mainly with CheW2, a third group with CheA and CheY, but none of the two CheWs, and a fourth group was not fished at all. An unexpected finding was the central position of CheD in the Che interaction network, indicating that this protein might play a key role in the halobacterial taxis signalling system.

In the PPI analysis, three proteins were identified that connect the chemotaxis system and the archaeal flagellar apparatus. These proteins interact with the chemotaxis proteins CheY, CheD, and CheC2, as well as the the flagella accessory (Fla) proteins FlaCE and FlaD. Two of the proteins are homologous to each other and belong to the protein family DUF439, the third is a HEAT_PBS family protein. In-frame deletion strains for all three proteins were generated and analysed for chemotaxis deficiencies: a) their photophobic responses were measured by a computer-based cell tracking system b) the flagellar rotational bias was determined by dark-field microscopy, and c) the chemotactic behaviour was analysed by a swarm plate assay.

Strains deleted for the HEAT_PBS protein or one of the DUF439 proteins proved unable to switch the direction of flagellar rotation. In these mutants, flagella rotate only clockwise, which results in exclusively forward swimming cells. Accordingly, chemo- and phototactic responses are disabled. Deletion of the second DUF439 protein had only minimal effects. In none of the deletion strains swimming motility was impaired.

By homology searches, HEAT_PBS proteins could be identified in the chemotaxis gene regions of all motile haloarchaea sequenced so far, but not of other archaeal species. DUF439 proteins, however, are inherent parts of archaeal chemotaxis gene regions, and they are restricted to this genomic context. Altogether, these results demonstrate that hitherto unrecognised archaea-specific Che proteins are essential for relaying taxis signalling to the flagellar apparatus in the archaeal domain.

# 1 Background

## 1.1 *H. salinarum*, an archaeal model organism

### 1.1.1 *Halobacterium salinarum*

*H. salinarum*, formerly called *H. halobium* and *H. salinarium*, is a rod-shaped, motile archaeon. The cells are usually between 2 and 6 μm long, approximately 0.4 - 0.7 μm wide, and monopolarly or bipolarly flagellated (Figure 1.1). The type of flagellation depends on the growth phase: during the logarithmic phase the majority of cells is monopolarly flagellated, while in the stationary phase bipolarly flagellated cells are dominating (Alam and Oesterhelt, 1984).

Among the archaea that have been cultured so far, the haloarchaea are the easiest to handle since they do neither require extreme temperatures for growth nor strict absence of oxygen. For a long time the haloarchaea have also been the only archaea which could be transformed. Thus they are predestined as a model for studying the archaeal domain (for review see Soppa, 2006).



Figure 1.1: **Electron micrograph of a *H. salinarum* cell.** A monopolarly inserted bundle of flagella is visible. 13.500 fold magnification. Image taken from Staudinger (2001).

After the discovery of bacteriorhodopsin, a light driven proton pump (Oesterhelt and Stoeckenius, 1973), and additional retinal proteins, much interest focused on *H. salinarum*. Bacteriorhodopsin can be easily isolated in large quantities, which allowed numerous studies of the structure and function of this protein (see for example Haupts *et al.*, 1999; Lanyi and Luecke, 2001; Lanyi, 2006, for review). It is still one of the best-characterised membrane proteins.

*H. salinarum* grows optimally in saturated brine environments like salt lakes and solar salterns. It survives in these harsh environments by responding appropriately

to different stimuli, using chemotactic and phototactic sensors connected to a sophisticated signal transduction network (Marwan and Oesterhelt, 2000). Halobacterial chemo- and phototaxis have been the subject of intensive research for a long time, resulting in excellent knowledge of the cellular responses (Nutsch *et al.*, 2003, 2005). The application and removal of light stimuli to trigger phototactic responses can be done in a perfectly controlled manner, which makes this an optimal system for modelling of biological processes. Combined with its relative simplicity this makes *H. salinarum* an important model organism for systems biology (e.g. Bonneau *et al.*, 2007; del Rosario *et al.*, 2007; Gonzalez *et al.*, 2008).

### 1.1.2 Archaea

The Archaea (Figure 1.2) were introduced as a distinct domain besides the Bacteria and Eukarya in the mid-1970s by Carl Woese on the basis of 16s rRNA sequences (Woese and Fox, 1977; Woese *et al.*, 1990). The name Archaebacteria and later Archaea was chosen as the first known members of this domain were found to live under extreme conditions. Such conditions might reflect the environmental situation on earth when life came into existence. By now, cultivation-independent approaches have shown that archaea do not exist exclusively in extreme habitats but are present in almost all environments examined to date (DeLong and Pace, 2001). For example, fluorescent *in situ* hybridisation experiments revealed that archaea represent 20% or more of all microbial cells in the oceans (DeLong *et al.*, 1999).



Figure 1.2: **Phylogenetic tree showing the three domains of life.** The tree is based on the sequences of the small-subunit rRNA. Figure taken from Allers and Mevarech (2005).

Archaea share similarities with both eukarya and bacteria, but they also exhibit a couple of unique features (for review see Allers and Mevarech, 2005). The most obvious similarity between archaea and bacteria is their prokaryotic morphology: both are single-celled, contain no nucleus, and are of approximately the same size. The core

metabolic functions of archaea like energy conversion and biosynthesis resemble those of bacteria, and both archaea and bacteria tend to organise their genes in polycistronic operons.

In contrast, the archaeal information-processing functions such as transcription and translation are similar to the eukaryotic ones, although considerably simpler. However, archaeal genomes possess numerous homologs of bacterial transcription regulators so that transcriptional regulation might be done in a bacterial-type mode (Geiduschek and Ouhammouch, 2005).

Among the unique features of archaea is their cell envelope, which became one of the earliest biochemical distinctions between the two prokaryotic domains (Woese *et al.*, 1978). The archaeal cell envelope is composed of a lipid bilayer, which is, depending on the genus, either surrounded by a proteinaceous surface layer (S-layer), or a rigid cell wall sacculus formed by polymers like pseudomurein or heteropolysaccharide, or nothing at all (for overview see Kandler, 1994). Unlike eubacterial and eukaryotic lipids, which usually consist of fatty acids that are linked to glycerol by ester bonds, the archaeal lipids are mainly isoprenyl glycerol ethers (Gambacorta *et al.*, 1994).

The archaeal domain can be subdivided into the phyla Euryarchaeota, Crenarchaeota, Korarchaeota, and Nanoarchaeota (Woese *et al.*, 1990; Barns *et al.*, 1996; Huber *et al.*, 2002). *H. salinarum* is a member of the family Halobacteriaceae, which belongs to the phylum Euryarchaeota (class Halobacteria, order Halobacteriales). All members of the Halobacteriaceae are obligate halophiles, that means they need elevated salt concentrations (2 M - 5.2 M for halobacteria) for growths (Oren, 1994).

### 1.1.3 Halophiles and their ecology

Most hypersaline environments originate either as a result of evaporation of seawater (thalassohaline environments) or they are formed by evaporative concentration of salts in lakes (athalassohaline environments, e. g. soda lakes or the dead sea). Whereas thalassohaline environments are markedly similar with respect to ion composition (mainly sodium and chloride) and pH (neutral to slightly alkaline), athalassohaline brines may differ greatly in their chemical composition. These differences, especially in the concentrations of the divalent cations calcium and magnesium as well as pH, require specific adaptations and therefore determine the range of organisms able to thrive in these brines (Oren, 1994; Kerkar, 2005).

Halophilic microorganisms are found in all three domains of life (Oren, 2002). Among eukaryotes, there are only few halophiles; the most important one is the green alga *Dunaliella*. This organism can be found almost ubiquitously in high-salt environments, and in many of them *Dunaliella* is the main or only primary producer of biomass. In contrast, the domain Bacteria contains a large number of halophilic microorganisms, spread over many different phylogenetic groups. Most of the halophilic bacteria are rather moderate halophiles, but also a few extreme halophilic species (e. g. *Halorhodospira*) are known. Among the Archaea, halophiles are found in the order Halobacteriales and in the methanogenic branch



Figure 1.3: **Crystalliser pond in the abandoned salterns of Sečovlje, Slovenia.** The red colour is caused by microorganisms like halobacteria.

of the euryarchaeota. Halophilic crenarchaeota have not yet been identified. Halobacteriales are the main component of biomass in many extremely hypersaline water bodies, and they are the main cause of the red coloration of such places due to large concentrations of C-50 carotenoid pigments in their membranes (Figure 1.3). Additional habitats of Halobacteriales are salted fish and hides treated with salt for preservation.

### 1.1.4 Adaptation to hypersaline environments

In order to live at high salt concentrations, halophiles have to maintain a cytoplasm that is at least isoosmotic with the outside medium; otherwise they would lose water to their environment since biological membranes are permeable to water. To achieve this, two different strategies are used (Oren, 1999):

1. Cells maintain low salt concentrations within their cytoplasm and compensate the osmotic pressure of the medium by organic compatible solutes (compatible-solute strategy).

2. The high extracellular ion concentration (normally mainly NaCl) is balanced by even higher intracellular salt concentrations (mainly KCl). This is called the salt-in strategy.

The compatible-solute strategy is found in organisms throughout all three domains of life. Compatible solutes are typically low-molecular-weight compounds that are soluble at high concentrations in water and are either uncharged or zwitterionic at physiological pH. Examples are polyols like glycerol and arabitol, sugars and their derivatives, amino acids and their derivatives, and quaternary amines such as glycine betaine (da Costa *et al.*, 1998). Compatible solutes do not interfere with protein function even at high concentrations so that the compatible-solute strategy does not require specially adapted proteins. However, the compatible solutes need either to be taken up from the medium actively or they must be synthesised. Therefore this strategy is energetically unfavourable in environments with salt concentrations exceeding 1.5 M (Dennis and Shimmin, 1997).

The salt-in strategy is only found in halophilic archaea of the order Halobacteriales and the anaerobic halophilic bacteria of the order Haloanaerobiales. The high intracellular salt concentrations raise the need for all enzymes and structural cell components to be adapted to ensure their function under these conditions (Lanyi, 1974; Eisenberg *et al.*, 1992). Such halophilic proteins differ in amino acid composition from their mesophilic counterparts (Figure 1.4 A). They contain an excess of acidic amino acids, especially on the surface of the protein (Figure 1.4 B). The frequency of the basic amino acid lysine is reduced (Lanyi, 1974; Fukuchi *et al.*, 2003). The acidic residues at the surface are highly hydrated and thereby maintain an extensive hydration network even under competition with a multitude of small cations for free water (Frolow *et al.*, 1996). The excess of acidic residues also reduces overall hydrophobicity and helps to prevent aggregation (salting out) of the protein. The downside is that halophilic proteins often lose their physiological interactions and even denature in solutions of low ionic strength (see Mevarech *et al.*, 2000, and references therein).

## 1.1.5 Bioenergetics

To survive in its challenging habitat, *H. salinarum* is capable of multiple ways of energy production and possesses a simple but highly effective energy storage system.

If sufficient oxygen is available, *H. salinarum* gains energy by aerobic respiration. Organic compounds, mainly amino acids, are oxidised via the tricarboxylic acid (TCA) cycle (Ghosh and Sonawat, 1998) in combination with a respiratory electron transport chain (Schäfer *et al.*, 1996) using oxygen as terminal electron acceptor. Since
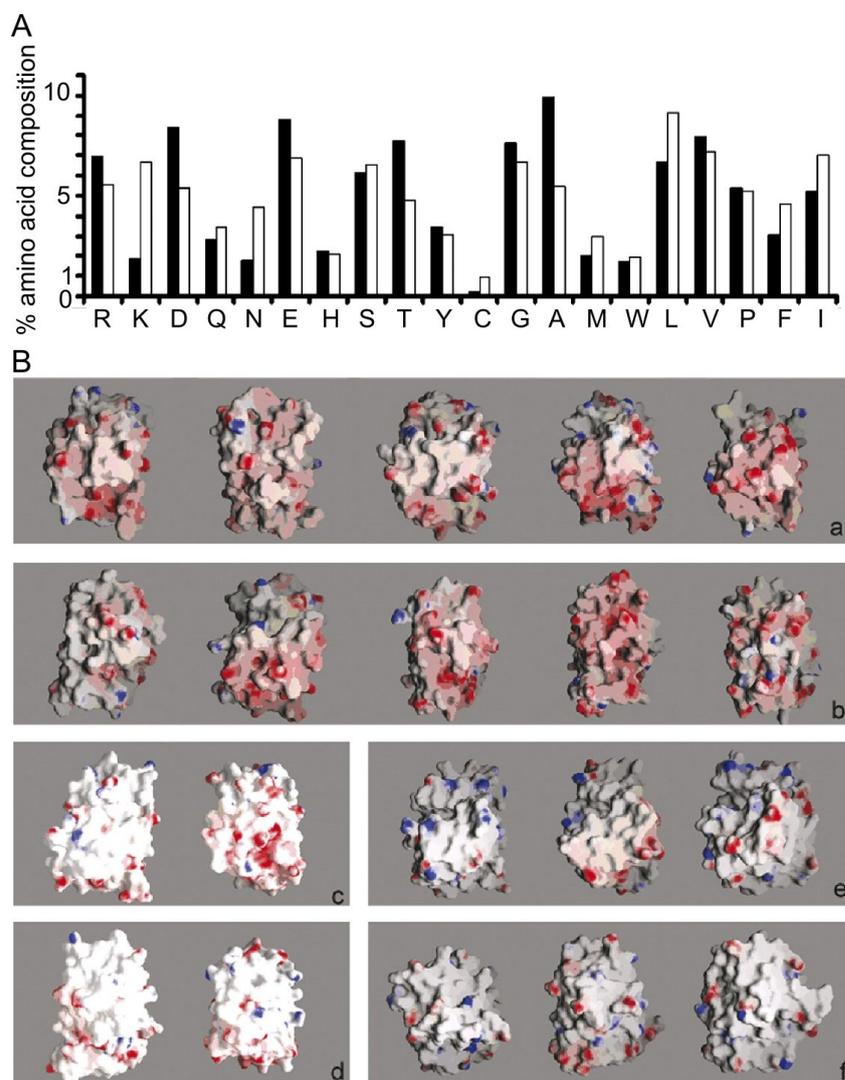
Figure 1.4: **Halophilic adaptation of proteins. A** Column charts indicating the average amino acid composition of halophilic and mesophilic DHFRs (Dihydrofolate reductase). The black bar corresponds to halophilic archaeal species and the white bar corresponds to the average amino acid composition of all other groups except the halophilic archaeal group. These groups include bacterial, fungal and the metazoan taxa. **B** Relative electrostatic surface potentials of DHFRs from halophilic and mesophilic organisms. It is clearly observed that all archaeal DHFRs of species that live in the Dead Sea (a and b) exhibit highly negatively charged surfaces (red means negative, blue positive surface charge). Only the DHFR from *Haloquadratum walsbyi* seems to differ (c and d) due to the extreme environment where this organism lives (high $Mg^{2+}$ concentrations). From left to right, halophilic enzymes correspond to sequences with Uniprot accession numbers Q5V600, Q5V3R2, P15093, Q9UWQ4, Q3IQP3 (a and rear view, b), Q18J41 and Q18HG9 (c and rear view, d). Non-halophilic DHFRs correspond to (from left to right) the apoenzymes of human DHFR (PDB ID: 1KMV), *Escherichia coli* DHFR (PDB ID: 7DFR) and *Candida albicans* DHFR (PDB ID: 1M7A) (e and rear view, f). Figures and caption slightly modified from Kastritis *et al.* (2007).

oxygen solubility is low in saturated brines, oxygen shortage is common in high salt environments. In such situations, *H. salinarum* can switch to anaerobic respiration using alternative electron acceptors such as dimethyl sulfoxide, triethyl amine N-oxide, nitrate, or fumarate (Oren and Trüper, 1990; Oren, 1991).

Alternatively, *H. salinarum* can produce energy via substrate level phosphorylation. L-arginine is fermented via the arginine deiminase pathway (Hartmann *et al.*, 1980; Ruepp and Soppa, 1996). Arginine deiminase converts arginine to citrulline, which, catalyzed by the catabolic ornithine transcarbamylase, reacts with inorganic phosphate to carbamoyl phosphate and ornithine. Carbamoyl phosphate is split into ammonia and carbon dioxide by carbamate kinase under phosphorylation of ADP to ATP.

Finally, *H. salinarum* is capable of photosynthesis. Light powers the retinal protein bacteriorhodopsin (BR) that pumps protons from the cytosol to the extracellular space. The resulting proton gradient can then be used by the ATPase to produce ATP (see Schäfer *et al.*, 1999, for review). Halorhodopsin (HR) is another retinal protein that uses light to pump ions against the membrane potential across the cell membrane. It pumps chloride ions into the cytosol, which helps to maintain the high intracellular chloride concentration required for osmotic balance. Furthermore, the electrogenic transport of chloride ions leads to a concomitant influx of potassium ions into the cytosol (Lindley and MacDonald, 1979; Schobert and Lanyi, 1982).

High intracellular potassium concentrations serve as energy storage to bridge periods of low energy supply ("potassium battery") (Wagner *et al.*, 1978; Schäfer *et al.*, 1999). The battery is charged as described above by the action of HR or by the proton gradient, which drives an efflux of sodium through a sodium:proton antiport with the resulting influx of potassium through the potassium uniport. When energy supply is low, the flux of the sodium:proton antiport reverses: sodium flows in and protons are pumped out, thereby strengthening the proton gradient, which is then used for ATP synthesis.

## 1.2 Signal transduction and taxis in prokaryotes

To sense environmental changes and respond appropriately is a basic prerequisite to survive in a dynamically changing environment. Changes in numerous intra- and extracellular parameters are monitored by sensory proteins. These proteins transmit

the signals via different signalling pathways to effectors which generate the response, e. g. alterations in gene expression or movement to more favourable locations.

Whereas the majority of sensory pathways in eukaryotic organisms are based on serine, threonine, or tyrosine phosphorylation, the most prominent pathways in prokaryotes are based on histidine-aspartate phosphorelay (HAP) systems (for review see Wadhams and Armitage, 2004). However, HAP systems were also identified in lower eukaryotes and plants (Wolanin *et al.*, 2002), and serine, threonine and tyrosine phosphorylation might play a considerable role in bacteria and archaea (see Kennelly, 2002, 2003, and references therein).

### 1.2.1 Two-component systems

HAP systems are also called two-component systems as they consist at least of a dimeric histidine kinase (HK) and a response regulator (RR). The basic workflow of HAP systems consists of trans-autophosphorylation of a histidine residue in one monomer of the HPK dimer by the γ-phosphoryl group of an ATP molecule that is bound to the kinase domain of the other monomer. The phosphoryl group is then passed to an aspartate residue of the RR where it alters the conformation and activity of the RR's output domain. Signal termination occurs by dephosphorylation of the RR, either spontaneously, by autodephosphorylation, or catalyzed by phosphatases. Transfer of the phosphoryl group back to the HK and then to another response regulator which acts as phosphate sink (Sourjik and Schmitt, 1998; Porter and Armitage, 2002) might be a further way for signal removal.

Sensed signals typically change the activity of the HPK and influence thereby the amount of phosphorylated RR. Signal reception in HAP systems involved in transcriptional regulation is generally done by a N-terminal sensory domain of the HPK, whereas the HAP systems controlling the motility behaviour utilise independent sensory proteins. The separation of sensor proteins and HPK allows sensing of many different parameters via different transducer proteins that act on one and the same HPK, generating an unambiguous output to the motility system (for review see Parkinson and Kofoid, 1992; Hoch, 2000; Stock *et al.*, 2000; Wadhams and Armitage, 2004; Szurmant and Ordal, 2004).

## 1.2.2 The principles of prokaryotic taxis

Although most of the prokaryotic HAP systems are involved in the regulation of gene expression, the most-studied pathway is the one that regulates flagella-driven taxis. Taxis is the biased movement in the direction of increasing concentrations of attractants or decreasing concentrations of repellents. Stimuli that trigger a tactic response include specific chemicals (chemotaxis), light (phototaxis), oxygen (aerotaxis), and salinity (osmotaxis) (see Armitage, 1999; Marwan and Oesterhelt, 2000, for review). Bacteria achieve motility either by swimming driven by flagella or as surface-mediated translocation (twitching motility, gliding motility), using type IV pili or other systems. Till now, in archaea only swimming motility driven by flagella has been observed (Bardy *et al.*, 2003), so the following section will focus on this type of motility.

As long as no changes in stimulation are detected, the cells perform a random walk: they change their direction of movement randomly without any preferred direction (Figure 1.5 A). In *H. salinarum*, this switching occurs on average every 10 seconds (Hildebrand and Schimz, 1990). Upon stimulation, the random walk becomes biased: if an overall improvement in the monitored parameters is detected, the cell prolongs the movement in this direction, whereas a worsening of the environment leads to quicker changing of the direction (Figure 1.5 B). Thus, bacteria and archaea do not find the optimal environment by straight following a concentration gradient but by biasing their random movement.

To bias their direction of movement, the organisms detect changes in the strength of a



Figure 1.5: **The (biased) random walk**. **A** If the cells do not detect changes in any stimuli, they perform a random walk. Overall, this kind of movement does not lead to any net displacement of the cell. **B** If a gradient of attractant or repellent stimuli is detected, the random walk becomes biased. Straight movement is prolonged when the conditions improve, and shortened when the environment becomes worse. This behaviour leads to a net movement towards places with more favourable conditions.

stimulus, e. g. the concentration of a chemical, and not the absolute stimulus strength. It is widely accepted that prokaryotes are too small to sense a concentration gradient along their cell size and therefore detect temporal changes (Macnab and Koshland, 1972; Berg and Purcell, 1977; see Thar and Kuhl, 2003, for a contrasting view). That

means that they need some kind of memory to be able to compare the actual strength of a stimulus with the previous one. This memory is achieved by adaptation, which lets the cell behave as if no stimulus were present after a short period of continuous exposure. The adaptation system is so efficient that a change in a few molecules can be sensed in the presence of background concentrations that can vary over at least five orders of magnitude (Kim *et al.*, 2001; Sourjik and Berg, 2002b).

Despite the overall similarity there are also fundamental differences between archaeal and bacterial swimming motility. The left-handed helical flagellar filaments of the bacterial prototypes *E. coli*, *B. subtilis*, and *S. typhimurium* form a cooperative bundle upon counterclockwise (CCW) rotation of the flagellar motor, leading to straight forward swimming of the cell. When the motor changes its direction of rotation to clockwise (CW), the flagellar bundle disassembles and the cell tumbles. When the motor turns back to clockwise rotation, the flagellar bundle reassembles and the cell starts swimming again, albeit in a different direction (Eisenbach, 1990).

In contrast, the right-handed flagellar bundle of *H. salinarum* stays intact independent of the rotational sense of the motor. CW rotation of the flagellum pushes the cell, whereas a CCW rotation pulls the cell, so the cell appears to swim with the flagella in front (Alam and Oesterhelt, 1984; Marwan *et al.*, 1991). When the cell switches from forward to backward movement or vice versa, there is a short stop phase of several 100 ms in between, in which the cell is slightly displaced, so it does not swim back on the same path (Marwan *et al.*, 1991).

## 1.3 Protein-protein interaction analysis

Protein-protein interactions (PPI) are fundamental for most biological processes, as nearly all proteins are functioning as part of larger complexes rather than working in isolation. The interactions between proteins may be static or transient, the latter often occurring in signalling and metabolic pathways. As elementary constituents of cellular complexes and pathways, protein interactions are intimately related to protein functions. Therefore studying protein interactions can help to infer the function of uncharacterised proteins ("guilt-by-association", Semple *et al.*, 2002). Knowledge of protein interactions is also invaluable for understanding a complex signal transduction network like the chemotaxis signalling system. The rationale of PPI analysis is

reviewed e. g. in Boulton *et al.* (2001); Titz *et al.* (2004); Collura and Boissy (2007).

With the availability of the first large-scale PPI datasets the network-based analysis of PPIs arised, with the aim to understand biological properties of the underlying system by studying network topology. For example, it has been found that PPI networks are scale-free (Barabasi and Albert, 1999), implying robustness to random component failure (Albert *et al.*, 2000; Goh *et al.*, 2002), and that the structure of the PPI network is related to whether or not a given protein is essential (Jeong *et al.*, 2001; Han *et al.*, 2004). However, newer studies demonstrated that the observed network topology might not necessarily represent the underlying "true" PPI network, but is heavily influenced by dataset biases (Hakes *et al.*, 2005; Han *et al.*, 2005; Stumpf *et al.*, 2005). Therefore it is important to be very cautious with inferring biological significance from network properties. Hakes *et al.* (2008) gives a critical commentary on this topic.

Several methods have been developed to investigate protein interactions. X-ray crystallography and NMR spectroscopy can characterise interactions at the atomic scale, producing very detailed data that show the precise structural relationship between interacting atoms and residues. In contrast, methods for studying interactions at the molecular scale do not reveal the precise chemical nature of the interactions but simply report that an interaction takes place. These methods include the yeast two-hybrid (Y2H) system and its derivatives for binary interactions (i. e. interactions between pairs of proteins) and affinity purification (AP) combined with mass spectrometry (MS) for complex interactions (i. e. interactions between multiple proteins). Both methods are reviewed in Causier (2004). A third class are prediction methods. These methods are either based on genomic information (e. g. domain fusions, phylogenetic profiles, gene neighbourhood), on interaction data on orthologous proteins in other species, on evolutionary information (for example conservation and variation of certain residues in an interaction site in the *in silico* two-hybrid system, or the similarity of phylogenetic trees), or just on the protein sequence information (machine learning on a large number of PPIs). An overview of the computational methods for PPI prediction is given for example in Valencia and Pazos (2002); Pitre *et al.* (2008).

Large scale protein interaction studies have been performed in *S. cerevisiae* using Y2H (Uetz *et al.*, 2000; Ito *et al.*, 2001) and AP-MS (Gavin *et al.*, 2002, 2006; Ho *et al.*, 2002; Krogan *et al.*, 2006), *C. elegans* and *D. melanogaster* with Y2H (Li *et al.*, 2004; Giot *et al.*, 2003; Formstecher *et al.*, 2005), in *E. coli* with AP-MS (Butland *et al.*, 2005; Arifuzzaman *et al.*, 2006), in *H. pylori*, *T. pallidum*, and *C. jejuni* with

Y2H (Rain *et al.*, 2001; Rajagopala *et al.*, 2007; Parrish *et al.*, 2007), and some viruses with Y2H (McCraith *et al.*, 2000; Uetz *et al.*, 2006).

The overlap between the different studies on the yeast interactome, the best studied interactome so far, is remarkably small (Bader *et al.*, 2004). This can be explained by considerable rates of false negatives and/or false positives in the single datasets (Hart *et al.*, 2006; Huang *et al.*, 2007). Furthermore, the different methods (Y2H vs. AP-MS) have dissimilar strengths and weaknesses (biases) (von Mering *et al.*, 2002): Y2H works rather well with transient PPI as those in signalling pathways, but, due to its binary character, loses interactions that need multiple proteins to participate. In contrast, the strength of AP-MS is the analysis of big, stable protein complexes, whereas short-lived, transient interactions might be lost. Neither of these methods is able to reproduce a comprehensive image of the underlying interactome.

The study of protein interactions in archaea is still at its beginning. To my knowledge, only one large-scale and a few mid-scale interaction studies have been carried out in archaeal organisms: Interactions of almost 1000 proteins from *P. horikoshii* were analysed using a mammalian two-hybrid system (Usui *et al.*, 2005), the RNA polymerase from *P. furiosus* by Far-Western blotting (Goede *et al.*, 2006), and RNase P subunits from *M. thermoautotrophicus* using Y2H (Hall and Brown, 2004). None of the commonly used techniques for PPI analysis has been described to be applicable for high-salt adapted proteins.

Computational methods for PPI analysis are much less powerful in archaea than in bacteria: till now only 52 archaeal genomes have been sequenced, compared to 626 bacterial genomes (numbers from NCBI in April 2008), limiting the usability of genome-based methods. Sequence-based approaches or inferring interactions from datasets from other organisms are also difficult, because no large experimental dataset exists for an archaeal organism, and the organisms with large datasets are rather different. Hence the analysis of PPI in *H. salinarum* should be based on an experimental rather than a computational strategy.

## 1.4 Objectives

The purpose of this study was to improve the understanding of the chemotaxis signal transduction system of *H. salinarum* through protein-protein interaction analysis.

Several aspects of this system are not fully understood, e. g. why this organism codes for two CheW and three CheC proteins. PPI analysis can help to recognise the role of a protein by pointing to its interaction partners, and it can identify previously overlooked proteins which are involved in a certain process.

The specific objectives of the research were to:

- Find a method for investigating the interactions of high-salt adapted proteins. None of the commonly used PPI analysis techniques has been shown to be applicable to halophilic proteins.

- Apply this method to the proteins known to participate in taxis signalling. By this, the roles of several Che proteins like CheW1 and CheW2, CheC1, CheC2, and CheC3, or CheD should be enlightened, and previously unrecognised proteins involved in or connected to the Che system identified. The unknown connection between the Che system and the archaeal flagellum was a further focus of this work.

- Perform functional studies to put the found interactions into context. PPI analysis will in most cases not be sufficient to elucidate the function of an unknown protein, but it is a valuable tool for generating hypothesis for follow-up experiments. These experiments were based on the deletion of proteins of unknown function and the subsequent characterisation of the resulting phenotype.

# 2 Materials and methods

## 2.1 General materials

### 2.1.1 Instruments

Devices related to specific methods are listed in the respective sections. Other instruments are listed in Table 2.1.

Table 2.1: **Instruments**

| Instrument | Distributor |
| --- | --- |
| Analytical balance HL52 | Mettler Toledo |
| Autoclave Varioklav 500 EP-Z | H+P Labortechnik |
| Balance PB3002-SDR | Mettler Toledo |
| Centrifuge 5417R, rotor FA45-30-11 | Eppendorf |
| Centrifuge RC5C Plus, rotor GS3 | Sorvall |
| Incubator BK5060E | Heraeus |
| PCR Thermocycler PCR System 9700 | GeneAmp |
| pH meter microprocessor pH 211 | Hanna Instruments |
| Power supply EPS 200 | Pharmacia Biotech |
| Shaker Unimax 2010 | Heidolph |
| Sonifier 250 | Branson |
| UV/Vis spectrometer Ultrospec 3000 | Pharmacia Biotech |
| Vacuum concentrator Speedvac Concentrator | Savant |

### 2.1.2 Chemicals and Kits

All chemicals used in this study were purchased from Sigma-Aldrich (St. Louis, MO, USA), Merck (Darmstadt, Germany), or Fluka (Buchs, Switzerland) at the highest purity grade available. Chemicals of particular importance for this study and exceptions are indicated within the respective chapters or listed in Table 2.2. Kits used in this study are listed in Table 2.3.

### 2.1.3 Enzymes

Restriction endonucleases were purchased from New England Biolabs (Ipswich, MA, USA). Other enzymes are listed in Table 2.4.

Table 2.2: **Chemicals**

| Chemical | Distributor |
| --- | --- |
| Anti Digoxigenin AP Fab fragments | Roche Diagnostics, Mannheim, Germany |
| Avicel PH-101 | Fluka, Buchs, Switzerland |
| Bacto™ agar | Difco, Detroit, MI, USA |
| Bacto™ tryptone | Difco, Detroit, MI, USA |
| Bacto™ yeast extract | Difco, Detroit, MI, USA |
| Blocking reagent for nucleic acid hybridisation | Roche, Mannheim, Germany |
| $C_{18}$, 3M Empore™ High Performance Extraction Disk | 3M, Neuss, Germany |
| Coomassie Brilliant Blue R-250 | Serva, Heidelberg, Germany |
| DIG-11-dUTP | Roche Diagnostics, Mannheim Germany |
| DNA Ladder, GeneRuler™ | Fermentas, St. Leon-Rot, Germany |
| DNA Molecular Weight Marker II, VII, DIG-labeled | Roche, Mannheim, Germany |
| $\alpha$-cyano-4-hydroxy-cinnamic acid | Sigma Aldrich, St. Louis, USA |
| L-Leucine, U-$^{13}C_6$ | Cambridge Isotope Laboratories, Andover, MA, USA |
| Nonidet P40 | Roche, Mannheim, Germany |
| $PEG_{600}$ | Sigma Aldrich, St. Louis, USA |
| Protein Marker, Prestained, Broad Range | New England Biolabs, Ipswich, MA, USA |
| Protein Marker, PageRuler™ unstained | Fermentas, St. Leon-Rot, Germany |
| Protease Inhibitor Cocktail Tablets, Complete Mini, EDTA-free | Roche, Mannheim, Germany |

Table 2.3: **Kits**

| Kit | Distributor |
| --- | --- |
| ABI Prism BigDye™ v3.1 | Applied Biosystems, Foster City, CA, USA |
| Gateway Vector Conversion System | Invitrogen, Karlsruhe, Germany |
| Gateway LR Clonase II Enzyme Mix | Invitrogen, Karlsruhe, Germany |
| In-Fusion™ Dry-Down PCR Cloning Kit | BD Biosciences, Heidelberg, Germany |
| Penta · His™ HRP Conjugate Kit | Qiagen, Hilden, Germany |
| pENTR™/D-TOPO® Cloning Kit | Invitrogen, Karlsruhe, Germany |
| QIAprep 8 Miniprep Kit | Qiagen, Hilden, Germany |
| QIAquick Gel Extraction Kit | Qiagen, Hilden, Germany |
| QIAquick PCR Purification Kit | Qiagen, Hilden, Germany |
| Rapid DNA Ligation Kit | Fermentas, St. Leon-Rot, Germany |

Table 2.4: **Enzymes**

| Enzyme | Distributor |
| --- | --- |
| Calf intestinal phosphatase | New England Biolabs, Ipswich, MA, USA |
| Phusion™ High-Fidelity DNA Polymerase | Finnzymes, Espoo, Finland |
| *Taq* polymerase | MPI for Biochemistry, Martinsried, Germany |
| T4 DNA ligase | Invitrogen, Karlsruhe, Germany |
| Trypsin, Modified Sequencing Grade | Promega, Madison, WI, USA |

## 2.1.4 Strains

Table 2.5: **Strains**

| Strain | Description | Source or Reference |
|---|---|---|
| *H. salinarum* R1 | *H. salinarum* wt (DSM 671) | DSM |
| *H. salinarum* S9 | Highly motile single colony isolate of the S9 strain | Stoeckenius *et al.* (1979); selected for motility by Weidinger (2007) |
| *E. coli* DH5$\alpha$ | F$^-$ $\Phi$80d*lacZ*$\Delta$M15 $\Delta$(*lacZY A – argF*) U169 *recA*1 *endA*1 *hsdR*17($r_k^-$ , $m_k^+$ ) *phoA suppE*44 $\lambda^-$ *thi*-1 *gyrA*96 *relA*1 | Invitrogen, Karlsruhe, Germany |
| *E. coli* BL21(DE3) | F$^-$, *ompT*, *hsd*S$_B$($r_B^-$m$_B^-$), dcm$^+$, Tet$^r$, *gal*$\lambda$(DE3) *endA*, Hte [*argU ileY leuW* Cam$^r$] | Stratagene, La Jolla, CA, USA |
| *E. coli* Mach1™T1$^R$ | F$^-$ $\Phi$80(*lac*Z)$\Delta$M15 $\Delta$*lac*X74 *hsd*R($r_k^-$m$_k^+$) $\Delta$*rec*A1398 *endA*1 *tonA* | Invitrogen, Karlsruhe, Germany |
| *E. coli* ccdBsurvival | F$^-$ *mcrA* $\Delta$(*mrr-hsd*RMS-*mcr*BC) $\Phi$80*lacZ*$\Delta$M15 $\Delta$*lac*X74 *recA*1 *ara*$\Delta$139 $\Delta$(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str$^R$) *endA*1 *nup*G *ton*A::P$_{trc}$-*ccd*A | Invitrogen, Karlsruhe, Germany |

## 2.1.5 Software

Software related to specific methods is mentioned in the corresponding chapters. Additionally used software is listed in Table 2.6.

Table 2.6: **Software**

| Software | Source |
|---|---|
| Gimp | http://www.gimp.org/ |
| Inkscape | http://www.inkscape.org/ |
| VectorNTI | Invitrogen, Karlsruhe, Germany |
| Emacs | http://www.gnu.org/software/emacs/ |
| LaTeX 2$_\varepsilon$ | http://www.latex-project.org/ |

## 2.2 General methods

### 2.2.1 Growth and storage of *E. coli*

*E. coli* cells were grown in LB (lysogeny broth, also known as Luria broth or Luria-Bertani broth) medium at 37 °C on a shaker at 250 rpm (see Table 2.7). When necessary, antibiotics were added to the medium at the indicated concentrations. For storage, 1 ml overnight culture was mixed with the same amount of glycerol and placed at $-78$ °C.

Table 2.7: **Media and antibiotics for *E. coli***

| | | |
|---|---|---|
| LB medium | 10 g bacto tryptone | 1 % (w/v) |
| | 5 g yeast extract | 0.5 % (w/v) |
| | 10 g NaCl | 1 % (w/v) |
| | ad 1 l $H_2O_{bidest}$ | |
| | autoclave | |
| | for agar plates 15 g agar were added to 1 l of medium | |
| | | |
| Antibiotics | ampicillin (100) | 100 µg/ml |
| | ampicillin (300) | 300 µg/ml |
| | kanamycin | 25 µg/ml |
| | chloramphenicol | 50 µg/ml |

## 2.2.2 Growth and storage of *H. salinarum*

*H. salinarum* cells were either grown in complete medium (Halomedium, HM) (Oesterhelt and Krippahl, 1983) or in synthetic medium (SM) (see Table 2.8). Cultures were grown at 37 °C or 40 °C at 100-250 rpm on a shaker. For storage, cells in HM were hermetically sealed and left in the dark at room temperature.

## 2.2.3 Separation of DNA fragments by agarose gel electrophoresis

DNA fragments were separated by electrophoresis in an 1 % agarose gel in 1 x TAE buffer (40 mM Tris/acetate, 2 mM EDTA, pH 8.3) at 80-120 V. For staining of DNA, the gels contained ethidium bromide at a concentration of 0.5 µg/ml. Gels were photographed on a UV light table with a CCD camera.

## 2.2.4 Purification of DNA fragments

After PCR reactions or restriction digests, DNA fragments were either purified directly from the reaction batch or they were extracted from an agarose gel after electrophoresis. In both cases a QIAquick™ Gel Extraction kit was used according to manufacturer's instructions, either applying the PCR purification or the gel extraction protocol. Elution was done in 30 µl or 50 µl buffer EB.

## 2.2.5 Analytical and preparative restriction digestion

0.5 to 1 µg DNA for an analytical restriction digest and 5 to 10 µg DNA for an preparative restriction digest were incubated with 3 to 5 units per µg DNA of an appropriate restriction endonuclease for 1 to 3 h. Reaction buffer and temperature were chosen according to manufacturer's recommendations. After digestion the DNA was subjected to agarose gel electrophoresis for analysis of fragment size or extraction

Table 2.8: **Media and antibiotics for *H. salinarum***

| | | |
|---|---|---|
| Halomedium | 250 g NaCl | 4.3 M |
| | 20 g $MgSO_4 \cdot 7H_2O$ | 80 mM |
| | 2 g KCl | 27 mM |
| | 3 g $Na_3$citrate $\cdot 2H_2O$ | 10 mM |
| | 10 g bacto peptone | |
| | ad 1 l $H_2O_{bidest}$, pH adjusted to 7.2, autoclaved | |
| | for agar plates 15 g agar were added to 1 l of medium | |
| | | |
| Synthetic medium | 235 g NaCl | 4 M |
| | 10 g $MgSO_4 \cdot 7H_2O$ | 50 mM |
| | 2.44 g KCl | 27 mM |
| | 0.1 g $KNO_3$ | 1.7 mM |
| | 178 mg $Na_2HPO_4 \cdot 2H_2O$ | 1 mM |
| | 0.05 mg $CuSO_4 \cdot 5H_2O$ | |
| | 2.3 mg $MgSO_4 \cdot 4H_2O$ | |
| | 0.3 mg $MnSO_4 \cdot H_2O$ | |
| | 0.44 mg $ZnSO_4 \cdot 7H_2O$ | |
| | 20 mg ascorbic acid | |
| | 2.41 mg $NaMoO_4 \cdot 2H_2O$ | |
| | 2.37 mg $CoCl_2 \cdot 6H_2O$ | |
| | 0.4 g L-arginine | |
| | 0.44 g L-isoleucine | |
| | 0.8 g L-leucine | |
| | 0.224 g L-lysine $\cdot H_2O$ | |
| | 0.2 g L-methionine | |
| | 0.5 g L-threonine | |
| | 0.25 g L-valine | |
| | 0.601 g L-serine | |
| | 0.1 g L-phenylalanine | |
| | 0.2 g L-tyrosine | |
| | 50 µg thiamin | |
| | 50 µg folic acid | |
| | 5 µg biotin | |
| | 7.5 ml glycerol | |
| | ad 1 l $H_2O_{bidest}$, pH adjusted to 7.2 | |
| | | |
| Antibiotics | novobiocin | 0.15 µg/ml |
| | mevinolin | 25 µM |

of certain fragments, or the DNA was purified using the QIAquick™ PCR Purification Protocol.

## 2.2.6 Dephosphorylation of linearised plasmids

If plasmids for a ligation reaction were cut with only one restriction enzyme or with two enzymes producing compatible ends, the phosphate residue at the 5'-end was

enzymatically removed to prevent recirculation of the plasmid by self-ligation. 5-10 μg linearised plasmid were incubated with 20-30 units calf intestine phosphatase (CIP) for 1 h at 37 °C. The reaction was set up according to manufacturer's instructions. After dephosphorylation, the DNA was purified with the QIAquick™ PCR purification kit.

## 2.2.7 Ligation

For ligation of a DNA fragment in a linearised plasmid T4 DNA ligase was used. 50-100 ng plasmid and an appropriate amount of the DNA fragment were mixed in a molar ratio of 1:1 to 1:3 and incubated with 2-4 units T4 DNA ligase for 1 h at room temperature. Alternatively, ligation reactions were performed with the Rapid DNA Ligation Kit (Fermentas) according to manufacturer's instructions.

## 2.2.8 In-Fusion™ cloning

In-Fusion™ (BD Biosciences) cloning allows the directional placement of genes into cloning vectors at almost any desired restriction site without introducing additional bases to insert or vector.

15 bp extensions were added to the PCR primers that match the ends of the linearised target vector. The In-Fusion™ enzyme fuses these homologous regions with the corresponding ends of the linearised vector by converting double-stranded extensions into single-stranded DNA. 1 μl of vector and 100-200 ng PCR product were mixed and $H_2O$ added to a total volume of 10 μl. The mixture was pipetted to an In-Fusion dry-down reaction tube, mixed by pipetting up and down several times, and then incubated at 42 °C for 30 min. After incubation, the reaction mixture was diluted with 40 μl TE buffer (10 mM Tris, 1 mM EDTA, pH 8) and 2.5 μl of the diluted mixture transformed to *E. coli.*

## 2.2.9 Gateway™ cloning

The Gateway® technology is a cloning system based on the site-specific recombination properties of bacteriophage lambda (Landy, 1989). It provides a fast way to clone DNA sequences into multiple different vectors (Hartley *et al.*, 2000).

In this study genes were first cloned into the entry vector pENTR™/D-TOPO via directional TOPO cloning. The resulting entry clones were then used to move the cloned genes to different destination vectors (e. g. pMS3-6).

### Directional TOPO cloning

Genes were cloned with the pENTR™/D-TOPO® Cloning Kit, which directionally clones a blunt-end PCR product into an entry vector for the Gateway system. TOPO Cloning is based on Topoisomerase I from *Vaccinia* virus (Shuman, 1991). To achieve

directional cloning, four bases (CACC) are added to the 5' end of the PCR product. A single-stranded overhang in the cloning vector (GTGG) invades the 5' end of the PCR product and anneals to the added bases. This stabilises the PCR product in the correct orientation resulting in a high probability for proper cloning.

The reaction was set up as follows:

| | |
|---|---:|
| PCR product | 1 µl |
| Salt Solution | 0.5 µl |
| H$_2$O | 1 µl |
| pENTR/D-TOPO vector | 0.5 µl |
| | 3 µl |

After 20 min of incubation at RT the tube was placed on ice and 2 µl of the mixture were transformed to *E. coli*. Selection of transformants was done by growth on LB agar plates containing 30 µg/ml kanamycin.

**Lambda recombination (LR)**

LR recombination was used to move cloned genes from the Gateway entry vector to one or more destination vectors.

Lambda recombination occurs between specific attachment (*att*) sites, which serve as the binding site for the recombination proteins. Upon lambda integration by the lysogenic pathway, recombination occurs between *att*B on the *E. coli* chromosome and *att*P on the lambda chromosome (BP recombination). The resulting prophage is bounded by hybrid *att*L and *att*R sites. Excisive recombination between *att*L and *att*R by the lytic pathway regenerates the original *att*B and *att*P sites (LR recombination). In the Gateway system, the wild-type lambda *att* sites have been modified to improve the efficiency of the recombination reactions and ensure specificity to maintain orientation and reading frame.

LR recombination reactions were catalyzed by LR Clonase™II enzyme mix, that contains the bacteriophage lambda Integrase (Int) and Excisionase (Xis), and the *E. coli* Integration Host Factor (IHF) protein.

The reaction was set up as follows:

| | |
|---|---:|
| Entry clone plasmid DNA | 0.5 µl |
| Destination vector | 0.5 µl |
| TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) | 3 µl |
| LR Clonase™II | 1 µl |
| | 5 µl |

The reaction was incubated for 1-2 h at 25 °C. Then 0.5 µl Proteinase K solution (2 µg/µl) were added and the mixture incubated at 37 °C for 10 min. This enzyme digests the recombination proteins and thereby improves transformation efficiency. *E. coli* cells were transformed with 1 µl of the reaction mixture.

## 2.2.10 Transformation of *E. coli*

Chemical competent *E. coli* cells for transformation were prepared by the method of Inoue (Inoue *et al.*, 1990).

**Preparation of competent *E. coli* cells**

Table 2.9: **Solutions for *E. coli* transformation.**

| | | |
|---|---|---|
| Inoue transformation buffer (TB) | 10.88 g $MnCl_2 \cdot 4H_2O$ | 55 mM |
| | 2.2 g $CaCl_2 \cdot 2H_2O$ | 15 mM |
| | 18.65 g KCl | 250 mM |
| | 20 ml 0.5 M PIPES (pH 6.7) | 10 mM |
| | ad 1 l $H_2O_{bidest}$ | |
| | | |
| SOB medium | 20 g tryptone | |
| | 5 g yeast extract | |
| | 0.5 g NaCl | |
| | 10 ml 250 mM KCl | |
| | ad 1 l $H_2O_{bidest}$ | |
| | adjusted pH to 7.0 with 5 M NaOH | |
| | before use 5 ml 2 M $MgCl_2$ added | |

*E. coli* cells were spread on a LB agar plate and incubated overnight at 37 °C. A single colony was picked, transferred to 25 ml SOB and incubated at 37 °C on a shaker at 250 rpm for 6-8 h. 10 ml of this starter culture were used to inoculate 250 ml SOB, and this main culture was then incubated at 20 °C with moderate shaking (120 rpm). When the culture reached an $OD_{600}$ of 0.55 it was transferred to an ice-water bath for 10 min and the cells harvested by centrifugation at 2500 x g for 10 min at 4 °C. The supernatant was discarded, the cells were resuspended in 80 ml ice-cold TB and centrifuged again. After discarding the supernatant the cells were resuspended in 20 ml TB, and 1.5 ml DMSO added. The bacterial suspension was mixed and stored on ice for 10 min.

50 µl aliquots of the suspension were dispensed into prechilled 1.5 ml reaction tubes and frozen in a bath of liquid nitrogen. The cells were stored at −78 °C until needed.

**Transformation**

The required amount of tubes was removed from the freezer and the cells thawed quickly by holding the tubes in the palm of the hand. When the cells started thawing the tubes were transfered to an ice bath immediately and stored on ice for 10 min.

The transforming DNA (2.5 µl of a ligation reaction, 2 µl of a TOPO cloning reaction, or 1 µl of a LR recombination reaction) was added, and the tubes gently swirled several times and placed on ice for 30 min. Thereafter the tubes were placed in a 42 °C water bath for 90 s and transfered back to ice for 1-2 min. 800 µl LB medium were

added to each tube and the cultures incubated in a shaking incubator (250 rpm, 37 °C) for 45 min. Up to 200 µl of the cultures were spread on LB agar plates containing the appropriate antibiotic. If a low yield was expected, the whole culture was centrifuged (1 min, 14000 rpm in a tabletop centrifuge) and the pellet resuspended in 150 µl LB and completely spread on LB agar plates. The plates were incubated overnight at 37 °C.

Alternatively, One Shot® Mach1™-T1$^R$ Chemically Competent *E. coli* were used for transformation according to manufacturer's instructions.

## 2.2.11 Transformation of *H. salinarum*

Transformation of *H. salinarum* was performed according to Cline *et al.* (1989) with some modifications.

Table 2.10: **Solutions for *H. salinarum* transformation**

| | |
|---|---|
| Spheroplasting solution (SPS) | 2 M NaCl |
| | 27 mM KCl |
| | 50 mM Tris-HCl (pH 8.75) |
| | 15 % (w/v) sucrose |
| | |
| 60 % PEG$_{600}$ | 60 % (v/v) PEG$_{600}$ |
| | 40 % (v/v) SPS |
| | freshly prepared |

Halobacterial cells were grown in 35 ml of complete medium at 37 °C on a shaker at 250 rpm to an OD$_{600}$ of 0.4-0.8. 1 ml of this culture was used to inoculate a fresh culture which was grown under the same conditions. When this culture reached an OD$_{600}$ of 0.5 to 0.8, 1.5 ml for each transformation reaction were transfered to a microfuge tube and centrifuged for 2 min at 10000 x g at room temperature. The supernatant was removed completely and the cells were resuspended in 150 µl SPS. A mixture of 15 µl 0.5 M EDTA (pH 8.0) and 15 µl SPS was added and the cells incubated for 10 min at RT. After this, a mixture of 5 µl plasmid DNA (circa 1 µg) and 5 µl SPS was added followed by 5 min incubation at RT. 190 µl 60 % PEG$_{600}$ were pipetted to the cells and immediately mixed by inverting the tube 3-4 times. After incubating for $20 - 30$ min at RT, 1 ml of complete medium + 15 % (w/v) sucrose was added and the tubes centrifuged for 2 min at 10000 x g at RT. The supernatant was discarded and the cells resuspended in 1 ml of complete medium + 15 % sucrose. The cultures were incubated overnight at 37 °C on a shaker (250 rpm) to allow the cells to recover. On the next day the cells were pelleted (10000 rpm, 2 min, RT), resuspended in 150 µl complete medium and spread on plates containing 0.15 µg/ml novobiocin or 10 µg/ml mevinolin and 80-100 µg/ml X-Gal. The plates were incubated at 40 °C until single colonies were visible (8-12 days).

## 2.2.12 Polymerase chain reaction (PCR)

The specific amplification of DNA stretches was performed by PCR (Saiki *et al.*, 1988). To guarantee a low error rate, all reactions were performed with Phusion™ DNA polymerase, that combines proofreading activity and a high processing speed. A typical PCR reaction was set up as follows:

**Reaction mixture:**

| | |
|---|---|
| Template | 1 µl |
| Primer, fo (10 pmol/µl) | 2.5 µl |
| Primer, re (10 pmol/µl) | 2.5 µl |
| dNTP-Mix (10 mM each) | 1 µl |
| DMSO | 1.5 µl |
| Phusion DNA-polymerase | 0.5 µl |
| Buffer (5 x) | 10 µl |
| H$_2$O | 31 µl |
| | 50 µl |

**Cycler program:**

| | | |
|---|---|---|
| 98 °C | 1 min | |
| 98 °C | 10 s | |
| 55 − 72 °C | 20 s | 30 cycles |
| 72 °C | 30 s / 1 kb | |
| 72 °C | 10 min | |
| 4 °C | ∞ | |

As annealing temperature the calculated annealing temperature of the lower melting primer (http://www.metabion.com/biocalc/) was chosen. The template was 50-500 ng genomic DNA or 10-100 ng plasmid DNA.

## 2.2.13 DNA sequencing

Sequencing of plasmid DNA or PCR products was performed with the chain terminator method of Sanger (Sanger *et al.*, 1977) using fluorescence-labeled didesoxynucleotides. The reaction was done with the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit vs 3.1 (Applied Biosystems), which contains buffer, dNTPs and labeled ddNTPs, and a thermostable polymerase. For sequencing, 0.5-1 µg plasmid DNA or 0.1-0.5 µg PCR product were used. A typical sequencing reaction was set up as follows:

**Reaction mixture:**

| | |
|---|---|
| Template | 2 µl |
| Primer (10 pmol/µl) | 1.5 µl |
| BigDye | 1.5 µl |
| BigDye Puffer | 1 µl |
| Betaine (5 M) | 2 µl |
| H$_2$O | 2 µl |
| | 10 µl |

**Cycler program:**

| | | |
|---|---|---|
| 94 °C | 1 min | |
| 94 °C | 30 s | |
| 60 °C | 4 min | 25 cycles |
| 4 °C | ∞ | |

After cycling, the samples were sent to the institute's DNA sequencing service where they were analysed on an ABI 3730 sequencer. Sequences were assembled and checked with the program VectorNTI (Invitrogen).

## 2.2.14 Isolation of plasmid DNA

For the preparation of plasmid DNA from *E. coli*, cells were grown in 3 ml LB medium containing the appropriate antibiotic at 37 °C overnight on a shaker (250 rpm). Isolation of plasmid DNA was done from 2 ml of the culture using the QIAprep 8 Miniprep Kit or the QIAprep spin Miniprep Kit (Qiagen, Hilden) according to manufacturer's instructions. DNA was finally eluted in 100 µl (QIAprep 8) or 50 µl (QIAprep spin) buffer EB.

## 2.2.15 Protein precipitation with TCA

Proteins were precipitated to remove interfering substances like salt and ethylene glycol, and to raise protein concentration. The protein solution was adjusted with the equal amount of 20 % (w/v) trichloroacetic acid (TCA) to a final concentration of 10 % TCA and placed on ice for 30 min. After this, the mixture was centrifuged for 30 min at 14000 rpm and 4 °C and the supernatant removed. The pellet was washed with 50 % (v/v) ice-cold acetone and centrifuged again at 14000 rpm, 4 °C for 15 min. Washing was repeated until no salt crystals were visible (normally 1-2 times).

## 2.2.16 SDS PAGE

SDS polyacrylamide gel electrophoresis (SDS PAGE) was done with 4-12 % Bis Tris gels from the NuPAGE® system (Invitrogen).

Protein samples were heated in LDS sample buffer containing sample reducing agent for 10 min at 70 °C and applied to the gels. Electrophoretic separation was performed in 1 x MES running buffer at a constant voltage of 180-200 V. As molecular weight standard the PageRuler™ Protein Ladder (Fermentas) or the Prestained Protein Marker (broad range) (NEB) was used.

## 2.2.17 Coomassie staining of protein gels

Gels were stained in Staining Solution (Table 2.11) for 1 h or overnight with gentle shaking. After this the Staining Solution was replaced by Destain I and the gel slowly shaken for 30 min. Then Destain I was replaced by Destain II followed by an additional hour of slow shaking. After destaining the gel was transferred to Storing Solution.

## 2.2.18 Silver staining of protein gels

Silver staining of protein gels was done by the method of Blum *et al.* (1987) with minor modifications. This method does not crosslink the proteins in the gel, which is an important requirement to allow mass spectrometric identification of proteins after staining. The silver staining protocol is given in Table 2.12.

Table 2.11: **Coomassie staining solutions**

| | |
|---|---|
| Staining Solution | 40 % (v/v) ethanol |
| | 10 % (v/v) acetic Acid |
| | 0.1 % (w/v) Coomassie Brilliant Blue R250 |
| | stirred overnight and filtered |
| | |
| Destain I | 40 % (v/v) ethanol |
| | 10 % (v/v) acetic acid |
| | |
| Destain II | 10 % (v/v) acetic acid |
| | |
| Storing Solution | 1 % (v/v) acetic acid |

Table 2.12: **Silver staining protocol**

| Step | Time | Solution |
|---|---|---|
| Fix | 2 x 30 min | 50 % (v/v) methanol, 12 % (v/v) acetic acid |
| Wash | 3 x 20 min | 50 % (v/v) ethanol |
| Sensitise | 1 min | 200 mg/l $Na_2S_2O_3$ |
| Wash | 2 x 1 min | $H_2O$ |
| Stain | 20 min | 2 g/l $AgNO_3$, 1 ml/l formaldehyde ($\geq 37\%$) |
| Wash | 20 s | $H_2O$ |
| Develop | until sufficient | 60 g/l $Na_2CO_3$, 5 mg/l $Na_2S_2O_3$, 0.75 ml/l formaldehyde ($\geq 37\%$) |
| Stop | 10 min | 12 % acetic acid |
| Store | | 1 % acetic acid |

## 2.2.19 Western blot

Western blotting is the transfer of separated proteins from a gel to the surface of a membrane in an electric field. The proteins are bound and immobilised on the membrane and can be detected subsequently by using antibodies.

**Membrane transfer**

Proteins were transferred to polyvinylidene difluoride (PVDF) membranes using the XCell II™ Blot Module (Invitrogen), a semi-wet transfer unit, according to manufacturer's instructions.

Before use, blotting pads and filter paper was soaked in transfer buffer. The PVDF membrane was pre-wetted in methanol for 30 s and then placed into transfer buffer. Two blotting pads were placed in the cathode core of the blot module and covered with a filter paper. The gel was sprinkled with blotting buffer and laid on the filter paper. The membrane was placed on the gel and covered with another filter paper. 3-4 blotting pads and the anode core were added and the blot module slided into the XCell II™ Mini-Cell. The blot module was filled with transfer buffer until the gel/membrane sandwich was covered. To dissipate heat the outer buffer chamber was

Table 2.13: **Buffers for western blot**

| | |
|---|---|
| Transfer buffer | 25 mM tris |
| | 192 mM glycine |
| | 20 % methanol |
| | adjusted pH to 8.6 |
| | |
| TBS buffer | 10 mM tris-HCl (pH7.5) |
| | 150 mM NaCl |
| | |
| TBS-Tween/Triton buffer | 20 mM tris-HCl (pH 7.5) |
| | 500 mM NaCl |
| | 0.05 % (v/v) tween 20 |
| | 0.2 % (v/v) triton X-100 |
| | |
| Blocking buffer | 0.1 g Blocking Reagent (Qiagen) |
| | 20 ml 1 x Blocking Reagent Buffer (heated to 70 °C) |
| | 0.1 % (v/v) tween-20 |

filled with deionised $H_2O$. The transfer was performed for 1-2 h at 25 V.

**Immunodetection**

His-tagged proteins were detected using the Penta · His™ HRP Conjugate Kit (Qiagen). The Anti · His HRP conjugate consists of a mouse monoclonal IgG1 Anti · His Antibody coupled to horseradish peroxidase. It can be used for direct detection of His-tagged proteins by chemiluminescent methods.

The membrane was washed twice for 10 min with TBS buffer at room temperature and incubated overnight in blocking buffer at 4 °C. After this, it was washed twice for 10 min each time in TBS-Tween/Triton buffer at room temperature and once for 10 min in TBS. Now it was incubated in Anti · His HRP Conjugate solution (1 : 3000 in blocking buffer) at room temperature for 1 h. It was washed again twice for 10 min in TBS-Tween/Triton buffer and once for 10 min in TBS at room temperature. The chemiluminescence detection reaction was performed with Lumi-Light Western Blot Substrate (Roche) according to manufacturer's recommendations.

## 2.2.20 Preparation of genomic DNA

Genomic DNA from *H. salinarum* for PCR and Southern Blot analysis was prepared by water lysis of the cells without further purification. 1 ml of a fresh culture were pelleted by centrifugation (14000 rpm, 2 min, RT), and the medium removed. The cells were lysed by adding 300 μl deionised water and pipetting up and down. The lysate was heated at 70 °C for 10 min to inactivate nucleases and stored at 4 °C.

## 2.3 Materials and methods for yeast two-hybrid screening

Yeast two-hybrid screening was done with the tools developed by James *et al.* (1996). Bait and prey proteins were cloned into the vector pGBDc1 or pGADc1, respectively. Baits and preys were transformed pairwise into the *S. cerevisiae* strain PJ69-4A. This strain contains three different reporter genes, each under the control of a different inducible promoter: the HIS3 gene under control of the GAL1 promoter, the ADE2 gene under control of the GAL2 promoter, and the *lacZ* gene under control of the GAL7 promoter. Protein interaction assays were performed by growth on SC minus His and growth on SC minus Ade.

### 2.3.1 Growth and storage of *S. cerevisiae*

Liquid cultures were inoculated from fresh plates and grown in YPD at 30 °C on a shaker (200 rpm). For storage, glycerol was added to a stationary culture (final concentration 20 %) and cells stored at −78 °C. The composition of the used culture media is given in Table 2.14.

### 2.3.2 Construction of two-hybrid expression plasmids

Genes to analyse were amplified from *H. salinarum* or *E. coli* genomic DNA by PCR using the primers shown in Table 2.15. All oligonucleotides were synthesised by Metabion (Martinsried, Germany).

PCR products were purified and digested with the restriction enzymes indicated in the primer names. The plasmids pGAD-C1 and pGBD-C1 were digested with the same pairs of enzymes, and the PCR products ligated into the linearised plasmids. In the case of Eco_cheA the vector was dephosphorylated before ligation since BamHI and BglII produce compatible ends. 3 μl of ligation mixture were transformed to *E. coli* DH5α and clones selected on LB agar plates containing ampicillin. Plasmid DNA of single clones was isolated and the insert size (and for Eco_cheA insert orientation) verified by restriction digestion. All plasmids used in this study are listed in Table 2.16.

### 2.3.3 Transformation of yeast

Yeast cells were transformed pairwise with bait and prey expression plasmids with the method described in Knop *et al.* (1999).

#### Preparation of competent cells

50 ml YPD were inoculated with 5 μl yeast cells from a fresh overnight culture and the cells grown at 30 °C to an $OD_{600}$ of 0.5 to 0.7. Cells were harvested by centrifu-

Table 2.14: **Media for *S. cerevisiae***

| YPD | 10 g Bacto-yeast extract | 1 % |
| | 20 g Bacto-peptone | 2 % |
| | 20 g dextrose | 2 % |
| | ad 1 l $H_2O_{bidest}$ | |
| | | |
| SC | 1.7 g yeast nitrogen base (w/o amino acids) | 0.67 % |
| | 20 g dextrose | 2 % |
| | 5 g ammonium sulfate | 0.5 % |
| | 2 g drop-out mix | 0.2 % |
| | ad 1 l $H_2O_{bidest}$ | |
| | | |
| Drop-out mix | 1 g adenine hemisulfate* | |
| | 1 g arginine · HCl | |
| | 5 g aspartic acid | |
| | 5 g glutamic acid | |
| | 1 g histidine · HCl* | |
| | 4 g isoleucine | |
| | 2 g leucine* | |
| | 3 g lysine · HCl | |
| | 1 g methionine | |
| | 2.5 g phenylalanine | |
| | 2 g serine | |
| | 2 g threonine | |
| | 1.5 g tryptophan* | |
| | 3 g tyrosine | |
| | 9 g valine | |
| | 1 g uracil | |

The appropriate components to prepare synthetic complete drop-out media were omitted (marked by asterisks). All media were autoclaved. For solid media, 2 % agar were added before autoclaving.

Table 2.15: **Primer for two-hybrid plasmids**

| Primer | Sequence |
|---|---|
| Hsa_cheA_EcoRI_fo | AGTCGAATTCATGGACGACTACCTCGAAGC |
| Hsa_cheA_PstI_re | AGTCCTGCAGTTACAGCGTAGCCACGTCC |
| Hsa_cheW1_EcoRI_fo | AGTCGAATTCATGAACCTCGAAGACGCCGAC |
| Hsa_cheW1_BamHI_re | AGTCGGATCCTCAGACCTGATTGTCGATGG |
| Hsa_dodecin_EcoRI_fo | AGTCGAATTCATGGTCTTCAAGAAGGTCC |
| Hsa_dodecin_BamHI_re | AGTCGGATCCCTACTGGGAGCCGTCGAG |
| Hsa_MDH_EcoRI_fo | AGTCGAATTCATGGGACTCGATGATGACG |
| Hsa_MDH_BamHI_re | AGTCGGATCCTCAGTCCTCGCGCTGCTG |
| Hsa_nuoA_EcoRI_fo | AGTCGAATTCATGAATCCATGGATCGCCATC |
| Hsa_nuoA_BamHI_re | AGTCGGATCCCTACTCATGTGTGGTCCTTTG |
| Hsa_nuoB_EcoRI_fo | AGTCGAATTCATGAGTAGTGACCAACCAAG |
| Hsa_nuoB_BamHI_re | AGTCGGATCCTCATGGTGAATCAGCCCAG |
| Eco_cheA_BamHI_fo | AGTCGGATCCGTGAGCATGGATATAAGC |
| Eco_cheA_BglII_re | AGTCAGATCTTCAGGCGGCGGTGTTCGC |
| Eco_cheW_EcoRI_fo | AGTCGAATTCATGACCGGTATGACGAATG |
| Eco_cheW_BamHI_re | AGTCGGATCCTTACGCCACTTCTGACGC |

Table 2.16: **Plasmids and strains for Y2H analysis**

| Plasmid | Description | Source |
|---|---|---|
| pGAD-C1 | pGADc1::*cheA* (Eco) | James *et al.* (1996) |
| pGBD-C1 | pGADc1::*cheA* (Eco) | James *et al.* (1996) |
| pMS31 | pGADc1::*cheA* (Eco) | This study |
| pMS32 | pGBDc1::*cheA* (Eco) | This study |
| pMS33 | pGADc1::*cheW* (Eco) | This study |
| pMS34 | pGBDc1::*cheW* (Eco) | This study |
| pMS35 | pGADc1::*cheA* (Hsa) | This study |
| pMS36 | pGADc1::*cheW1* (Hsa) | This study |
| pMS37 | pGBDc1::*cheW1* (Hsa) | This study |
| pMS38 | pGADc1::*dodecin* (Hsa) | This study |
| pMS39 | pGBDc1::*dodecin* (Hsa) | This study |
| pMS40 | pGADc1::*mdh* (Hsa) | This study |
| pMS41 | pGBDc1::*mdh* (Hsa) | This study |
| pMS42 | pGADc1::*nuoA* (Hsa) | This study |
| pMS43 | pGBDc1::*nuoA* (Hsa) | This study |
| pMS44 | pGADc1::*nuoB* (Hsa) | This study |
| pMS45 | pGBDc1::*nuoB* (Hsa) | This study |

| Strain | Description | Source |
|---|---|---|
| PJ69-4A | *MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ* | James *et al.* (1996) |

gation at $500 \times g$ for 5 min at room temperature, washed once with 0.1-0.5 volumes of sterile water and once with 0.1 volumes SORB (Table 2.17). The SORB was removed completely by aspiration and the cells were finally resuspended in 360 µl SORB. 40 µl of carrier DNA (salmon sperm DNA, denatured at 95 °C for 10 min and put on ice) were added and the cells aliquoted into 50 µl portions and stored at $-78$ °C.

**Transformation**

Table 2.17: **Solutions for yeast transformation**

| SORB | 100 mM lithium acetate | PEG | 100 mM lithium acetate |
|---|---|---|---|
| | 10 mM Tris-HCl pH 8.0 | | 10 mM Tris-HCl pH 8.0 |
| | 1 mM EDTA/NaOH pH 8.0 | | 1 mM EDTA/NaOH pH 8.0 |
| | 1 M sorbitol | | 40 % (w/v) Polyethyleneglycol 3350 |
| | adjust pH to 8.0 with acetic acid | | adjust pH to 8.0 with acetic acid |
| | filter sterilise, store at 4 °C | | filter sterilise, store at 4 °C |

Competent cells were thawed at room temperature. 1 µl of each plasmid DNA and 90 µl PEG (Table 2.17) were added to 15 µl of cells, the suspension mixed by 2 min vortexing and incubated at room temperature for 30 min. 12 µl DMSO were added and the cells placed in a water bath at 42 °C for 15 min. The cells were sedimented (3 min, $500 \times g$, RT), the supernatant removed, and the cells resuspended in 100 µl $H_2O$. Transformation reactions were plated on SC minus Trp, Leu and incubated for 2-3 days at 30 °C.

### 2.3.4 Protein interaction assay

To identify interactions between bait and prey proteins, transformants were transferred to plates containing selective media that allow growth only in the case of reporter gene activation. To reduce the risk of false negatives due to clones containing corrupted plasmids, two colonies from each transformation plate were merged and streaked together. The plates were incubated at 30 °C for 1-2 days.

First screens were performed with selection on SC minus His, Leu, Trp. To reduce the growth of false positives, rescreening was done with selection on SC minus His, Trp, Leu in the presence on 5 mM, 15 mM, and 30 mM 3-aminotriazol, or with selection on SC minus Ade, Leu, Trp.

## 2.4 Materials and methods for AP-MS of halobacterial protein complexes

### 2.4.1 Construction of vectors

Vectors for the expression of CBD fusion proteins in *H. salinarum* are based on the pVT shuttle vector (Tarasov *et al.*, 2008). This vector contains an *E. coli* replication origin as well as ampicillin and novobiocin resistance genes for selection of transformants in *E. coli* and *H. salinarum*, respectively. Additionally, the vector contains the *bgaH* gene coding for an halophilic β-galactosidase to allow for blue/white screening of transformants in *H. salinarum*.

**Vectors for In Fusion™ cloning**

For In Fusion™ cloning, the vectors pMS1 and pMS2 were constructed that can be linearised with NsiI at the position where the gene should be inserted.

In the first step, a NsiI site in the parent plasmid pVT was removed and the first His-tag and a new NsiI site added. The region between the NsiI site and the XbaI site was cut out and replaced by the same region that was PCR-amplified from pVT using the

Table 2.18: **Primer for establishing AP-MS.**

| Primer | Sequence |
|---|---|
| PrR16CBD-fo | TATGGATCCGTGGATAAAAC |
| PrR16CBD-re | CATCTAGATCTCCCATATGGCTAGCTACTACACTGCCACCGGGTTC |
| ANX-PstI-fo | ATGCCTGCAGACCGACGCGAAAGCCGAC |
| ANX-XNH-re | AGTCTCTAGATGCATGGTGGTGATGGTGGTGGGATCCCGGTTCGAGGTTC |
| Bopterm-XbaI-fo | AGTCTCTAGAACCGCGTCGATCGAAGGTC |
| Bopterm-HindIII-re | GAAGAAGCTTGCATGCCTGC |
| CBD-L-NsiI-fo | AGTCATATGCATCGGCGCCGTCGAGGAGCGCatgGCAAATACACCGGTATC |
| CBD-HIS-XbaI-re | AGTCTCTAGAGGTGGTGGTGGTGGTGATGTACTACACTGCCACCGGGTTC |
| CBD-L-BamHI-re | AGTCGGATCCGCGCTCCTCGACGGCGCCGATTACTACACTGCCACCGGGTTC |
| VC-NsiI-fo | AGATGCATGCATCAACAAGTTTGTACAAAAAAGCTG |
| VC-NsiI-re | GATGCGATGCATACCACTTTGTACAAGAAAGC |
| VCBspHI_for | TCTATCTCATGATCGAACCCGGTGGCAGTGTAGTAATC |
| VCXbaI_rev | TAGAGTTCTAGAGGTGGTGGTGGTGGTGATGTGATACCGGTGTATTTGCCATG |
| cheA-IF-fo | ACCACCATCACCACCATATGGACGACTACCTCGAAG |
| cheA-IF-re | CGACGGCGCCGATGCACAGCGTAGCCACGTCCAG |
| cheW1-IF-fo | ACCACCATCACCACCATATGAACCTCGAAGACGC |
| cheW1-IF-re | CGACGGCGCCGATGCAGACCTGATTGTCGATGGTC |
| dod-IF-fo | ACCACCATCACCACCATATGGTCTTCAAGAAGGTCC |
| dod-IF-re | CGACGGCGCCGATGCACTGGGAGCCGTCGAGTTCG |
| rpoK-IF-fo | ACCACCATCACCACCATATGAGCGACTCACAGCATTTC |
| rpoK-IF-re | TCCTCGACGGCGCCGATGCAGGTCCCCTCCCGCCGAAC |
| OE2638F_fo | CACCATGAGCGACTCACAGCATTTCAGC |
| OE2638F_re | GGTCCCCTCCCGCCGAACAGTGAAC |

Table 2.19: **Plasmids and strains for establishing AP-MS.**

| Plasmid | Comment | Source |
|---|---|---|
| pET15 | *E. coli* expression vector | Novagen |
| pET15-CBD | CBD cloned into pET15 | This study |
| pVT | *E. coli* - *H. salinarum* shuttle vector | Tarasov *et al.* (2008) |
| pENTR/D-TOPO | Gateway entry vector | Invitrogen |
| pMS1 | C-terminal CBD-tagging vector, In-Fusion cloning | This study |
| pMS2 | Double CBD-tagging vector, In-Fusion cloning | This study |
| pMS3 | C-terminal CBD-tagging vector, Gateway cloning | This study |
| pMS4 | Double CBD-tagging vector, Gateway cloning | This study |
| pMS5 | SILAC control for pMS3, Gateway cloning | This study |
| pMS6 | SILAC control for pMS4, Gateway cloning | This study |
| pMS51 | pMS1::*cheA* | This study |
| pMS52 | pMS2::*cheA* | This study |
| pMS53 | pMS1::*cheW1* | This study |
| pMS54 | pMS2::*cheW1* | This study |
| pMS55 | pMS1::*rpoK* | This study |
| pMS56 | pMS2::*rpoK* | This study |
| pMS57 | pMS1::*dodecin* | This study |
| pMS58 | pMS1::*dodecin* | This study |
| pMS208 | pENTR::*rpoK* | This study |
| pMS286 | pMS3::*rpoK* | This study |
| pMS295 | pMS4::*rpoK* | This study |
| pMS291 | pMS5::*rpoK* | This study |
| pMS293 | pMS6::*rpoK* | This study |

The SILAC plasmids (pMS3-6) of the Che proteins are listed in Table 2.5.3. The *H. salinarum* R1 strains transformed with the pMS plasmids (pMSxy) were designated in the same manner (MSxy) as the plasmids.

primers ANX-PstI-fo and ANX-XNH-re. The NsiI-produced overhang of the vector is compatible to the PstI-produced overhang of the insert allowing easy ligation and preventing future cleavage with NsiI at this site. Then the transcriptional terminator of the halobacterial *bop* gene, PCR-amplified from the plasmid pHUSbrfus (Besir, 2001) with the primers Bopterm-XbaI-fo and Bopterm-HindIII-re, was inserted between the new plasmid's XbaI and HindIII sites. In the next step, a construct consisting of a linker (IGAVEER, the linker of the two $\beta$-sheets of halobacterial dodecin), CBD, His tag, and a stop codon, obtained by PCR amplification of the CBD from the plasmid pWL-CBD (Ortenberg and Mevarech, 2000) with the primers CBD-L-NsiI-fo and CBD-HIS-XbaI-re, was inserted between the (new) NsiI site and the XbaI site. The resulting plasmid was checked by sequencing of all manipulated regions and designated pMS1.

To construct pMS2, the promoter PrR16 and the CBD were amplified from the plasmid pWL-CBD with the primers PrR16-CBD-fo and CBD-L-BamHI-re, and inserted in the BamHI site of pMS1. Correct orientation of the insert was verified by restriction digestion with NcoI and XbaI, and the resulting plasmid checked by sequencing of the newly added region.

**Vectors for Gateway™ cloning**

The vectors pMS3 and pMS4 for Gateway™ cloning were derived from pMS1 and pMS2, respectively, by insertion of the Gateway Vector conversion cassette, which contains the *att* sites necessary for Gateway cloning, the *ccdB* gene and a chloramphenicol resistance gene. The cassette was PCR-amplified with the primers VC-NsiI-fo and VC-NsiI-re and cloned into the NsiI site. Correct orientation of the insert was verified by restriction digestion with NotI and the resulting plasmids checked by DNA sequencing.

Plasmids pMS5 and pMS6 for background control in SILAC experiments were derived from pMS3 and pMS4, respectively, by removing the CBD(s). Both pMS3 and pMS4 were cut with NcoI and XbaI. For pMS3, this removes part of the vector conversion cassette and the CBD, for pMS4 both CBDs and the whole vector conversion cassette. To reconstitute the Gateway features, the vector conversion cassette was PCR-amplified with the primers VCBspHI_for and VCXbaI_rev. For the missing part in the digested pMS3, the PCR product was digested with NcoI and XbaI, the respective fragment extracted from a gel and ligated into the digested plasmid, gaining the vector pMS5. For pMS4, the PCR product was digested with BspHI (giving an overhang compatible to the NcoI-derived overhang in the plasmid) and XbaI, and ligated into the digested plasmid. This new vector was designated pMS6. Cells transformed with the Gateway destination vectors were grown in LB medium with chloramphenicol, as after growth in LB Amp defective plasmids were prepared.

## 2.4.2 Generation of bait expression and control strains

Bait proteins were cloned into the expression vectors with the appropriate cloning method: In Fusion cloning (2.2.8) for pMS1 and pMS2, and Gateway cloning (2.2.9) for pMS3-6. Plasmids were verified by restriction digestion, and the inserts sequenced after In-Fusion cloning and in the entry vector after Gateway cloning.

Expression plasmids were transformed in *H. salinarum* R1 (2.2.11). Expression of the tagged bait protein was verified by affinity purification. Control strains transformed with pMS5 and pMS6 were checked by western blot with an anti-penta-his antibody (2.2.19).

The background strain for indirect bait fishing, expressing the plain CBD, was generated by transformation of empty pMS4 (without any bait) into *H. salinarum* R1. The expression of the CBD was verified by affinity purification. The site of integration into the genome was not determined.

### 2.4.3 Establishing the affinity purification procedure

#### 2.4.3.1 Purification from *E. coli*

The CBD was PCR-amplified from pWL-CBD using the primers PrR16CBD-fo and PrR16CBD-re. The PCR product was digested with NcoI and NdeI, and the CBD cloned into the respective sites of pET15. The resulting plasmid was verified by sequencing the manipulated regions and transformed to *E. coli* strain BL21.

Purification was done according to the following protocols:

**Protocol Eco1**  A preculture (25 ml LB with ampicillin) was grown overnight on a shaker (150 rpm) at 37 °C. 3 ml preculture were used to inoculate the main culture (100 ml LB Amp), which was grown under identical conditions. When an $OD_{600}$ of 0.6 was reached, the expression was induced by addition of IPTG to a final concentration of 1 mM. After 4 h the cells were pelleted by centrifugation at 4000 x g, 4 °C for 20 min. They were resuspended in 1 ml resuspension buffer (RB: 300 mM KCl, 100 mM NaCl, 2 mM $P_i$, 0.5 mM PMSF, pH 7.5), and sonified (6 x 20 s, output control 5, duty cycle 50 %) on ice-water. The lysate was cleared by centrifugation (14000 rpm, 20 min, 4 °C), and the supernatant transfered to a fresh tube. 50 µl 10 % (w/v) cellulose suspension (fibrous medium, Sigma) in RB were added and the mixture incubated for 1 h in an overhead shaker. The cellulose was spun down (3000 rpm, 3 min, RT), and the pellet washed with 200 µl RB. The cellulose was pelleted as before and the supernatant removed. The washing step was repeated three times. Now the pellet was resuspended in 40 µl LDS sample buffer and the suspension boiled for 5 min at 100 °C. The tube was centrifuged (5000 rpm, 5 min) and the supernatant transfered to a fresh tube. 7.5 µl lysate and supernatant (after cellulose incubation) and 15 µl of the washing and elution (boiling) fractions were loaded on a gel.

**Protocol Eco2**  Expression and lysate preparation were done as described above. A cellulose column was prepared by pipetting 350 µl cellulose suspension (10 % (w/v) Sigma fibrous medium in RB) into a Mobicol empty spin column. The column was centrifuged (300 x g, 1 min, RT), washed with 500 µl RB to remove fines, and centrifuged again.

600 µl lysate were applied to the column and the cellulose resuspended. After 1 min the column was centrifuged (2000 x g, 1 min), and the flowthrough discarded. This step was repeated with the rest of the lysate. The column was washed five times with 500 µl CFE and centrifuged as before. For elution, 100 µl ethylene glycol were added, the cellulose resuspended, and after 1 min incubation the column was centrifuged as before. The eluate was collected in a fresh reaction tube. The elution was repeated twice. 15 µl of the washing and elution fractions were loaded on a gel.

**Protocol Eco3**  Essentially the same as protocol 2. The differences were the use of microcrystalline cellulose (Avicel PH-101, Fluka) as matrix, and the column was centrifuged only with 300 x g for 20 s in all centrifugation steps.

### 2.4.3.2 Purification from *H. salinarum*

**Protocol Hsa1**  The bait expression strain was precultured in 35 ml Halomedium containing 0.15 μg/ml novobiocin at 37 °C on a shaker (150 rpm) until an $OD_{600}$ of 0.5-1.0 was reached. 1 ml of this preculture was used to inoculate 100 ml Halomedium. The main culture was incubated on a shaker (110 rpm) at 37 °C. When the main culture had reached an $OD_{600}$ of 0.6 to 1.0, cells were harvested by centrifugation (8000 rpm, 15 min, 8 °C) and resuspended in 1-2 ml CFE buffer (3 M KCl, 1 M NaCl, 400 mM $NH_4Cl$, 40 mM $MgCl_2$, 10 mM Tris/HCl, pH 7.5) + 0.01 % Triton X100 + 0.5 mM PMSF. Cells were lysed by sonication on ice water (2 x 20 s, Branson sonifier 250, 3 mm disruptor horn, output level 2, constant) and the lysate cleared by centrifugation at 14000 rpm, 4 °C for 20 min in a tabletop centrifuge.

In the meantime a cellulose column was prepared by pipetting 300 μl cellulose suspension (10 % (w/v) Avicel PH-101 in CFE) into a Mobicol empty spin column. The column was centrifuged (300 x g, 1 min, RT), washed with 600 μl CFE to remove fines, and centrifuged again.

The cleared lysate was applied to the column in 600 μl portions and the cellulose resuspended by vortexing. After 1 min incubation at room temperature the column was centrifuged (300 x g, 1 min, RT) and the flow-through discarded. The cellulose was washed twice with 600 μl CFE. After each washing step the column was centrifuged (300 x g, 1 min, RT) and the flow-through discarded. An additional centrifugation (770 x g, 30 s, RT) was performed after the last washing step to reduce the amount of retained buffer. For elution, 200 μl ethylene glycol were applied to the column, the cellulose resuspended, and the column centrifuged. The eluate was collected in a fresh microfuge tube and elution repeated. Proteins were precipitated with ice-cold acetone. 3 μl lysate and flow-through, 15 μl of the washing fraction, and the total eluted protein were loaded on a gel.

**Protocol Hsa2**  Similar to protocol Hsa1. The main culture had a volume of 1 l, and the resulting cell pellet was resuspended in 3 ml CFE + 0.1 % Triton X100. Elution fractions were pooled.

**Protocol Hsa3**  Similar to protocol Hsa2. Elution was performed twice with 400 μl ethylene glycol.

**Protocol Hsa4**  A main culture of 200 ml Halomedium was grown as described above. At an $OD_{600}$ of around 1.0, cells were harvested by centrifugation (8000 rpm, 15 min,

15 °C) and resuspended in 1.6 ml CFE + 0.1 % Triton X100. Cells were lysed by sonication on ice water (2 x 20 s, Branson sonifier 250, 3 mm disruptor horn, output level 2, constant), and the lysate cleared by centrifugation at 14000 rpm, 18 °C for 20 min in a tabletop centrifuge. Cellulose columns were prepared as described before.

The cleared lysate was applied to the column in 600 μl portions and the cellulose resuspended by pipetting up and down. After 1 min incubation at room temperature the column was centrifuged (300 x g, 1 min, RT) and the flow-through discarded. The cellulose was washed twice with 600 μl CFE + 0.1 % Triton X100 and once with CFE. After each washing step the column was centrifuged (300 x g, 1 min, RT) and the flow-through discarded. An additional centrifugation (770 x g, 30 s, RT) was performed after the last washing step to reduce the amount of retained buffer. For elution, 600 μl ethylene glycol were applied to the column, the cellulose resuspended by pipetting up and down, and the column centrifuged. The eluate was collected in a fresh microfuge tube and proteins were precipitated with ice-cold acetone. The total eluted protein was loaded on a gel.

## 2.4.4 Affinity purification of CBD-tagged proteins

The bait expression strain was precultured in 35 ml Halomedium containing 0.15 μg/ml novobiocin at 37 °C on a shaker (150 rpm) until an $OD_{600}$ of 0.5-1.0 was reached. This preculture was used to inoculate 100 ml Halomedium to an $OD_{600}$ of 0.01. The main culture was incubated on a shaker (110 rpm) at 37 °C. When the main culture had reached an $OD_{600}$ of 0.6 to 1.0, cells were harvested by centrifugation (8000 rpm, 15 min, 15 °C) and resuspended in 1-2 ml CFE buffer (3 M KCl, 1 M NaCl, 400 mM $NH_4Cl$, 40 mM $MgCl_2$, 10 mM Tris/HCl, pH 7.5) plus complete protease inhibitor (Complete Mini, EDTA-free, Roche). Cells were lysed by sonication on ice water (2 x 20 s, Branson sonifier 250, 3 mm disruptor horn, output level 2, constant), and the lysate cleared by centrifugation at 14000 rpm, 18 °C for 20 min in a tabletop centrifuge.

In the meantime a cellulose column was prepared by pipetting 300 μl cellulose suspension (10 % (w/v) Avicel PH-101 in CFE) into a Mobicol empty spin column. The column was centrifuged (300 x g, 1 min, RT), washed with 600 μl CFE to remove fines, and centrifuged again.

The cleared lysate was applied to the column in 600 μl portions and the cellulose resuspended by pipetting up and down. After 1 min incubation at room temperature the column was centrifuged (300 x g, 1 min, RT) and the flow-through discarded. The cellulose was washed three times with 600 μl CFE + 0.5 % NP40 and once with CFE. After each washing step the column was centrifuged (300 x g, 1 min, RT) and the flow-through discarded. An additional centrifugation (770 x g, 1 min, RT) was performed after the last washing step to reduce the amount of retained buffer. For elution, 600 μl ethylene glycol were applied to the column, the cellulose resuspended, and the column centrifuged. The eluate was collected in a fresh microfuge tube and proteins were precipitated with TCA (2.2.15).

## 2.4.5 CBD-AP and SILAC: Direct bait fishing

The bait expression strain and the control strain were precultured in 35 ml Halomedium containing 0.15 µg/ml novobiocin at 37 °C on a shaker (150 rpm) until an $OD_{600}$ of 0.5-1.0 was reached. 500 µl of these first precultures were used to inoculate second precultures that were grown under identical conditions. When the second precultures had reached an $OD_{600}$ of 0.8-1.0, the main cultures were inoculated. For the bait expression strain, 100 ml synthetic Halomedium containing $^{13}C_6$-leucine were inoculated to an $OD_{600}$ of 0.01, for the control culture 100 ml synthetic Halomedium containing $^{12}C_6$-leucine. To guarantee identical conditions for both the bait and the control culture, the inoculum for both cultures was brought to a total volume of 1.5 ml with complex medium.

The main cultures were incubated on a shaker (110 rpm) at 37 °C in the dark until they had reached an $OD_{600}$ of around 0.8. In order to work with roughly the same number of cells from both cultures, differences in the OD of bait and control culture were compensated by reducing the volume used from the culture with higher density accordingly. Cells were harvested by centrifugation (8000 rpm, 15 °C, 15 min) and pellets resuspended in 1 ml CFE with complete protease inhibitor. Cells were lysed by sonication on ice water as described above, and the lysate cleared by centrifugation at 14000 rpm, 18 °C for 20 min in a tabletop centrifuge. The supernatant was transfered to a fresh reaction tube.

Cellulose columns were prepared as described in 2.4.4. 300 µl lysate from each culture were applied to the column, the cellulose resuspended, and after 1 min incubation the column centrifuged (300 x g, 1 min, RT). This step was repeated two times, followed by washing, elution, and protein precipitation as described in 2.4.4. The protein pellet was stored at −78 °C.

## 2.4.6 CBD-AP and SILAC: Indirect bait fishing

*H. salinarum* R1 was precultured in 35 ml Halomedium at 37 °C on a shaker (150 rpm) until an $OD_{600}$ of 0.5-1.0 was reached. 500 µl of this first preculture were used to inoculate a second one that was grown under identical conditions. When the second preculture had reached an $OD_{600}$ of 0.8-1.0, it was used to inoculate two cultures with 100 ml synthetic medium, one containing $^{13}C_6$-leucine, the other one containing $^{12}C_6$-leucine, to an $OD_{600}$ of 0.01. The inoculum was brought to a total volume of 1.5 ml with complex medium. The cultures were incubated on a shaker (110 rpm) at 37 °C in the dark until they had reached an $OD_{600}$ of around 0.8.

In parallel, the bait expression strain and the pMS4 control strain were precultured in 35 ml Halomedium containing 0.15 µg/ml novobiocin at 37 °C on a shaker (150 rpm) until an $OD_{600}$ of 0.5-1.0 was reached. A second preculture was grown as described before. When an $OD_{600}$ of 0.8-1.0 was reached, the main cultures (200 ml Halomedium; the culture volumes were chosen larger to ensure saturation of the cel-

lulose in each column by bait or CBD, respectively) were inoculated to an $OD_{600}$ of 0.01 and incubated at 37 °C on a shaker (110 rpm). The main cultures were harvested at an $OD_{600}$ of around 1.0.

Cells of all four cultures were pelleted by centrifugation (8000 rpm, 15 °C, 15 min) and pellets resuspended in 1 ml CFE with complete protease inhibitor. Cells were lysed by sonication on ice water as described above, and the lysate cleared by centrifugation at 14000 rpm, 18 °C for 20 min in a tabletop centrifuge. The supernatant was transfered to a fresh reaction tube.

Two cellulose columns were prepared as described in 2.4.4. 600 μl lysate from the bait expression culture and the pMS4 control culture were applied to the columns, the cellulose resuspended, and after 1 min incubation the column was centrifuged (300 x g, 1 min, RT). This step was repeated, and the columns washed three times with CFE + 1 % NP40 + 20 % ethylene glycol and once with CFE.

Lysate from the *H. salinarum* R1 wt cells was applied to these columns in 600 μl portions, the cellulose resuspended, and after 1 min incubation the column centrifuged (300 x g, 1 min, RT). Washing, elution and protein precipitation was done as described in 2.4.4. The protein pellet was stored at −78 °C.

### 2.4.7 Mass spectrometry: Sample preparation

#### Cutting of gel slices

For MALDI-TOF PMF analysis, single bands were removed from 1 D gels with a scalpel and cut in pieces of circa 1 mm$^3$. Depending on the number of samples, gel pieces were either transfered to 0.5 ml reaction tubes or into 96 well microtiter plates for subsequent in-gel digestion.

For LC-MS/MS analysis, the whole lane was cut out of the gel and divided into 10-15 slices. Size of the slices was chosen according to the estimated amount of tryptic peptides derived from the respective part of the lane. Additionally, very thick bands were separated from weaker ones to prevent masking of low-abundance proteins. Slices were cut into pieces of circa 1 mm$^3$ and transfered to 0.5 ml reaction tubes.

#### Tryptic in-gel digestion

Samples from silver-stained gels were destained by oxidation of the silver before digestion (Gharahdaghi *et al.*, 1999). The gel pieces were shaken in 30 μl destaining solution (15 mM potassium hexacyanoferrate III, 50 mM sodium thiosulfate) until they were colourless. Subsequently, they were washed three times for five minutes with 100 μl $H_2O$. Coomassie stained gels were used directly for digestion.

Tryptic in-gel digestion was performed by a protocol modified from Shevchenko *et al.* (1996). Gel pieces were washed alternately in 50 μl 50 % acetonitrile and 50 μl 50 mM ammonium bicarbonate, each three times for 10 min. After this, they were incubated

in 50 µl 10 mM dithiothreitol (DTT), 50 mM ammonium bicarbonate for 45 min at 56 °C and in 50 µl 55 mM iodacetamide, 50 mM ammonium bicarbonate for 30 min at room temperature in the dark. The gel pieces were washed again alternately in 50 µl 50 % acetonitrile and 50 µl 50 mM ammonium bicarbonate, each three times for 10 min. To digest proteins, the gel pieces were incubated overnight in 25 µl trypsin solution (20 µg trypsin solved in 20 µl storage buffer and diluted with 5.2 ml 50 mM ammonium bicarbonate) at 37 °C on a shaker.

The supernatant was transfered to a 0.5 ml reaction tube and remaining peptides were eluted in three steps by incubation in 50 µl $H_2O$, 50 µl 50 % acetonitrile, and 50 µl 50 % acetonitrile, 0.1 % TFA on a shaker for 20 min each time. The supernatant of the digestion and all elution steps was pooled, frozen in liquid nitrogen, and dried down in a vacuum concentrator.

### Desalting

Salts remaining from the tryptic digestion (mainly ammonium bicarbonate) need to be removed from the sample prior to mass spectrometric measurements.

Samples for MALDI-TOF PMF were desalted by repeated dissolving in $H_2O$ and drying down in the vacuum concentrator, until no salt was visible (normally 2-3 times).

Samples for Nano-LC MS/MS require a higher purity and therefore they were desalted by reversed-phase (RP) chromatography using self-packed Stage tips (STop And Go Extraction, Rappsilber *et al.*, 2003). A small disk (app. 0.5 mm diameter) was punched out of Teflon embedded C18 material (C18 Empore™ Extraction Disk, 3M) and placed in a GELoader® pipette tip. Solutions were pressed through this column by applying pressure with a 1 ml syringe or by centrifugation (1500 x g for binding peptides, 3000 x g for all other steps). The C18 material was equilibrated with 10 µl isopropanol and washed with 10 µl 10 % formic acid. The peptides were dissolved in 1 µl formic acid, diluted with 9 µl $H_2O$, and passed through the column. The tips were washed twice with 10 µl 10 % formic acid and peptides eluted with 5 µl 80 % methanol, 10 % TFA into a 0.5 ml reaction tube. Eluted peptides were dried down in a vacuum concentrator.

### MALDI target preparation

Peptides from tryptic digests were spotted on 384 spot MALDI targets using a MAP II pipetting robot (Bruker Daltonics). Peptides were dissolved in 10 µl 33 % acetonitrile, 0.1 % TFA in an ultrasound bath for 30 s. 0.5 µl sample were mixed with 0.5 µl matrix solution (saturated $\alpha$-cyano-4-hydroxy-cinnamic acid in 40 % acetonitrile, 0.1 % TFA) and pipetted to the target. Cocristallisation was achieved by drying at the air (dried droplet method).

In every ninth position a peptide standard was spotted that was used for calibration of the mass spectrometer. The peptide standard solution was a mixture of

the following peptides (each 100 pmol/µl; the numbers in brackets are the molecular mass in g/mol): bradykinin fragment 1-7 (757.3998), angiotensin 2 acetate human (1046.542), angiotensin 1 acetate human (1296.6853), substance P (1347.7361), bombesin (1619.823), ACTH (1-17) (2093.0868), ACTH (18-39) (2465.199), somatostatin 28 (3147.4714), insulin chain B oxidised f. bovine (3494.651) (all purchased from Sigma Aldrich, St. Louis, USA). 3 µl of peptide standard solution were diluted with 18 µl 33 % acetonitrile, 0.1 % TFA and spotted in the same way as the samples.

### 2.4.8 Mass spectrometry: Data acquisition

#### MALDI-TOF peptide mass fingerprinting (PMF)

MALDI-TOF peptide mass fingerprint spectra were acquired on a Bruker Reflex III mass spectrometer in reflex mode with a detector voltage of 1851 V. Spectra acquisition was done automatically using *Fuzzy Control*. The output of the laser (337 nm) was set to 15 % in the beginning and 50 % at maximum. For each sample 200 shots and for each standard 100 shots were summed up, with a maximum of 20 shots on one target position. The mass window was set between 800 and 4000 m/z.

#### Nano-LC MS/MS (Q-TOF)

Peptides were chromatographically separated on a CapLC system (Waters) and the eluate directly injected into a Q-TOF ultimate mass spectrometer (Waters). The dried peptides were dissolved in 20 µl 5 % formic acid, and 1-6 µl (depending on the amount of protein estimated by the intensity of the Coomassie stained gel) were loaded into the CapLC using an auto sampler. They were bound to the precolumn (self-packed, 100 µm x 25 mm ReproSil-Pur 200 $C_{18}$-AQ, 5 µm, Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) with a flow of 2 µl/min and analysed on the main column (self-packed, 75 µm x 150 mm ReproSil-Pur 200 $C_{18}$-AQ, 3 µm) with a flow of 200 nl/min. Bound peptides were eluted in an acetonitrile gradient (Table 2.20) and injected into the mass spectrometer.

Mass spectrometric analysis was performed in the *positive ion mode* with a capillary voltage of 2.3 kV. The mass window was set to 300-2000 Da in MS mode and 50-2000 Da in MS/MS mode. Survey scans were acquired for 1.5 s. From each survey scan up to two peptides were chosen for fragmentation by CID; selection criteria were the signal intensity and the charge state (at least two-fold). CID was performed with a collision voltage between 16 and 40 kV (depending on peptide mass and charge) and helium as collision gas.

Table 2.20: **nano-HPLC gradient for LC-MS/MS analysis**

| Time | Percent A | Percent B | Flow |
|---|---|---|---|
| 3 | 100 | 0 | 2 |
| 13 | 100 | 0 | 5 |
| 21 | 100 | 0 | 5 |
| 25 | 90 | 10 | 5 |
| 30 | 85 | 15 | 5 |
| 40 | 80 | 20 | 5 |
| 60 | 75 | 25 | 5 |
| 80 | 65 | 35 | 5 |
| 95 | 50 | 50 | 5 |
| 100 | 0 | 100 | 5 |
| 105 | 100 | 0 | 5 |
| 110 | 0 | 100 | 5 |
| 115 | 100 | 0 | 5 |
| 125 | 100 | 0 | 5 |

A: 2 %ACN, 0.5 % FA; B: 80 %ACN, 0.5 % FA

## 2.4.9 Mass spectrometry: Data processing

### Processing of MALDI-TOF PMF data

MALDI PMF spectra were annotated using the program Xmas (version 5.1.1.16, Bruker Daltonics) and a peak list generated. This peak list was used for a database search with Mascot (Matrix Science) against a *Halobacterium salinarum* R1 protein sequence database. Carbamidomethylation of cysteine was set as a required modification. The peptide mass tolerance was set to 200 ppm, the charge state to 1+.

### Processing of Nano-LC MS/MS data

Peak lists were extracted from the raw data with Mascot Distiller (see Table 2.21 for parameters) and submitted to the Mascot server for search against a *Halobacterium salinarum* R1 protein sequence database. Carbamidomethylation of cysteine was set as a required modification and oxidation of methionine and acetylation of the protein's N-terminus as variable modification. Up to three missed cleavage sites were allowed. For SILAC experiments, $^{13}C_6$-Leucine were additionally set as variable modification. Mass tolerance was set to 1.5 Da for MS and 0.6 Da for MS/MS. If several samples had to be analysed, the process of peak list generation and search submission was automated by use of the Mascot Daemon.

## 2.4.10 Determination of SILAC ratios

Protein ratios in SILAC experiments were determined with the tool ASAPRatio (Li *et al.*, 2003) embedded in the Trans-Proteomic Pipeline (TPP; Keller *et al.*, 2005).

Table 2.21: **Parameters for Mascot Distiller**

| |
|---|
| MS.UncentroidingHalfWidth = 0.2 |
| MS.UncentroidingPointsPerDa = 20 |
| MS.RegriddingPointsPerDa = 20 |
| MS.AggregationMethod = 0 |
| MS.MinPeakCount = 1 |
| MS.MaxPeakCharge = 3 |
| MSMS.UncentroidingHalfWidth = 0.2 |
| MSMS.UncentroidingPointsPerDa = 20 |
| MSMS.RegriddingPointsPerDa = 20 |
| MSMS.AggregationMethod = 2 |
| MSMS.MinPeakCount = 10 |
| MSMS.MaxPeakCharge = 2 |
| MSMS.UsePrecursorAsMaxCharge = 1 |
| MSMS.PrecursorChargeSources = 1;0 |
| MSMS.PrecursorDefaultCharges = 2;3 |
| MSMS.RedeterminePrecursorMZ = 1 |
| MSMS.PrecursorMZTolerance = 3 |
| MSMS.IgnoreSingleChargedPrecursor = 0 |
| TimeDomain.MinPrecursorMass = 300 |
| TimeDomain.MaxPrecursorMass = 16000 |
| TimeDomain.PrecursorGroupingTolerance = 0.3 |
| TimeDomain.MaxIntermediateScans = 1 |
| TimeDomain.MinScansInGroup = 1 |
| PeakSelection.FilteringThreshold = 0.7 |
| PeakSelection.MinFilteringPeakMZ = 50 |
| PeakSelection.MaxFilteringPeakMZ = 100000 |
| PeakSelection.FilteringMinSNRatio = 2 |
| PeakSelection.MinPeakWidth = 0.01 |
| PeakSelection.ExpectedPeakWidth = 0.1 |
| PeakSelection.MaxPeakWidth = 1 |
| PeakSelection.MaxIterations = 500 |
| PeakSelection.RejectWidthOutliers = 0 |
| PeakSelection.BaselineCorrection = 0 |

Raw data were converted to mzXML files using Masswolf. Corresponding database search result files (Mascot ".dat" files) were renamed accordingly to the mzXML files and converted to pepXML with the MascotConverter. pepXML files derived from the same experiment were combined and processed through the Trans-Proteomic Pipeline. ASAPRatioPeptideParser was used with the options "lL" (set leucine as labeled residue), "-C" (quantitate only the charge state where the CID was made), "B" (return a ratio even if the background is high), and "-F" (use fixed scan range for light and heavy peptide). All other tools were run with the default parameters. Batch processing of raw data and renaming of Mascot ".dat" files were automatised by self-written Perl scripts. Ratios of all proteins with at least two peptides identified and a protein probability higher than 0.85 were checked manually on the basis of

the extracted ion chromatograms and adjusted if necessary (e. g. background level or scan range). To accomplish a better presentability of the protein ratios a symmetrical measure, called ASAP Score, was introduced. ASAP Score was calculated as follows:

$$ASAPScore = \begin{cases} ASAPRatio(H/L) - 1 & \text{if } ASAPRatio(H/L) \geq 1 \\ 1 - \frac{1}{ASAPRatio(H/L)} & \text{if } ASAPRatio(H/L) < 1 \end{cases}$$

### 2.4.11 Thresholds and statistics

In the results from SILAC experiments only proteins were included that were identified with at least two different peptides, had a ProteinProphet probability of 0.95 or higher (5 % false identification rate), and were quantifiable by ASAPRatio (at least one of the identified peptides had to contain leucine, and the extracted ion chromatograms had to be utilisable).

The SILAC ratios were tested for standard normal distribution with the Shapiro-Wilk normality test, implemented as module shapiro.test in the statistic package R (R Development Core Team, 2008).

## 2.5 Materials and methods for the chemotaxis protein interaction network

### 2.5.1 Generation of expression and control strains

The complete coding regions of all bait proteins were amplified with the primers listed in Table 2.22 and cloned into pMS3 - pMS6 as described in 2.4.2. The resulting plasmids (Table 2.5.3) were transformed into *H. salinarum* R1. At least two clones of the pMS3 and pMS4 strains were verified by a test affinity purification according to the protocol given in 2.4.4, and subsequent evaluation of the eluted proteins on a silver-stained gel. Clones which have shown a band in the expected height (pMS3) or the characteristic double band (bait + CBD, bait + 2 x CBD for pMS4) were used for further analysis. At least two clones of the pMS5 and pMS6 strains were verified by western blotting (2.2.19) with an anti-His antibody (anti-His HRP conjugate, QIAGEN). Clones which showed a band in the height of the bait protein were used in further experiments.

### 2.5.2 Bait fishing, mass spectrometry, data analysis

Direct and indirect bait fishing experiments were performed as described in 2.4.5 and 2.4.6, respectively. Eluted proteins were separated by gel electrophoresis, and the lanes cut into 10 - 15 slices. Samples were prepared for mass spectrometric analysis as described in 2.4.7 and measured on a QTOF mass spectrometer (see 2.4.8 for

details). Mass spectra were searched against a *H. salinarum* R1 protein database, and the results processed through the Trans-Proteomic Pipeline (TPP) for protein quantification and validation of results (2.4.9, 2.4.10). The identifications meeting the criteria given in 2.4.11 were uploaded into the Result DB and the experiments evaluated with the Result Viewer (5.3.3). Finally, the interactions were exported from the ResultViewer and illustrated as networks using Cytoscape (Shannon *et al.*, 2003).

### 2.5.3 Che protein interactions in other organisms

Interactions from other organisms were fetched from the STRING (von Mering *et al.*, 2007) and BIND (Alfarano *et al.*, 2005) databases. STRING was queried in COG mode with the COGs: COG0632 (CheA), COG2201 (CheB), COG3143 (CheZ), COG0835 (CheW), COG1352 (CheR), COG1776 (CheC/FliY), COG1871 (CheD), COG0784 (CheY and other RRs), COG0840 (MCP). Only interactions with experimental evidence were accepted. The prediction methods were not useful for the purpose of this study because they connect all Che proteins (all in gene neighbourhood, strong genomic co-occurrence). The same is true for evidence from databases (mainly from KEGG: all in the same pathway), and textmining (registers co-mentioning in one abstract). The BIND database was queried by keyword search for CheA, CheB, CheZ, CheW, CheR, CheC, CheD, CheY, and MCP. Functional interactions were collected by literature search in PubMed.

Table 2.22: **Oligonucleotides for Che protein interaction analysis**

| Primer | Sequence |
| --- | --- |
| OE1428F_fo | CACCGTGCTGAGTTCCAGCAGCGACGC |
| OE1428F_re | CGCGCGAGTGAACGCGTCGAGTC |
| OE1620R_fo | CACCATGTCCACCATCGCTGGTCTGG |
| OE1620R_re | GTCGTGGCGGAACGCCCGGTGG |
| OE2374R_fo | CACCATGTCCGATGACGAGACGGAC |
| OE2374R_re | TACGAGGGCGTCGGGTTCGAC |
| OE2378R_fo | CACCATGAGCGAGGCCGGTCGGGAG |
| OE2378R_re | CGAGAGGCGTGACTTCACG |
| OE2401F_fo | CACCGTGCCATCGCTGTACGGGCTGG |
| OE2401F_re | CGTTTTTCCGCCCAGCTTCGAGATC |
| OE2402F_fo | CACCATGAGCGAGTCAGAGTACAAGATAG |
| OE2402F_re | TTCCTCGTTGATCGCTTCGCTGGC |
| OE2404R_fo | CACCATGTCGGAATCAGCAATCGCAGAC |
| OE2404R_re | GAGGTTGACGTCCTCGATGTGTTC |
| OE2406R_fo | CACCTTGACTGACTTCCAAACCCTG |
| OE2406R_re | CGTGTCAGCGACCCGACTGTAG |
| OE2408R_fo | CACCGTGACGATCCGCGTTGGCGTG |
| OE2408R_re | AATTACGTGCACCTCGGCATC |
| OE2410R_fo | CACCATGCGTGTCGATCTCGACGC |
| OE2410R_re | CGACTCCTCACTCGCGTCGG |
| OE2414R_fo | CACCATGAGCACAATGATCGACATTC |
| OE2414R_re | GATACTGTTGATCATCGAGAC |
| OE2415R_fo | CACCATGGACGACTACCTCGAAGCGTTCG |
| OE2415R_re | CAGCGTAGCCACGTCCAGGATGGTC |
| OE2416R_fo | CACCATGACAGAGGCACTGGTGGTC |
| OE2416R_re | CGTCGTCCTCCGTATCGAATC |
| OE2417R_fo | CACCATGGCGAAGCAGGTCTTACTGGTC |
| OE2417R_re | TGCGGTGAGCACGTCCGAAATAGCG |
| OE2419R_fo | CACCATGAACCTCGAAGACGCCGAC |
| OE2419R_re | GACATGATTGTCGATGGTCTC |
| OE3280R_fo | CACCATGAGTGCCACGATCGAGCTG |
| OE3280R_re | CGAGAGGTCGTTGAGCAGCGAC |
| OE4643R_fo | CACCATGAACGCCGACATCGACGCGGTG |
| OE4643R_re | CATCTCCCATCACCTCCCGGAT |

Table 2.23: **Plasmids and strains for Che protein interaction analysis**

| Plasmid | Comment | Source | Plasmid | Comment | Source |
|---------|---------|--------|---------|---------|--------|
| pMS340 | pENTR::*OE1428F* | This study | pMS101 | pMS3::*cheA* | This study |
| pMS393 | pMS4::*OE1428F* | This study | pMS106 | pMS4::*cheA* | This study |
| pMS395 | pMS6::*OE1428F* | This study | pMS246 | pMS5::*cheA* | This study |
| pMS280 | pENTR::*purH/N* | This study | pMS254 | pMS6::*cheA* | This study |
| pMS282 | pMS5::*purH/N* | This study | pMS124 | pENTR::*cheB* | This study |
| pMS285 | pMS6::*purH/N* | This study | pMS157 | pMS3::*cheB* | This study |
| pMS316 | pMS3::*purH/N* | This study | pMS241 | pMS4::*cheB* | This study |
| pMS331 | pMS4::*purH/N* | This study | pMS328 | pMS6::*cheB* | This study |
| pMS136 | pENTR::*cheW2* | This study | pMS102 | pENTR::*cheY* | This study |
| pMS151 | pMS3::*cheW2* | This study | pMS105 | pMS3::*cheY* | This study |
| pMS166 | pMS4::*cheW2* | This study | pMS156 | pMS4::*cheY* | This study |
| pMS271 | pMS5::*cheW2* | This study | pMS247 | pMS5::*cheY* | This study |
| pMS272 | pMS6::*cheW2* | This study | pMS255 | pMS6::*cheY* | This study |
| pMS114 | pENTR::*parA1* | This study | pMS103 | pENTR::*cheW1* | This study |
| pMS400 | pMS4::*parA1* | This study | pMS191 | pMS3::*cheW1* | This study |
| pMS401 | pMS6::*parA1* | This study | pMS195 | pMS4::*cheW1* | This study |
| pMS426 | pENTR::*cpcE* | This study | pMS284 | pMS5::*cheW1* | This study |
| pMS451 | pMS4::*cpcE* | This study | pMS290 | pMS6::*cheW1* | This study |
| pMS452 | pMS6::*cpcE* | This study | pMS107 | pENTR::*cheC2* | This study |
| pMS370 | pENTR::*OE2402F* | This study | pMS164 | pMS4::*cheC2* | This study |
| pMS387 | pMS4::*OE2402F* | This study | pMS327 | pMS6::*cheC2* | This study |
| pMS388 | pMS6::*OE2402F* | This study | pMS281 | pENTR::*OE4643R* | This study |
| pMS390 | pENTR::*OE2404R* | This study | pMS294 | pMS6::*OE4643R* | This study |
| pMS414 | pMS4::*OE2404R* | This study | pMS319 | pMS4::*OE4643R* | This study |
| pMS415 | pMS6::*OE2404R* | This study | | | |
| pMS126 | pENTR::*cheR* | This study | | | |
| pMS187 | pMS4::*cheR* | This study | | | |
| pMS234 | pMS3::*cheR* | This study | | | |
| pMS288 | pMS5::*cheR* | This study | | | |
| pMS292 | pMS6::*cheR* | This study | | | |
| pMS146 | pENTR::*cheD* | This study | | | |
| pMS154 | pMS3::*cheD* | This study | | | |
| pMS163 | pMS4::*cheD* | This study | | | |
| pMS264 | pMS5::*cheD* | This study | | | |
| pMS269 | pMS6::*cheD* | This study | | | |
| pMS139 | pENTR::*cheC3* | This study | | | |
| pMS235 | pMS4::*cheC3* | This study | | | |
| pMS337 | pMS6::*cheC3* | This study | | | |
| pMS127 | pENTR::*cheC1* | This study | | | |
| pMS205 | pMS3::*cheC1* | This study | | | |
| pMS240 | pMS4::*cheC1* | This study | | | |
| pMS322 | pMS6::*cheC1* | This study | | | |
| pMS100 | pENTR::*cheA* | This study | | | |

The *H. salinarum* R1 strains transformed with the pMS plasmids (pMSxy) were designated in the same manner (MSxy) as the plasmids.

# 2.6 Materials and methods for identification of archaea-specific Che proteins

## 2.6.1 Construction of in frame deletion mutations

Table 2.24: **Primer for construction of deletion mutations**

| Primer | Sequence | Site |
|---|---|---|
| OE2401F_us_fo | ATTCGAGGATCCGGTGAGGAAGCCTCCGCACAGG | BamHI |
| OE2401F_us_re | AAACCGCCACCAGTGGCCTAGACAGACGACACGCGAC | |
| OE2401F_ds_fo | GTCGTCTGTCTAGGCCACTGGTGGCGGTTTCAGTCGC | |
| OE2401F_ds_re | TCCTATAAGCTTGCGGGATGCTTCACCTGTAGC | HindIII |
| OE2402F_us_fo | TTAACGGGATCCACACCAGCGCCGGAACGCCACG | BamHI |
| OE2402F_us_re | GGAGCAGCCCCACGATCCTCACCTACGATCGGCGCTTC | |
| OE2402F_ds_fo | GATCGTAGGTGAGGATCGTGGGGCTGCTCCCCTCGG | |
| OE2402F_ds_re | TATCGGAAGCTTCTCGCGCTCTCGCCGGGTGACC | HindIII |
| OE2404R_us_fo | TTAACGCTGCAGCGTCTCCATTCTGGCCACCGAC | PstI |
| OE2404R_us_re | TCCCCTCGGACGCCACACCCCCGTATACTCAATTATAC | |
| OE2404R_ds_fo | TGAGTATACGGGGGTGTGGCGTCCGAGGGGAGCAGCC | |
| OE2404R_ds_re | TTAAGCGGATCCGCATCAAGGTCGACGAATCGG | BamHI |
| OE2404_us_fo_neu | TTAACGTCTAGACGTCTCCATTCTGGCCACCGAC | XbaI |
| OE2402F-OE2404R_us_re | TGAGTATACGGGGGTTCCTCACCTACGATCGGCGCTTC | |
| OE2402F-OE2404R_ds_fo | GATCGTAGGTGAGGAACCCCCGTATACTCAATTATAC | |

In-frame deletion plasmids were constructed using the vectors pMKK100 (Koch and Oesterhelt, 2005) and pMS3 (2.4.1). The reason for the use of two vectors was oppositional experience with the used antibiotic resistance genes – pMS3 contains novobiocin resistance, pMKK100 mevinolin resistance. Whereas our lab has positive experience with the use of novobiocin resistance for the transformation of *H. salinarum* R1 and had occasionally problems with mevinolin resistance (high background), another lab, mainly working with *H. salinarum* S9, had more success with the use of mevinolin for selection. Finally, in this study no advantage of either selection marker in one of the used strains was found (for details see Miller, 2007).

All PCR reactions were done with Phusion™ Polymerase according to supplier's instructions and genomic DNA of *H. salinarum* strain R1 as template. 500 bp of sequence upstream and downstream of the targeted gene were amplified by PCR using the primers listed in Table 2.24 (us_fo and us_re for the upstream sequence, ds_fo and ds_re for the downstream sequence). The corresponding PCR products were fused in a second PCR using the external primers (us_fo and ds_re). For the double deletion ΔΔOE2402F OE2404R, the external primers OE2402F_us_fo and OE2404R_us_fo (note the different orientation of this genes in the genome), were used in both PCR rounds. The fusion products were purified by gel extraction and digested with the enzymes indicated in Table 2.24 and ligated into both pMS3 and pMKK100 digested with the same enzymes. The resulting deletion plasmids were verified by DNA sequencing

Table 2.25: **Strains and plasmids**

| Plasmid | Comment | Source |
|---|---|---|
| pMKK100 | *E. coli* / *H. salinarum* shuttle vector | Koch and Oesterhelt (2005) |
| pMS3 | see 2.4.1 | This study |
| pAM1 | pMKK100 with ΔOE2401F fragment | Miller (2007) |
| pAM2 | pMS3 with ΔOE2401F fragment | Miller (2007) |
| pAM3 | pMKK100 with ΔOE2402F fragment | Miller (2007) |
| pAM4 | pMS3 with ΔOE2402F fragment | Miller (2007) |
| pAM5 | pMKK100 with ΔOE2404R fragment | Miller (2007) |
| pAM6 | pMS3 with ΔOE2404R fragment | Miller (2007) |
| pAM7 | pMKK100 with ΔΔOE2402F-OE2404R fragment | Miller (2007) |
| pAM8 | pMS3 with ΔΔOE2402F-OE2404R fragment | Miller (2007) |

| Strain | Comment | Source |
|---|---|---|
| R1Δ1 | R1 with in-frame deletion of OE2401F | Miller (2007) |
| S9Δ1 | S9 with in-frame deletion of OE2401F | Miller (2007) |
| R1Δ2 | R1 with in-frame deletion of OE2402F | Miller (2007) |
| S9Δ2 | S9 with in-frame deletion of OE2402F | Miller (2007) |
| R1Δ4 | R1 with in-frame deletion of OE2404R | Miller (2007) |
| S9Δ4 | S9 with in-frame deletion of OE2404R | Miller (2007) |
| R1Δ2-4 | R1 with in-frame deletion of OE2402F-OE2404R | Miller (2007) |
| S9Δ2-4 | S9 with in-frame deletion of OE2402F-OE2404R | Miller (2007) |
| S9Δ1/1$^+$ | *in cis* complementation of S9Δ1 | This study |
| S9Δ2/2$^+$ | *in cis* complementation of S9Δ2 | This study |
| S9Δ4/4$^+$ | *in cis* complementation of S9Δ4 | This study |
| S9Δ2-4/2-4$^+$ | *in cis* complementation of S9Δ2-4 | This study |

of the insert.

Deletion mutants were generated by transformation of the deletion plasmids into the wild type strains R1 and S9 and subsequent cultivation without selection pressure as described in Koch and Oesterhelt (2005). Briefly, after the transformation and plating on X-gal and antibiotic containing plates two blue clones were picked and grown in complex medium without antibiotics. After three passages of the culture, roughly 600 cells were plated on X-gal containing plates without antibiotics. Red colonies (red colour indicates that these cells have lost the integrated plasmid) were inoculated into complex medium and screened for the loss of the target gene by PCR using the primers spanning the flanking regions.

## 2.6.2 Southern blot analysis

Deletions were verified by Southern blot analysis. Genomic DNA of wild type and deletion strains was isolated as described in 2.2.20 and digested with BglI. DIG-labeled DNA probes (one for the gene's upstream region and one for the gene) were generated via PCR amplification of the upstream or the gene sequence from genomic DNA in the presence of DIG-11-dUTP (Roche). The digested DNA was subjected to 1% agarose

gel electrophoresis, blotted to Hybond-N nylon membrane (Amersham Biosciences) and then UV-crosslinked. The use of blocking reagent, hybridisation procedure and chemiluminescent detection with CSPD® chemiluminescent substrate (Roche) was according to standard protocols. For a detailed description see Miller (2007).

### 2.6.3 Complementation of deletions

Deleted genes were reintroduced into all deletion strains *in cis*. Complementation plasmids for each deletion were constructed by PCR amplification of the deleted gene(s) together with the flanking regions from *H. salinarum* R1 genomic DNA using the external primers (us_fo, ds_re). For the complementation of $\Delta 4$ and $\Delta 2$-4, the primer OE2404R_us_fo_neu had to be used due to a PstI cutting site in the gene. Inserts were digested with the respective restriction enzymes and cloned into pMS3, and the resulting plasmids were verified by sequencing of the insert.

Each deletion strain was transformed with the corresponding complementation plasmid and a double crossover triggered as described above. Red colonies were inoculated into complex medium and screened for reintroduction of the target gene by PCR using the primers spanning the flanking regions.

### 2.6.4 Swarm plates

Semi-solid agar plates were prepared from complex medium with $0.25\%$ agar. Wild type and deletion cultures were grown to an $OD_{600}$ of $0.6$-$0.8$ and reinoculated twice with equal amounts of cells to achieve equal cell densities in the final cultures. $10\ \mu l$ of culture with an $OD_{600}$ of $0.6$-$0.8$ were injected with a pipette tip into the soft agar. The plates were incubated for $3$-$6$ days at $37\ °C$ in the dark.

### 2.6.5 Computerised cell tracking (Motion analysis)

#### The HaloTrack system

Reversal frequencies of unstimulated cells and after application of a photophobic stimulus were measured with a computerised cell tracking system (Stefan Streif, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg; unpublished). The system consists of a phase-contrast microscope equipped with a CCD camera, which allows stimulus-free observation of the cells using infrared light. The camera transmits the stream of images to a data-processing computer, which constructs a time-resolved cell track for each cell in the frame.

To measure the response to light stimuli, the light from two computer-controlled light sources can be applied to the cells. In this study, two photophobic stimuli were used: a blue light pulse and a orange light step-down. The blue light was applied through the objective, thereby only stimulating the cells in the visual field. This

Table 2.26: **Instruments for motion analysis**

| Instrument | Model and Distributor |
| --- | --- |
| Microscope | Olympus BX51 |
| TV adapter | U-TV1X-2, Olympus |
| C mount adapter | UCMAD-3, Olympus |
| Camera | CCD camera 4912-4000-0000, COHU, San Diego, CA, USA |
| Excitation light sources | MT20-SPA illumination system, Olympus, equipped with |
| | 150 W Hg/Xe mixed gas arc burner |
| Excitation filters | $480 \pm 50$ nm, AHF Analysentechnik, Tübingen, Germany |
| | $580 \pm 50$ nm, AHF Analysentechnik, Tübingen, Germany |
| Infrared filter | RG780 |
| Beam splitters | DCLP650, Olympus |
| Stage | Olympus Biosystems, custom made and equipped with peltier element |
| Peltier Element | PE 94, Linkam Scientific Instruments Ltd., Surrey, England |

allowed repeated measurements on the same slide without stimulating the same cells more than once. Orange light was applied through the condenser, illuminating a much broader area. Thereby also cells swimming into the visual field during the measurement were illuminated and thus adapted to the light.

**Specimen preparation**

Microscopic slides (76 x 26 mm, Menzel Gläser, Braunschweig) were washed three times for 1 min in deionised water. After that, they were washed for 2 h in acetone and for 2 h in ethanol. All washing steps were done on a shaker at 100 rpm. The slides were wiped dry with lint-free paper cloth (Kimwipes® lite precision wipes, Kimberly-Clark®).

Cover slips were washed for 2 min in acetone with vigorous shaking and then four times for 2 min in deionised water. The cover slips were dried in an oven at 80 °C.

Cells were grown in 35 ml Halomedium for 2 - 3 days until an $OD_{600}$ of 0.6 - 0.9 was reached. Cells were diluted with Halomedium and arginine to an $OD_{600}$ of 0.32 and a final arginine concentration of 0.1 % (w/v). After that the cells were incubated in the dark at RT for at least 20 min. For measurement, 5 µl cell suspension were pipetted on a slide and sealed under a cover slip with a molten 2:1 (w/w) mixture of paraffin wax and vaseline.

**Measurement of reversal rates**

Before starting the measurements, the specimen was incubated for 25 min on the heated stage (25 °C). With each culture three experiments were performed: measurement of spontaneous reversals (without stimulation), measurement of reversals after a blue-light pulse, and after a step-down of orange light. For each experiment, a fresh slide was prepared and a series of 20 single measurements was performed. One mea-

surement included 5 s recording of cell movement of which finally a 4 s interval was analysed for cell reversal.

For measuring the blue light response, a blue light pulse (0.5 s duration, 5 % intensity) was applied at the beginning of the tracking interval. After each measurement the position on the slide was changed to avoid repeated stimulation of the same cells.

For measurement of the response to an orange light step-down, the cells were initially adapted to orange light for 5 min. At the beginning of the tracking interval, the orange light was switched off for 4 s. Prior to the next measurement, the cells were adapted again for 45 s. Orange light was applied with maximum intensity. Since the orange light is applied through the condenser, it illuminates almost the whole specimen. Thereby it is ensured that no cells that have not been completely adapted enter the visual field during the tracking interval. The drawback is that the cells are repeatedly stimulated during the 20 measurements, which might cause unwanted side effects. Such effects have, however, not yet been observed.

### Detection of reversals

Reversals are detected by an algorithm based on a Kalman filter (Stefan Streif, unpublished). Briefly, for each time point, a prediction of the cell position for some time span in the future is made based on the last measurements. The prediction is compared with the actual position after the time span has elapsed. Reversals are detected by this comparison (see also Marwan and Oesterhelt, 1990) with a false positive and false negative rate of 2 and 2.5%, respectively (Stefan Streif, unpublished). The 95% confidence intervals were calculated assuming a binomial distribution according to Lorenz (1996).

By measuring known straight-swimming mutants (cheY**; del Rosario *et al.*, 2007), the false positive detection of reversal events (tracking error) was determined to be maximally 2.5-5 % in a 4 s observation interval (Stefan Streif, unpublished).

### Calculation of confidence intervals for the reversal rates

The 95 % confidence interval for the reversal rates were calculated. As reversal events follow the binomial distribution, the borders of the confidence interval could be calculated with the following equation:

$$\left.\begin{array}{c} P_U \\ P_L \end{array}\right\} = \frac{1}{n+c^2}\left[k \pm \frac{1}{2} + \frac{c^2}{2} \pm c\sqrt{\left(k \pm \frac{1}{2}\right)\left(1 - \frac{k \pm \frac{1}{2}}{n}\right) + \frac{c^2}{4}}\right]$$

$P_U$ and $P_L$ are the upper and lower limit of the confidence interval. $k$ is the number of cells that reversed, $n$ is the total number of evaluated cells. $c$ is a coefficient depending on the confidence level. For the chosen confidence level of 95 %, $c = 1.96$.

## 2.6.6 Dark-field microscopy

Table 2.27: **Instruments for dark-field microscopy**

| Instrument | Model and Distributor |
| --- | --- |
| Microscope | Olympus BX50 |
| Light source | 100 W mercury lamp USH-120D, Olympus |
| Condenser | Olympus U-DCW cardioid immersion dark-field condenser |
| | N.A. 1.40 - 1.20 |
| Objective | UPlanFL 40x, Olympus, N.A. 0.75 |
| Photo tube | U-SPT, Olympus |
| Photo eye peace | PE 5 x 125 |
| TV adapter | U-PMTV, Olympus |
| C mount adapter | UCMAD-2, Olympus |
| Camera | CCD C-5405, Hamamatsu Photonics, Herrsching, Germany |
| Camera Controller | C2400, Hamamatsu Photonics, Herrsching, Germany |

Dark-field microscopy was used to investigate the flagellar rotational bias (the fraction of CW or CCW swimming cells). Dark-field microscopy was done essentially as described by Staudinger (2007).

Cell culture and preparation of microscopic specimens was done as described in 2.6.5. Cells were diluted to an $OD_{600}$ of 0.1 with Halomedium, and arginine added to a final concentration of 0.1 %. The camera controller was set to AGC on, high/low off, gain 0 - 4, contrast 0 - 1.5, all shading options 50 %, and everything else off. 50 μl immersion oil ($n_e = 1.5180$, Leitz, Wetzlar, Germany) were pipetted on the condenser, the slide put onto the stage, and the condenser adjusted to maximal height. The cells were focused and the condenser gradually lowered again (with permanent refocusing) until the flagella became visible. Each specimen was used for one hour at the most.

The flagellar rotational bias was determined by counting the cells swimming with the flagellum in front of the cell body (CCW) and cells swimming with the flagellum behind the cell body (CW). Counting was assisted by the software "Halo-counter", written by Stefan Streif (Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany) in MATLAB®.

## 2.6.7 Bioinformatic analysis

To collect information about unknown proteins, the databases COG (Tatusov *et al.*, 1997, 2003), Pfam (Finn *et al.*, 2006, 2008), InterPro (Mulder *et al.*, 2007), SMART (Schultz *et al.*, 1998), and STRING (von Mering *et al.*, 2007) were queried with the protein sequence.

The multiple alignment of the DUF439 proteins was calculated using ClustalX (Thompson *et al.*, 1994, 1997) using standard parameters. For phylogenetic analysis, a neighbour-joining tree was calculated from the multiple alignment applying the Phylip package (Felsenstein, 2005). Again, standard parameters were used.

**Table of Che and Fla orthologs**

For identification of Che and Fla orthologs in archaeal genomes, a combination of homology search, genome region analysis, and cluster analysis based on pairwise similarity was applied.

First, *che* and *fla* gene regions were identified by Psi-Blast against an archaeal protein sequence database using *H. salinarum* CheD and FlaH as queries. The database contained the predicted proteins from all complete archaeal genomes available in Gen-Bank in October 2007, except for *Halobacterium salinarum* NRC-1, which contains the same *che* and *fla* genes as *H. salinarum* R1 (Ng *et al.*, 2000; Pfeiffer *et al.*, 2008b). CheD and FlaH were chosen as queries because a former study had demonstrated that they are highly conserved throughout all chemotactic or motile archaea, and they have no close non-orthologous homologs, which would make result evaluation difficult (Klein, 2005). All hits with an e-value of $10^{-8}$ or lower were accepted. This cut-off, however, was not critical, as there were no hits with e-values between $10^{-10}$ and $10^{-2}$ for CheD and $10^{-10}$ and $10^{-5}$ for FlaH.

Second, the genes in the neighbourhood of the identified *cheD* and *flaH* genes were examined by BlastP against the archaeal protein sequence database, and querying CDS and Pfam. Based on homology or identified domains, the genes were assigned to pools. The pools were: CheA, CheB, CheC, CheD, CheR, CheW, CheY, DUF439, FlaC, FlaD/E (FlaD and FlaE could not be distinguished), FlaF, FlaG, FlaH, FlaI, FlaJ. The examination of neighbouring genes was repeated until on each side of the region three genes with no obvious relation to chemotaxis or flagellation were found.

Third, the pools were extended to identify homologs located apart from the main genome regions. For this, each member of a pool was used as query in a Blast search against the archaeal protein sequence database. All hits with an e-value of $10^{-3}$ or smaller were included into the extended pools. Fetching the query sequences from an archaeal protein database, performing the blast runs, and writing the extended pools was automated using self-written Perl scripts.

Fourth, the extended pools were clustered based on pairwise similarity. This was done with the CLANS application (Frickey and Lupas, 2004). CLANS takes a set of sequences (in this case all sequences from the extended pools) as input, performs all-against-all BLAST searches, and displays the pairwise similarities as 3D graph. Sequences are represented as vertices in the graph. BLAST high-scoring segment pairs (HSPs) are shown as edges between the vertices and provide attractive forces proportional to the negative logarithm of the HSPs *P*-value. The attractive forces are balanced by a mild repulsive force between all vertices. Vertices positions are determined in an iterative process, in which the vertices, after initial random placement, are moved along the force vectors resulting from all pairwise interactions. By this approach, sequence groups with high similarities between all sequences come close together (form clusters), whereas sequences from different groups are separated even when single pairs between the groups build high-scoring alignments. Iteration was run

until movement of vertices became negligible. The cluster in which the members of the non-extended pools were found was extracted and the members considered as the final group of orthologs. Proteins which were not included into this ortholog cluster but did also not cluster with any other proteins and had only connections to the ortholog cluster, were included into the final ortholog group as well (marked with an asterisk in Table S4). The applicability of the method was supported by the fact that in all cases the members of the non-extended pools were found in one cluster.

# 3 Yeast two-hybrid analysis of halobacterial proteins

## 3.1 Introduction

The yeast two-hybrid (Y2H) system (Fields and Song, 1989) is based on the fact that several eukaryotic transcription factors consist of a DNA-binding domain (DBD) and a transcription activation domain (AD), which are functionally and physically separable (Hope and Struhl, 1986; Keegan *et al.*, 1986). The DBD localises the protein to specific DNA sequences, whereas the AD activates transcription by contacting the transcription machinery. To give a functional transcription factor, both domains do not need to be present on the same polypeptide (Ma and Ptashne, 1988), but it is sufficient to bring them into close proximity to each other (Figure 3.1 A and B). Fields and Song (1989) realized that this can be achieved by fusing each of the domains to one of a pair of physically interacting proteins (Figure 3.1 C and D).

In the original work, Fields and Song (1989) used the transcription factor GAL4, and the GAL4 system is, together with LexA (Vojtek *et al.*, 1993), still one of the most used Y2H systems. To report an interaction, the fusion constructs (the bait:
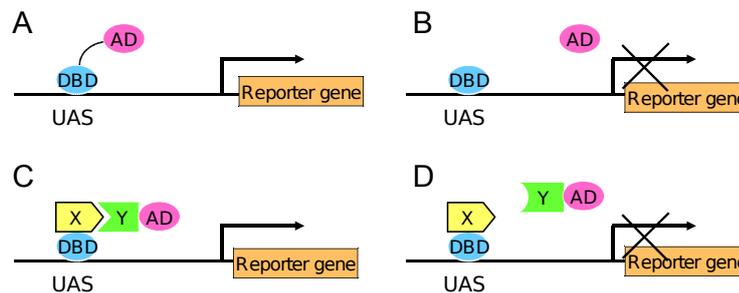


Figure 3.1: **The Y2H system. A, B** Several eukaryotic transcription factors consist of a DNA-binding domain (DBD) and a transcription activation domain (AD). Both domains must be linked together to activate transcription. **C, D** DBD and AD can be fused to a pair of proteins (X, Y). If these proteins interact, the functional transcription factor is reconstituted and transcription activated. (UAS: upstream activating sequence).

protein X with the DBD; the prey: protein Y with the AD) are transformed to a yeast strain which carries appropriate reporter genes (commonly used are for example HIS3, LEU2, ADE2, lacZ, URA3, LYS2, GFP) under the control of the split transcription factor. The readout is normally growth of the cells on selective media and/or a colour reaction (blue/white screening).

## 3.2 Results and Discussion

### 3.2.1 Analysis of a test set of proteins

A set of known interactors from *H. salinarum* and, as control, two interacting proteins from *E. coli* were selected to test if the Y2H system is suited to analyse interactions of halophilic proteins (Table 3.1). The test proteins include CheA and CheW1, whose interaction was demonstrated in other species (see Szurmant and Ordal, 2004, and references therein), Dodecin, which self-associates to dodecamers (Bieger *et al.*, 2003), malate dehydrogenase (MDH), which forms dimers of dimers (shown for the enzyme from *H. marismortui* in Mevarech *et al.*, 2000), and NuoA and NuoB, two components of the NADH dehydrogenase complex. NuoA is a transmembrane protein, and the only protein in the test set with a non-acidic pI value. The control proteins from *E. coli* were CheA and CheW.

Table 3.1: **Proteins for Y2H.**

| Pair | Bait | Prey | |
|------|-------|--------|----|
| 1 | CheA | CheW | |
| 2 | CheW | CheA | |
| 3 | CheA | - | *E. coli* |
| 4 | CheW | - | |
| 5 | - | CheA | |
| 6 | CheW1 | CheA | |
| 7 | - | - | |
| 8 | CheW1 | - | |
| 9 | Dodecin | Dodecin | |
| 10 | Dodecin | - | *H. salinarum* |
| 11 | MDH | MDH | |
| 12 | MDH | - | |
| 13 | NuoA | NuoB | |
| 14 | NuoB | NuoA | |
| 15 | NuoA | - | |

These proteins were cloned into the Y2H vectors and transformed pairwise into yeast (Table 3.1). The bait plasmid for *H. salinarum* CheA was not obtained in three cloning attempts and thus this bait was omitted. The transformants were assayed for growth on SC minus His (Figure 3.2).

The interaction between *E. coli* CheA and CheW was detectable in both bait-prey combinations. However, CheA as bait produced growth also in the absence of a prey, albeit in a reduced form (autoactivation). All halobacterial proteins except NuoA and NuoB led to growth of the yeast cells independent of the presence of a prey. NuoA and NuoB did not result in growth in any of the tested combinations.

Figure 3.2: **Yeast two-hybrid screen of a set of known interactors from *H. salinarum* and *E. coli*.** Selection on SC minus histidine after 48h of growth. The numbers reference the bait-prey pairs given in Table 3.1. (1, 2) *E. coli* cheA/cheW led to growth in both bait-prey combinations. (3) *E. coli* cheA also produced growth in absence of a prey. (6-12) All halobacterial baits, except nuoA and nuoB, produced growth independent of a prey. (13-15) nuoA, nuoB did not lead to growth in any combination.

Hence the interaction of *E. coli* CheA and CheW was detectable (with the limitation of the autoactivation by CheA), whereas for none of the halobacterial proteins a conclusion about an interaction could be drawn. For the Nuo proteins the screen was negative in both combinations. These proteins are membrane proteins and therefore the Y2H system is generally not the best tool to test them for interactions since Y2H displays interactions that take place in the yeast nucleus. However, for several membrane proteins interactions have been reported in Y2H studies (e.g. Uetz *et al.*, 2000).

All other tested halobacterial proteins were autoactivating as baits. Autoactivating baits, i.e. baits that activate transcription independent of the AD, thereby leading to false positives, are a common problem of Y2H screens (Bartel *et al.*, 1993). For this reason it is generally necessary to include a negative control for each bait, that means to test it without any prey. Nevertheless succesful screening of autoactivating baits is possible in many cases. The stringency of selection can be raised so that the transcription of the reporter gene caused by autoactivation is insufficient for growth, while the strong transcriptional activation caused by an interaction is still detectable.

## 3.2.2 Rescreening with higher stringency

To reduce the impact of the autoactivating baits, the screen was repeated with higher stringency (Figure 3.3). The stringency was increased by adding 5 mM, 15 mM, or 30 mM 3-aminotriazol (3-AT) to the medium for selection on SC minus His, or by selection on SC minus Ade. 3-AT is a competitive inhibitor of the HIS3 reporter gene
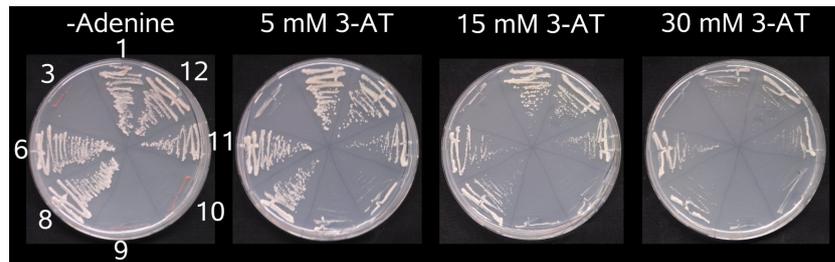
Figure 3.3: **Y2H screen with higher stringency.** Selection was performed on SC minus adenine or SC minus histidine in the presence of different concentrations of 3-aminotriazole (3-AT). The numbers reference the bait-prey pairs given in Table 3.1. (1, 3) Cells with *E. coli* cheA as bait could only grow if the prey cheW was present. For the halobacterial proteins growth was either suppressed completely (dodecin: 9, 10), or growth was possible independent of the presence of a prey (cheW1/cheA: 6, 8; MDH: 11, 12).

used for selection on SC minus His and thus suppresses growth when the reporter gene is only weakly activated. Selection on SC minus Ade is per se more stringent for the used Y2H system.

Both the addition of 3-AT as well as selection on SC minus Ade suppressed the autoactivation of *E. coli* CheA, while in the presence of the prey CheW growth was possible. That means this interaction could be clearly demonstrated in this assay. For the halobacterial proteins, growth was either suppressed completely or much slower but still possible, again independent of the presence of a prey. Once again the demonstration of an interaction between the halobacterial proteins was impossible.

### 3.2.3 Halobacterial proteins and yeast transcriptional activation

The results of the Y2H screens raised the question why most of the tested halobacterial proteins were such strong activators of transcription in yeast when fused to a DNA-binding domain. Ruden *et al.* (1991) demonstrated that around 1 % of short peptides coded by *E. coli* can activate transcription in yeast when fused to a DBD. All of these peptides bore an excess of acidic amino acids. Later studies confirmed that the acidic activation domain is a wide-spread feature of transcriptional activators in yeast and other eukaryotes (for review see Triezenberg, 1995).

The main adaptation of halobacterial proteins to their environment containing multimolar concentrations of salt (mainly K Cl and Na Cl) is an excess of acidic residues on the surface (see 1.1.4). All tested proteins (except for the Nuos) had a pI of 3.8 to 4.3. Hence it is not surprising that these proteins were strong autoactivating baits.

## 3.3 Conclusions

For all tested halobacterial proteins the Y2H screening failed, either due to autoactivation or just by a negative test. 82% of the proteins coded by *H. salinarum* have a pI between 3.5 and 5.5 (Tebbe *et al.*, 2005), so most proteins will probably be autoactivating baits. Even if stringent selection is able to suppress the autoactivation, the question remains how halobacterial proteins fold in the yeast nucleus. It has been described for several halophilic proteins that they need elevated salt concentrations to maintain their fold and their interactions (see for example Mevarech *et al.*, 2000).

Therefore, even if a test set of six proteins is not representative, these findings led to the conclusion that the Y2H assay is not suited to analyse the interactions of halobacterial proteins.

# 4 Affinity purification and mass spectrometry of halobacterial protein complexes

## 4.1 Introduction
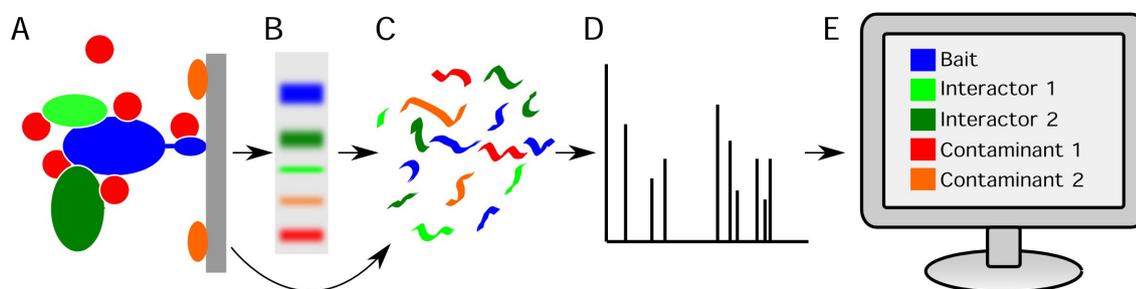


Figure 4.1: **Affinity purification of protein complexes combined with mass spectrometry.** **A** The protein of interest (the bait, blue) is expressed in-frame with an epitope tag. This tag can then be used as an affinity handle to purify the bait along with its interaction partners (the preys, green) on an appropriate affinity matrix (grey). Contaminants (red, orange) can bind unspecifically to the isolated proteins or the affinity matrix. **B** Optionally, the isolated proteins can be separated by gel electrophoresis. For samples with low complexity, gel separation may not be necessary so that the complete sample can be digested at once. **C** Isolated proteins or the whole complex are digested (usually using trypsin) to generate a mixture of peptides that can be identified by MS. **D** Normally the peptides are separated by reversed-phase chromatography followed by MS/MS analysis. **E** Database searching identifies the proteins present in the sample, and some evaluation is necessary to discriminate interaction partners and contaminants. Figure redrawn after Gingras *et al.* (2007).

The affinity purification of protein complexes combined with mass spectrometry (AP-MS, Figure 4.1), also called bait fishing, is a method used to enlighten protein complex composition. The protein complexes are isolated directly from cell lysate in one or more purification steps using affinity tags, allowing the analysis under near physiological conditions. Thereafter the purified proteins are identified by mass spectrometry. Commonly used tags in PPI studies are the TAP (tandem affinity purification) tag (Gavin *et al.*, 2002, 2006; Krogan *et al.*, 2006; Butland *et al.*, 2005), the Flag tag (Ho *et al.*, 2002), and the His tag (Arifuzzaman *et al.*, 2006).

Two major problems with AP-MS are transient interactions, which are too short-lived to allow copurification, and contaminants, which bind unspecifically to the proteins of interest or the affinity matrix. The first problem can be handled by modified purification protocols allowing shorter handling time (e. g. Gloeckner *et al.*, 2007) or cross-linking (e. g. Guerrero *et al.*, 2006) of the proteins. To tackle the second problem, either improved purification protocols are applied (usually tandem affinity purification protocols with two consecutive affinity steps, see Puig *et al.* (2001) for review), or relative quantitation can be used (e. g. Schulze and Mann, 2004; Ranish *et al.*, 2003) to discriminate interactors and background proteins in mass spectrometry.

Relative quantitation can be applied to discriminate true interaction partners from unspecific background, because specific interaction partners are enriched when fished by a bait protein as compared to a control without tagged bait. In contrast, the quantities of background proteins that bind to the affinity matrix should be independent of the presence of a tagged bait. An easy and reliable way to make mass spectrometric measurements quantitative is labelling with stable isotopes. The proteins from the bait sample are labelled with a certain stable isotope (e. g. $^{13}$C) and the proteins from the control sample with a different isotope (e. g. $^{12}$C). Thereby it is possible to measure the relative quantity of the peptides (and thus proteins) that are derived from the bait and the control sample, giving a measure if the identified protein was specifically bound to the bait or unspecific background (Figure 4.2). SILAC (**s**table **i**sotopic **l**abeling by **a**mino acids in **c**ell culture, Ong *et al.*, 2002) is a preferable labelling technique as it introduces the labels as early as possible in the course of the experiment, thereby reducing the risk of errors due to separate handling of sample and control.

An AP-MS method for halobacterial protein complexes should maintain high-salt conditions throughout the binding and washing steps, since halophilic proteins might lose their interactions when the salt concentration decreases (see Mevarech *et al.*, 2000, and references therein). Except for the His tag, none of the common tags can be used when working with multimolar salt concentrations. Purification of His-tagged proteins from *H. salinarum* was hampered by low yields and high background levels in several cases (personal communication of Christoph Schwarz, MPI of Biochemistry, Martinsried). Probably the highly negative surface charges of halobacterial proteins result in unspecific binding to the Ni$^{2+}$-matrix.

The groups of Moshe Mevarech and Jerry Eichler purified proteins from the halophilic archaeon *Haloferax volcanii* by using the cellulose-binding domain (CBD) from the
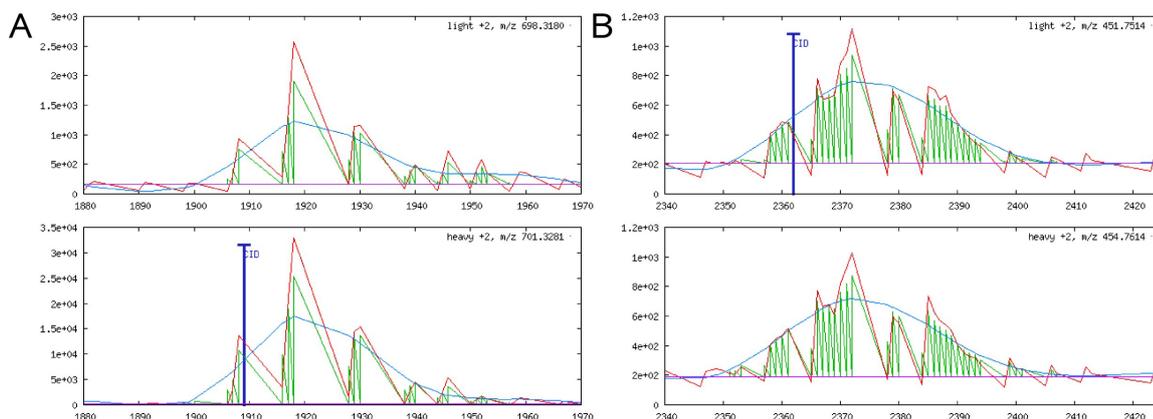
Figure 4.2: **Extracted ion chromatograms from a SILAC AP-MS experiment. A** Extracted ion chromatograms (XIC) of the light (upper panel) and heavy (lower panel) isotopic form of a peptide from an interaction partner. Note the different scaling on the Y-axis. **B** The respective chromatograms for a peptide from a background protein. The red line gives the intensity as measured by the mass spectrometer. The blue line is a smoothed curve derived from the raw signal. The horizontal purple line indicates the background level. The blue bar (CID) indicates the MS/MS scan that identified the respective peptide. The area under the elution curve (green, calculated from the average of raw and smoothed curve) is proportional to the amount of peptide. The quotient of the areas of both isotopic forms (SILAC ratio) allows to draw conclusions if an identified protein was specifically bound by the bait or unspecific background. The graphs were generated by ASAPRatio (Li *et al.*, 2003).

CipB protein from *Clostridium thermocellum* as affinity tag (Ortenberg and Mevarech, 2000; Irihimovitch *et al.*, 2003; Irihimovitch and Eichler, 2003). Cellulose-binding domains are essential components of the majority of cellulose degrading enzymes (Tomme *et al.*, 1998). They have an affinity for cellulose, but do not exhibit any hydrolytic activity. The CBD from *C. thermocellum* consists of 167 amino acids and has a molecular weight of 18 kDa (Morag *et al.*, 1995), making it a rather large tag. The use of a large tag results in a higher risk of unwanted effects on the bait protein like preventing protein complex formation due to steric hindrance. However, it was demonstrated that the binding of the CBD to cellulose is insensitive to high salt concentrations. Therefore this study focused on the use of the CBD as affinity handle for the purification of protein complexes from *H. salinarum*.

The purification procedure for halophilic protein complexes developed in this study allows reasonable short handling time (roughly 30 min between cell lysis and elution), which should help to prevent loss of short-lived interactions. Therefore no cross-linking was applied. A TAP procedure could not be developed due to the lack of appropriate affinity tags and proteases that can be used under high-salt conditions.

Hence SILAC was applied to identify specific interaction partners regardless of high levels of background proteins.

In a first step, vectors were created that enable the expression of CBD-tagged proteins in *H. salinarum*. In parallel, *E. coli*-expressed CBD was used to establish a basic purification protocol. This protocol was then adapted for the expression in and purification from *H. salinarum*. With such a protocol in hands, the focus was shifted to the identification of interaction partners. First, MALDI TOF peptide mass fingerprinting was used for protein identification. Since this method showed major problems with the tested samples, the more powerful ESI LC MS/MS technique was chosen as identification method. Finally, the procedure was modified to make use of SILAC for discrimination of interaction partners from background. For the determination of SILAC ratios, the tool ASAPRatio (Li *et al.*, 2003) embedded in the Trans-Proteomic Pipeline (Keller *et al.*, 2005) was used.

## 4.2 Results and Discussion

### 4.2.1 Construction of vectors

Vectors needed to be constructed that allow the expression of CBD-fused bait proteins in *H. salinarum* (Figure 4.3 A, B). The commonly used vectors for transformation of *H. salinarum* are not maintained as plasmids but integrated into the genome. This fact can be used to achieve the expression of the tagged bait protein by integrating an appropriate plasmid into the genome at the site of the bait gene (Figure 4.3 C, D). Thereby the expression of C-terminally tagged bait protein under the control of the endogenous promoter, N-terminally tagged bait protein under control of an exogenous promoter, or both versions at once in one cell can be achieved. In this study, only the expression of the C-terminally tagged bait (single tagging) and of C- and N-terminally tagged bait at once (double tagging) was done.

For the expression of the N-terminally tagged construct the modified ferredoxin promoter PrR16 is used (Danner and Soppa, 1996). This promoter leads to high expression levels so that most proteins are overexpressed when under its control. Bait overexpression can be problematic since it might lead to misfolding and thereby to the association with non-physiological interaction partners. Furthermore, the stoichiometry between bait and interactors is lost. On the other hand, bait overexpression might
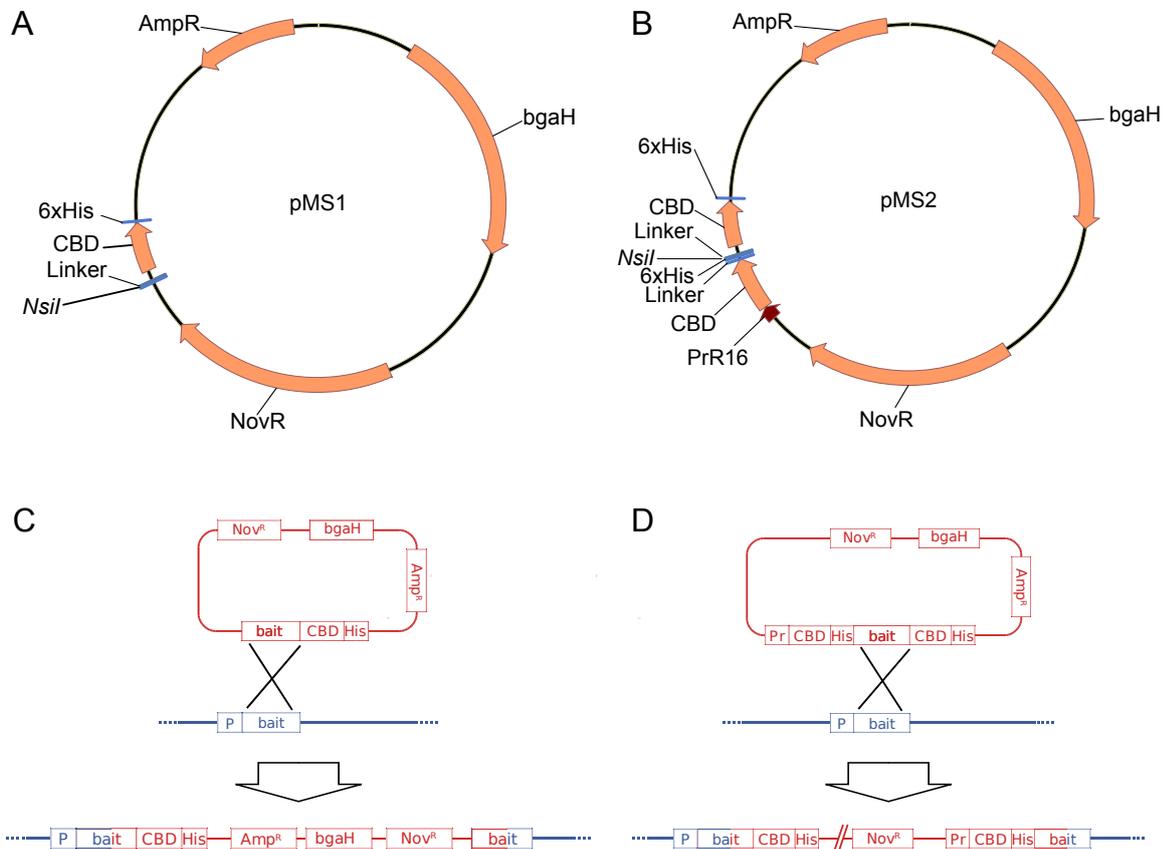
Figure 4.3: **Plasmids for expression of CBD fusion proteins in *H. salinarum* (In-Fusion cloning). (A, B)** Schematic representation of the plasmids pMS1 and pMS2. Both plasmids contain a pUC origin (not indicated) and an ampicillin resistance (AmpR) for amplification in *E. coli*. The novobiocin resistance (NovR) and β-galactosidase (*bgaH*) are for selection of transformants in *H. salinarum*. Baits are cloned into the NsiI site. Baits are fused to the CBD(s) with a short linker in between (IGAVEER, this is the linker of the two β-sheets in halobacterial dodecin). His tags (6 x His) are added to the bait-CBD fusion to allow easy detection in western blots. Downstream of the fusion protein is a transcriptional terminator (not indicated) from the *H. salinarum bop* gene. **C** Integration of a pMS1 construct (red) into the genome (blue) leads to the expression of the bait C-terminally fused to the CBD under control of the bait's endogenous promoter. **D** The integration of pMS2 constructs leads both to the expression of the bait C-terminally fused to the CBD under control of the bait's endogenous promoter and the expression of the bait, N-terminally fused to the CBD under control of the promoter PrR16 (a modified ferredoxin promoter developed by Danner and Soppa, 1996).

also be an advantage because it allows to work with smaller cultures (especially for low-abundance proteins). Even though the preys are not overexpressed, the higher concentration of an overexpressed bait can help to fish weak interactors by shifting the dissociation equilibrium towards associated (of course, this can again lead to the identification of non-physiological interactions).

All expression plasmids were derived from pVT (Tarasov *et al.*, 2008) as described in 2.4.1. The first generation of vectors (Figure 4.3) was designed for In-Fusion Cloning™ (2.2.8). The plasmids can be linearised by cutting with the restriction endonuclease NsiI. The PCR-amplified bait, bearing 15 bp overhangs homologous to the ends of the linearised vector, can then be recombined into the linear vector to create the desired plasmid.

These vectors were used in the experiments performed for establishing the method, but overall the efficiency of In-Fusion Cloning™ was low and most baits needed several attempts for succesful cloning. The problem were not clones with the wrong insert, but no clones at all. The reason for the low success rate is not known, but other people in the lab had problems as well. Possibly, the In-Fusion system does not work well with the GC-rich DNA from *H. salinarum*. To improve cloning, the plasmids were converted to Gateway™ destination vectors (see 2.2.9 for Gateway cloning). The resulting vectors, pMS3 for single tagging and pMS4 for double tagging, are shown in Figure 4.4 A and B. From these vectors appropriate SILAC control plasmids were derived by removing the CBD(s) and designated pMS5 and pMS6.

## 4.2.2 The purification procedure

### 4.2.2.1 Expression in *E. coli*

To have a fast and effective expression system to test the cellulose-based purification procedure, the CBD was cloned into the expression vector pET15 (Novagen) and transformed to *E. coli* strain BL21. The generated expression strain was then used to establish a basic CBD purification protocol.

Starting point for protocol development were the methods used by the groups of Mevarech and Eichler (Ortenberg and Mevarech, 2000; Irihimovitch and Eichler, 2003), and the protocol for the formerly commercially available Cbind™ kit (Novagen). The major development steps are given in the protocols Eco1-3 (2.4.3.1). Corresponding results are shown in Figure 4.5. The most important improvements were the use of
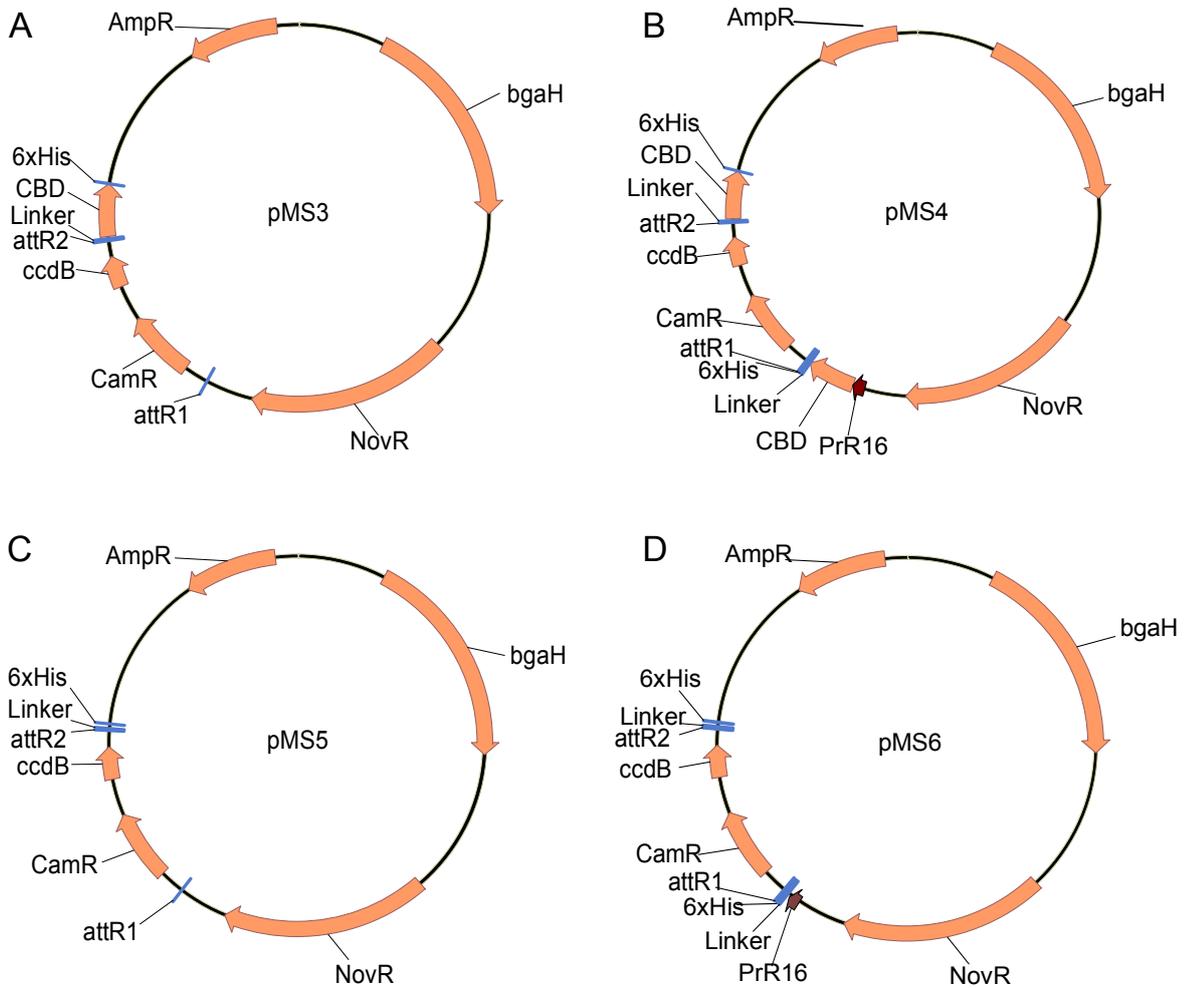
Figure 4.4: **Plasmids for expression of CBD fusion proteins in *H. salinarum* (Gateway cloning). A, B** The plasmids for single and double tagging, respectively. **C, D** The corresponding SILAC control plasmids. See the legend of Figure 4.3 for general description. The Gateway conversion added the *attR*1 and *attR*2 sites for recombination, a chloramphenicol resistance (CamR), and the cytotoxic *ccdB* gene for counterselection after LR recombination.
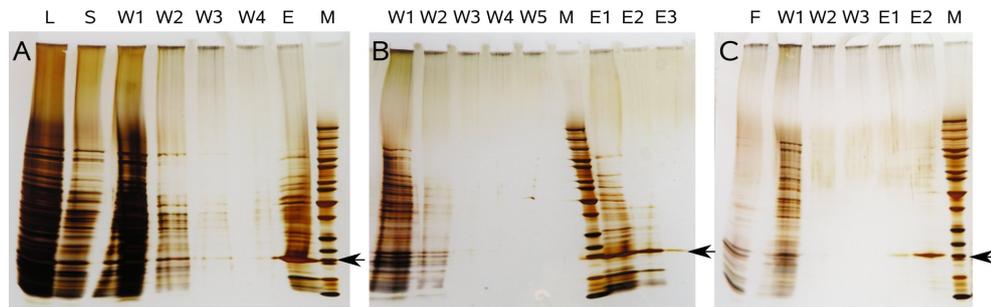
Figure 4.5: **Establishing the purification protocol with *E. coli*-expressed CBD. A** The first protocol achieved some enrichment of the CBD (arrow) by purification on amorphous cellulose, but the amount was low and the level of background high. **B** The second protocol used columns instead of batch purification and ethylene glycol for elution. The amount of bound CBD was slightly bigger (in A 15 of 40 µl were loaded to the gel, in B 15 of 100 µl), but the background level was still too high. **C** The major improvement was achieved with the use of microcrystalline cellulose (Avicel PH-101) instead of amorphous cellulose. L lysate, S supernatant, W1-W4 washing fractions, E1-E3 elution fractions, M marker.

columns instead of batch purification, elution with ethylene glycol, and the use of microcrystalline cellulose (Avicel PH-101, Fluka). Protocol 3 was the starting point for purification of CBD-tagged proteins expressed in *H. salinarum*.

### 4.2.2.2 Expression in *H. salinarum*

CheA (OE2415R), CheW1 (OE2419R), Dodecin (OE3073R), and RpoK (OE2638F, a subunit of the RNA polymerase) were cloned into the vectors pMS1 and pMS2 (Figure 4.3), and the resulting plasmids transformed into *H. salinarum* R1. Expression of the fusion proteins was checked by western blotting with an antibody against the His-Tag(s) included in the vectors. Figure 4.6 A shows a representative result. For all tested clones a band at the expected size (bait + CBD) is visible. For baits in pMS2 (double tagging) there was always a second band with approximately 20 kDa higher molecular weight present. Mass spectrometric analysis has revealed that this band corresponds to the bait bearing the CBD both at the N- and the C-terminus. The reason for expression of this fusion protein is unknown. It might be due to additional integration of the plasmid into the genome at a different site. The western blots confirmed the expression of single- and double-tagged CBD-fusion proteins of all tested baits.

The expression strains were used for affinity purification according to protocol Hsa1 (2.4.3.2). A representative result is shown in Figure 4.6 B. Binding of the bait protein
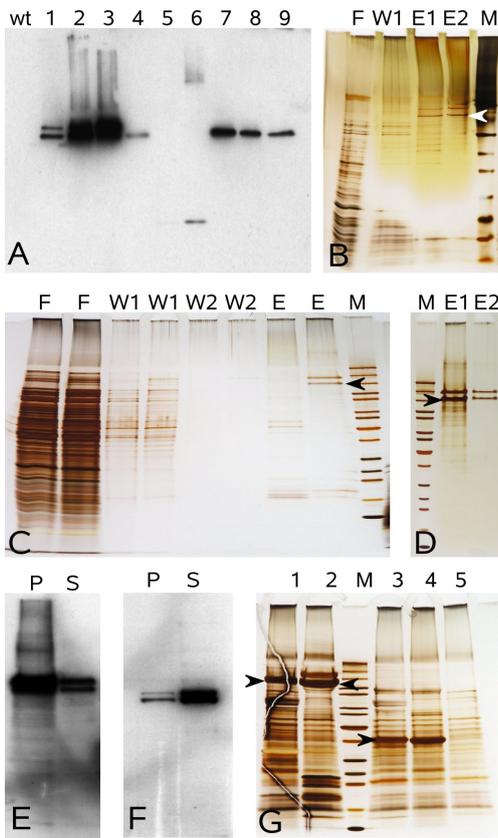
Figure 4.6: **Purification of CBD fusion proteins from *H. salinarum*.** **A** Western blot to detect the expression of CBD fusion proteins of nine clones. wt wildtype; 1-3 pMS2::*cheA*; 4,7-9 pMS1::*cheA*; 5,6 pMS1::*dodecin*. **B** Affinity purification of pMS2::*cheA* clone 3 according to protocol Hsa1. F Flow-through; W1 first washing fraction; E1,E2 eluted proteins; M marker. **C** The same clone (right lane each) and wildtype purified according to protocol Hsa2 (larger culture volume). **D** The same clone eluted according to protocol Hsa3 to test if elution was incomplete. **E, F** Western blots of the pellet (P) and the supernatant (S) after centrifuging the lysate at 4 °C (E) or 18 °C (F). **G** Eluted proteins after purification according to protocol Hsa4. 1 pMS1::*cheA*; 2 pMS2::*cheA*; 3 pMS1::*cheW*; 4 pMS2::*cheW*; 5 wt. The arrowheads point to the expected bands.

was demonstrated (two bands as double tagging was used), but the amount of protein was very low. Figure 4.6 C shows purification from a larger culture of the same expression strain (protocol Hsa2), and surprisingly the amount of bound protein did hardly increase. Earlier experiments with *E. coli*-expressed CBD have shown that the used amount of cellulose has a much higher binding capacity (compare with Figure 4.5 B and C, where only 15 of 100 µl eluate were loaded to the gel; after purification from *H. salinarum* total eluted protein was loaded). In D, another experiment with d-cheA is shown, which was done to check if eventually elution of the protein was incomplete. The amount of eluant (ethylene glycol) was raised and again two elution steps were performed (protocol Hsa3). Indeed, the second elution step revealed some residual bound protein, but the main part (and all coeluting proteins) was found in the first elution fraction. A western blot (Figure 4.6 E) of the pellet and supernatant after centrifugation of the cell lysate (performed as described in protocol Hsa1) revealed that the majority of the tagged protein was in the pellet. It was observed that during the centrifugation step at 4 °C salt of the CFE buffer precipitates, which might cause

the proteins to coprecipitate. The blot was repeated with all steps carried out at room temperature, and the major part of the protein was detected in the supernatant (Figure 4.6 F). The affinity purification protocol was modified (centrifugation at 18 °C, protocol Hsa4) to avoid protein precipitation. Figure 4.6 G shows affinity purifications of single and double tagged CheA and CheW1.

The main development step was to do the centrifugation without cooling, which seemed to lead to protein precipitation under the initial conditions. Several more parameters have been varied (e.g. the detergent concentration in the buffer, number and volume of the washing steps, amount of culture, cellulose and lysate), but the results did not change significantly (data not shown). Hence, the final CBD affinity purification protocol (see 2.4.4) was developed from protocol Hsa4 with only minor modifications.

### 4.2.3 Identification by mass spectrometry

#### 4.2.3.1 Identification by MALDI TOF PMF
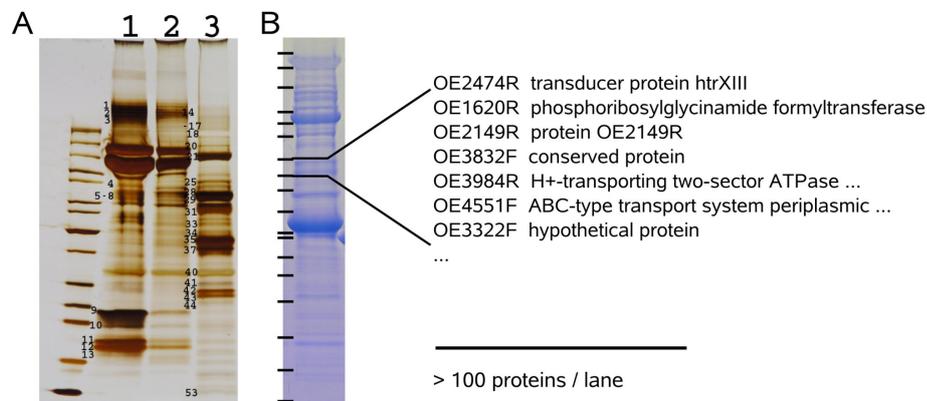


Figure 4.7: **Identification of proteins after affinity purification. A** Silver-stained gel of eluates after affinity purification used for MALDI TOF PMF analysis. 1 pMS2::*cheA*; 2 pMS2::*cheA* after different lysis protocol (not analysed); 3 pMS2::*rpoK*. The numbers indicate the slices excised for analysis. Lane 3 was cut completely into slices, so the numbers are just "landmarks". **B** AP of pMS2::*cheW1* used for LC MS/MS analysis. The slices prepared for mass spectrometry analysis and some identified proteins are indicated.

Affinity purifications with single and double tagged bait proteins were performed and eluted proteins identified by MALDI TOF PMF (matrix-assisted laser desorption-ionisation time-of-flight mass spectrometry peptide mass fingerprinting). Figure 4.7 A

Table 4.1: **Proteins identified by MALDI TOF PMF.**

| Slice | Protein |
|---|---|
| 1-4,7 | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA |
| 4-7 | (OE2205F) chitinase (EC 3.2.1.14) |
| 5,8 | (OE2201F) chitinase (EC 3.2.1.14) |
| 8 | (OE2206F) probable chitinase (EC 3.2.1.14) |
| 9,10 | (OE2419R) purine-binding chemotaxis protein cheW1 |
| 11,12 | (OE4643R) conserved protein |
| 15,17,18, 20-22 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A1 |
| 18 | (OE4759F) cell surface glycoprotein precursor |
| 24 | (OE1737R) dnaK-type molecular chaperone hsp70 |
| 25 | (OE5212F) SMC-like protein sph1 |
| 27 | (OE2205F) chitinase (EC 3.2.1.14) |
| 27 | (OE4741R) DNA-directed RNA polymerase (EC 2.7.7.6) chain B1* |
| 29 | (OE4742R) DNA-directed RNA polymerase (EC 2.7.7.6) chain B2 |
| 32 | (OE2648F) conserved protein |
| 34 | (OE4739R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A2 |
| 35 | (OE2631F) DNA-directed RNA polymerase (EC 2.7.7.6) chain D |
| 37-39 | (OE3542R) protein OE3542R |
| 42 | (OE1279R) DNA-directed RNA polymerase (EC 2.7.7.6) epsilon chain |
| 46 | (OE2628F) ribosomal protein S4 |
| 47 | (OE3817R) ribosomal protein S19.eR |
| 50 | (OE3491R) heat shock protein homolog |

Proteins with Mascot score > 47 (p<0.05) are shown. * The RNA polymerase chain B1 was only identified with a Mascot score of 32 (not significant).
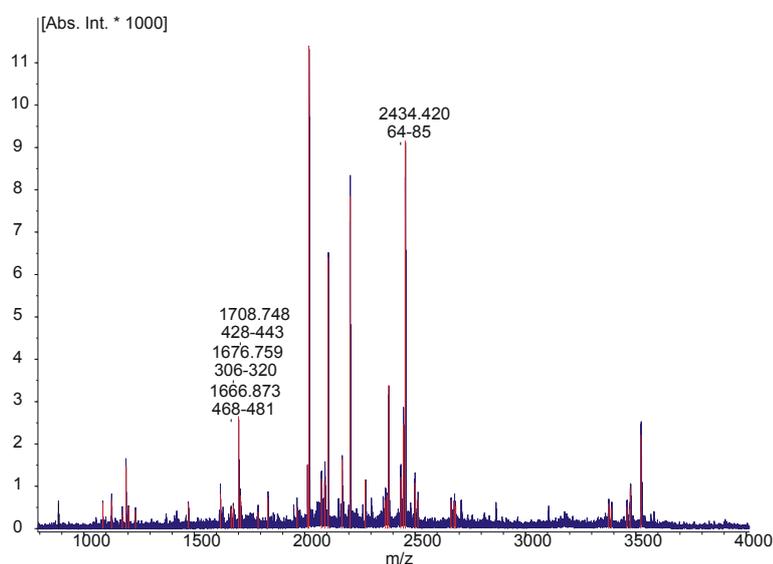


Figure 4.8: **Spectrum from MALDI TOF PMF.** The peaks labelled with their mass match to peptides from RpoB1. The numbers below the mass indicate the respective peptide position in the protein sequence.

shows a representative result of such experiments with double tagged CheA and RpoK. From lane 1 selected bands were excised and subjected to MALDI TOF analysis. Lane 3 was cut completely into slices and analysed. Identified proteins are shown in Table 4.1 (see Supplementary Table S1 for the complete list of identifications).

These results demonstrates that MALDI TOF PMF enabled the identification of several copurified proteins including expected interaction partners: CheW1 with CheA and five RNA polymerase subunits with RpoK. Noteworthy, the RNA polymerase epsilon chain was also identified as a prey of RpoK. This protein is special because it has no homologs in any other sequenced genome.

However, also the limitations of this identification method became visible: (1) Most of the spectra contained a lot of unassigned peaks, suggesting additional proteins in this slice. But only in few cases the reliable identification of a protein mixture was possible. An example for this is slice 27 (Figure 4.8), where the RNA polymerase chain B1 was only identified with a Mascot score of 32 (not significant). This was probably due to another prominent protein present in this slice (the chitinase OE2205F). (2) All of the identified subunits of the RNA polymerase have a molecular mass of more than 25 kDa. No subunit with lower molecular weight was identified with significant score. Problems with the identification of small proteins are intrinsic in MALDI TOF PMF, since smaller proteins lead to less identifiable peptides, and therefore give a less characteristic fingerprint.

### 4.2.3.2 Identification by LC MS/MS

To improve the identification of small proteins and protein mixtures, additional experiments were performed and the eluted proteins subjected to nano-LC MS/MS analysis. Figure 4.7 B shows an affinity purification with double tagged CheW1. The whole lane was cut into 17 slices, each containing several proteins. Compared to Figure 4.7 A, bigger slices were chosen, since the presence of multiple proteins in one sample is unproblematic for LC-MS/MS analysis. By this approach, more than 100 proteins were identified from this lane, and the identification of small proteins was strikingly improved. The experiment with rpoK was repeated, and LC-MS/MS analysis identified five RNA polymerase subunits with a molecular weight of less than 25 kDa (subunits P, K, F, E1, H) as well as all bigger subunits. (In this experiment, the overall identification rate was rather low due to instrumental problems. A repetition would possibly

have identified even more subunits, but experiments with rpoK were discontinued in favour of the Che proteins.)

Hence LC-MS/MS was the method of choice for this project. However, the more than 100 proteins identified from one lane are definitely not all interaction partners of the respective bait. Some protocol optimisation like increasing the detergent concentration or adding ethylene glycol to the washing buffer was tried to reduce the amount of background proteins (gels not shown). Less background was always accompanied by the loss of interactions (assayed by the CheA-CheW1 interaction) or even by leaking of bait from the column. Therefore the high background levels were accepted and SILAC applied to discriminate true interaction partners from background proteins.

## 4.2.4 SILAC: Discrimination of interaction partners from background

### 4.2.4.1 Direct bait fishing

The direct bait fishing procedure (Figure 4.9 A) was the first protocol used to specifically identify interaction partners. This protocol requires for each bait expression strain a corresponding control strain which is transformed with a similar plasmid like the bait strain, just without CBD. Those control strains are necessary because the overexpression of the baits in pMS4 or the chromosomal integration of the plasmids might influence the expression of other proteins, which would result in shifted SILAC ratios when compared to wildtype. One disadvantage is that the control strains do not express the CBD, and therefore proteins that bind to the CBD will appear as interaction partners. Such proteins, however, will always be the same and thus can be easily identified in a control experiment (see 6.2.1).

It was found that in many cases direct bait fishing allows a clear differentiation between specifically enriched proteins (potential interaction partners) and background proteins. A representative example is given in Figure 4.10 A. Four potential interaction partners of CheC3 are clearly separated from a high number of putative background proteins.
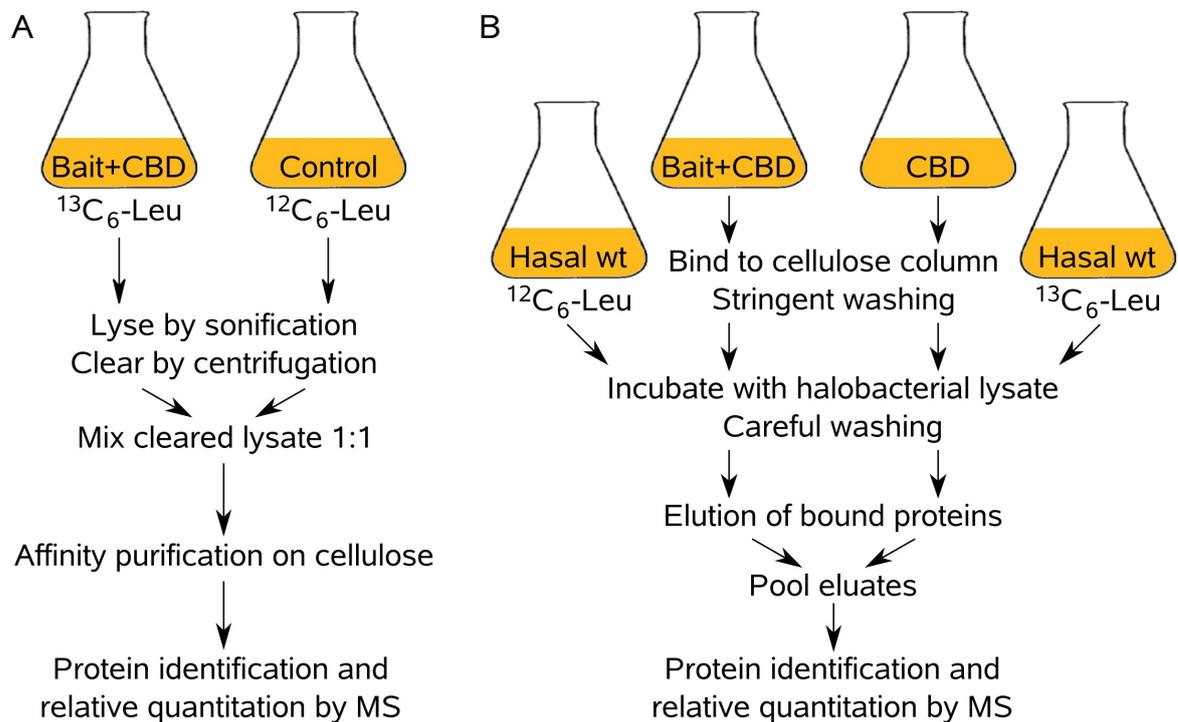
Figure 4.9: **Schematic of purification procedures applying SILAC. A** Direct bait fishing. The bait expression strain (transformed with the bait's pMS3 or pMS4 construct) and the control strain (transformed with the bait's pMS5 or pMS6 construct) are grown in synthetic medium containing $^{13}$C-leucine or $^{12}$C-leucine, respectively. The lysate from both strains is mixed and purification done on one cellulose column. **B** Indirect bait fishing. The bait expression strain (transformed with the bait's pMS4 construct) and the control strain (expressing the plain CBD) are grown in complex medium. Bait and CBD are bound to separate cellulose columns and stringently washed in order to remove all proteins except bait or CBD. The columns are incubated with halobacterial lysate from cells grown in SM containing $^{12}$C-leucine (bait) or $^{13}$C-leucine (pMS4), respectively. After elution, the eluates are pooled.

### 4.2.4.2 The exchange problem

In a few cases certain expected interaction partners showed an SILAC ratio close to one (e. g. CheW1 when fished with CheA, Figure 4.10 B). This was even more surprising as these proteins were identified with very high sequence coverage with the corresponding baits (and with very low coverage or not at all with other baits), indicating a specific enrichment of these preys. The reason for this is probably exchange of the prey protein from the bait and control culture after mixing the lysate and before the unbound proteins are washed away.
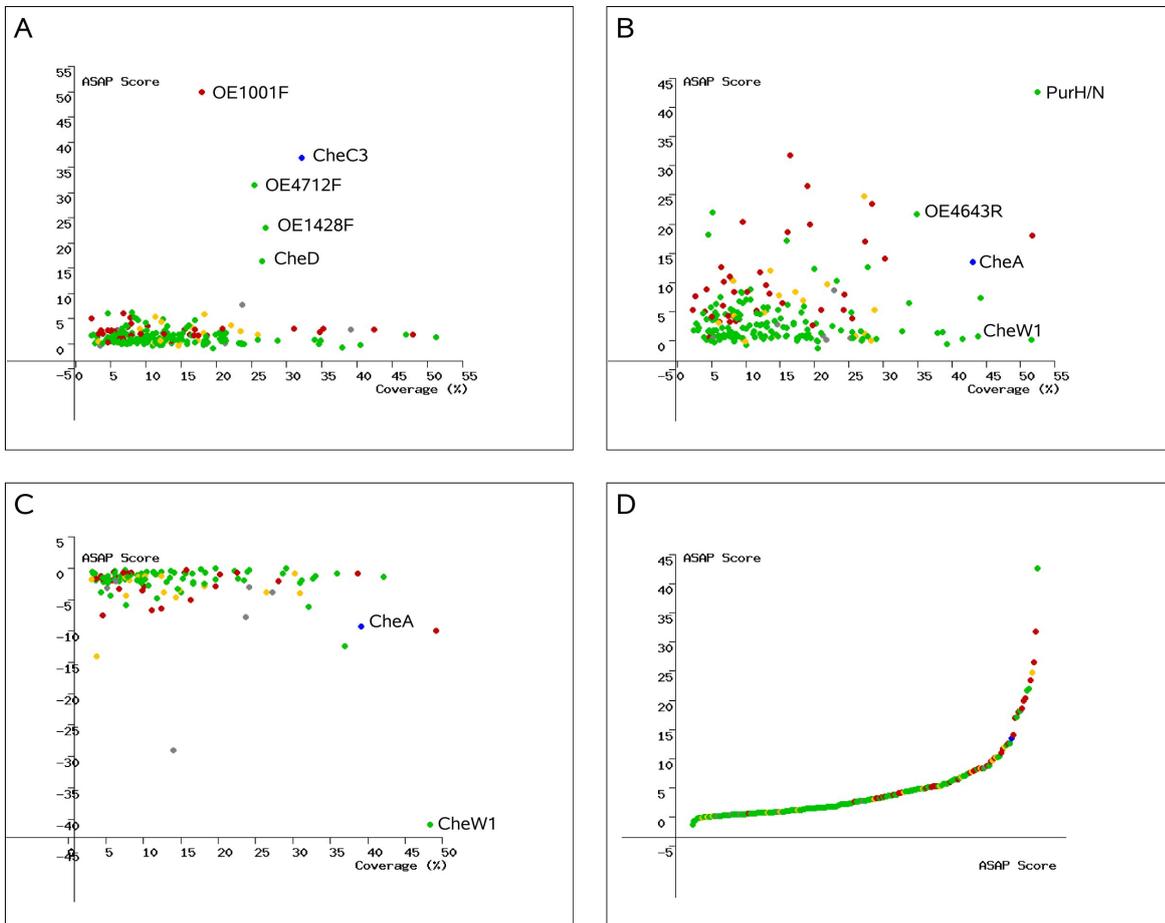
Figure 4.10: **Discrimination of interaction partners and background proteins with SILAC.** The plots show the distribution of the proteins identified in bait fishing experiments. Each dot represents one protein. Colours: Blue bait, green cytosolic proteins, red transmembrane proteins, yellow proteins with lipid anchor, grey proteins considered as background (see Table 6.3). The Y-axis represents the ASAP Score (a symmetrical representation of the SILAC ratio, see 2.4.10). In direct bait fishing experiments a high ASAP score means a high probability for an interaction, in indirect bait fishing this is indicated by a strong negative ASAP score (reversed labelling, see text). The X-axis shows the sequence coverage of the protein identification (A-C) or the ASAP Score in increasing order (D). (Note that the bait protein is not correctly quantified, so its position on the plot is meaningless.) **A** Direct bait fishing with CheC3. A clear separation of a few potential interaction partners from several putative background proteins can be seen. **B** Direct bait fishing with CheA. No clear separation of interaction partners and background is possible. Note that the expected interaction partner CheW1 was identified with high sequence coverage, but the SILAC ratio is close to 1 (exchange problem). Most of the other proteins with a high SILAC ratio are halobacterial transducers and associated proteins. **C** Indirect bait fishing with CheA. In this experiment CheW1 has a strong negative ASAP score, indicating an interaction with CheA. **D** Different representation of the direct bait fishing experiment with CheA. The distribution of the SILAC ratios of the identified proteins makes it hard to define a threshold for interaction vs. background.

### 4.2.4.3 Indirect bait fishing

To tackle this problem, the indirect bait fishing method (Figure 4.9 B) was developed. Here the bait and control purification are performed in separate columns, preventing any exchange between the proteins. Furthermore, the prey proteins are purified from genetically unmodified cells. This is an advantage because it is possible that the chromosomal integration of the tagging vector at the locus of the bait protein interferes with prey protein expression (interacting proteins are often found directly adjacent or even in one operon, see Dandekar *et al.*, 1998).

In the first indirect bait fishing experiments labelling was done as in direct fishing: $^{13}$C-leucine for the bait, and $^{12}$C-leucine for control. In these experiments some proteins, that were identified as interaction partners in direct fishing, came up with a strong negative ASAPScore (data not shown). That means the SILAC ratio was shifted towards the control and not the bait. The reason for this was probably residual bound prey protein after the stringent washing step, before the lysate was applied to the column (remember that the bait for indirect fishing is expressed in complex medium and therefore always $^{12}$C-labelled). To circumvent this problem, the labelling was reversed: in indirect experiments the bait was incubated with $^{12}$C-labelled and the control with $^{13}$C-labelled lysate. The drawback of this approach is that proteins bound to the CBD in the control culture will appear as interactors. But again these proteins should always be the same and thus easily be identified in control experiments.

With indirect bait fishing, the CheA - CheW1 interaction could be clearly demonstrated (Figure 4.10 C). However, direct fishing (Figure 4.10 B) revealed PurH/N and OE4643R as potential interaction partners of CheA. Both were later confirmed as interactors by reciprocal fishing (i. e. using a prey as bait in an additional experiment). These proteins were not even identified in the indirect experiment. So the indirect method has also intrinsic disadvantages compared to the direct fishing. Table 4.2 gives a comparison of the advantages and disadvantages of both methods. As a consequence both methods should be applied to each bait to identify as many interactions as possible.

In principle also an intermediate method would be possible: direct fishing with the purification on two columns. However, for this approach the development of an adequate SILAC control would be rather difficult (such control strain should have roughly the same genomic modifications at the bait locus like the bait strain, and

Table 4.2: **Advantages and disadvantages of the bait fishing methods.**.

| Method | Pro | Contra |
|---|---|---|
| direct | Protein complexes assemble *in vivo* | Fishing from genetically modified cells |
| | Bait and control handled together | Exchange between bait and control possible |
| indirect | Fishing from wildtype cells | Complexes with tagged bait not assembled *in vivo* |
| | No exchange possible | Bait and control handled separately |

it should produce roughly the same amount of CBD for binding to the column). Furthermore, this approach would lose the advantage of early sample pooling to reduce handling errors, and would not allow to fish from unmodified cells. Hence only both "extreme" methods were developed and the intermediate approach was rejected.

#### 4.2.4.4 Thresholds

The evaluation of the SILAC AP-MS experiments required the definition of some kind of threshold, if an identified protein should be considered as contaminant or as putative interaction partner. However, this was hampered by the following aspects:

In some experiments, no clear separation of the putative interaction partners from the crowd of background proteins was seen, but a more continuous distribution of SILAC ratios (Figure 4.10 D). This might be due to "sticky" baits that tend to bind unspecifically to several proteins. This effect was also observed with proteins that bind to membrane proteins (like CheW1 and CheA, which bind to membrane-bound transducers), possibly due to indirect interactions via the membrane. These experiments made it difficult to define a fixed threshold.

Other studies apply a statistical measure like the Grubbs outlier test (Selbach and Mann, 2006) or the *z*-test (Dobreva *et al.*, 2008) to measure if the deviation of a SILAC ratio is significant. Those tests presuppose that the parameter to be tested follows a standard normal distribution in the test population (or at least in the fraction considered as background). This prerequisite is not met for most datasets produced in this study (checked by Shapiro-Wilk normality test), so these tests are not applicable.

To have nonetheless the possibility to interpret the data, a new measure called "association score" and a simple benchmark called "association rating" were defined. The examination of several experiments with a clear separation (like Figure 4.10 A) revealed that the main part of proteins had an ASAPScore of less than three. Some proteins had a score between three and five, and few between five and seven. Then

followed a gap with hardly any proteins, and then the clearly enriched proteins with ASAPScores of ten or more. For sticky baits or membrane associated proteins these values were slightly higher (see Figure 4.10 D for an example). Most proteins had an ASAPScore below five, quite a lot between five and seven, and only for ASAPScores of nine or above the amount of proteins dropped significantly. The association score is basically the ASAPScore adjusted for the different isotopic labelling (positive for both direct and indirect experiments), and it is adjusted if a bait was sticky (including membrane association). In such case, the association score is diminished by two (corresponding to one grade of the association rating, see below). A bait is considered as sticky if 20 or more proteins have an ASAPScore above 3. From the association score the association rating was defined as shown in Table 4.3. In the following study only proteins with an association rating of high or very high were considered as interaction partners.

Without a doubt such an arbitrarily defined benchmark will neither give the right rate in all cases nor can it measure the "quality of its guess" like a probability-based score. But for the following reasons it is sufficient for the purpose of this study:

Table 4.3: **Definition of association rating.**

| Association Score | Association Rating |
|:-----------------:|:------------------:|
| $\leq 3$ | none |
| $>3, \leq 5$ | low |
| $>5, \leq 7$ | medium |
| $>7, \leq 9$ | high |
| $>9$ | very high |

First, almost all interactions discussed in chapter 6 were found with an association score of ten or higher, and only very few interactions were in the grey zone between seven and nine. Hence in most cases there is a clear discrimination, and to a certain extent a different threshold would not affect the results. Second, PPI analysis should be considered mainly as hypothesis generation, because those experiments alone can never elucidate the biological significance of an interaction. (The best statistical test will not identify an interaction as false positive which does never take place *in vivo* because the presumed interaction partners are not co-localised). Since follow-up experiments are required anyhow an elevated error-rate is tolerable in such a study.

## 4.3 Conclusions

With a cellulose-binding domain from *Clostridium thermocellum* as tag it was possible to establish an affinity purification procedure for protein complexes from *H. salinarum*.

Both the expression of the bait under its endogenous promoter (close to physiological expression levels) as well as bait overexpression is possible. The developed protocol maintains high-salt concentrations throughout the whole purification process. The processing time of 30 min is reasonably short, which increases the chance to copurify low-affinity interaction partners.

Because the single-step purification procedure lead to a high level of contaminant proteins, relative quantitation by isotopic labelling was applied to enable the discrimination of specific interaction partners from background. Two protocols were developed for this: direct and indirect bait fishing. Both of these procedures have intrinsic advantages and disadvantages, and both were demonstrated to work for certain interactions and fail for others. That means both direct and indirect fishing should be performed for each bait.

Overall, the established procedure enabled the copurification of several expected interaction partners.

# 5 The bioinformatics environment

## 5.1 Introduction

The scope of the project made a bioinformatics infrastructure necessary to manage baits and plasmids, store information about experiments, assist data analysis, and make results accessible. Furthermore, it was desirable to have the ability to integrate the experimental results with existing knowledge, for example on protein function or protein properties. Hence the developed system was integrated into HaloLex (http://www.halolex.mpg.de/; Pfeiffer *et al.*, 2008a). HaloLex is the department's information system for genomic and experimental data on *H. salinarum*, other haloarchaea, and by now also other prokaryotes.

The information is stored in three databases: the Bait DB, the Experiment DB, and the Result DB. The databases are accessed via web applications implemented as CGI scripts in Perl. The functionalities of the implemented system will be presented in the next sections.

## 5.2 The databases

To store the required information, three databases were created: the Bait DB, storing information on baits, vectors, primers, and plasmids; the Experiment DB for information about experiments, protocols, MS runs, and the generated result files; and the Result DB that collects the generated results. The implementation was done as relational databases in MySQL. Figure 5.1 shows a relationship scheme for the databases.

### 5.2.1 Bait DB

The central tables of this DB are bait and bait_simple. bait stores only the bait name and sequence. bait_simple stores additional information: which gene the bait corresponds to (parent), the chromosomal position (nstart, nstop), if it is the complete
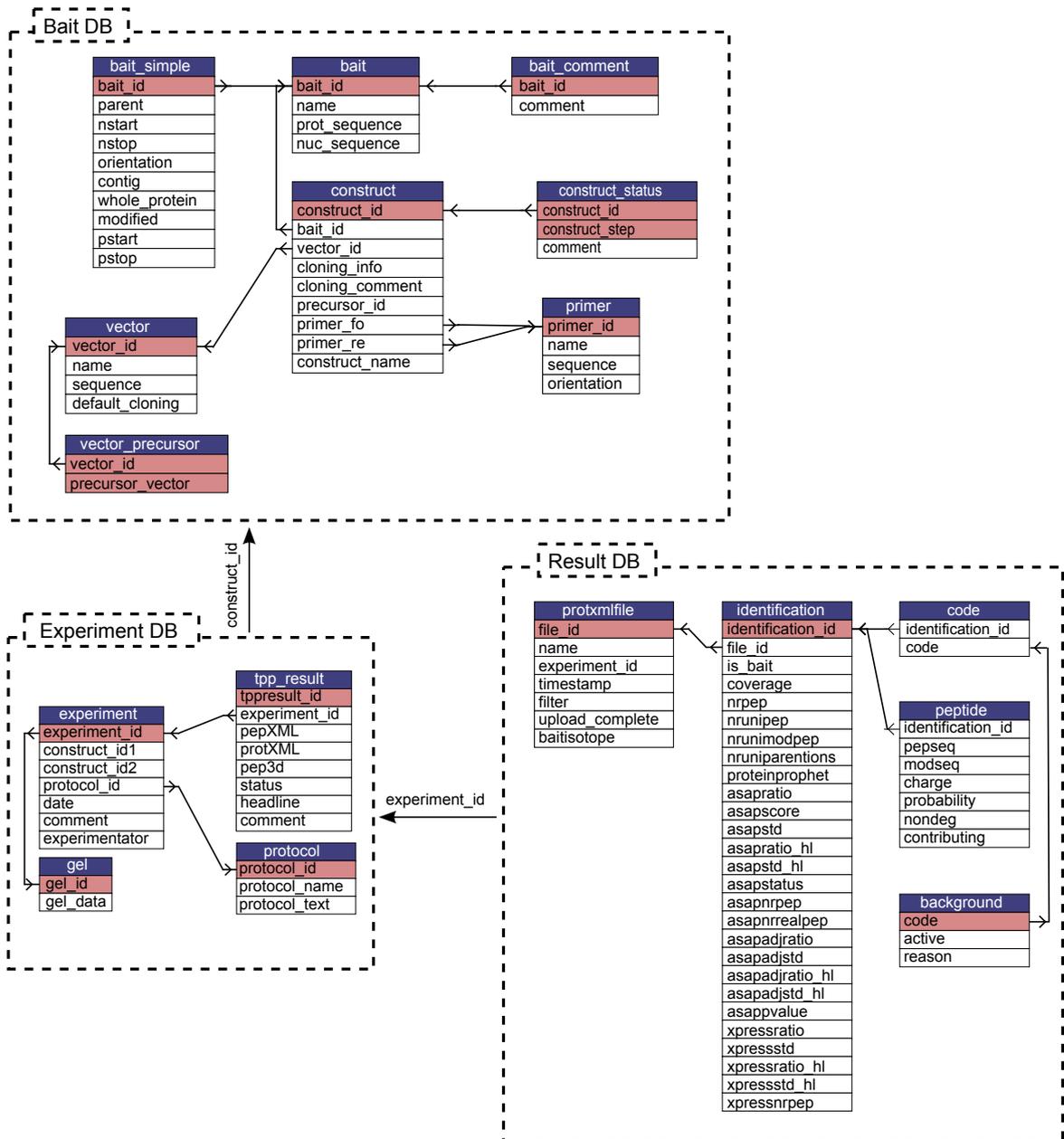
Figure 5.1: **Relationship schema of the project's databases.** The table names are highlighted blue, the primary keys of the tables red. For description of the databases see text.

protein, contains mutations, and the position of the parent on the bait (the bait can contain additional sequences like an extra linker or tag). With this table baits corresponding to a complete halobacterial protein, or to a part of a protein, can be represented. The tables bait and bait_simple can also account for mutated baits and baits containing additional sequence. Not representable with these tables are artificial baits (e.g. domain fusions). (For this a table bait_complex was foreseen, but not implemented as there was no need for such baits).

Each bait can have one or more constructs (i.e. plasmids). Constructs are just stored as a reference to the bait (bait_id) and to a vector (vector_id), and the used cloning method. With these informations, the nucleotide sequence of the construct can be assembled if required. A construct can also store links to two cloning primers. Each construct can be linked to $n$ cloning steps (stored in construct status). Vectors are stored in the table vector with name, sequence, and the vectors default cloning method (this allows automation when entering constructs). The additional table vector_precursor is important for gateway cloning: it links gateway destination vectors to all possible entry vectors. Again, this allows automation when entering constructs, because the software can guess which construct was the entry clone used in LR recombination. Overall this structure allows the reconstruction of every cloning step performed in the lab: the entry clone refers to the primers used for PCR amplification of the bait and thus takes into account any sequence variations introduced with the primers. All follow-up constructs refer to the entry clone so that they inherit such variations.

### 5.2.2 Experiment DB

The main table of the experiment DB is experiment. An experiment refers to one or two constructs from the Bait DB (the plasmid(s) for transformation of the strain(s) used in the experiment). Additionally, the date of the experiment, the experimentator (both together make it easy to find the experiment in a lab journal), and a comment can be entered. Each experiment refers to one protocol (stored as HTML-formated text in protocol), and one gel image (stored in gel). An experiment can be linked to one or more results of the trans-proteomic pipeline (TPP). These are referenced in the table tpp_result just by the file names of the generated files, and furnished with additional information (status shows for example if the result was already manually

verified and should be used for further analysis, and a comment – divided into headline and comment – can be added).

### 5.2.3 Result DB

The Result DB stores the output generated by the TPP. The information is parsed from the resulting protxml file. The table protxmlfile takes some general information on the uploaded file: the file name, the identifier of the corresponding experiment, the time of upload (timestamp), the filtering criteria applied during upload (e. g. only proteins above a certain identification probability), the isotopic state of the bait (light or heavy), and a tag which is set after the upload is completed to prevent the work with incomplete files (for example, if the connection was aborted during upload).

The table identifications lists the identified proteins. The columns are identification_id (unique id), file_id (refers to the corresponding result file), is_bait (if the protein was bait in this experiment), coverage (sequence coverage), nrpep, nrunipep, nrunimodpep, nruniparentions (the number of total peptide identifications including multiple identifications of the same peptide, the number of unique peptides discarding modifications and charge state, the number of unique peptides considering modifications, but not charge state, and the number of unique peptides considering both modifications and charge state), proteinprophet (the identification probability), and the results of the two quantification tools ASAPRatio and XPress.

The table code links an identification to one or more codes (i. e. gene identifiers like OE1234F; more than one code can occur for duplicated gene regions). Peptide stores the peptides which were identified, and background is a user-editable table with the purpose to classify certain proteins as contaminants.

## 5.3 The applications

The tools to access and modify the content of the databases were implemented as web applications with CGI scripts in Perl. With this approach the use of the tools is platform-independent and does not require the installation of additional software. The scripts make use of the module CGI.pm (written by Lincoln D. Stein), BioPerl (Stajich *et al.*, 2002), and, for graphical output, of the GD library.

Figure 5.2: **The Bait Browser.**

## 5.3.1 Bait management

For entering bait data to the database, the bait editor was developed. It allows entering new baits (proteins from *H. salinarum* or part of it) in a semi-automated manner. After identifying a bait by its code (e. g. OE1234F), the relevant information (gene name, protein and nucleic acid sequence, chromosomal position) are automatically fetched from the HaloLex database. If desired, the bait sequence can be manipulated to account for mutations, and a comment can be entered. Vectors used for cloning are entered through the Vector Editor.

The Bait Browser (Figure 5.2) is the web interface to access the bait database. It provides two working modes: In "info" mode, it displays a list of the baits in the database. In "LIMS" mode, additionally information about the generated plasmids and their cloning status, including the used primers, is given. This additional information can be entered and edited through the Bait Browser. Some features to assist the laboratory work have been implemented: the cloning of the plasmids is reconstructed *in silico*, so that the sequence of all generated plasmids is available. Various output formats for the sequence can be chosen. Most useful was the GenBank format that

includes features (like coding sequences) required to draw plasmid maps. The Bait Browser offers the simulation of restriction digests of the plasmids. That is a fast and convenient way to evaluate the analytical restriction digest usually done after cloning. A third feature is assistance in the generation of cloning primers. The Bait Browser can automatically propose a forward and reverse primer for amplification of the bait, and calculates the primer's melting temperature ($T_M$).

## 5.3.2 Experiment management
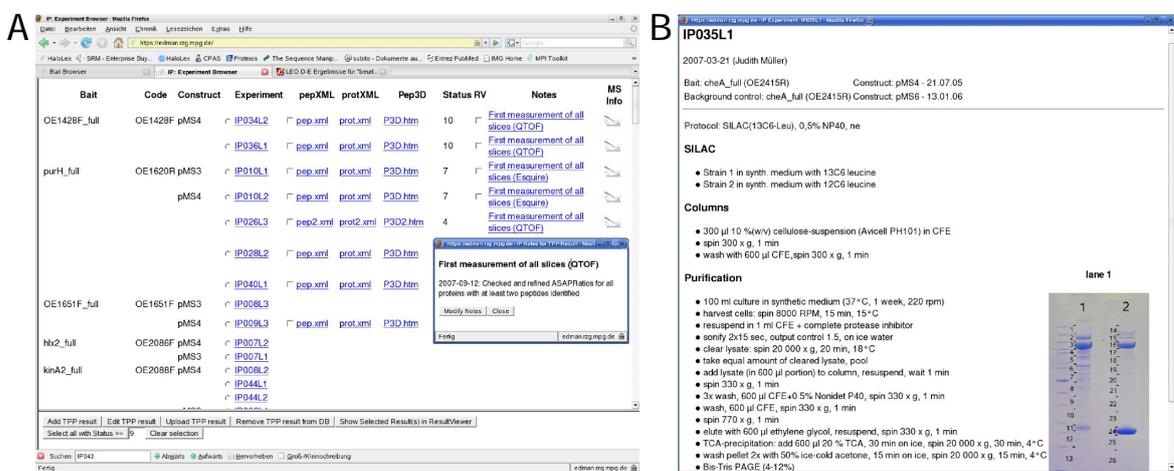


Figure 5.3: **The Experiment Browser. A** The main page. **B** Details page of an experiment.

Information about affinity purification experiments, including the used bait, protocol, date, and an image of the resulting gel, can be entered through the experiment editor. The protocols to which the experiments refer are entered and modified in a separate protocol editor.

This information can be accessed through the Experiment Browser (Figure 5.3). The Experiment Browser provides a list of the experiments with links to experimental details, the generated result files, comments on the results, and quick access to the measuring parameters of the mass spectrometer. Overall, this provides the possibility to easily track back all steps that were performed to obtain the experimental results. One experiment corresponds to exactly one gel lane, which is cut into a variable number of slices. Each experiment can have multiple result files (the output of the TransProteomic Pipeline) to account for repeated measurements or database searches

Figure 5.4: **The Result Viewer.**

with different parameters. The connection to the result files is managed through the experiment browser.

The experiment browser is also the entrance for the Result Viewer. One or more results of the experiments can be selected for display and further analysis in this tool.

### 5.3.3  Result evaluation

The purpose of the Result Viewer is to integrate information from the HaloLex database (e.g. protein function, transmembrane helices, ...) and experimental results. Therefore the experimental results are parsed from the protXML file generated by the Trans-Proteomic pipeline and written to the Result DB by the Upload Result tool. This tool allows the application of some filtering criteria like a threshold for the identification probability. It is also possible to update a result which is already in the database, for example after the refinement of the quantitation ratio.

The Result Viewer displays a table containing the proteins identified in certain experiments (Figure 5.4). It is possible to have either one table per experiment or to merge several experiments into one table. The merged table can be used to compare

results from different experiments or to identify proteins that were identified in a suspicious high number of experiments (promiscuous preys). The table of the Result Viewer is highly configurable. The user can choose which columns should be displayed in order to identify a correlation between the experimental results and certain properties of the identified proteins (function class, domains, size, ...).

The Result Viewer can also generate scatter plots of the results from one experiment. This is useful to get an overview of the distribution of SILAC ratios of the identified proteins, or to see correlations between some properties of the identified proteins. The variables to be plotted on the x- and y-axis can be chosen by the user, and different colour schemes can be applied. Examples for plots generated by the Result Viewer can be seen in Figure 4.10.

## 5.4 Conclusions

The developed environment is a valuable resource for both performing the AP-MS experiments as well as accessing and evaluating the results. The Bait Browser is most useful for laboratory work since it allows to keep track of the cloning progress, and it assists the work with some special features like construct sequence assembly, primer design, and restriction analysis. The Experiment Browser makes the raw experimental results accessible via links to the TPP result files, and it also provides required background information (protocols, instrumental settings). Thus the main task of the Experiment Browser is to point to the results and make transparent how they were obtained. The Result Viewer assists the evaluation of the results by providing additional information on the identified proteins, by giving the possibility to compare results from multiple experiments, and by graphical representation.

# 6 Chemotaxis protein interaction network

## 6.1 Introduction

### 6.1.1 The Che system: a specialised two-component system for taxis signalling

The two-component system mediating tactic responses is commonly referred to as the Che system. The Che system in Archaea is basically similar to the corresponding bacterial system, and the proteins involved are homologs of the respective bacterial proteins. However, there are minor and major variations in this system, both between different bacterial clades and between Bacteria and Archaea (for review see Szurmant and Ordal, 2004). A universal scheme of the Che system is shown in Figure 6.1.

The overall workflow of taxis signalling can be divided into four steps – signal reception and transduction, excitation, adaptation, and signal termination. These steps are described in the next sections. An additional section is about fumarate, which is also involved in flagellar motor switching.

#### 6.1.1.1 Signal reception and transduction

Signals are recognised and transduced by a certain class of proteins called halobacterial transducer proteins (Htrs) in *H. salinarum* or methyl-accepting chemotaxis proteins (MCPs) in other archaea and bacteria. The transducers control the activity of the histidine kinase CheA and thereby the level of CheY-P, the output molecule of the Che system. MCPs/transducers usually contain a cytoplasmic signalling region where the histidine kinase CheA and the coupling protein CheW bind, a methylation/demethylation region that is critical for adaptation, and one or two HAMP (histidine kinase, adenylyl cyclase, methyl-accepting chemotaxis protein, and phosphatase) domains. The HAMP domains convey the signal of ligand binding from the sensor domain or protein to the output module. The way the signal is transduced

across the membrane is still not completely understood - most experimental data support the sliding-piston model, in which ligand binding induces a piston-like sliding of the signalling helix toward the cytoplasm (Falke and Hazelbauer, 2001).

In contrast to many eukaryotic receptors that dimerise upon ligand binding, prokaryotic transducers form stable dimers even in the absence of ligand (Milligan and Koshland, 1988). The receptors in prokaryotes are not distributed evenly around the cell body but form large clusters where thousands of sensory complexes are thought to come together (Maddock and Shapiro, 1993; Sourjik and Berg, 2000; for review see Kentner and Sourjik, 2006; Hazelbauer *et al.*, 2008). Whereas the clustering of receptors as essential feature of prokaryotic signal processing is widely accepted, the arrangement of receptors in the clusters is still under discussion. The crystal structure of the cytoplasmic fragment of the *E. coli* serine receptor *Tsr* shows trimers of receptor dimers (Kim *et al.*, 1999), which can be the basic building blocks of receptor arrays. Hexagonal cluster structures can be formed by connecting the trimers of dimers by CheA dimers (Shimizu *et al.*, 2000). However, newer findings suggest that CheA and CheW



Figure 6.1: **General model of prokaryotic chemotaxis systems.** Biochemical processes in the two-component chemotaxis pathway are shown. Hexagons represent response regulator domains. Components found throughout all species are in red, components found in almost all species are in orange; components which are only present in certain species are in yellow. Figure and caption slightly modified from Szurmant and Ordal (2004).

are not required for cluster formation (Kentner and Sourjik, 2006), and the crystal structure of the cytoplasmic part of a *Thermotoga maritima* receptor (Park *et al.*, 2006) revealed *hedgerows* of dimers instead of trimers of dimers. The different models of receptor cluster formation are reviewed in Kentner and Sourjik (2006).

### 6.1.1.2 Excitation

The input, usually sensed by the receptors, influences the autophosphorylation activity of the histidine kinase CheA. After autophosphorylation of a particular histidine

residue, the phosphoryl group is immediately transferred from CheA to the response regulator CheY. Phosphorylated CheY (CheY-P) is the output signal for the flagellar motor. Hence CheA integrates the different stimuli to generate an unambiguous output to the flagellar motor. In *E. coli* (Borkovich *et al.*, 1989), *S. meliloti* (Schmitt, 2002), *R. spheroides* (Shah *et al.*, 2000), and *H. salinarum* (Rudolph and Oesterhelt, 1996) attractants decrease and repellents increase CheA activity, in *B. subtilis* CheA regulation is reversed (Garrity and Ordal, 1997).

CheA consists of five domains (P1-P5) (Bilwes *et al.*, 1999). The P1, or Hpt (histidine phosphotransfer), domain contains the histidine residue that is phosphorylated. P2 is the docking site for CheY and CheB, which receive the phosphoryl group from CheA-P. P3 is the dimerisation domain. Dimerisation of CheA is crucial since CheA autophosphorylates *in trans*. On P4 (HATPase_c: Histidine kinase-like ATPases) ATP binding and catalysis occurs, and P5 (also called CheW domain due to homology to this protein) is where CheA binds the receptors and the coupling protein CheW.

For activation of CheA at the receptors, the coupling protein CheW is required. CheW proteins are found in all bacterial and archaeal species with a chemotaxis system (judged by the presence of CheA and receptors). *H. salinarum* contains two CheW paralogues. The exact role of these is not clear, but deletion of CheW1 and CheW2 results in different phenotypes (Aregger, 2003).

As implied above, the output of the chemotaxis signalling system is CheY-P. Elevated CheY-P levels cause CCW flagellar rotation (smooth swimming) in *B. subtilis* and CW flagellar rotation (tumbling) in *E. coli*. Therefore the overall outcome is the same in both species – attractants cause smooth swimming, repellents cause tumbling and reorientation. In *H. salinarum*, CheY-P causes reversals and is required for CCW swimming as were shown by the constantly CW swimming phenotype of CheY and CheA deletion strains (Rudolph and Oesterhelt, 1996).

In flagellated bacteria, the target of CheY-P is the protein FliM (Welch *et al.*, 1993), which builds together with FliN and FliG the flagellar motor switch complex. The binding site of CheY-P is the highly conserved N-terminal peptide of FliM (Bren and Eisenbach, 1998). Archaea, although using CheY-P as switch factor in a similar fashion, do not possess FliM homologs. Also no equivalent to the CheY-P binding peptide has been identified. The site of interaction of CheY-P in Archaea is unknown (Nutsch *et al.*, 2003; Szurmant and Ordal, 2004; Ng *et al.*, 2006).

### 6.1.1.3 Adaptation

Prokaryotic taxis requires a memory to decide whether during a move the conditions improved or worsened. That means that the actual stimulus strength is permanently compared to the stimulus strength as it was before (Koshland, 1977). This is achieved by the adaptation system(s). The best understood adaptational mechanism is the methylation system of CheR and CheB, but other systems, e. g. involving CheC and CheD (Muff and Ordal, 2007), exist.

The methyltransferase CheR transfers methyl groups from S-adenosylmethionine to certain glutamate residues in the methylation region of the receptors (Kehry and Dahlquist, 1982; Nowlin *et al.*, 1987). Methylation of receptors increases and demethylation decreases the activity of the signalling complex (Ninfa *et al.*, 1991; Borkovich *et al.*, 1992). CheB is a methylesterase that demethylates the same residues which are methylated by CheR. The methyl groups are released as methanol. Whereas CheR is constitutively active, the activity of CheB is regulated via phosphorylation of its response regulator domain by CheA: CheB-P is 100fold as active as unphosphorylated CheB (Lupas and Stock, 1989). That means that CheB forms a feedback loop between CheA and the receptors. In *H. salinarum* and *E. coli*, but not in *B. subtilis*, CheB functions also as receptor glutamine deamidase. Glutamines have roughly the same effect on the signalling complex activity as methyl esterified glutamates (Rollins and Dahlquist, 1981; Kehry *et al.*, 1983; Koch, 2005; Koch *et al.*, 2008). In *B. subtilis*, this reaction is catalysed by CheD (Kristich and Ordal, 2002).

Several bacteria like *B. subtilis* and *H. pylori* contain CheV, a two-domain protein consisting of a N-terminal domain homologous to CheW and a C-terminal response regulator domain (Fredrick and Helmann, 1994). In CheV, phosphorylation of the response regulator domain seems to affect the conformation of the coupling domain, thereby decoupling CheA and the receptors. Thus the signalling complexes of ligand-bound receptors can reassume their prestimulus activity (Karatan *et al.*, 2001). A gene coding for CheV was not yet found in any archaeal genome.

In *B. subtilis*, a third way of adaptation is described involving the proteins CheC and CheD (Muff and Ordal, 2007). CheC is a CheY-P phosphatase (Szurmant *et al.*, 2003), CheD catalyses the deamidation of glutamines at the receptors (Kristich and Ordal, 2002). Both proteins were shown to form a heterodimer (Rosario and Ordal, 1996). This interaction increases the CheC phosphatase activity (Szurmant *et al.*, 2004) and

inhibits the deamidation activity of CheD (Chao *et al.*, 2006). CheY-P was shown to stabilise the CheC:CheD interaction, and thus CheC and CheD form a third feedback loop to the receptors (Muff and Ordal, 2007). *H. salinarum* possesses both a CheD and CheC homologs. However, receptor deamidation activity in this organism has been demonstrated for CheB and not CheD (Koch, 2005; Koch *et al.*, 2008), and the function of CheD is unclear. It remains to be elucidated if other adaptation systems than the CheR/CheB methylation system play a role in archaeal chemotaxis.

### 6.1.1.4 Signal termination

The chemotaxis system must be able to respond to changing stimuli within seconds to effectively direct the movement towards the best places. This is in part achieved by a short half-life of CheY-P, which is, depending on the species, in the range of a few seconds (Rudolph *et al.*, 1995) to almost one minute (Hess *et al.*, 1988; Stock *et al.*, 1988). CheY was found to actively catalyse autodephosphorylation (Lukat *et al.*, 1990; Silversmith *et al.*, 1997).

To further accelerate signal removal, several chemotactic organisms express CheY-P phosphatases. In the γ- and β-proteobacteria, this is done by the protein CheZ (Hess *et al.*, 1988; Zhao *et al.*, 2002). In other chemotactic eubacteria and all chemotactic archaea, CheC in combination with CheD seems to be involved in CheY-P dephosphorylation (Szurmant *et al.*, 2004). Furthermore, in *B. subtilis* the flagellar motor switch protein FliY, a distinct homolog of CheC and the CheX protein (a CheY-P phosphatase present for example in *T. maritima*), was found to have CheY-P phosphatase activity (Szurmant *et al.*, 2004).

In α-proteobacteria like *S. meliloti*, a third mechanism of signal termination is present. A second CheY acts as phosphate sink and thereby possibly assists the phosphate removal from the "main" CheY (Sourjik and Schmitt, 1998).

### 6.1.1.5 Fumarate as switch factor

The first evidence for fumarate as switch factor was found in *H. salinarum*, where it restored wild-type behaviour in a straight-swimming mutant (Marwan *et al.*, 1990). It was demonstrated that fumarate is released to the cytoplasm from membrane-bound pools after light stimulation (Marwan *et al.*, 1991; Montrone *et al.*, 1993). Almost to the same time, Barak and Eisenbach (1992b) observed that fumarate and CheY

are required for switching in cell envelopes of *E. coli* and *S. typhimurium*. Using an *E. coli* strain with increased cytoplasmic fumarate concentrations due to a deletion of succinate dehydrogenase (SDH), which acts on fumarate, a correlation between the cytoplasmic fumarate level and both the switching frequency and the fraction of cells rotating clockwise could be established (Montrone *et al.*, 1996, 1998; Prasad *et al.*, 1998). Prasad *et al.* (1998) also demonstrated that the target of fumarate is the switch and not CheY, and that it acts, at least in part, by lowering the free energy difference between the CW and the CCW state of the motor. In a recent work, fumarate reductase (FRD) was identified as target of fumarate at the motor in *E. coli* (Cohen-Ben-Lulu *et al.*, 2008). This enzyme, otherwise functioning in anaerobic respiration, interacts with the flagellar motor switch protein FliG. However, *H. salinarum* does neither code for FRD (which is mainly found in obligate or facultative anaerobic bacteria) nor for FliG. So in this species fumarate must act by a different, till now unknown, mechanism. The excitation part of fumarate signalling, i.e. when and how it is released, has not yet been identified either.

### 6.1.2 The components of the Che system of *H. salinarum*

Regarding the coded proteins, the taxis pathway of *H. salinarum* is more similar to that from the gram-positive soil bacterium *B. subtilis* than to that from the gram-negative enterobacterium *E. coli*. Functionally, however, this is not always the case. As mentioned above, CheA in *H. salinarum* is activated by repellents, similar to *E. coli* and different from *B. subtilis*.

In the genome of *H. salinarum*, 18 Htrs have been identified due to homology of their signalling region to eubacterial MCPs (Ng *et al.*, 2000; Pfeiffer *et al.*, 2008b). The transducers either include an own sensing domain so that they act as receptors and transducers in one molecule, or they interact with separate receptor proteins (Kokoeva and Oesterhelt, 2000; Kokoeva *et al.*, 2002). Five of the Htrs were predicted to contain no transmembrane domain, so they are thought to recognise their signals intracellularly. The other transducers contain two or more transmembrane helices and recognise their signals at the membrane or extracellularly.

Until now the function of only eight of these transducers could be assigned: The transducers HtrI and HtrII are coupled to the sensory rhodopsins I and II (SRI, SRII), respectively (Spudich and Spudich, 1993; Zhang *et al.*, 1996). The sensory rhodopsins
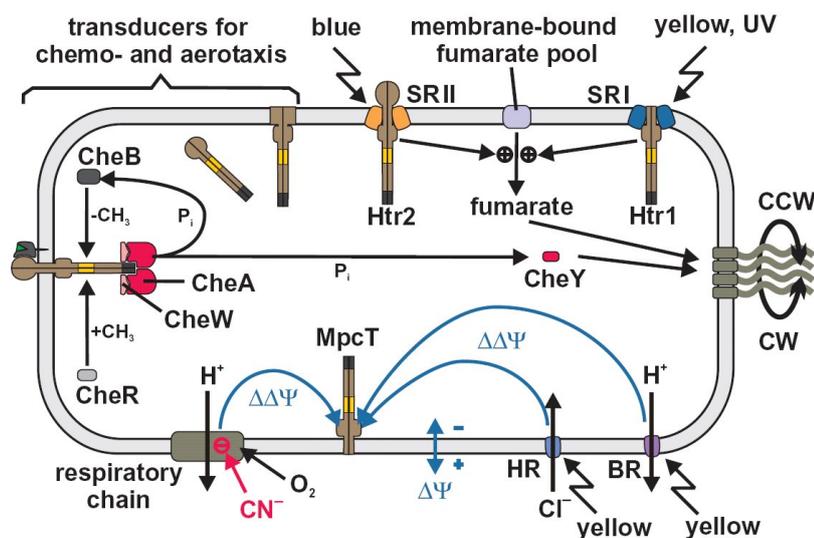
Figure 6.2: **Overview of the Che system of *H. salinarum*.** Htrs are depicted as dimers (brown) and are shown in their expected topology. The Htr regions involved in adaptation (yellow) and in signal relay (dark grey) to the flagellar motor via Che proteins are indicated. The proteins CheC1-3 and CheD are omitted and CheW1 and CheW2 not differentiated since their role is unclear. Htr1 and Htr2 transduce light signals via direct interaction with their corresponding receptors, SRI and SRII. Repellent light signals mediated by SRI and SRII elicit the release of the switch factor fumarate from a membrane-bound fumarate pool. MpcT senses changes in $\Delta\Psi$, generated via light-dependent changes in ion-transport activity of BR and HR. Signalling via MpcT occurs either in the absence of oxygen or in the presence of cyanide. Both conditions inhibit the respiratory chain and produce a decreased level of membrane energisation (low $\Delta\Psi$). Figure and part of caption taken from Koch and Oesterhelt (2005).

are retinal proteins which sense the intensity of light at three wavelengths: SRI senses orange light as attractant and UV light as repellent stimulus, SRII is sensitive for blue light, which is also a repellent stimulus. Hou *et al.* (1998) have demonstrated that HtrII also acts as chemotransducer for serine taxis. BasT (Htr3) mediates taxis towards the **b**ranched **a**nd **s**ulfur containing amino acids leucine, isoleucine, valine, methionine, and cysteine (Kokoeva and Oesterhelt, 2000), and CosT (Htr5) towards **c**ompatible **os**molytes of the betaine family (Kokoeva *et al.*, 2002). Both the soluble HemAT (Htr10) and the membrane bound Htr8 sense oxygen (Brooun *et al.*, 1998; Hou *et al.*, 2000) – the first one produces a phobic reaction, the second one an attractive response. MpcT (Htr14) senses changes in the membrane potential, and therefore gives feedback about the energy supply of the cell (Koch and Oesterhelt, 2005). Finally, Car was found to be an intracellular sensor for arginine (Storch *et al.*, 1999).

Table 6.1: **Functions of the Che proteins of *H. salinarum*.**

| Protein | Demonstrated / expected functions in *H. salinarum* | Functions in other organisms | Deletion phenotype | | | |
|---|---|---|---|---|---|---|
| | | | Spont | Photo | Swarm | CCW |
| CheA | Phosphorylation of CheY and CheB | Phosphorylation of CheY and CheB | − − −[1] | − − −[1] | − − −[1] | − − −[1] |
| CheW1 | Coupling of CheA to receptors | Coupling of CheA to receptors | 0[2] | 0[2] | −[2] | nd |
| CheW2 | | | 0[2] | − − −[2] | − − −[2] | nd |
| CheY | Essential for switching and CCW swimming | Switching/CCW (CW) rotation in Bsu (Eco) | − − −[1,3] | − − −[1,3] | − − −[1,3] | − − −[1,3] |
| CheB | Receptor demethylation, deamidation | Receptor demethylation, in Eco also deamidation | 0[3]/++[1] | −[3]/NA[1] | − −[1,3] | +[3]/0[1] |
| CheR | Receptor methylation | Receptor methylation | − − −[3] | − − −[3] | −[1] | −[1] |
| CheC1 | CheY-P phosphatase? | CheY-P phosphatase, CheD inhibition (feedback) | −[3] | −[3] | −[3] | −[3] |
| CheC2 | | | 0[3] | 0[3] | 0[3] | −(−)[3] |
| CheC3 | | | − − −[3] | − −(−)[3] | −[4] | −(−)[3*] |
| CheD | Enhancer of CheC(s)? | Receptor deamidase and enhancer of CheC in Bsu, receptor deamidase and methylesterase in Tma | − − −[4] | − − −[4] | − −[4] | − −[4] |

The phenotype was analysed for Spont: spontaneous switching; Photo: photophobic response (increase of switching after repellent light stimulus); Swarm: swarm ring formation; CCW: Fraction of counterclockwise swimming cells. The rating is by comparison to wild-type: 0 like wt; − switching/ photophobic response/ swarming reduced; − − strongly reduced; − − − (almost) no switching/ photophobic response/ swarming; + switching/ photophobic response/ swarming increased; ++ strongly increased. (nd not determined; NA not applicable). Functions in other organisms are thought to be universal, unless certain organisms are indicated (Eco: *E. coli*, Bsu: *B. subtilis*, Tma: *T. maritima*). The questionmark behind some expected functions in *H. salinarum* indicates that these are just predictions by homology to the respective protein from *B. subtilis* without experimental evidence.
[1] Rudolph and Oesterhelt (1996), [2] Aregger (2003), [3] Weidinger (2007) (* The CCW ratio of ΔCheC3 had to be corrected after reevaluation of the results from Weidinger, 2007), [4] Staudinger (2007)

*H. salinarum* codes for ten homologs of bacterial Che proteins (Ng *et al.*, 2000; Pfeiffer *et al.*, 2008b): *cheR*, *cheD*, *cheC1*, *cheC3*, *cheB*, *cheA*, *cheY*, and *cheW1* are organised in one operon (Aregger, 2003). A second CheW homolog, *cheW2*, is located close to the *fla* gene region, and a third *cheC*, *cheC2*, somewhere else in the genome. An illustration of the chemotaxis system of *H. salinarum* is shown in Figure 6.2. The functions and experimental results for the halobacterial Che proteins are summarised in Table 6.1.

## 6.2 Results and Discussion

All ten Che proteins were subjected to direct and indirect bait fishing. Six further proteins that were found as interaction partners were used as baits in order to confirm the detected interactions and to extend the interaction network. Additionally, ParA1 (OE2378R) was used as bait because a former study demonstrated some relation to chemotaxis for this protein (Staudinger, 2001). Table 6.2 gives an overview of the experiments of this project.

Table 6.2: **Bait fishing experiments for the Che interaction network.**

| Gene | Protein | Direct | Indirect |
|---|---|---|---|
| OE2374R | CheW2 | IP024L2 | IP017L2*, IP043L1 |
| OE2406R | CheR | IP015L1 | (IP034L1), (IP043L3) |
| OE2408R | CheD | IP024L1 | IP038L1 |
| OE2410R | CheC3 | IP015L2 | IP023L2 |
| OE2414R | CheC1 | IP016L1 | IP025L1 |
| OE2415R | CheA | IP005L2, IP035L1 | IP035L2 |
| OE2416R | CheB | IP016L2, IP036L2 | IP025L2 |
| OE2417R | CheY | IP013L1 | IP018L2 |
| OE2419R | CheW1 | IP014L2 | IP027L2 |
| OE3280R | CheC2 | IP019L2 | (IP027L1), IP043L2 |
| OE1428F | | IP034L2 | (IP036L1) |
| OE1620R | PurN/PurH | IP040L1 | IP028L2 |
| OE2401F | CpcE | IP030L1 | IP030L2 |
| OE2402F | | IP026L2 | IP026L1 |
| OE2404R | | IP029L1 | IP029L2 |
| OE4643R | | IP023L1 | IP028L1 |
| OE2378R | ParA1 | IP042L2 | IP042L1 |

Experiments in brackets were not included in the final dataset because of too many bound proteins (more than 20 unexpected interactors with an association score > 7); * IP017L2 was not done with reversed labelling (see 4.2.4). Therefore some putative interactors (found in the direct experiment) appear with negative ASAPScore.

Direct bait fishing experiments with pMS3/pMS5 were stopped early in the course of the project, because it was found in test expressions that several proteins were expressed at too low levels (e.g. CheB, the CheCs) to allow succesful purification from 100 ml cultures. Larger culture volumes would have been necessary so that the method were too expensive with SILAC labelling. And, for CheA and CheW1, hardly any differences were detectable between experiments performed with either pMS3 or pMS4 (data not shown). So only experiments with pMS4/pMS6 were continued.

Overall, in the experiments 597 unique proteins were identified (overall 5505 protein identifications) according to the criteria given in 2.4.11. Of the total identifications, 267 were rated with an association score of "high" or "very high" and thus classified as interactions. Merging multiply identified interactions results in a final set of 201 interactions. A list of all detected interactions is given in Supplementary Table S2, the complete data set can be accessed via HaloLex (`http://www.halolex.mpg.de/`; Pfeiffer *et al.*, 2008a).

## 6.2.1 Evaluation of experimental results

### 6.2.1.1 Contaminants

To identify contaminants that bind to the CBD and therefore appear as interactors, a direct and an indirect bait fishing experiment were performed in which the CBD expression strain (MS4) was tested against wildtype. The proteins detected as "interactors" of the CBD in these controls were included in the background table (see 5.2.3) and thus ignored in the results of all further experiments. Furthermore, some other proteins were added to the background table because they either bound to the cellulose column when incubated with wildtype cell lysate (the chitinases OE2201F, OE2205F, and OE2206F), or they were highly promiscuous (found in almost every experiment) and are involved in protein folding or degradation. As the bait is overexpressed and the CBD is from a non-halophilic organism, the bait might be partially misfolded so that proteins involved in protein folding and degradation bind to it. The complete list of background proteins is shown in Table 6.3.

Additionally, five preys were marked as promiscuous as they were interactors of at least three not directly interacting baits. These proteins could either be hubs in the interaction network (and thus be particular important) or unspecific binders. See 6.2.2.5 for details on these proteins.

Table 6.3: **Proteins considered as background.**

| Protein | Reason |
|---------|--------|
| OE1275F | Highly promiscuous, involved in protein degradation (probably due to misfolded bait) |
| OE1736R | Promiscuous, involved in protein folding (probably due to misfolded bait) |
| OE1737R | Found in background experiment |
| OE2201F | Binds to cellulose column |
| OE2205F | Binds to cellulose column |
| OE2206F | Binds to cellulose column |
| OE2296F | Highly promiscuous, involved in protein degradation (probably due to misfolded bait) |
| OE2998R | Found in background experiment |
| OE3642F | Found in background experiment |
| OE3925R | Promiscuous, involved in protein folding (probably due to misfolded bait) |
| OE4122R | Promiscuous, involved in protein folding (probably due to misfolded bait) |
| OE4674F | Found in background experiment |

### 6.2.1.2 Reproducibility

Of the 201 interactions listed in Supplementary Table S2, 23 were reproduced, that means they were detected in more than one experiment. This low number is owing to the fact that most interactions had no chance to be identified more than once. Most baits were indeed used in two experiments (direct and indirect bait fishing), but the different methods identify different interactions (see 4.2.4).

Only 8 interactions (that means four protein pairs) were reciprocally confirmed (i. e. found in both bait-prey combinations: protein X was a prey of bait Y, and Y was a prey of bait X). There are several possible reasons for that: First, only some of the preys have later been used as baits for reciprocal fishing. Hence for all others a reciprocal confirmation was *per se* not possible. Second, steric hindrance by the tag (the CBD is a relatively large tag with a molecular weight of 18 kDa) might prevent the formation of interactions in some cases. Third, some proteins are more difficult to identify than others, so that certain preys are missed. This is, for example, the case for small proteins, which might be lost during the sample preparation procedure (Klein *et al.*, 2007). An example for this is the CheA-CheY interaction: with the small CheY (13.4 kDa) as bait, it is easy to identify the big protein CheA as prey. The other way around it did not work, CheY as prey was not identified in any experiment. Additionally, preys that are located in the same gel slice as the bait might be missed, because their identification is hampered by the huge amount of bait in the sample.

**6.2.1.3 How to interpret the results**

Protein interactions form the basis for such different things like structural components that shape organelles or the whole cell, molecular machines like ribosomes, and signal transduction systems like the Che system. Thus it is not surprising that they can be of very different type: weak or strong, and permanent or transient.

Quantitatively, these properties are described by the mass action law:

$$\frac{[A][B]}{[AB]} = K_d = \frac{1}{K_a} = \frac{k_d}{k_a}$$

| | | |
|---|---|---|
| $K_d$ / $K_a$ | = | equilibrium constant for dissociation / association |
| $k_d$ | = | ($k_{off}$) first-order rate constant for the uni-molecular dissociation reaction |
| $k_a$ | = | ($k_{on}$) second-order rate constant for the bimolecular association reaction |

The range of values observed for $K_d$ in biologically relevant protein-protein interactions ("interaction strength") varies over at least 12 orders of magnitude form $10^{-4}$ to $10^{-16}$ M (Uetz and Vollert, 2006). Also the kinetics varies significantly: some protein interactions dissociate in split seconds, whereas others last for hours.

None of the currently available techniques for PPI analysis can cover this huge range of properties. Affinity purification methods have their detection limit at a $K_d$ somewhere between 1-50 µmol/l (Castagnoli *et al.*, 2004), depending on parameters like the kinetics of association and dissociation and protein abundance. That means, if an interaction is not detected, there are three possible reasons:

- The interaction does not take place at all.

- The interaction does not exist under the tested conditions, e. g. because the participating proteins are not expressed, or the interaction requires posttranslational modifications which are only present under certain circumstances.

- The interaction is there, but not detectable with the used method. This can be, because the interaction is too weak, too short-lived, the protein abundance is too low, or the interaction partner cannot be identified.

That means it is difficult to draw conclusions from a not identified interaction. What can be done to some extent is comparison: if prey X was clearly identified as interaction

partner of bait A, but not fished with bait B, it is very unlikely that this prey interacts with bait B with the same strength than with bait A – under the tested conditions.

Similarly, it is not possible to unambiguously judge an interaction that was detected. Again, there are three possible explanations:

- The bait and the prey are direct and specific interactors.

- The interaction is indirect: that can be if the prey is a component of the same protein complex as the bait, or if the proteins are hold together by something else like a membrane patch.

- The prey is a contaminant which is bound unspecifically. This can happen due to misfolding, or because bait or prey are "sticky".

Proteins falling into the third category can, to some extent, be excluded by methodological approaches like the use of SILAC and the removal of promiscuous preys. Distinguishing between the first two possibilities is not possible in AP-MS experiments. That means, when in the following sections is stated that two proteins interact, it is always possible that this interaction is indirect.

## 6.2.2 The interaction network

The detected interactions were used to draw a protein interaction network (Figure 6.3). All parts of the network are connected, with the exception of ParA1 (OE2378R) with its two interaction partners (see 6.2.2.7). The features of the interaction network are presented in the next sections. To give a clear view on the discussed aspects, a simplified version of the network is shown in Figure 6.4. In this simplified representation all preys that were fished by only one bait and preys fished by two or more not interacting baits (thus including the promiscuous preys), were omitted.

### 6.2.2.1 The core: CheA, CheY, CheW1

The core of the chemotaxis signal transduction system is the histidine kinase CheA, which is bound to the transducers by the coupling protein CheW and phosphorylates the response regulator CheY to generate the output signal CheY-P.

This assumed organisation of the core was confirmed by the detected interactions. CheA was found to strongly interact with CheW1, and both were found to interact

with a couple of transducers (for details on the transducers see 6.2.2.3). The second CheW protein, CheW2, exhibited a different interaction pattern than CheW1 and is only weakly connected to the core (see 6.2.2.2). Both CheA and CheW1 as well as several Htrs were detected as interaction partners of CheY. It should be noted again that AP-MS does not allow to determine the exact complex topology, so the interactions between CheY and CheW1 or the Htrs might be indirect via CheA.

Two proteins were identified as unexpected interaction partners of the core. PurH/N (OE1620R), annotated as phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
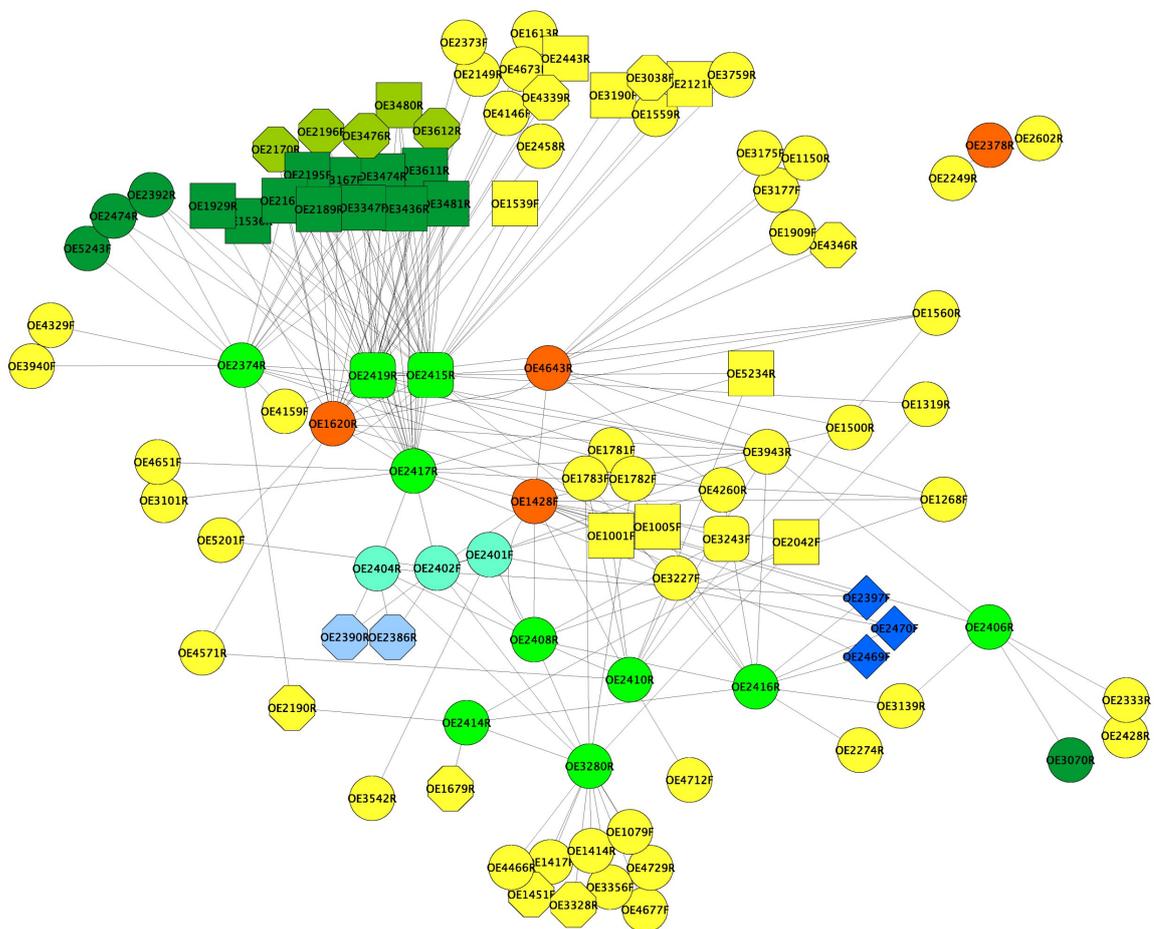


Figure 6.3: **Chemotaxis protein interaction network.** Colours: Light green Che protein, dark green Htr, yellow-green Htr-associated protein, cyan new Che protein (chapter 7), light blue Fla protein, blue Flagellin, orange other bait, yellow remaining protein; shape indicates predicted membrane association: box transmembrane protein, octagon membrane associated (see Supplementary Table S2), rhomb extracellular, box with round corners cytosolic/membrane associated, circle cytosolic.
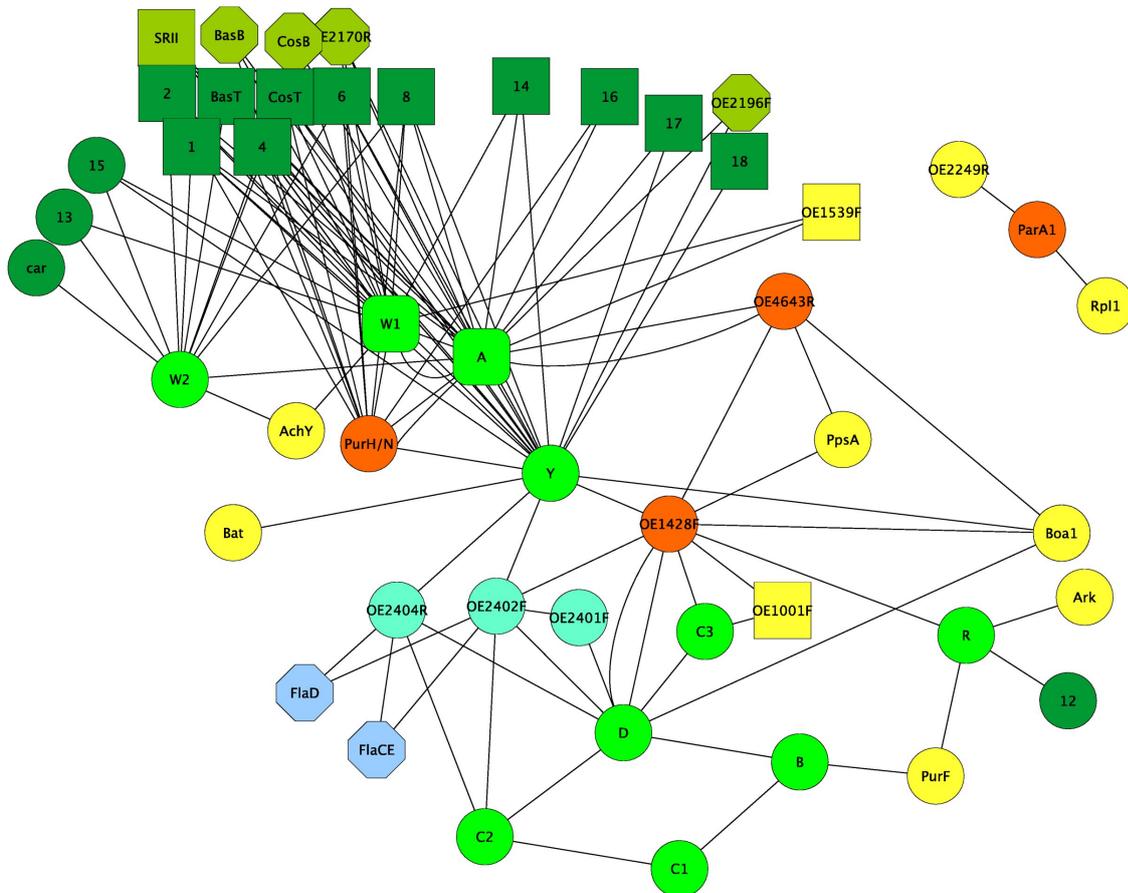
Figure 6.4: **Simplified chemotaxis protein interaction network.** For colours and shape of nodes see legend of Figure 6.3. All preys detected by only one bait, and preys detected by two or more not interacting baits were removed, except for some proteins discussed in the text. In the protein labels, the prefixes "Che" and "Htr" were omitted.

/ phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3), was detected as interaction partner of CheA and CheW1. PurH/N was used as bait in additional experiments, and the interaction with CheA could be reciprocally confirmed. Generally, with fishing several membrane-bound transducers, the association pattern of PurH/N was quite similar to the one of CheA (see Figure 6.5 C). PurH/N fulfils two essential enzymatic activities in purine metabolism. Its relation to the chemotaxis system is completely unclear.

The second unexpected interactor is OE4643R, a conserved protein of unknown function. Unlike PurH/N, this protein was only fished by CheA and not with CheW1 and CheY. When used as bait, OE4643R fished CheA, but it did not reveal the typical
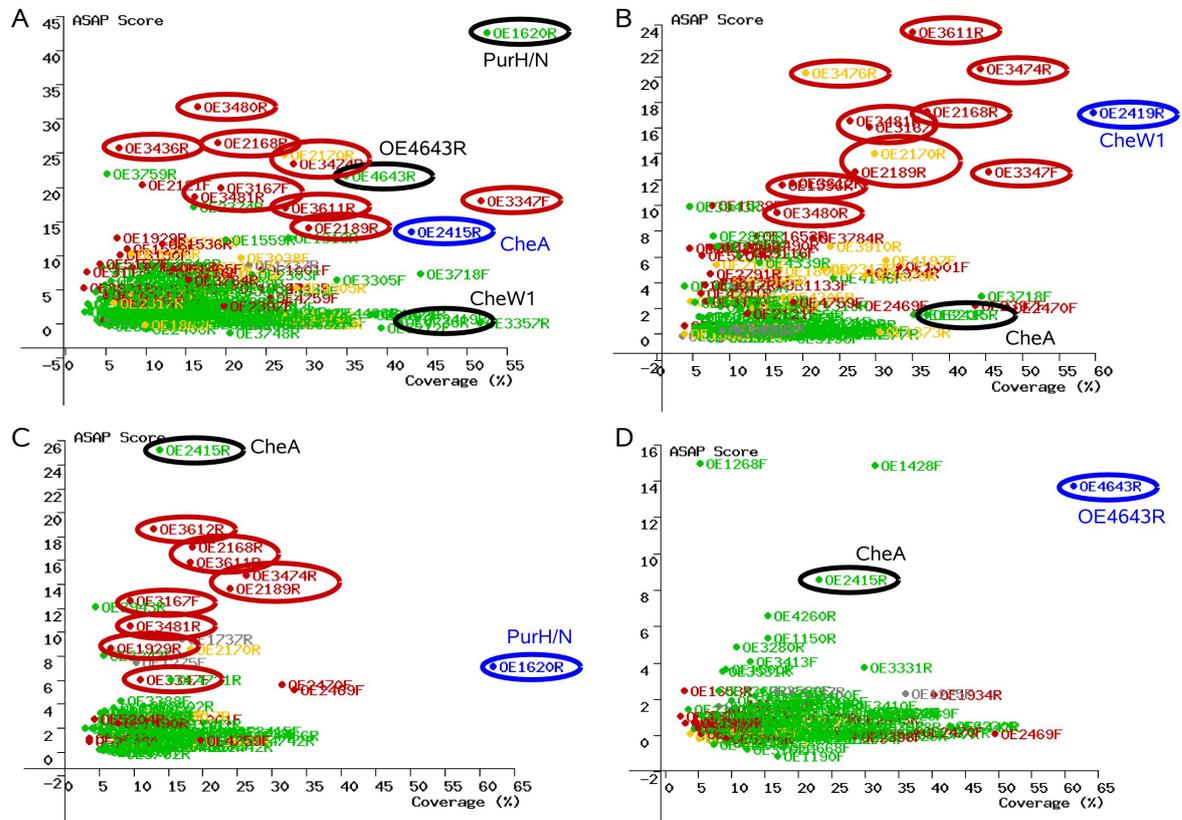
Figure 6.5: **Association patterns of the core proteins.** Plots indicating the SILAC ratios of proteins identified in bait-fishing experiments with CheA. See Figure 4.10 for explanation of such plots. (**A**), CheW1 (**B**), PurH/N (**C**), and OE4643R (**D**). Colours: Blue Bait protein, red Htr or Htr-associated protein, black Core or core-associated protein.

```
MDDYLEAFVR EGEEHVTSLN NALLELESDP GNEEAMDEIF RTAHTLKGNF GAMGFEDASD LAHAVEDLLD EMRQGNLEVT
SDRMDRIFEG IDGIEACLDE IQATGDVDRD VTGTIESVRA VLDEVDGDGG SGTTTSSGDA GSPAGDGDVD ATRVVDADTI
DAAEDPVYHI HIDMGDSQMK GVDGMFVLEE ATEAFDLLGA EPSPDAINDG EYGDGFELVV ATPSDEVSDT VAAFPKLSDA
TVTAVGDDEH APDADSGTEA DASADDDADD AGTTADSGSS SGGSSAIDNT DTEIQSVRVD VDQLDELHGL VEQLVTTRIK
LRRGMEESDR EVLDELDELD KITSSLQDTV MDMRLVPMKK IVGKFPRLVR DLAREQDKDI DFVVEGDDVE LDRTILTEIS
DPLMHLLRNA VDHGIEKPAV REDNGKDREG TITLSAERDR DHVLIQVRDD GAGIDHDTMR EKAIEKGVKT REEVQDMPDD
DVEDLVFHPG FSTNDEVTDV SGRGVGMDVV RDTVTRLDGS VSVDSTPGEG TTFTMTLPVT VAIVKVLFVE SGGEEYGIPI
KTVDEISRMK SVKSVDGEEV ITYDETVYPL VRLGDALNVP DETRNGDGML VRIRDSERQV AVHCDDVRGQ EEVVVKPFEG
ILSGIPGLSG AAVLGEGDVV TILDVATL
```

Figure 6.6: **Identification of CheA in bait fishing experiments with CheW1 and OE4643R.** Peptides shown in bold were identified in bait fishing experiments with OE4643R, underlined peptides with CheW1.

association pattern of the core proteins (see Figure 6.5 D, no CheW1, no Htrs with their associated proteins). So this protein seems to interact with a not transducer-associated form of CheA. In enterobacteria, two forms of the CheA protein exist: CheA$_L$, the full length protein, and CheA$_S$, a N-terminally truncated form, which has an alternative translation initiation site (McNamara and Wolfe, 1997). A similar scenario as reason for the different interaction pattern can be rejected here, as for both the Htr-bound form (fished with CheW1) and the cytosolic form (fished with OE4643R) the N-terminal peptide was identified (see Figure 6.6).

OE4643R belongs to the Pfam (Finn *et al.*, 2006, 2008) protein family DUF151 (DUF means "domain of unknown function"), and the cluster of orthologous groups (Tatusov *et al.*, 1997, 2003) COG1259 ("uncharacterised conserved protein"). A homolog of this protein from *Thermotoga maritima*, TM0160, has been crystallised and the structure solved to 1.9 Å resolution (Spraggon *et al.*, 2004). Unfortunately, even with the structure in hands it was not possible to deduce the function of this protein. Spraggon *et al.* (2004) propose that there is putatively a novel type of active site in this protein, so it might possess an unknown enzymatic activity. By the evaluation of the genomic context of homologs of TM0160 in 19 species, these authors found a preponderance of enzymes involved in amino acid metabolism, which indicates a possible role in this process for this protein family.

The background of the interactions of these two proteins with the core or CheA, respectively, is not yet known. Similar interactions have not been described in any other organism. In several bacterial species, the phosphoenolpyruvate-dependent carbohydrate phosphotransferase systems (PTSs) can mediate positive chemotaxis towards PTS carbohydrates (see Postma *et al.*, 1993, and references therein). Responses towards carbohydrates are mediated by two pathways: The first pathway is based on carbohydrate binding to periplasmic binding proteins which are components of the ATP-binding cassette transporters for the carbohydrates. With their bound ligands, the periplasmic binding proteins bind to certain MCPs, leading to a tactic response (Hazelbauer and Adler, 1971). The second pathway uses the PTSs, which relay taxis signalling when the substrate is transported (for review see Lengeler and Jahreis, 1996). A PTS consists of a substrate-specific, membrane-bound Enzyme II (EII) complex, which is phosphorylated by a cytoplasmic donor phosphorelay and phosphorylates the substrate during transport. The phosphorelay is made from Enzyme I (EI), a phosphoenolpyruvate (PEP)-dependent histidine kinase, and a phosphohistidine carrier

protein (HPr). Taxis signalling through PTSs in *E. coli* requires CheA, CheW, and CheY, but no MCPs (Rowsell *et al.*, 1995). It was demonstrated *in vitro* that EI, but not EI-P, inhibits CheA autophosphorylation (Lux *et al.*, 1995). In a more recent work, Lux *et al.* (1999) found that the PTS signal acts to inhibit MCP-bound CheA. Hence the unexpected interactions of CheA described in this study might belong to similar, yet unknown, taxis signalling pathways that target CheA.

### 6.2.2.2 Different interactions of the two CheW proteins

In none of the bait fishing experiments with CheW2, CheA was identified as prey. The other way around, with CheA as bait, CheW2 was found as prey in one experiment. This means, unlike CheA-CheW1, this interaction was not reciprocally reproducible, nor seemed CheW2 to be such an outstanding interactor (in terms of SILAC ratio and sequence coverage) of CheA as CheW1 is. These findings suggest that CheW1 is the main coupling protein to tie CheA to several transducers. This is inconsistent with the phenotypic characterisations of deletion mutants by Aregger (2003). In that study, the deletion of CheW2 led to severe defects in phototaxis and chemotaxis (assayed by computer-assisted cell tracking and swarm plates), whereas the deletion of CheW1 had hardly any impact on phototaxis and only small impact on chemotaxis.

To further study the roles of the two CheW proteins, a comparative bait fishing experiment with both proteins as bait was done (Figure 6.7). This experiment was performed as indirect bait fishing in which instead of the control pMS4 the second CheW was used – CheW1 was bound to one cellulose column and incubated with light ($^{12}$C) cell lysate, CheW2 to a second column and incubated with heavy ($^{13}$C) cell lysate.

In this experiment, the light form ($^{12}$C) of CheA and PurH/N (might be indirect via CheA) was present in a relatively high amount, whereas the heavy form ($^{13}$C) was hardly detectable (see Figure 6.7 B for a representative chromatogram of a CheA peptide). This means, these proteins were strongly bound to CheW1, whereas the binding to CheW2 was very weak or short-lived. It is even possible that the interaction of CheW2 and CheA detected here and with CheA as bait was just indirect via the Htrs.

The three (soluble) transducers Htr11 (Car), Htr13, and Htr15 showed the opposite behaviour. From these proteins the heavy form was detected in large quantities,
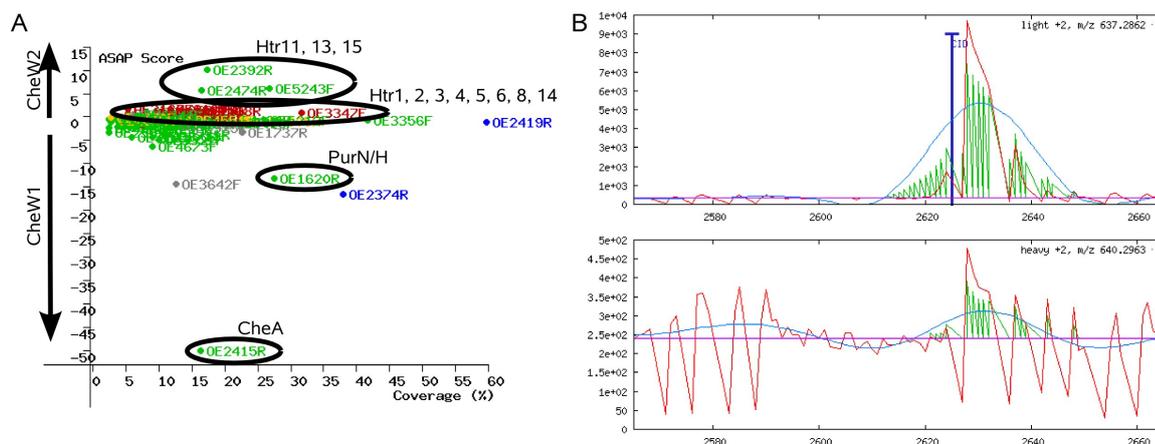
Figure 6.7: **Comparative bait fishing with the two CheW proteins. A** Plot of the SILAC ratios of proteins identified in the CheW comparison experiment. Proteins which are bound to a higher extent to CheW2 are shifted up, proteins bound to higher extend to CheW1 are shifted down. Proteins bound to both baits to the same extent as well as background proteins appear with an ASAP Score close to zero. **B** Extracted ion chromatograms of a representative peptide of CheA (the N-terminal peptide MDDYLEAFVR). The upper panel shows the light (bound to CheW1), the lower the heavy (bound to CheW2) form. Note the different scaling on the Y axis.

whereas the amount of light peptides was relatively low (albeit clearly detectable). This is congruent with the single bait fishing experiments with CheW1 and CheW2, where these transducers showed a strong binding to CheW2, whereas the association with CheW1 seemed to be much weaker. Htr11 (Car) was not even identified when fishing with CheW1.

The identified membrane-bound transducers (Htr1, 2, 3, 4, 5, 6, 8, 14) exhibited a SILAC ratio of nearly one, meaning they were bound to both CheWs to the same extent. However, it should be noted that the transducers form clusters in the membrane (see 6.1.1.1). It is possible that both CheWs just fish similar amounts of the whole clusters, so the association of certain Htrs with one or the other CheW might only be indirect. This can also play a role for the cytosolic transducers, for which clustering was also demonstrated in other organisms (Thompson *et al.*, 2006).

Overall, the role of the two CheW proteins remains unclear. CheW1 seemed to be the main coupling protein for the formation of signalling complexes by binding to CheA and to several membrane-bound transducers. CheW2, in contrast, exhibited no strong binding to CheA. Binding to the membrane-bound transducers was similar for both CheWs, but binding to three cytosolic Htrs was stronger for CheW2 than for CheW1. A possible explanation for the presence of two CheW homologs could be that they

allow a different weighting of certain signals under different conditions (for example, when light and/or oxygen are present in sufficient amounts, arginine taxis should be much less important than in the dark with oxygen shortage). This weighting could be regulated e. g. by different expression levels or posttranslational modifications. It is also possible that CheW2 is the connection to an additional, not yet elucidated part of the taxis signalling system (compare 6.1.1.5). A third explanation may be that CheW2 forms an alternative adaptational mechanism. Thereby it would decouple the signalling complexes on certain transducers by competition with CheW1. Additional experiments will be needed to enlighten the functions of the two CheW proteins.

### 6.2.2.3 The transducers exhibit nonuniform interaction patterns

The 18 transducers of *H. salinarum* show a quite nonuniform image when compared through all experiments (Table 6.4). As mentioned above, eight membrane-bound transducers (Htr1, 2, 3, 4, 5, 6, 8, 14) were fished by CheW1 and CheW2 to a similar extent, and they were also fished by CheA and CheY. Three cytosolic transducers, Htr11 (car), Htr13, and Htr15, were fished by CheW2 and to lesser extent also by CheW1. Htr15 was also a prey of CheY. Htr16, Htr17, and Htr18 were only found associated with CheA and CheY, and Htr7, Htr9, Htr10 were not identified as interaction partners at all (see note in the legend of Table 6.4 regarding CheW2). Htr12 was only fished with CheR, and this was the only transducer found with this bait. An interaction between the transducers and CheB was not detected. The interaction of CheR and CheB with the Htrs is probably too short-lived to be detected by AP-MS.

The interpretation of the different interaction patterns is difficult, especially since the signals for only eight Htrs are known. A simple correlation (e. g. attractant vs. repellent signalling, or phototaxis vs. chemotaxis) could not be found. Again, different affinities could give the possibility for different weighting of certain signals under different conditions. It should be investigated in further experiments, if Htr16, Htr17, and Htr18 indeed form signalling complexes without one of the CheWs, as well as the role of the Htr12-CheR interaction.

With the transducers Htr2, BasT, CosT, Htr6, and Htr18, the putative associated proteins sensory rhodopsin II (SRII), BasB, CosB, OE2170R, and OE2196F (the latter two are putative periplasmic substrate binding proteins like BasB and CosB) were copurified.

Table 6.4: **The halobacterial transducers as preys.**

| Htr | Gene | Name | Signal | TM | W2 (1/2) | R (1/0) | A (2/1) | Y (1/1) | W1 (1/1) | Affinity |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | OE3347F | | Orange/UV light | 2 | 0(1)/1(2) | 0(1)/- | 1(2)/1(1) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 2 | OE3481R | | Blue light, serine | 2 | 0(1)/1(2) | 0(1)/- | 1(2)/0(0) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 3 | OE3611R | BasT | Leu, Ile, Val, Met, Cys | 2 | 0(1)/1(2) | 0(1)/- | 1(2)/0(0) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 4 | OE2189R | | | 2 | 0(1)/1(2) | 0(1)/- | 2(2)/0(1) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 5 | OE3474R | CosT | Compatible osmolytes | 2 | 0(1)/1(1) | 0(1)/- | 2(2)/0(0) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 6 | OE2168R | | | 2 | 0(1)/1(2) | 0(1)/- | 2(2)/0(1) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 7 | OE3473F | | | 3 | 0(1)/0(0) | 0(0)/- | 0(0)/0(0) | 0(0)/0(0) | 0(0)/0(0) | W2?* |
| 8 | OE3167F | | $O_2$ (attractant) | 6 | 0(1)/1(2) | 0(1)/- | 2(2)/0(1) | 0(1)/1(1) | 1(1)/1(1) | W1, W2, A, Y |
| 9 | OE2996R | | | 0 | 0(1)/0(0) | 0(0)/- | 0(0)/0(0) | 0(0)/0(0) | 0(0)/0(0) | W2?* |
| 10 | OE3150R | HemAT | $O_2$ (repellent) | 0 | 0(1)/0(0) | 0(0)/- | 0(0)/0(0) | 0(0)/0(0) | 0(0)/0(0) | W2?* |
| 11 | OE5243F | Car | Arg | 0 | 0(1)/1(1) | 0(0)/- | 0(1)/0(1) | 0(1)/0(0) | 0(1)/0(1) | W2, W1*** |
| 12 | OE3070R | | | 0 | 0(0)/0(0) | 1(1)/- | 0(0)/0(0) | 0(0)/0(0) | 0(0)/0(0) | R |
| 13 | OE2474R | | | 0 | 0(1)/1(1) | 0(0)/- | 0(1)/0(0) | 0(0)/0(1) | 0(1)/1(1) | W2, W1** |
| 14 | OE1536R | MpcT | $\Delta\Psi$ | 2 | 0(1)/0(1) | 0(0)/- | 0(2)/0(0) | 0(1)/1(1) | 1(1)/1(1) | W1, A, Y, W2*** |
| 15 | OE2392R | | | 0 | 0(1)/1(2) | 0(1)/- | 0(1)/0(1) | 0(1)/1(1) | 0(1)/1(1) | W2, W1**, Y |
| 16 | OE1929R | | | 2 | 0(1)/0(0) | 0(0)/- | 1(1)/0(0) | 0(0)/0(1) | 0(1)/0(0) | A, W2?* |
| 17 | OE3436R | | | 3 | 0(0)/0(0) | 0(0)/- | 1(1)/0(0) | 0(1)/1(1) | 0(0)/0(0) | A, Y |
| 18 | OE2195F | | | 2 | 0(1)/0(0) | 0(0)/- | 0(0)/0(0) | 0(0)/1(1) | 0(1)/0(0) | Y, A****, W2?* |

The columns are: TM # transmembrane helices; W2, R, A, Y, and W1 give identifications in bait fishing experiments with CheW2, CheR, CheA, CheY, and CheW1, respectively. The numbers in brackets behind the bait show the number of direct/indirect experiments. The numbers given for the transducers are a(b)/c(d) with a/c being the number of direct/indirect experiments where this transducer was found as interaction partner, and b/d being the numbers of respective experiments where this transducer was identified at all. The column affinity summarises the baits which fished the respective transducer as prey. * The soluble transducers clearly associated with CheW2 (Car, Htr13, Htr15) showed the exchange problem in direct fishing. The other soluble and two membrane-bound transducers were also identified in direct fishing with CheW2, and also with a SILAC ratio of nearly one. In indirect fishing, they were not identified at all. So it is possible that they are also interactors of CheW2 and exhibit the exchange problem. ** were also found with CheW1 in indirect fishing, but the affinity for CheW2 is much higher (comparison experiment). *** was not identified in single bait fishing with this CheW, but showed some affinity in the CheW comparison experiment. **** Htr18 was not identified with CheA, but its putative associated protein OE2196F.

OE1539F, a conserved protein of unknown function, was fished by the same baits as Htr14 to which it is in genomic proximity, so it might be a Htr14-associated protein. But, unlike the other transducer-associated proteins, it is not directly adjacent in the genome or exhibits even an overlap. OE1539F homologs in *N. pharaonis* and *H. walsbyi* are not in proximity to any transducer (*H. walsbyi* is neither chemotactic nor motile at all, so it does not code for chemotaxis proteins), which further decreases the likelihood that OE1539F is a transducer-associated protein.

### 6.2.2.4 Other Che Proteins

The other Che proteins are not directly connected to the core. Indirect connections to the core occur from CheD and CheC2 to OE2402F and OE2404R, which interact with CheY, and via some connectors (see 6.2.2.5).

CheD appeared not only as connector to the core, but also as a hub for the other Che proteins. It was found to interact with CheC2, CheC3, CheB, and the unknown proteins OE2401F, OE2402F, and OE2404R. This important place in the network is in agreement with the high conservation of CheD throughout chemotactic bacteria and archaea (Szurmant and Ordal, 2004), and also with the severe phenotype of a CheD deletion (compare Table 6.1). However, the function of the CheD protein remains to be elucidated – in *B. subtilis* and *T. maritima*, receptor deamidase activity of CheD was demonstrated (Kristich and Ordal, 2002; Chao *et al.*, 2006), but this is not the case for *H. salinarum* (Koch, 2005). The study performed in *T. maritima* demonstrated also receptor methylesterase activity for CheD. This function of CheD has not been described for any other organism, and it is not clear if CheD demethylates receptors in *H. salinarum* (Koch *et al.*, 2008).

*H. salinarum* expresses three CheC proteins. From these, CheC1 and CheC2, both consisting of one CheC domain, were found to interact. CheC3, which consists of two CheC domains, was not detected to interact with another CheC. So it can be speculated that there are two functional CheC units in *H. salinarum*: one is the heterodimer built from CheC1 and CheC2, the other one is CheC3. The deletion of either CheC1 or CheC2, however, results in different phenotypes (see Table 6.1), meaning that the remaining partner is still able to perform some function. Both CheC2 and CheC3, but not CheC1, were shown to interact with CheD, so that both putative CheC units interact with CheD.

An interaction between CheC2 and the proteins OE2402F and OE2404R was detected. These proteins were demonstrated to belong to a new class of archaeal chemotaxis proteins that might act at the interface between the Che system and the archaeal flagellum (see chapter 7).

CheC1 was shown to interact with CheB. CheB is as receptor methylesterase a key player in adaptation. The methylesterase activity is controlled by the phosphorylation status of its response regulator domain. So it can be speculated that one of the CheCs, which are thought to be CheY-P phosphatases, also dephosphorylates CheB-P. Additionally, CheB exhibited an interaction with CheD. Together, these interactions might provide some regulation of or feedback to the reception and adaptation system. Such a feedback mechanism in *B. subtilis* is described in 6.2.3.

CheR showed no direct connection to any of the other Che proteins. This is not that surprising since CheR is thought to be constitutively active (Simms *et al.*, 1987). As mentioned above, an interaction of CheR and CheB with the halobacterial transducers was, with the exception of CheR-Htr12, not detectable.

The proteins OE2401F, OE2402F, and OE2404R, which were identified as interactors of CheY, CheD, and CheC2, have been characterised further by follow-up experiments and bioinformatic analysis. These results are presented in chapter 7.

### 6.2.2.5 Connectors: Hubs or sticky background

Some preys came up as interactors in several experiments. The proteins OE1428F, OE1560R, OE1783F, OE3227F, OE3943R, and OE4260R were found with at least three not directly interacting baits and therefore considered as promiscuous. These proteins might either be hubs in the network and therefore are of particular importance, or they are unspecific binders and possibly totally unrelated. One of these proteins, OE1428F, was used as bait for reciprocal fishing, and the interaction with one Che protein, CheD, could be confirmed. The confirmed interaction is a strong evidence that OE1428F is indeed connected to the Che proteins. The other proteins have not been tested as baits, and their place in the network should be viewed with caution.

OE1428F was fished with CheY, OE2402F, CheD, CheC3, CheR, and OE4643R. The interaction with CheD could be reciprocally confirmed. OE1428F is a protein of unknown function, without any known domain or motif. Homologs were only detected

in the halophilic archaea *Haloarcula marismortui* and *Halorubrum lacusprofundi*. The total lack of information about this protein makes it impossible to draw conclusions about its role in the Che network.

OE1560R was associated with CheC3, OE1620R, OE4643R, and CheW1. It is annotated as conserved hypothetical protein. Homologs are found in several archaeal species, including non-chemotactic ones like *H. walsbyi*. Domain search with the protein sequence of OE1560R against InterPro (Mulder *et al.*, 2007) and Pfam (Finn *et al.*, 2006, 2008) finds significant hits to DUF516 (InterPro; DUF means domain of unknown function) and the family tRNA_deacylase (Pfam entry PF04414).

OE1783F was a prey of CheW2, CheC3, CheB, and CheC2. It is annotated as SufB domain protein. SufB is the permease component of an ABC-type transport system involved in Fe-S cluster assembly. In the genome, OE1783F is located next to the other components of this transport system. Homologs are found in many archaea and bacteria, including non-chemotactic archaea.

OE3227F is annotated as "homolog to nicotinate-nucleotide dimethylbenzimidazole phosphoribosyltransferase". It was fished by CheA, CheB, CheC3, and CheY. By homology search (Psi-Blast), it was not possible to find any confirmation for this annotation, but OE3227F is located at the end of the cobalamin biosynthesis operon, and the enzymatic activity in the annotation is involved in cobalamin biosynthesis. Homologs are found in *H. marismortui*, *N. pharaonis*, and the three *Methanosarcina* species. All of these species contain chemotaxis genes.

OE3943R was fished with OE1620R, CheW2, OE2401F, CheR, CheC3, CheB, CheY, and CheW1. It is annotated as conserved hypothetical protein. Domain searching revealed that it contains the Pfam domain DUF1743, which is "found in many hypothetical proteins and predicted DNA-binding proteins such as transcription-associated proteins" (Pfam). This protein also belongs to the cluster of orthologous groups (Tatusov *et al.*, 1997, 2003) COG1571 ("Predicted DNA-binding protein containing a Zn-ribbon domain"). Homologs are found in many archaea and bacteria, including non-chemotactic archaea.

OE4260R was a prey of CheW2, OE2402F, CheB, and OE4643R. Its annotation is "probable N-acetyltransferase (EC 2.3.1.-)". OE4260R belongs to COG1670 ("RimL, Acetyltransferases, including N-acetylases of ribosomal proteins") and the Pfam family Acetyltransf_1 (PF00583). Homologs are found in a variety of bacteria and, with weaker homology, in some archaea, including non-chemotactic species.

OE1268F, annotated as bacterio-opsin activator-like protein 1 (Boa1), was fished with OE1428F, CheD, CheY, and CheC2. It is not classified as promiscuous prey according to the above mentioned criterion, because CheD, CheY, and CheC2 are interaction partners of OE1428F. Boa1 contains a PAS domain, a GAF domain, and a helix-turn-helix domain, so it is probably both sensor and transcriptional regulator. Neither the genes it regulates nor the conditions under which this happens are known.

Overall, the role of none of these proteins in the Che system is clear, if they are involved at all. Follow-up experiments like reciprocal bait fishing and characterisation of deletion strains should be done to explore the function of these proteins and possibly enlighten hitherto unrecognised aspects of taxis signalling.

### 6.2.2.6 Unexpected interactors

Several other proteins for which the relation to the chemotaxis system is totally unclear were fished with one or more baits. A few potentially interesting binders will be presented in this section.

In the indirect bait fishing experiment with the chemotaxis response regulator CheY, the transcriptional regulator bat (bacterioopsin-activator of transcription, OE3101R) came up as interaction partner. Bat induces BR synthesis in case of low oxygen tension (Gropp and Betlach, 1994). It contains a PAS and a GAF domain, so it is probably both, sensor and transcriptional activator.

Both CheW1 and CheW2 fished the protein OE4159F (AchY). This protein was, except for several htrs, CheA, and the promiscuous prey OE3943R, the only interaction partner of both CheWs. It is annotated as adenosylhomocysteinase, catalysing the conversion between S-adenosyl-L-homocysteine and L-homocysteine + adenosine. According to the KEGG database, this enzymatic function is involved in methionine metabolism and selenoamino acid metabolism. It should be noted that S-adenosyl-L-homocysteine is formed by the action of CheR (Springer and Koshland, 1977; Simms and Subbaramaiah, 1991), so there exists a known link between this substance and taxis signalling.

OE4643R and OE1428F were found to interact with OE1500R (PpsA), annotated as pyruvate-water dikinase (phosphoenolpyruvate synthase). The reaction catalysed by this enzyme, the synthesis of phosphoenolpyruvate from pyruvate and ATP, is an essential step in gluconeogenesis, and, in certain organisms, also for glycolysis via a

modified Embden-Meyerhof pathway (Imanaka *et al.*, 2006).

CheR as bait fished the signalling histidine kinase Ark (OE2333R). This histidine kinase probably forms a two-component system with the adjacent response regulator hrg (OE2334R). The interaction with CheR could be an indication for cross-talk between different signalling pathways.

Both CheR and CheB, the antagonists in methylation-dependent adaptation, fished the protein OE3139F, annotated as amidophosphoribosyltransferase PurF. PurF catalyses the interconversion between 5-phospho-$\beta$-D-ribosylamine + diphosphate + L-glutamate and L-glutamine + 5-phospho-$\alpha$-D-ribose 1-diphosphate. On the one hand, this reaction forms a connection between glutamate metabolism and purine metabolism. On the other hand, the deamidation of glutamines on the receptors was shown to be essential for maturation of certain receptors (Saulmon *et al.*, 2004) and is part of an alternative adaptational system (Chao *et al.*, 2006; Muff and Ordal, 2007). Hence this enzyme, catalysing the interconversion of glutamate and glutamine, might be a till now unrecognised part of the reception/adaptation system.

Several other unexpected interactors have been detected for one or more of the baits (see Figure 6.3 and Supplementary Table S2). Neither for the proteins discussed here nor for the other proteins the relation to the Che system is clear. Reciprocal fishing should be done to confirm – if possible – the interactions. After confirmation, functional studies, like the characterisation of deletion mutants, are required to unravel the role of these proteins in taxis signalling.

### 6.2.2.7 Not connected: ParA1

The protein OE2378R (ParA1) was included in the PPI analysis because it is located adjacent to the *fla* gene region and a former study had shown that the deletion of this gene leads to deficiencies in chemotaxis (Staudinger, 2001). Unfortunately, for this bait no connection to chemotaxis- or motility-related proteins could be found. Two putative interaction partners were fished with OE2378R: the ribosomal protein L1 (Rpl1, OE2602R) and OE2249R, a conserved protein of unknown function.

OE2249R was formerly annotated in HaloLex as "transducer protein weak homolog lacking transduction domain", but this annotation was replaced by "conserved hypothetical protein". The former annotation was based on weak homology to the MCPsignal domain and the region around one HAMP domain of several Htrs. This
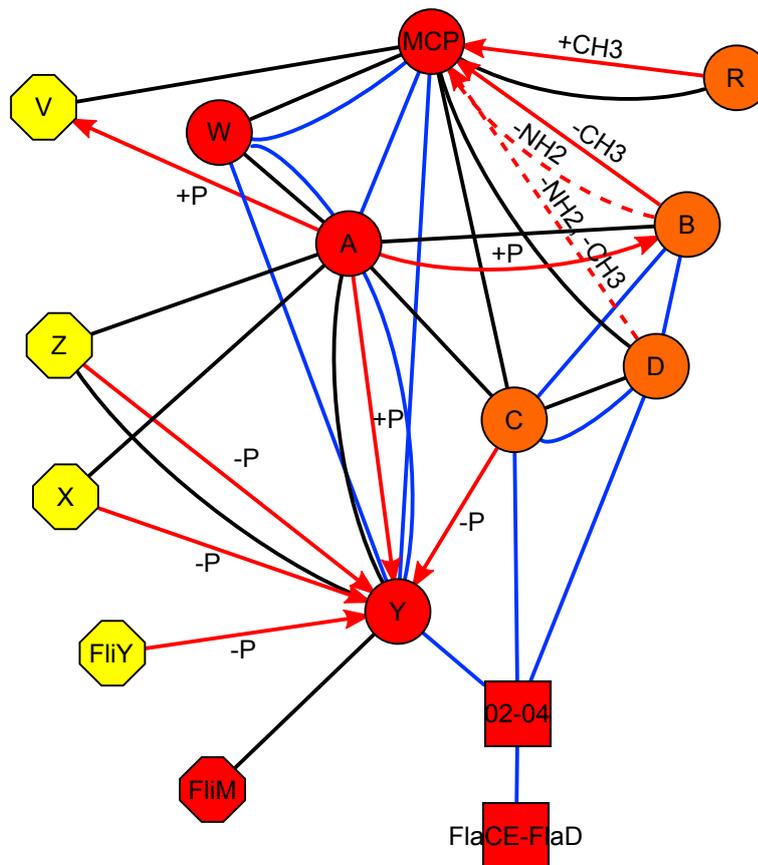
Figure 6.8: **Comparison of the Che interaction network with data from other organisms.**
Circles indicate proteins found in Archaea and Bacteria, boxes proteins found only in Archaea, and octagons proteins found only in Bacteria. Universal (in the respective domain) proteins are shown in red, proteins found universally in chemotactic archaea and optionally in chemotactic bacteria in orange, and proteins found in some bacteria in yellow. Black line: demonstrated physical interaction from literature; blue line: interaction in *H. salinarum* described in this study; red line: functional interaction from literature. Dashed lines indicate that this interaction occurs only in some species.

homology is, however, so low that searching against Pfam or SMART does not detect these domains (or any other domain) when queried with OE2249R. The function of OE2249R, as well as the role of ParA1 and its interaction with a ribosomal protein, remain to be elucidated.

## 6.2.3 Comparison with data from other organisms

The interactions detected in this study were compared to interactions between the Che proteins from other organisms. Of particular interest for comparison was the mammalian-two hybrid dataset from *Pyrococcus horikoshii* (Usui *et al.*, 2005), be-

cause this is the only dataset from an archaeal organism. The Che proteins tested in this study were CheB, CheC1, CheC2, CheD, CheR, CheW, CheY, DUF439 (see chapter 7), and one MCP. Unfortunately, the only interaction found with these proteins

Table 6.5: **Physical and functional interactions between Che proteins described in literature.**

| Interaction | Species |
|---|---|
| CheW-MCP | HP[1], EC[2,3], TD[4], CJ[39] |
| CheR-MCP | ST[5] |
| CheV-MCP | CJ[39] |
| CheA-CheW | HP[1], EC[2,6,3], TD[4], CJ[39] |
| CheA-CheB | EC[7] |
| CheA-CheX | TD[4] |
| CheA-CheZ | EC[8,9] |
| CheA-CheY | HP[1], EC[10,11,12,13,3,7], TM[14], TD[4], CJ[39] |
| CheC-CheD | PH[15], TM[16], BS[17] |
| CheC-CheA | BS[18] |
| CheC-MCP | BS[18] |
| CheD-MCP | BS[18], TM[16] |
| CheY-CheZ | EC[19,20,21] |
| CheY-FliM | EC[22,23] |
| CheR methylates MCP | EC[24], BS[25], HS[38] |
| CheB demethylates MCP | EC[24], BS[26], HS[38] |
| CheB deamidates MCP | EC[27], HS[28,38] |
| CheD deamidates MCP | BS[29], TM[16] |
| CheD demethylates MCP | TM[16] |
| CheA phosphorylates CheY | EC[30], BS[31], HS[32] |
| CheA phosphorylates CheB | EC[33,30], BS[34], HS(hom) |
| CheA phosphorylates CheV | BS[34] |
| CheX dephosphorylates CheY | BB[35] |
| CheZ dephosphorylates CheY | EC[36] |
| CheC dephosphorylates CheY | BS[37], HS(hom) |
| FliY dephosphorylated CheY | BS[37] |

HP: *H. pylori*, EC: *E. coli*, TD: *T. denticola*, ST: *S. typhimurium*, BS: *B. subtilis*, HS *H. salinarum*, BB: *B. burgdorferi*. References: [1] Rain *et al.* (2001), [2] Boukhvalova *et al.* (2002), [3] Schuster *et al.* (1993), [4] Sim *et al.* (2005), [5] Djordjevic and Stock (1998), [6] Gegner and Dahlquist (1991), [7] Yamamoto *et al.* (2005), [8] Wang and Matsumura (1996), [9] Kott *et al.* (2004), [10] Shukla and Matsumura (1995), [11] McEvoy *et al.* (1998), [12] Welch *et al.* (1998), [13] Gouet *et al.* (2001), [14] Park *et al.* (2004), [15] Usui *et al.* (2005), [16] Chao *et al.* (2006), [17] Rosario and Ordal (1996), [18] Kirby *et al.* (2001), [19] Zhao *et al.* (2002), [20] Sourjik and Berg (2004), [21] Blat and Eisenbach (1996), [22] Sourjik and Berg (2002a), [23] Lee *et al.* (2001), [24] Sherris and Parkinson (1981), [25] Kirsch *et al.* (1993b), [26] Kirsch *et al.* (1993a), [27] Kehry *et al.* (1983), [28] Koch (2005), [29] Kristich and Ordal (2002), [30] Hess *et al.* (1988), [31] Bischoff *et al.* (1993), [32] Rudolph *et al.* (1995), [33] Stewart *et al.* (1990), [34] Karatan *et al.* (2001), [35] Motaleb *et al.* (2005), [36] Silversmith *et al.* (2003), [37] Szurmant *et al.* (2004), [38] Koch *et al.* (2008), [39] Parrish *et al.* (2007). HS(hom) means that this reaction was not experimentally verified in *H. salinarum*, but concluded to occur by homology.

was CheC1-CheD, so that a detailed comparison with the *H. salinarum* dataset is not possible. Generally, the interaction detection in that study was rather low; the mammalian two-hybrid system might not be the optimal method to analyse interactions of proteins from a hyperthermophilic archaeon.

In the two large-scale studies carried out in *E. coli* by Butland *et al.* (2005) and Arifuzzaman *et al.* (2006), only interactions between CheW and two MCPs were identified, and several proteins not expected to be related to chemotaxis fished with some Che proteins. There was, however, no overlap with the unexpected proteins fished in the present study.

Another large-scale dataset was produced in *H. pylori* by Y2H analysis. This organism has a rather simple chemotaxis system, with only four MCPs, CheA, CheW, CheY, and three CheVs (Tomb *et al.*, 1997; O'Toole *et al.*, 2000). In this study the interactions MCP-CheW-CheA-CheY were detected, as well as interactions with some unexpected proteins (no overlap to *E. coli* and *H. salinarum* unexpected proteins). Further large-scale Y2H datasets exist for *Campylobacter jejuni* (Parrish *et al.*, 2007) and *Treponema pallidum* (Rajagopala *et al.*, 2007). In *C. jejuni*, interactions between MCPs-CheV and MCPs-CheW-CheA-CheY were detected. Although this organism codes for CheR and CheB (Parkhill *et al.*, 2000), no interactions for these proteins were reported. CheC and CheD are not present in *C. jejuni*. The *T. pallidum* dataset does not contain interactions between Che proteins.

Since large-scale studies did not deliver adequate data for comparison, the STRING (von Mering *et al.*, 2007) and BIND (Alfarano *et al.*, 2005) databases were queried for respective data from smaller studies carried out in any prokaryotic organism, and literature searching was done. Additionally, functional interactions (i.e. enzymatic reactions) between the Che and related proteins were collected from PubMed. Resulting data (Table 6.5) was used to draw a general Che protein interaction network (Figure 6.8). Most of the reported interactions were found in *E. coli* and the spirochaete *T. denticola*, only four physical interactions were reported for *B. subtilis*. This weakens the comparison with the *H. salinarum* network, which is with regard to the used proteins more closely related to the one from *B. subtilis*. For example, neither *E. coli* nor *T. denticola* code for a CheC or CheD protein.

The interactions of the core are generally in agreement between *H. salinarum* and other organism's data. The *H. salinarum* dataset contains probably indirect interactions (CheY-CheW, CheY-MCP) because it was generated by AP-MS. In the data-

bases, no direct interaction between CheA and MCPs is deposited. In literature, it is reported that the CheA-MCP association generally depends on CheW (Gegner *et al.*, 1992). A direct (weak) CheA-MCP interaction was suggested for inhibitory signalling (Ames and Parkinson, 1994) in *E. coli*. A direct interaction with CheA, however, is likely to occur in *H. salinarum*, because some transducers were fished with CheA, but not with either of the CheWs (see 6.2.2.3).

Interactions of CheR and CheB with Htrs could not be demonstrated in *H. salinarum* (except CheR-Htr12). In the databases, there is only one reference for a physical interaction between CheR and a MCP, based on a crystal structure after co-crystallisation. That means that these interactions seem to be hard to detect with PPI analysis methods.

CheD plays an important role in the *H. salinarum* interaction network: it was found to interact with CheC2, CheC3, CheB, and OE2401F, OE2402F, and OE2404R (see chapter 7). Of these, only the interaction with CheC has been described before. Instead of this, in *B. subtilis* an interaction of CheD with the MCPs was identified by Y2H analysis (Kirby *et al.*, 2001). Such an interaction was not detected in *H. salinarum*. This might be due to different functions of CheD in both organisms (see subsubsection 6.1.1.3).

The interaction between CheC and CheD was earlier demonstrated in *B. subtilis* and interpreted as feedback loop to the transducers via CheD's deamidase activity, which is decreased by CheC binding (Muff and Ordal, 2007). Furthermore, the interaction with CheD increased the CheY-P dephosphorylation activity of CheC 5-fold (Szurmant *et al.*, 2004). CheY-P stabilises the CheC-CheD complex, thus closing the feedback circuit. The role of the CheC-CheD interaction in *H. salinarum* remains unclear. An effect on the activity of CheC is possible, whereas a feedback loop to the transducers in unlikely due to the lack of receptor deamidase activity of CheD.

Such a feedback loop could be formed by the CheD-CheB and CheC1-CheB interactions detected in this study. CheC1-CheB-CheD in *H. salinarum* might thus be analogue to CheD-CheC in *B. subtilis*.

In this study, an interaction between CheC2 and the proteins OE2402F and OE2404R was detected. These proteins were found to interact both with Che proteins (CheY, CheD, CheC2) and with the flagella-accessory proteins FlaCE and FlaD (these proteins are discussed in detail in chapter 7). Hence they might be constituents of or be associated with the archaeal flagellar motor switch. Provided that this hypothesis

holds true, the interaction with CheC2 might reflect a situation similar to *B. subtilis*: In this organism, FliY, the main CheY-P phosphatase, is localised at the flagellar motor switch, and CheC, the second CheY-P phosphatase, at the signalling complex (Szurmant *et al.*, 2004). A direct or indirect interaction of one of the CheCs with the signalling complex was, however, not identified in *H. salinarum*. Generally, phosphatase localisation turned out to be a conserved and important principle in bacterial chemotaxis systems (Rao *et al.*, 2005).

## 6.3 Conclusions

The protein interaction study generally confirmed the expected topology of the core of the taxis signalling system: interactions between CheA, CheW1, CheY, and several Htrs could be detected. With PurH/N and OE4643R two unexpected interactors of the core or CheA, respectively, have been identified, whose functional role should be enlightened in follow-up experiments. For the two CheW proteins, CheW1 and CheW2, different interactions have been detected. CheW1 seems to be the main coupling protein for the formation of stable signalling complexes between the Htrs and CheA. The different Htrs as preys revealed nonuniform interaction patterns: some Htrs were associated to CheW1, CheW2, CheA, and CheY, others mainly with CheW2, a third group with CheA and CheY, but none of the two CheWs, and a fourth group was not fished at all. The underlying principle behind these different affinities remained unclear. The interactions of the other Che proteins were difficult to interprete because for most of these proteins the function is not or not completely known. An unexpected finding was the central position of CheD in the Che interaction network, indicating that this protein might play a key role in the halobacterial taxis signalling system.

Overall, the results of this study demonstrate that the halobacterial taxis signalling system exhibits numerous facets that have not been discovered so far. Even though the PPI analysis alone did not allow to explain any of these new aspects, it delivered a great amount of starting points for follow-up experiments, and it contributed valuable knowledge for the understanding of this signal transduction network.

# 7 Identification of archaea-specific chemotaxis proteins which interact with the flagellar apparatus

## 7.1 Introduction

In both, Bacteria and Archaea, taxis signalling is based on a modified two-component signal transduction system. Even though several variations of this Che system exist in different bacterial and archaeal species, the overall mechanism as well as the proteins involved are conserved (see 6.1.1 for details). In both domains, the output of the Che system is the phosphorylated response regulator CheY (CheY-P) that regulates the direction of rotation of the flagellar motor (Barak and Eisenbach, 1992a; Rudolph *et al.*, 1995).

The target of the Che system is the flagellum, a rotating, propeller-like structure. Unlike the Che system, the flagellum in archaeal species is only superficially similar to the bacterial one. It exhibits neither the same morphology nor are its constituting proteins homologs of the respective bacterial proteins. Whereas the mode of action of CheY-P at the flagellar motor is well understood in Bacteria, in Archaea the analogous mechanism remained elusive.

In the previous chapter (chapter 6), the interaction network of the halobacterial Che proteins was described. In this network, three proteins occur as interaction partners of CheY, CheD, and CheC2, which due to their genomic location between the *che* and the



Figure 7.1: **Chemotaxis and motility gene cluster of *H. salinarum.*** Genes involved in chemotaxis are shown in blue, motility genes in green. The proteins investigated in this chapter are shown in light blue (the homologs OE2402F and OE2404R) and cyan. A protein of unknown function is coloured grey.

*fla* gene region (Figure 7.1) seemed to be promising candidates for further analysis. All three proteins are of unknown function. These proteins were subsequently used as baits for interaction screening, and the flagella accessory proteins FlaCE and FlaD identified as interaction partners. To characterise the unknown proteins, deletion mutants were generated and assayed for their ability to perform chemotaxis and phototaxis, and their flagellar rotational bias analysed.

### 7.1.1 The archaeal and bacterial flagellum are distinct structures

#### 7.1.1.1 The bacterial flagellum

Bacteria use a variety of structures to achieve motility, including the bacterial flagellum that drives swimming motility, type IV pili mediating twitching motility, or the contractile cytoskeleton of *Spiroplasma*.

The bacterial flagellum (Figure 7.2) is composed of three main parts (see Berg, 2003, and references therein): the basal body, the hook, and the filament. The basal body contains the motor, which uses proton influx to generate rotational motion, the secretion and assembly apparatus, and it anchors the whole flagellum to the cell membrane and cell wall. The motor is subdivided in a stator, built from the proteins MotA and MotB that form the proton-conductive channel, and a rotor. The rotor contains several proteins, including the switch proteins FliM, FliN, and FliG (and FliY in *B. subtilis*; Bischoff and Ordal, 1992), which allow the motor to rotate in CW and CCW direction. The hook, a flexible structure, is the connection between the basal body and the flagellar filament. It is made from a single protein, FlgE. At the junction between the hook and the filament, the two hook-associated proteins FlgK and FlgL are located. The filament is composed of the flagellin FliC and the capping protein FliD. For assembly, flagellins are transported through the hollow core of the growing filament and attached at the tip.

#### 7.1.1.2 The archaeal flagellum

In archaea, so far only swimming motility driven by flagella (and buoyancy by gas vesicles) has been reported. Although the archaeal flagellum (Figure 7.3) is superficially similar to its bacterial counterpart, it is strikingly different in composition and in assembly (see Thomas *et al.*, 2001a, and references therein). The archaeal flagel-
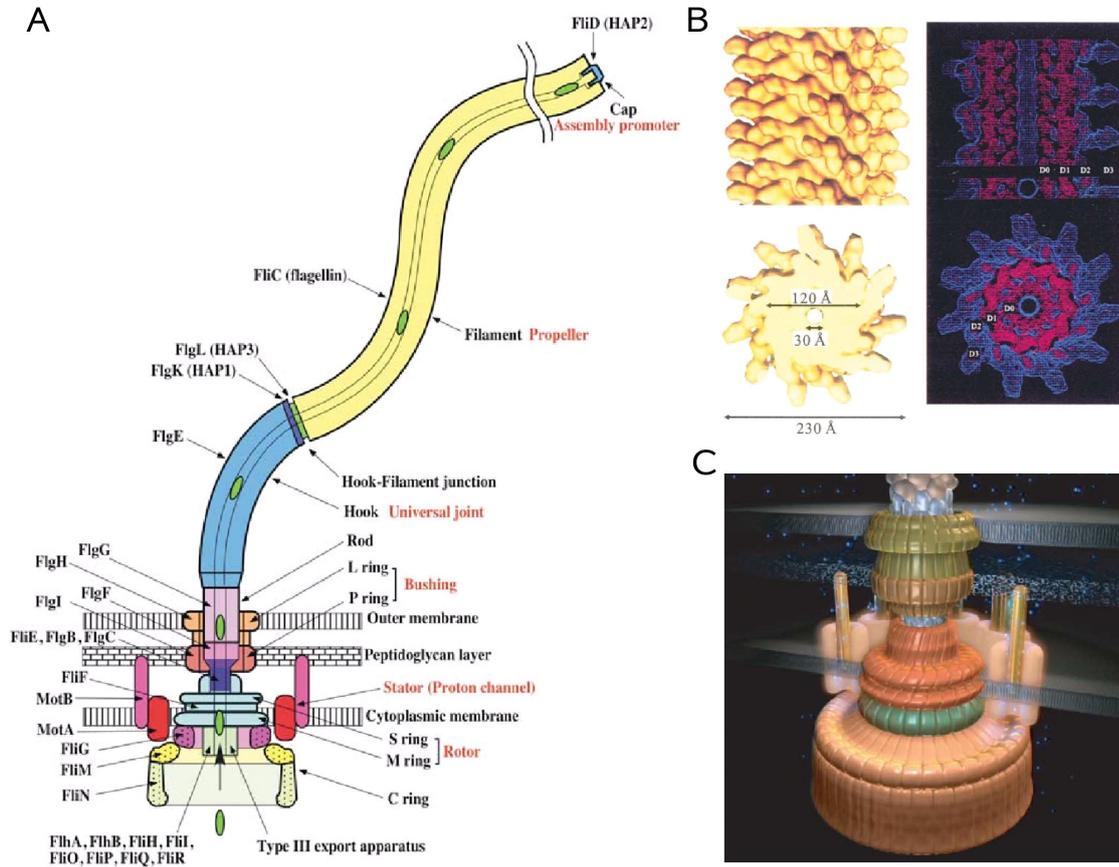
Figure 7.2: **The bacterial flagellar apparatus.** **A** Schematic showing the architecture of the bacterial flagellum. See text for details. **B** Three-dimensional density map of the bacterial flagellar filament. The upper panels show a side view, the lower panels the end-on view from the proximal end. Left: solid surface representation, right: wire frame representation of a 50 Å thick cross section and a 30 Å thick longitudinal section. Note the central channel through which the flagellins are transported to the tip where they are assembled to the growing filament. **C** An image derived from cryo-EM of the bacterial flagella motor. Grey represents the propeller, the driving shaft is cyan, the bushing is yellow and dark orange, the stator is peach with an orange cylinder, the rotor is dark orange and green, and the switch regulator is in pale orange. (Image of K. Namba at Osaka University, Osaka.) **A** and **B** taken from Yonekura *et al.* (2002), **C** from Chiu *et al.* (2006).
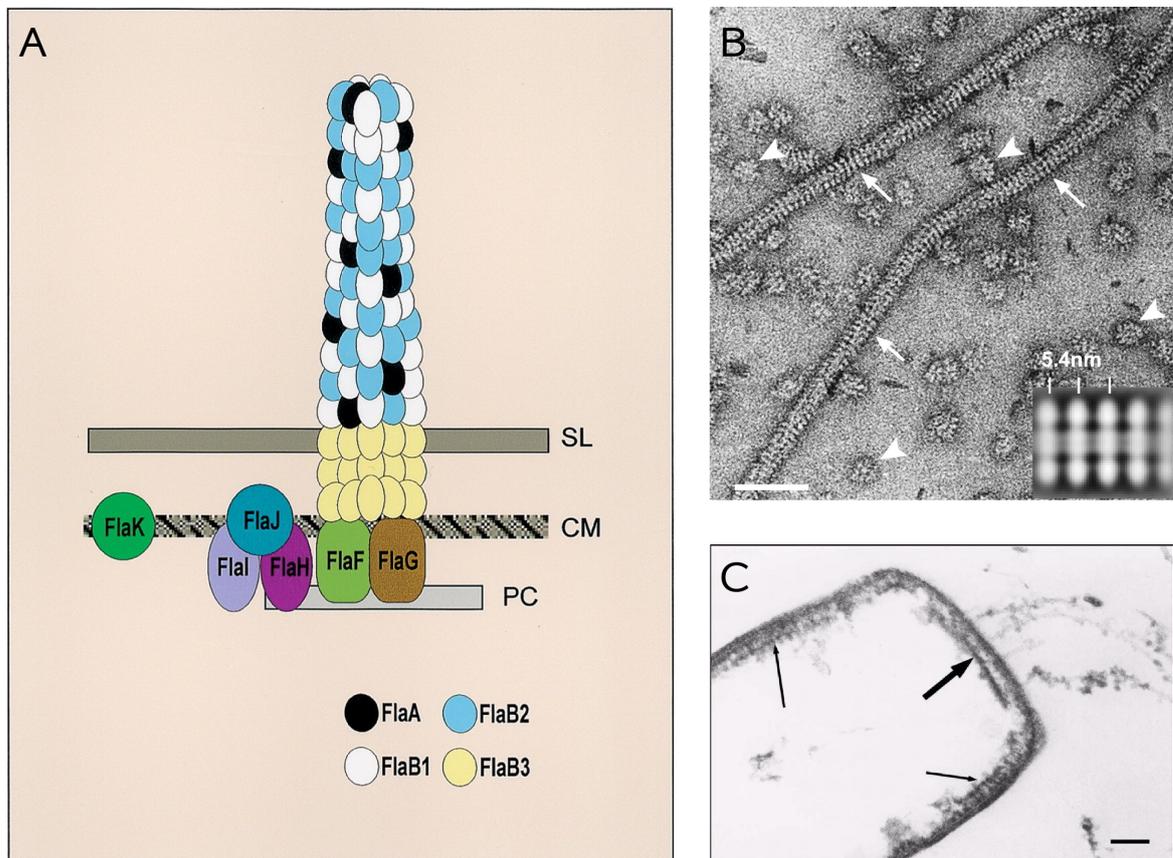
Figure 7.3: **The archaeal flagellar apparatus. A** Speculative scheme showing the architecture of the archaeal flagellum. SL: S layer; CM: cytoplasmatic membrane; PC: polar cap. **B** Electron micrograph of disassembling flagellar filaments of the thermophilic crenarchaeon *Sulfolobus shibatae B12*. The filament is a stacked disk structure. Stacked (arrows) and isolated (arrowheads) disks are seen. The scale bar represents 50 nm. The disk spacing in the averaged image of stacked disks is approximately 5.4 nm (inset). The overall structure and symmetry is the same for the halobacterial flagellar filament (Cohen-Krausz and Trachtenberg, 2008). **C** Electron micrograph of a ultra-thin section of *Halobacterium salinarum* strain VKMB-1231. The bold arrow indicates the discoid lamellar structure (DLS), thin arrows point to polar organelles. The scale bar represents 100 nm. **A** taken from Bardy *et al.* (2003), **B** from Cohen-Krausz and Trachtenberg (2008), and **C** from Metlina (2004).

lar filament is thinner (10-14 nm diameter compared to approximately 24 nm) and lacks a central channel (Figure 7.3 B) (Cohen-Krausz and Trachtenberg, 2002, 2008). That means that it must be assembled by a completely different mechanism than the bacterial filament, because the flagellins cannot be transported through the growing filament to the tip. The observation that the bacterial and archaeal flagellum are different structures was confirmed by genome analysis: For the proteins constituting the bacterial flagellar apparatus, no homologs have been detected in any archaeal genome (Thomas *et al.*, 2001a), suggesting very strongly that the archaeal motility apparatus must be built from different components (Ng *et al.*, 2006). Furthermore, the archaeal flagellar motor is not driven by proton-motive force like most bacterial motors, but either by ATP directly or by an ATP-dependent ion gradient (Streif *et al.*, 2008).

All attempts to isolate the complete archaeal flagellar apparatus have failed. Solubilisation of *H. salinarum* cell envelopes yielded so called "polar cap" structures with many flagella attached (Kupper *et al.*, 1994). In ultra-thin sections of *H. salinarum* cells, a "discoid lamellar structure" was observed at the cell poles in close proximity to the flagella (Figure 7.3 C) (Speranskii *et al.*, 1996; Metlina, 2004). The composition and function of these structures remained unknown.

In some respect, archaeal flagella resemble more the bacterial type IV pili than the bacterial flagella (Cohen-Krausz and Trachtenberg, 2002; Bardy *et al.*, 2004): (i) archaeal flagellins are synthesised as preproteins with short signal peptides which are processed by a specific signal peptidase (called FlaK), similar to bacterial pilins (ii) archaeal flagellins are glycosylated (iii) flagellins are probably added to the filament at the base (iv) the flagella accessory protein FlaI is homologous to the bacterial PilT and PilB proteins, which are involved in pilin export and pilus retraction.

Known components of the archaeal flagellar apparatus are the flagellins, which compose the filament, and a number of conserved proteins which are coded by genes located close to the flagellin genes in archaeal genomes: the flagella accessory proteins FlaC, FlaD, FlaE, FlaF, FlaG, FlaH, FlaI, and FlaJ (Kalmokoff and Jarrell, 1991; Thomas and Jarrell, 2001). The exact role of the Fla proteins is not understood, but it has been shown that they are required for flagellar motility (Patenge *et al.*, 2001; Chaban *et al.*, 2007; Staudinger, 2007). Some of them (FlaH, FlaI, FlaJ) are thought to be involved in flagellin secretion and assembly, similar to the type IV pili system in bacteria. Which proteins constitute the archaeal flagellar motor and its switch remains to be elucidated.

### 7.1.2 The flagellar motor switch is the target of CheY-P

In bacteria, CheY-P binds to the flagellar motor switch protein FliM (Welch *et al.*, 1993), which forms together with FliN and FliG, and in *B. subtilis* also FliY, the motor switch complex. The binding site of CheY-P is the highly conserved N-terminal region of FliM (Bren and Eisenbach, 1998). Binding of CheY-P raises the probability that the motor switches to rotation in the opposite direction (reviewed in Berg, 2003).

CheY-P is the flagellar motor switch factor also in *H. salinarum* and probably also other archaea (Rudolph *et al.*, 1995; Rudolph and Oesterhelt, 1995). However, the interaction site of CheY-P is unknown, since for its target protein in bacteria, FliM, just as for all other proteins constituting the bacterial flagellar apparatus, no homologs can be found in archaeal genomes (Nutsch *et al.*, 2003; Szurmant and Ordal, 2004; Ng *et al.*, 2006). No equivalent to the CheY-P binding region has been identified either.

Besides CheY-P, fumarate is a further factor involved in flagellar motor switching, both in Archaea and Bacteria (see 6.1.1.5). In a recent work, fumarate reductase (FRD) was identified as the target of fumarate at the motor in *E. coli* (Cohen-Ben-Lulu *et al.*, 2008). This enzyme, otherwise functioning in anaerobic respiration, interacts with the flagellar motor switch protein FliG. However, *H. salinarum* does neither code for FRD (which is mainly found in obligate or facultative anaerobic bacteria) nor for FliG. So in this species fumarate must act by a different, till now unknown, mechanism.

## 7.2 Results and Discussion

### 7.2.1 Interaction analysis revealed connectors of Che and Fla proteins

Protein interaction analysis of the halobacterial Che proteins revealed two proteins of unknown function, OE2402F and OE2404R, as interaction partners of CheY, CheD, and CheC2. These proteins are homologous to each other and are coded by adjacent genes, located between the *che* genes and the type B flagellins (Figure 7.1).

To clarify the role of OE2402F and OE2404R, these proteins were used as baits in additional bait fishing experiments. Both proteins were shown to interact with the flagellar accessory proteins FlaCE, and OE2404R also with FlaD (Figure 7.4). The third protein coded by a gene located between the *che* gene region and flagellins

(Figure 7.1), OE2401F, was also subjected to protein interaction analysis, although it was not detected as prey in previous experiments. OE2401F was shown to interact with CheD and OE2402F.

These results indicate that all three proteins play a role in the chemotaxis signalling pathway of *H. salinarum.* Due to their interaction with Che proteins as well as with Fla proteins, OE2402F and OE2404R build a link from the chemotaxis signal transduction system to the archaeal flagellar apparatus. In bacteria, this link is built by the interaction of CheY-P with the flagellar motor switch protein FliM. Hence it can be speculated that OE2401F, OE2402F,



Figure 7.4: **Interactions of the newly identified proteins.** The arrows indicate the direction bait - prey in the bait fishing experiments.

and OE2404R either are part of the archaeal flagellar motor switch, or they are adapters which fit the bacterial-like Che system to the yet unidentified switch. It should be taken into account that CheY in the PPI analysis is with the utmost probability unphosphorylated, which, in analogy to bacteria, should decrease the affinity for the flagellar motor switch. In *E. coli*, CheY binds FliM with a five times lower affinity than CheY-P (McEvoy *et al.*, 1999). However, the overexpression of the bait CheY should, at least partially, compensate the lower affinity, so that fishing switch proteins with unphosphorylated CheY should be possible.

The function of the flagellar accessory proteins is not known, but their critical role in flagellation has been demonstrated (Patenge *et al.*, 2001; Thomas and Jarrell, 2001; Thomas *et al.*, 2001b, 2002; Staudinger, 2007). The FlaE part of FlaCE is homologous to FlaD, both proteins contain a FlaD/E domain. Thus it is possible that the FlaD/E domain is the site of interaction with OE2402F and OE2404R. Deletion of *flaCE* and *flaD* in *H. salinarum* results in cells with a reduced number of flagella that are hardly (Δ*flaD*) or not (Δ*flaCE*) motile (Staudinger, 2007). In *Methanococcus maripaludis*, the deletion of *flaC* resulted in non-motile and non-flagellated cells (Chaban *et al.*, 2007). These findings were interpreted as indicating that FlaCE and FlaD either fulfil essential functions in flagellar secretion and assembly, or that they are part of the flagellar machinery itself. A role only in flagellar assembly would make it rather difficult to interpret the connection to CheY via the proteins identified in this study. Thus, considering the results of the protein interaction study, it seems more likely that FlaCE and FlaD are components of the flagellar motor or associated structures.
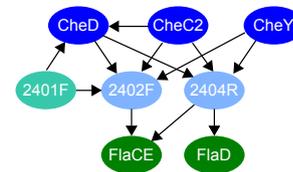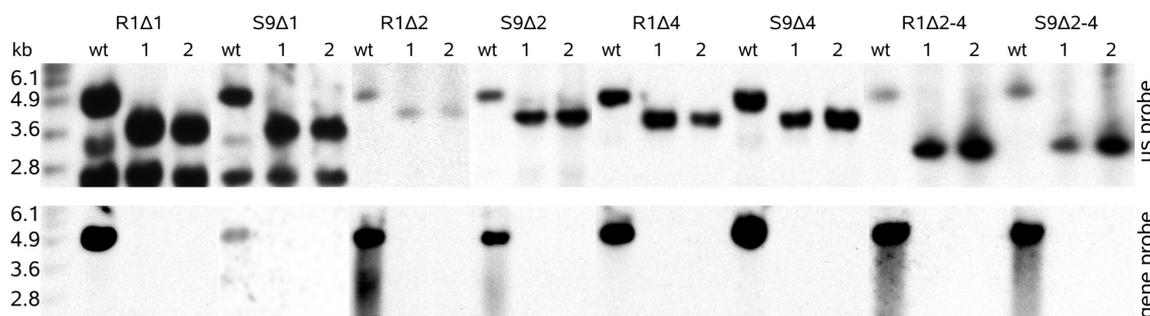
Figure 7.5: **Confirmation of deletion strains by Southern blot analysis.** Each deletion strain was probed with DIG-labelled 500 bp upstream sequence of the target gene(s) (us probe) and DIG-labelled target sequence (gene probe). 1 and 2 indicate the clones of the respective deletion that showed the expected bands and were used for further analysis. The upstream probe for OE2401F revealed an additional band, probably due to unspecific binding. This band, however, did not affect the significance of the blot.

Another possibility to explain the interaction with CheY would be that OE2402F and OE2404R are CheY-P phosphatases. It was demonstrated in *E. coli* that the phosphatase CheZ binds also unphosphorylated CheY (McEvoy *et al.*, 1999). Alternatively, these proteins could also add or remove a different modification to or from CheY, like acetylation (see for example Yan *et al.*, 2008).

OE2401F, OE2402F, and OE2404R also interact with CheD, and OE2402F and OE2404R with CheC2. Since not much is known about the function of these proteins in *H. salinarum* (see Table 6.1), the interpretation of these interactions is hardly possible. It can only be speculated that CheD and the CheCs are, similar to *B. subtilis* (Muff and Ordal, 2007), involved in some kind of feedback loop. The interactions of CheD and CheC2 were discussed in the previous chapter (6.2.3).

None of the interactions described here has been reciprocally confirmed. Whereas the two Fla proteins were not used as baits, in all other cases reciprocal confirmation would have been theoretically possible. However, several reasons can prevent the detection of an interaction with the swapped bait-prey combination. 6.2.1 in the previous chapter gives details on this.

### 7.2.2 Construction of in-frame deletion mutants

To clarify the role of the proteins, in-frame deletion strains for OE2401F-OE2404R (referred to as Δ1, Δ2, and Δ4) and a double deletion ΔΔOE2402F OE2404R (Δ2-4)

were created using a two-step recombination method (Koch and Oesterhelt, 2005). As host, two *H. salinarum* strains were used: Strain R1 was used because it is considered as wildtype, and this strain was used for PPI analysis. Additionally, the deletion mutations were done in strain S9, because S9 cells are better suited for motion analysis (see 2.6.5) and determination of the flagellar rotational bias (see 2.6.6), whereas R1 cells tend to stick to the glass surface of the microscope slides (Spudich and Stoeckenius, 1979). Clones that had undergone the second recombination event were screened for the absence of the target gene(s) by PCR and confirmed by Southern blot analysis using probes for the target gene and its upstream region (Figure 7.5). DNA from the deletion strains did not hybridise with the gene probe, and showed the expected size decrease when probed with the gene's upstream region.

Since the deletions in both parent strains exhibited the same phenotype, they will be discussed together in the following sections. As independent biological replicates, the use of two parent strains gives a high degree of certainty for the phenotypic findings.

### 7.2.3 OE2401F and OE2402F are essential for chemotaxis and phototaxis

To examine the effect of the deletions on chemotaxis and motility, the deletion strains were analysed by swarm plate assays. A swarm plate is a semi-solid agar plate (Halomedium + 0.25 % agar), in which the cells are inoculated. The agar concentration is low enough to allow movement of the cells in the agar. When the cells grow, they metabolise various nutrients, and create a concentration gradient. Cells which are motile and capable of chemotaxis move along this gradient away from the inoculation site, forming extending rings, called swarm rings. Figure 7.6 shows representative swarm plates for each deletion
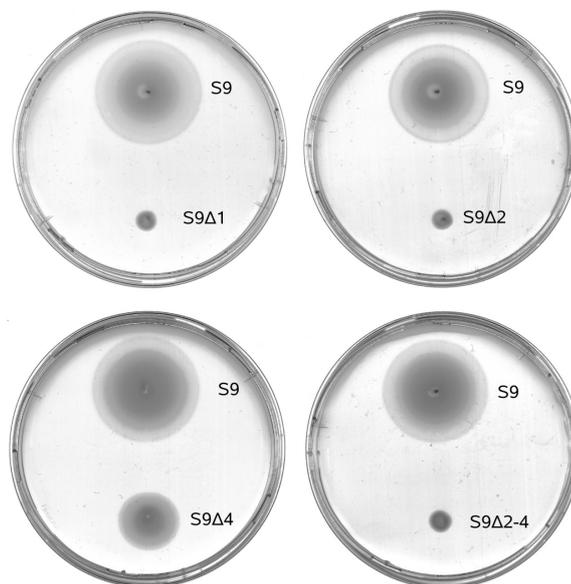


Figure 7.6: **Swarming ability of the deletion strains.** Representative swarm plate for each deletion in S9 after three days of growth at 37°C.
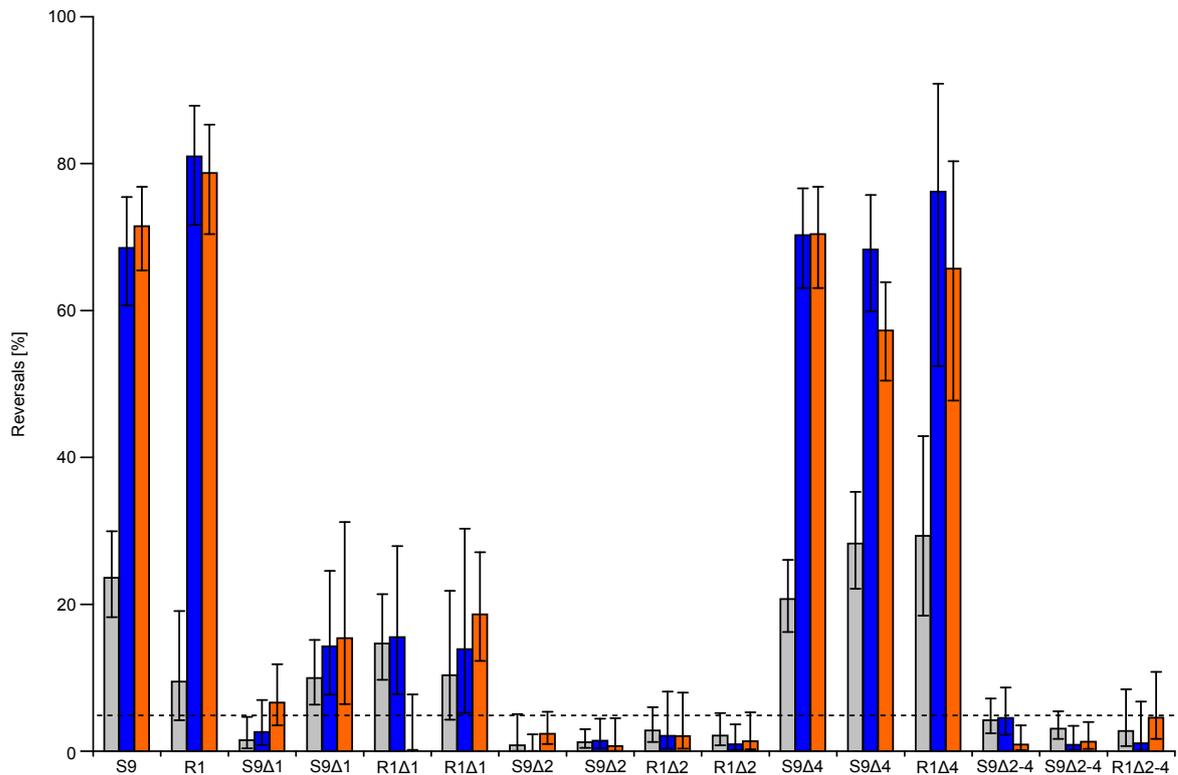
Figure 7.7: **Reversals of the wild type and deletion strains as measured by computer-based cell-tracking.** The percent reversal in a 4 second interval was determined either without stimulation (spontaneous, grey bar), after a blue light pulse (blue bar), or after a step down in orange light (orange bar). Error bars represent the 95 % confidence interval. The dashed line indicates the estimated tracking error of 5 %. Two clones of each deletion strain were measured, except for R1Δ4 and R1Δ2-4.

in S9, compared to wildtype (see Supplementary Figure S1 for all swarm plates). After three days of growth, the wild type strains formed large swarm rings. The deletion strains Δ1, Δ2, and Δ2-4 did not show any swarming. Δ4 cells produced swarm rings, but of a reduced size.

Reduced or impaired ring formation on swarm plates can be due to problems in signal transduction or flagellar motility. In order to determine the defects of the deletion strains, their swimming ability was evaluated by microscopy, and the frequency of reversal of their swimming direction was measured with a computer-based cell-tracking system (Figure 7.7; see Supplementary Table S3 for details). This system automatically determines the rate of reversing cells over a certain observation time.

Visual inspection clearly demonstrated that all deletion strains were motile without detectable swimming defects. The wild type strains showed in a 4 s observation interval

a reversal rate of 10 % (R1) and 25 % (S9) in the unstimulated state. Upon stimulation with a blue light flash or orange light step down (both are repellent stimuli), wild type cells responded effectively with reversal rates of 70-80 %.

In the strains Δ2 and Δ2-4 very low reversal rates of up to 5 % were measured, both spontaneous and after stimulation. These low levels of reversals are mostly due to tracking errors. These strains displayed a smooth-swimming phenotype with hardly any switching, similar to *cheY* and *cheA* deletion strains (Rudolph and Oesterhelt, 1996; del Rosario *et al.*, 2007).

Similar results were obtained for the Δ1 strains. The reversal rates for three of the Δ1 clones were slightly higher than the estimated tracking error of 5 %, but this may have been due to the low number of cells evaluated for these clones, which is also reflected by the broader confidence intervals. A significant increase of reversals after repellent stimulation could not be detected, indicating that this deletion has disabled the response to repellent stimuli. It leads to a strongly reduced switching frequency or even also to a smooth-swimming phenotype.

For Δ4, no significant difference to wild type was visible, both with and without stimulation.

## 7.2.4 Δ1, Δ2, and the double deletion Δ2-4 show almost 100% CW rotational bias

To further characterise the defects of the deletion strains, the flagellar rotational bias was measured. Cells with clockwise (CW) rotating flagella are pushed forward by the right-handed flagellar bundle, whereas cells with counterclockwise (CCW) rotating flagella are pulled backward. These two swimming modes can be distinguished by dark-field microscopy (Alam and Oesterhelt, 1984; Rudolph and Oesterhelt, 1996).

These measurements were only done with the S9 strains since R1 cells tend to stick on the glass slides which makes counting of motile cells and determining their swimming direction cumbersome. Excecpt for Δ1, only one clone for each deletion was analysed, because the results were in complete agreement with the other phenotypic findings, and this experiment is very

Table 7.1: **Flagellar rotational bias of the deletion strains**

| Strain | CW | CCW | % CW |
|--------|-----|------|------|
| S9 | 290 | 210 | 58 |
| S9Δ1 C1 | 494 | 6 | 99 |
| S9Δ1 C2 | 481 | 19 | 96 |
| S9Δ2 | 500 | 0 | 100 |
| S9Δ4 | 511 | 498 | 51 |
| S9Δ2-4 | 499 | 1 | 100 |

time-consuming. The two S9$\Delta$1 clones were investigated because they showed a slightly different phenotype in the phototaxis measurements (smooth-swimming vs. some residual switching).

The numbers of cells observed swimming in each direction are shown in Table 7.1. Wildtype cells show a distribution between forward and backward swimming of close to 50:50, as expected (Marwan *et al.*, 1991; Rudolph and Oesterhelt, 1996). Cells of the deletion strain $\Delta$1, $\Delta$2, and the double deletion $\Delta$2-4 show a bias toward forward swimming of almost 100%. Again, this corresponds to the phenotype of *cheY* and *cheA* deletion strains (Rudolph and Oesterhelt, 1996; del Rosario *et al.*, 2007). The slight discrepancy of both S9$\Delta$1 clones found in the cell tracking assay also showed up in this experiment, proving the reliability of the applied methods. Cells lacking OE2404R exhibit a rotational distribution of nearly 50:50, similar to wildtype.

## 7.2.5 Interpretation of deletion phenotypes

The phenotypic characteristics of the deletion strains (see Table 7.2 for an overview) demonstrated that OE2401F and OE2402F are essential for the ability to control the direction of flagellar rotation. Without these proteins, the flagella rotate only clockwise and switching is not possible, making the cells incapable of any tactic response. The deletion of OE2404R resulted only in a weak phenotype and the role of this protein remained unclear.

Cells of the strains $\Delta$1, $\Delta$2, $\Delta$2-4 did not or hardly show spontaneous switching, did not respond to repellent light stimulation, and were unable to form swarm rings. This means the deletion of OE2401F and OE2402F resulted in

Table 7.2: **Summary of phenotypes**

|              | $\Delta$1 | $\Delta$2 | $\Delta$4 | $\Delta$2-4 |
|--------------|-----------|-----------|-----------|-------------|
| Motility     | +         | +         | +         | +           |
| Chemotaxis   | -         | -         | (+)       | -           |
| Phototaxis   | -         | -         | +         | -           |
| CCW rotation | -         | -         | +         | -           |

similar phenotypes. This is noteworthy because these proteins were demonstrated to interact, indicating that they probably act cooperatively to perform their function. None of these strains exhibited defects in flagellar motility. The cells of these deletion mutants rotate their flagella almost exclusively clockwise. Thus they behave exactly like *cheY* and *cheA* deletion strains (Rudolph and Oesterhelt, 1996; del Rosario *et al.*, 2007).

These findings suggest that without OE2402F or OE2401F the Che system and the flagellum are decoupled. This can be, if either there is no CheY-P, or CheY-P

is present, but not effective. The first of these two possibilities seems less likely, because the PPI data suggest a role for OE2401F and OE2402F between CheY and the flagellum, and not upstream of CheY. Additionally, the homology of the archaeal Che system to the bacterial one argues against the first hypothesis: Current understanding is that the Che system of *H. salinarum*, with the ten known Che proteins, is complete up to CheY-P. Only for the part downstream of CheY-P have no homologs to bacterial proteins been found. However, it would be possible that the deletion of OE2401F and OE2402F do not affect the level of CheY-P directly but via a yet unknown side mechanism. In *B. subtilis*, for example, cells deleted for CheD exhibit a very tumbly phenotype, similar to CheA mutants (Rosario and Ordal, 1996). In *H. salinarum*, however, the deletion of no other Che protein than CheA and CheY actually leads to cells with smooth-swimming phenotype (compare Table 6.1). In all other cases some residual switching and swarm ring formation is present (Staudinger, 2007), supporting the idea that the defect of the OE2401F and OE2402F deletion strains is located between CheY-P and the flagellum.

Besides the two above mentioned deletions, a smooth-swimming phenotype has also been observed for the CheY** strain (Staudinger, 2007; del Rosario *et al.*, 2007). In this strain, the CheY protein carries two point mutations: the aspartate in position 10 is replaced by lysine, and the tyrosine in position 100 by tryptophane. In *E. coli*, a similarly modified CheY mimics CheY-P. Hence *E. coli* cells expressing this mutated CheY protein exhibit a tumbly phenotype (Scharf *et al.*, 1998). In *H. salinarum*, the in the first moment unexpected phenotype of the CheY** strain (expected was an increased switching frequency) was explained by introducing asymmetry in the motor switch model (del Rosario *et al.*, 2007; Staudinger, 2007). Since CheY** causes a smooth-swimming phenotype, the results obtained with Δ1, Δ2, Δ2-4 could also be explained with (drastically) elevated CheY-P concentrations. This could be due to missing CheY-P phosphatase activity or hyperactivation of CheA. This hypothesis seems rather unlikely because even in the absence of a phosphatase the level of CheY-P should be limited by the short half-life of CheY-P, which is in the range of few seconds (Rudolph *et al.*, 1995). The idea of hyperactivation of CheA is not supported by the obtained PPI data. This could only be explained by some feedback mechanism via the additional Che proteins, but then, again, remains the question why the deletion of these additional Che proteins results in a less severe phenotype. For the whole speculation based on the phenotype of CheY** it should be noted that the effect of

the CheY doublemutation was just deduced from the *E. coli* protein – in *H. salinarum* it cannot be ruled out that the mutated protein is just non-functional instead of constitutively active (Staudinger, 2007).

A further possibility to explain the behaviour of Δ1, Δ2, Δ2-4 is an influence of the deleted proteins on the switch factor fumarate, which might act independently of the Che system (see 6.1.1.5 for details). The first evidence for fumarate as switch factor was found in a straight-swimming mutant which could be reverted to wild-type behaviour by introducing fumarate into the cells (Marwan *et al.*, 1990), demonstrating that a defect in fumarate signalling can cause a phenotype similar to the one observed for Δ1, Δ2, Δ2-4. However, the detected protein interactions with CheY provide strong evidence that the proteins examined in the current study play a role in the action of CheY and not exclusively in fumarate switching.

The role of OE2404R remained unclear. The Δ4 strains were not distinguishable from wildtype strains in the phototaxis measurement and with respect to the flagellar rotational bias but produced significantly smaller swarm rings. Reduced swarm ring size is generally considered as outcome of diminished chemotaxis capability (given that the motility in itself is not affected). Several alternative hypotheses can be envisioned to explain the differences in the phototaxis measurements (no difference to wt) and swarming (reduced when compared to wt). First, OE2404R might only be involved in chemotaxis and not phototaxis signalling. Second, this protein might be required for fast and effective adaptation. Third, OE2404R might be required for fine tuning of the response. The first hypothesis seems rather unlikely since no other evidence exists that chemical and light signals utilise separate pathways in *H. salinarum*. The other possibilities are related to the phototaxis assay: This assay monitors the reaction after one strong and sudden change in light intensity, but does not report the adaptation efficiency or the reaction to more subtle stimuli. Further experiments should be done to test these three hypotheses. Dose-response curves for the phototactic behaviour would be a promising approach to discriminate between hypothesis one and three. Monitoring the cellular response after repeated phototactic stimulation could be used to test hypothesis two.

Finally, it should be mentioned that it is not exactly known what determines the swarming capability of *H. salinarum* cells. The widely accepted explanation for swarm ring formation is the formation of chemical gradients due to consumption of nutrients and excretion of metabolic end- or by-products by the cells at the site of inoculation.

Cells capable of chemotaxis and motility sense these gradients and bias their movement from the site of inoculation to the periphery. However, straight-swimming mutants do not form swarm rings at all, although straight movement in any direction is the fastest way to reach the periphery. Thus switching seems to be crucial for movement in semi-solid agar. Theoretically random switching should also lead to spreading of the cells, similar to diffusion driven by Brownian motion with its randomly occurring turnarounds after collisions. To what extent biasing the switching events in response to chemical concentrations increases the swarm ring size has not been investigated for *H. salinarum*. For *E. coli*, Wolfe and Berg (1989) demonstrated that nonchemotactic cells can form swarm rings if they rotate their flagella both CW and CCW. The cells spread more effectively when tumbles were more frequent. This behaviour was explained by the observation that cells that do not tumble tend to get trapped in the agar. If these findings can be transferred to *H. salinarum* cells, which do not tumble at all but swim forward and backward, remains to be investigated.

Overall, it can be said that OE2404R is involved in taxis signal transduction in *H. salinarum*, but it either fulfils a non-essential function or it can be replaced by its homolog, OE2402F, with only minor constraints.

## 7.2.6 Complementation of deletions reverted their phenotype to that of wildtype

All deletions in S9 background were complemented by reintroducing the deleted gene(s) *in cis*. The phenotype of the complementations was examined by swarm plates, and for the single deletions by phototaxis measurements. All complementations behaved exactly like the wild-type strains (see Figure 7.8), confirming that the phenotypes observed in the mutants were a direct result of their gene deletions.

## 7.2.7 Bioinformatic analysis

Bioinformatics analysis was done to collect information on the three unknown proteins. Starting point for this was a homology search, and querying databases like COG (Tatusov *et al.*, 1997, 2003), Pfam (Finn *et al.*, 2006, 2008), InterPro (Mulder *et al.*, 2007), SMART (Schultz *et al.*, 1998), and STRING (von Mering *et al.*, 2007). The goal of homology search was to identify orthologs from other organisms for which some
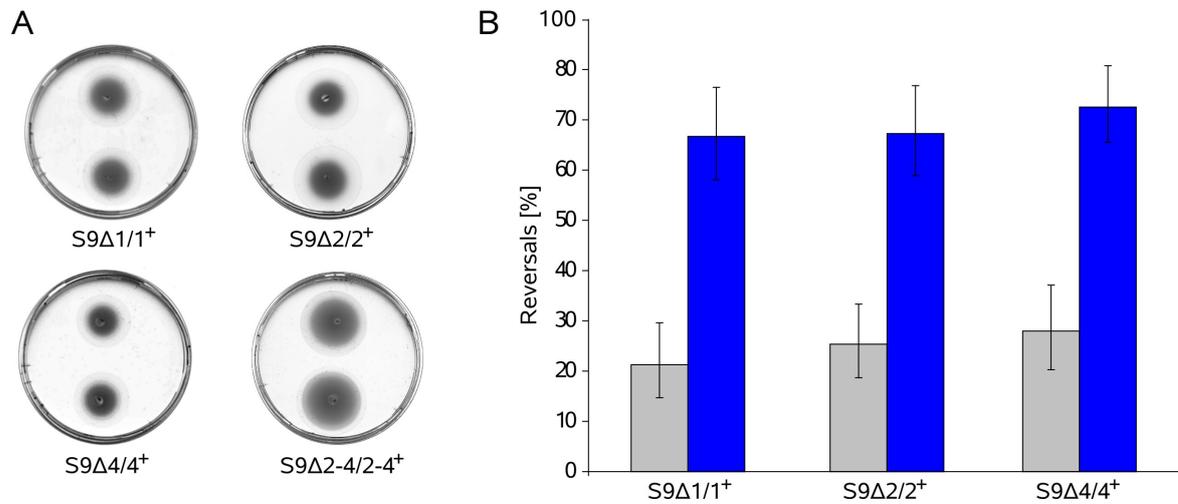
Figure 7.8: **Phenotype of complementations.** **A** Swarm plate assay. On each plate the complementation strain (bottom) is compared to the respective wildtype strain (top). **B** Computer-based cell tracking for the complementations of each single deletion. The percent reversal in a 4 second interval was determined either without stimulation (spontaneous, grey bar) or after a blue light pulse (blue bar). Error bars represent the 95 % confidence interval.

knowledge might exist, and to unravel correlations between the occurrence of the here investigated proteins and Che and Fla proteins.

### 7.2.7.1 Occurrence of *che* and *fla* genes in archaeal genomes

To have a reference for such co-occurrence comparisons, an exhaustive search for orthologs of Che and Fla proteins in all completely sequenced archaeal genomes published until October 2007 was done. Since homology searches using Psi-Blast were not sufficient to comprehensively identify homologs of some proteins (especially small proteins with rather low conservation like FlaC, D, E, F, and G were problematic), and did not allow the discrimination between orthologs and other homologs for other proteins (e. g. CheY and other response regulators), a combination of different methods was used for ortholog identification (see 2.6.7 for details). The resulting table of orthologs is shown in Supplementary Table S4.

As observed before (Klein, 2005), no *che* genes were detected in any archaeal genome without *fla* and flagellin genes. In contrast, several archaeal species contain *fla* genes and flagellins, but no *che* genes, leading to the conclusion that these species are motile, but their motility is not controlled by a Che system. *che* genes have not been detected in a crenarchaeal genome. If *che* genes were found, there is always the whole set

consisting of *cheA*, *cheB*, *cheC*, *cheD*, *cheR*, *cheW*, and *cheY* present. An exception is *Methanosarcina barkeri*, which has lost the *cheC* gene (the genomic position where *cheC* is located in the other *Methanosarcina* species still contains remnants of the N-terminus of *cheC*). Several archaeal species contain multiple copies of various *che* genes; the front-runner is *Methanospirillum hungatei* with 35 genes classified as *che* orthologs by the used method. A noteworthy finding is a CheA-CheC fusion protein detected in the genome of *M. hungatei*.

If in a crenarchaeal genome *fla* genes were identified, there were at least one flagellin, *flaG*, *flaH*, *flaI*, and *flaJ*, and usually also *flaF* (except in *Aeropyrum pernix*) present. The *flaG* gene in the sequenced strain of *Sulfolobus solfataricus* is interrupted by a transposase, but this insertion is neither stable under laboratory conditions nor is it found in a closely related strain (Szabó *et al.*, 2007). In euryarchaeota, there is additionally always a *flaD/E* gene present, if they possess *fla* genes. *flaD* and *flaE* genes could not be discriminated by the applied method, so they were merged into one ortholog group. Two versions of *flaD/E* genes can be distinguished: The species of the classes Methanomicrobia and Archaeoglobi code for a version of a FlaD/E protein (referred to as FlaD/E$^M$ in the following) with only low homology to the FlaD and FlaE proteins found in other euryarchaeal genomes (Ng *et al.*, 2006; Desmond *et al.*, 2007). A special case is *Methanococcoides burtonii* (class Methanomicrobia), which possesses both the versions of FlaD/E proteins, each in a complete *fla* gene region. This is an indication of lateral gene transfer (LGT) of a whole *fla* gene region. Such an LGT has also been proposed in a detailed study of the phylogenomics of the archaeal flagellum (Desmond *et al.*, 2007). In genomes with *flaD/E$^M$* (or in the case of *Methanococcoides burtonii* the genome region with *flaD/E$^M$*), no *flaC* gene, or *flaC* domain fused to a *flaE* gene, was found. In all other euryarchaeota with *fla* genes, FlaC is either coded as separate protein or as domain fused to an FlaD/E domain. Like the *che* genes, also the *fla* genes and flagellins are present in multiple copies in some genomes.

### 7.2.7.2 Only few findings for OE2401F

OE2401F was formerly annotated as "phycocyanin alpha phycocyanobilin lyase homolog" due to homology with this protein from cyanobacteria and red algae. The annotation was now changed to "conserved *che* operon protein". By Pfam, it is classi-

fied as a HEAT_PBS or HEAT family protein. These proteins are predicted to contain short bi-helical repeats. Several, but not all members of this family are thought to have lyase activity (Finn *et al.*, 2008).

For OE2401F, homology search turned out to be difficult because the repeats led to a high number of non-significant matches. Thus it was not possible to identify a reliable set of orthologs from other organisms and no conclusions about co-occurrence of this protein family with *che* or *fla* genes could be drawn. Close homologs were identified in the *che* and *fla* gene regions of the halophilic archaea *N. pharaonis* and *H. marismortui.* The idea that OE2401F and OE2402F act cooperatively to perform their function is also supported by the genomic location of OE2401F and its homologs in the haloarchaeal *che* gene regions, where it is always adjacent to a DUF439 protein. Additionally, HEAT-like repeat proteins are present in all sequenced haloarchaeal genomes (the above mentioned, *H. walsbyi*, and *H. salinarum*) in other genomic context. For none of these proteins any functional knowledge could be obtained. In the chemotaxis gene regions of other archaeal species, no homologs of OE2401F were found. Hence it remains to be investigated if these proteins are restricted to haloarchaea, or if similar proteins, coded elsewhere in the genome, play a role in taxis signalling also in other archaeal species.

Although homology search revealed no correlation between OE2401F homologs and *che* or *fla* genes, the examination of the archaeal *fla* gene regions resulted in a noteworthy finding. Adjacent to the flagellin genes in several archaeal genomes (several Methanococci and Thermococci), a protein belonging to the Adaptin_N family is located. Adaptin_N belongs to the same superfamily as HEAT/HEAT_PBS, the Armadillo repeat superfamily (Finn *et al.*, 2008). If the Adaptin_N proteins adjacent to the flagellins fulfil a similar function as OE2401F and its homologs and, if so, which function this is, remains elusive.

### 7.2.7.3 OE2402F and OE2404R belong to a family of unique archaeal Che proteins

OE2402F and OE2404R, both annotated as conserved hypothetical protein, are homologous to each other and belong to the protein family DUF439 (Finn *et al.*, 2008) and the cluster of orthologous groups COG2469. DUF439 is described as "archaeal protein of unknown function", COG2469 as "uncharacterized conserved protein".

Figure 7.9: **Organisation of chemotaxis genes in known archaeal genomes.** Known chemotaxis genes are shown in blue. Genes coding for proteins of the family DUF439 are shown in light blue, genes coding for HEAT domain proteins in cyan. Grey indicates that, where no name is given, the function of the coded protein is unknown, or the protein is probably unrelated to chemotaxis (S6: 30S ribosomal protein S6e) A // sign indicates separated genome regions. The asterisk indicates that this protein is interrupted by a frame-shift mutation.
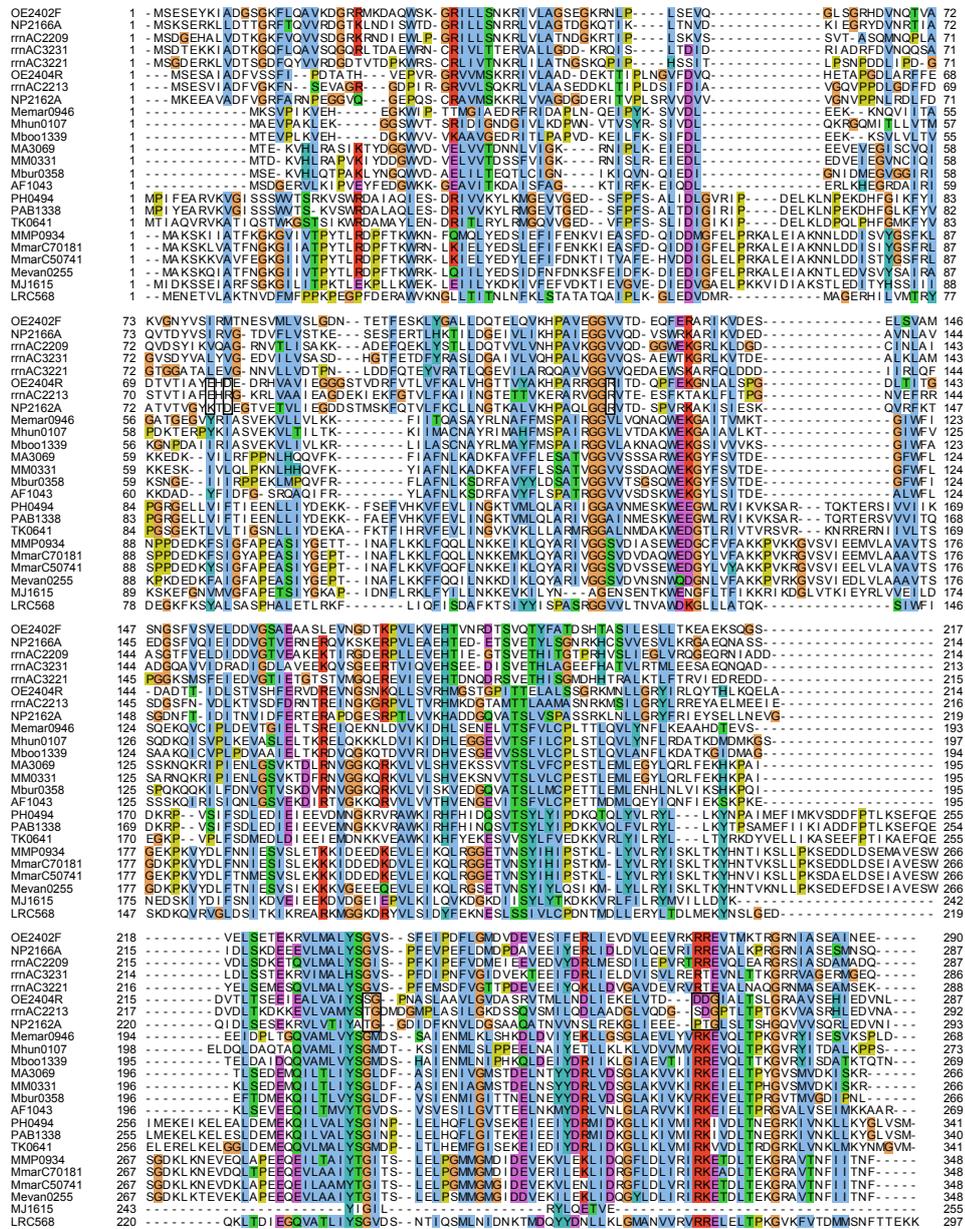
Figure 7.10: **Multiple alignment of the members of the protein family DUF439.** The species are: OE *Halobacterium salinarum*, NP *Natronomonas pharaonis*, rrn *Haloarcula marismortui*, Memar *Methanoculleus marisnigri*, Mhun *Methanospirillum hungatei*, Mboo *Candidatus Methanoregula boonei*, MA *Methanosarcina acetivorans*, MM *Methanosarcina mazei*, Mbur *Methanococcoides burtonii*, AF *Archaeoglobus fulgidus*, PH *Pyrococcus horikoshii*, PAB *Pyrococcus abyssi*, TK *Thermococcus kodakaraensis*, MMP *Methanococcus maripaludis S2*, MmarC7 *Methanococcus maripaludis C7*, MmarC5 *Methanococcus maripaludis C5*, Mevan *Methanococcus vannielii*, MJ *Methanococcus jannaschii*, LRC *uncultured methanogenic archaeon RC-I.* Colours are according to the ClustalX colouring scheme. The boxes point to peculiarities of the second DUF439 protein of the haloarchaea.

Homology searches have shown that no members of the family DUF439 can be found outside the domain archaea, and among the archaea, the presence of genes coding for such a protein strictly correlates with the presence of *che* genes (see Supplementary Table S4). The only exceptions are *Methanocaldococcus jannaschii*, which does not posses *che* genes but a DUF439 homolog, and *Methanosarcina barkeri*, which has *che* genes but no DUF439.

Examination of the genomic context revealed that the genes coding for DUF439 proteins are always located in the chemotaxis gene regions (Figure 7.9). The exceptions are *Methanocaldococcus jannaschii*, of course, and two of the four paralogs in *H. marismortui*. In 10 of 17 species the DUF439 protein is adjacent to CheY, which supports the interaction found between these proteins (Dandekar *et al.*, 1998).

The only archaeal *che* gene regions without DUF439 homolog are the *che2* regions of the Methanosarcina species. In *Methanosarcina barkeri* this is the only *che* region, as this species does not contain the part of the genome where the *che1* region in *M. mazei* and *M. acetivorans* is located (Galagan *et al.*, 2002; Deppenmeier *et al.*, 2002; Maeder *et al.*, 2006). The *che* region of *M. barkeri* is special in that it has lost *cheC*, which is present in all other archaeal *che* regions, so it might be not functional at all. For none of the Methanosarcina species flagellar motility was observed (Garrity *et al.*, 2001), although they probably have this capability since their genomes contain flagellins and a complete set of *fla* genes (see Supplementary Table S4). If the Methanosarcina *che2* region plays a role in controlling flagellar motility remains to be elucidated.

A multiple alignment of all members of the family DUF439 revealed only few conserved residues and several weakly conserved regions (Figure 7.10). Since no conserved motif could be detected the multiple alignment gave no hint on the function of the proteins. It is noteworthy that the protein from *Methanocaldococcus jannaschii* (no Che proteins) is less conserved and truncated at the C-terminus while this is well conserved in all other species. Hence it is likely that this protein is either non-functional or fulfils a different function. The presence of a DUF439 protein in the genome of *M. jannaschii*, while *che* genes are absent, can be explained by two scenarios: Either this gene is the remnant of a former *che* gene region which was lost. Since the DUF439 protein is located at the boundary of the *che* gene region in the other *Methanococcus* species (Figure 7.9), such an incomplete gene loss would have been possible and could also explain the C-terminal truncation. Alternatively, this gene could have been gained by horizontal gene transfer.

Two or more copies of DUF439 proteins were only found in the motile haloarchaea *H. salinarum, N. pharaonis,* and *H. marismortui.* All three species contain a second homolog in or adjacent to the *che* gene region. These second homologs lack several residues conserved in all other proteins of the family DUF439 (see boxes in Figure 7.10). Hence they fulfil probably a different function than the "main" DUF439 protein. This is consistent with the phenotypic results obtained for the deletions: The deletion of OE2404R resulted, other than the deletion of OE2402F, only in a weak pheno-



Figure 7.11: **Phylogenetic analysis of DUF439 proteins.** Unrooted phylogenetic tree by neighbour-joining, calculated from the multiple alignment shown in Figure 7.10. Species can be derived from the prefix of the protein name as explained in the legend of Figure 7.10.

type. Phylogenetic analysis (Figure 7.11) revealed that the second homologs in the *che* gene region of the haloarchaea (OE2404R, NP2162A, rrnAC2213) form a separate branch in the phylogenetic tree. That means that they either arose by a gene duplication prior to the divergence of the haloarchaea, or they arose later, and were distributed by lateral gene transfer. However, the second explanation seems unlikely, because it cannot explain the conserved localisation in the *che* gene region. *H. marismortui* contains two additional DUF439 homologs located elsewhere in the genome. These two paralogues resemble more the "main" DUF439 proteins than the second homolog of the haloarchaea as can be seen in the multiple alignment and the phylogenetic tree. If they also fulfil a function in taxis signalling remains elusive.

Overall, the presence of a DUF439 protein in (almost) all archaeal *che* gene regions indicates that these proteins are not only essential for chemo- and phototaxis in *H. salinarum*, but constitute a hitherto unrecognised class of archaeal chemotaxis proteins. The Che proteins in archaea were identified by homology to their bacterial counterparts (Rudolph and Oesterhelt, 1995; Rudolph *et al.*, 1995; Szurmant and Ordal, 2004, and references therein). The absence of DUF439 in bacteria might explain why these proteins were not recognised earlier. Since these proteins connect the chemotaxis system to the archaeal flagellum we propose the name CheF for this protein family.

# 7.3 Conclusions

Although the taxis signal transduction system in Archaea is similar to its bacterial counterpart, its target, the archaeal flagellar apparatus, is a completely different structure than the bacterial one. The connection between the bacterial-like signal transduction system and the unique archaeal flagellar motor has remained elusive (see Thomas *et al.*, 2001a; Szurmant and Ordal, 2004, for review).

Using protein-protein interaction analysis, we have identified three proteins in *Halobacterium salinarum* that connect the chemotaxis system and the archaeal flagellar apparatus. These proteins interact with the chemotaxis proteins CheY, CheD, and CheC2, as well as the flagella accessory proteins FlaCE and FlaD. Thereby they constitute the first known link between these systems. Two of the proteins belong to the protein family DUF439, the third is a HEAT_PBS family protein.

Strains deleted for the HEAT_PBS protein or one of the DUF439 proteins proved unable to switch the direction of flagellar rotation. In these mutants, flagella rotate only clockwise, which results in exclusively forward swimming cells that are unable to respond to tactic signals. Deletion of the second DUF439 protein had only minimal effects.

By homology searches, HEAT_PBS proteins could be identified in the chemotaxis gene regions of all motile haloarchaea sequenced so far, but not of other archaeal species. DUF439 proteins, however, are inherent parts of archaeal chemotaxis gene regions, and they are restricted to this genomic context. Altogether, these results demonstrate that we have identified hitherto unrecognised archaea-specific Che proteins that are essential for relaying taxis signalling to the flagellar apparatus in the archaeal domain.

# 8 Concluding remarks

The aim of this study was to improve the understanding of the taxis signal transduction system of *H. salinarum* by protein protein interaction analysis.

To achieve this goal, it was first necessary to find an appropriate method which enables the detection of interactions between halophilic proteins. It was tried if the yeast two-hybrid system can be used for this purpose. This assay was used to screen a test set of known interactors from *H. salinarum*, but here it failed in all cases. The reason was probably the halophilic adaptation of the proteins. Hence it was concluded that the yeast two-hybrid system is not the tool of choice to approach the problem of this study. Next, an affinity purification method for halobacterial protein complexes was developed which enables the identification of protein interaction partners by mass spectrometry. This method allowed the succesful detection of known and new interactions between halobacterial proteins. Therewith evidence was provided that the developed method can be used to address the main question of this study.

With this method, the protein interactions of the ten known chemotaxis proteins from *H. salinarum* were investigated and an interaction network of the halobacterial taxis signal transduction system constructed. This network points to several new aspects and components of the system. However, the knowledge of protein interactions alone was not sufficient to unravel the exact meaning of these new facets.

For this, a more in-depth analysis is needed. This was done for three proteins which were found to interact both with proteins of the taxis signal transduction system as well as the motility apparatus. Examination of deletion mutants revealed that two of them are essential for relaying taxis signalling to the flagellum. By homology searches, it was demonstrated that two of the new interaction partners belong to a family of proteins which are inherent components of archaeal chemotaxis gene regions and which are restricted to this genomic context. Overall, the combination of a top-down approach to examine a complete cellular module, followed by a bottom-up approach to have a closer look on some details proved succesful. With this strategy, it was possible

to identify overlooked archaeal chemotaxis proteins, which build the first known link from the taxis signal transduction system to the archaeal flagellar apparatus.

# 9 Supplementary material

Table S1: **Identification by MALDI TOF.**

| Slice | Protein | Mascot Score |
|---|---|---|
| 1 | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA | 118 |
| 2 | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA | 128 |
| 3 | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA | 136 |
| 4 | (OE2205F) chitinase (EC 3.2.1.14) | 70 |
|  | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA | 51 |
| 5 | (OE2205F) chitinase (EC 3.2.1.14) | 85 |
|  | (OE2201F) chitinase (EC 3.2.1.14) | 54 |
| 6 | (OE2205F) chitinase (EC 3.2.1.14) | 105 |
| 7 | (OE2205F) chitinase (EC 3.2.1.14) | 85 |
|  | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA | 52 |
| 8 | (OE2201F) chitinase (EC 3.2.1.14) | 93 |
|  | (OE2206F) probable chitinase (EC 3.2.1.14) | 91 |
| 9 | (OE2419R) purine-binding chemotaxis protein cheW1 | 85 |
| 10 | (OE2419R) purine-binding chemotaxis protein cheW1 | 57 |
| 11 | (OE4643R) conserved protein | 116 |
| 12 | (OE4643R) conserved protein | 57 |
| 13 | (OE2303F) DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) | 37 |
| 14 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 32 |
| 15 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 71 |
| 16 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 35 |
| 17 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 107 |
| 18 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 119 |
|  | (OE4759F) cell surface glycoprotein precursor | 68 |
| 19 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 30 |
| 20 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 104 |
| 21 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 241 |
| 22 | (OE4740R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A' | 230 |
| 23 | (OE1500R) pyruvate, water dikinase (EC 2.7.9.2) (phosphoenol | 25 |
| 24 | (OE1737R) dnaK-type molecular chaperone hsp70 | 85 |
| 25 | (OE5212F) SMC-like protein sph1 | 68 |
| 26 | (OE2205F) chitinase (EC 3.2.1.14) | 31 |
| 27 | (OE2205F) chitinase (EC 3.2.1.14) | 49 |
| 28 | (OE4742R) DNA-directed RNA polymerase (EC 2.7.7.6) chain B" | 33 |
| 29 | (OE4742R) DNA-directed RNA polymerase (EC 2.7.7.6) chain B" | 87 |
| 30 | (OE2415R) taxis sensor histidine kinase (EC 2.7.3.-) cheA | 47 |
| 31 | (OE3762R) glycerol kinase (EC 2.7.1.30) | 36 |
| 32 | (OE2648F) conserved protein | 79 |
| 33 | (OE4268F) glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3 | 39 |
| 34 | (OE4739R) DNA-directed RNA polymerase (EC 2.7.7.6) chain A" | 57 |

Table S1: (continued)

| Slice | Protein | Mascot Score |
|---:|---|---:|
| 35 | (OE2631F) DNA-directed RNA polymerase (EC 2.7.7.6) chain D | 64 |
| 36 | (OE2631F) DNA-directed RNA polymerase (EC 2.7.7.6) chain D | 42 |
| 37 | (OE3542R) protein OE3542R | 51 |
| 38 | (OE3542R) protein OE3542R | 56 |
| 39 | (OE3542R) protein OE3542R | 72 |
| 40 | (OE3541R) probable heat shock protein | 44 |
| 41 | (OE3541R) probable heat shock protein | 45 |
| 42 | (OE1279R) DNA-directed RNA polymerase (EC 2.7.7.6) epsilon c | 148 |
| 43 | (OE4735R) ribosomal protein S7 | 47 |
| 44 | (OE2874F) hypothetical protein | 33 |
| 45 | (OE3126F) protein OE3126F | 37 |
| 46 | (OE2628F) ribosomal protein S4 | 99 |
| 47 | (OE3817R) ribosomal protein S19.eR | 52 |
| 48 | (OE4736R) ribosomal protein S12 | 33 |
| 49 | (OE1352F) hypothetical protein | 27 |
| 50 | (OE3491R) heat shock protein homolog | 60 |
| 51 | (OE4696R) hypothetical protein | 28 |
| 52 | (OE2084R) transcription initiation factor TFB | 36 |
| 54 | (OE1279R) DNA-directed RNA polymerase (EC 2.7.7.6) epsilon c | 64 |
| 55 | (OE1279R) DNA-directed RNA polymerase (EC 2.7.7.6) epsilon c | 124 |

Best hits are shown and further proteins with a Mascot score > 47.

Table S2: **Interactions of the Che proteins.**

| Bait | Prey | Loc. | #d | #i | R | MAS | Prey Protein Name | Comment |
|---|---|---|---|---|---|---|---|---|
| OE1428F | OE1001F | TM | 1 | 0 | N | 12.01 | CHP | |
| OE1428F | OE1005F | TM | 1 | 0 | N | 10.88 | ABC-type transport system permease protein | |
| OE1428F | OE1268F | C | 1 | 0 | N | 13.6 | probable transcription regulator boa1 | |
| OE1428F | OE1500R | C | 1 | 0 | N | 13.2 | pyruvate, water dikinase | |
| OE1428F | OE2042F | TM | 1 | 0 | N | 8.73 | probable copper-transporting ATPase | |
| OE1428F | OE2397F | E | 1 | 0 | N | 9.61 | flagellin B1 | |
| OE1428F | OE2408R | C | 1 | 0 | Y | 10.9 | cheD | Trusted (reciprocally confirmed) |
| OE1428F | OE2469F | E | 1 | 0 | N | 13.57 | flagellin A1 | |
| OE1428F | OE2470F | E | 1 | 0 | N | 16.95 | flagellin A2 | |
| PurH/N OE1620R | OE1560R | C | 0 | 1 | N | 17.17 | CHP | promiscuous prey |
| PurH/N OE1620R | OE1929R | TM | 1 | 0 | N | 8.68 | htr16 | |
| PurH/N OE1620R | OE2168R | TM | 1 | 0 | N | 17.18 | htr6 | |
| PurH/N OE1620R | OE2170R | MA(L) | 1 | 0 | N | 8.63 | probable periplasmic substrate-binding protein | Probably indirect via htr6 |
| PurH/N OE1620R | OE2189R | TM | 1 | 0 | N | 13.69 | htr4 | |
| PurH/N OE1620R | OE2415R | C/MA(P) | 1 | 0 | Y | 25.3 | cheA | Trusted (reciprocally confirmed) |
| PurH/N OE1620R | OE2458R | C | 0 | 1 | N | 7.98 | IMP dehydrogenase | |
| PurH/N OE1620R | OE3167F | TM | 1 | 0 | N | 12.66 | htr8 | |
| PurH/N OE1620R | OE3243F | C (MA(P)?) | 1 | 0 | N | 8.1 | cobyrinic acid a,c-diamide synthase | |
| PurH/N OE1620R | OE3474R | TM | 1 | 0 | N | 14.81 | htr5 (cosT) | |
| PurH/N OE1620R | OE3481R | TM | 1 | 0 | N | 10.56 | htr2 | |
| PurH/N OE1620R | OE3611R | TM | 1 | 0 | N | 15.87 | htr3 (basT) | |

| Bait | Prey | | Loc. | #d | #i | R | MAS | Prey Protein Name | Comment |
|---|---|---|---|---|---|---|---|---|---|
| PurH/N | OE1620R | OE3612R | TM, MA(L) | 1 | 0 | N | 18.69 | basB | Probybly indirect via BasT |
| PurH/N | OE1620R | OE3943R | C | 1 | 1 | N | 12.83 | CHP | promiscuous prey |
| PurH/N | OE1620R | OE4571R | C | 0 | 1 | N | 20.9 | probable leucyl aminopeptidase | |
| CheW2 | OE2374R | OE1783F | C | 1 | 0 | N | 11.8 | sufB domain protein | promiscuous prey |
| CheW2 | OE2374R | OE2168R | TM | 0 | 1 | N | 7.98 | htr6 | |
| CheW2 | OE2374R | OE2189R | TM | 0 | 1 | N | 9.29 | htr4 | |
| CheW2 | OE2374R | OE2190R | MA(L) | 1 | 0 | N | 9.94 | aldehyde dehydrogenase | |
| CheW2 | OE2374R | OE2392R | C | 0 | 1 | N | 16.22 | htr15 | Trusted (exchange problem in direct fishing) |
| CheW2 | OE2374R | OE2474R | C | 0 | 1 | N | 14.79 | htr13 | Trusted (exchange problem in direct fishing) |
| CheW2 | OE2374R | OE3167F | TM | 0 | 1 | N | 14.55 | htr8 | |
| CheW2 | OE2374R | OE3347F | TM | 0 | 1 | N | 8.52 | htr1 | |
| CheW2 | OE2374R | OE3474R | TM | 0 | 1 | N | 9.56 | htr5 (cosT) | |
| CheW2 | OE2374R | OE3481R | TM | 0 | 1 | N | 7.57 | htr2 | |
| CheW2 | OE2374R | OE3611R | TM | 0 | 1 | N | 8.4 | htr3 (basT) | |
| CheW2 | OE2374R | OE3940F | C | 1 | 0 | N | 14.64 | CHP | |
| CheW2 | OE2374R | OE3943R | C | 1 | 0 | N | 16.94 | CHP | promiscuous prey |
| CheW2 | OE2374R | OE4159F | C | 1 | 0 | N | 7.31 | adenosylhomocysteinase | |
| CheW2 | OE2374R | OE4260R | C | 1 | 0 | N | 22.8 | probable N-acetyltransferase | promiscuous prey |
| CheW2 | OE2374R | OE4329F | C | 1 | 0 | N | 7.36 | IMP cyclohydrolase, archaeal type | |
| CheW2 | OE2374R | OE5243F | C | 0 | 1 | N | 14.85 | htr11 (car) | Trusted (exchange problem in direct fishing) |
| ParA1 | OE2378R | OE2249R | C | 0 | 1 | N | 52 | CHP | Trusted (exchange problem in direct fishing) |
| ParA1 | OE2378R | OE2602R | C | 0 | 1 | N | 12.71 | ribosomal protein L1 | |
| | OE2401F | OE2402F | C | 1 | 0 | N | 12.37 | CHP | |
| | OE2401F | OE2408R | C | 1 | 0 | N | 37.69 | cheD | |
| | OE2401F | OE2470F | E | 0 | 1 | N | 10.77 | flagellin A2 | |
| | OE2401F | OE3542R | C | 0 | 1 | N | 17.64 | glutamine-rich alkaline protein | suspicious prey |
| | OE2401F | OE3943R | C | 0 | 1 | N | 11.04 | CHP | promiscuous prey |
| | OE2402F | OE1428F | C | 1 | 0 | N | 22.64 | CHP | |
| | OE2402F | OE2386R | MA | 1 | 1 | N | 20.04 | flaCE | |
| | OE2402F | OE2390R | MA | 0 | 1 | N | 24.97 | flaD | |
| | OE2402F | OE4260R | C | 1 | 0 | N | 16.24 | probable N-acetyltransferase | promiscuous prey |
| | OE2402F | OE5201F | C | 1 | 1 | N | 18.3 | aspartate carbamoyltransferase catalytic subunit | |
| | OE2404R | OE2386R | MA | 0 | 1 | N | 9.33 | flaCE | |
| | OE2404R | OE2390R | MA | 0 | 1 | N | 9.07 | flaD | |
| | OE2404R | OE2397F | E | 0 | 1 | N | 52 | flagellin B1 | |
| CheR | OE2406R | OE1428F | C | 1 | 0 | N | 13.6 | CHP | |
| CheR | OE2406R | OE2333R | C | 1 | 0 | N | 17.94 | probable signal-transducing histidine kinase | |
| CheR | OE2406R | OE2428R | C | 1 | 0 | N | 11.26 | CHP | |
| CheR | OE2406R | OE3070R | C | 1 | 0 | N | 17.88 | htr12 | |
| CheR | OE2406R | OE3139R | C | 1 | 0 | N | 8.92 | amidophosphoribosyltransferase | |
| CheR | OE2406R | OE3943R | C | 1 | 0 | N | 8.82 | CHP | promiscuous prey |
| CheD | OE2408R | OE1268F | C | 1 | 0 | N | 7.58 | probable transcription regulator boa1 | |
| CheD | OE2408R | OE1428F | C | 1 | 0 | Y | 18.13 | CHP | Trusted (reciprocally confirmed) |
| CheD | OE2408R | OE2402F | C | 1 | 0 | N | 10.58 | CHP | |
| CheD | OE2408R | OE2404R | C | 1 | 0 | N | 14.62 | CHP | |

Table S2: (continued)

| | Bait | Prey | Loc. | #d | #i | R | MAS | Prey Protein Name | Comment |
|---|---|---|---|---|---|---|---|---|---|
| CheD | OE2408R | OE3243F | C (MA(P)?) | 1 | 0 | N | 8.14 | cobyrinic acid a,c-diamide synthase | |
| CheC3 | OE2410R | OE1001F | TM | 1 | 0 | N | 50 | CHP | |
| CheC3 | OE2410R | OE1428F | C | 1 | 0 | N | 22.99 | CHP | |
| CheC3 | OE2410R | OE1560R | C | 0 | 1 | N | 7.27 | CHP | promiscuous prey |
| CheC3 | OE2410R | OE1783F | C | 0 | 1 | N | 8.17 | sufB domain protein | promiscuous prey |
| CheC3 | OE2410R | OE2408R | C | 1 | 0 | N | 16.46 | cheD | |
| CheC3 | OE2410R | OE3227F | C | 0 | 1 | N | 19.06 | homolog to nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase | promiscuous prey |
| CheC3 | OE2410R | OE3943R | C | 0 | 1 | N | 12.35 | CHP | promiscuous prey |
| CheC3 | OE2410R | OE4571R | C | 0 | 1 | N | 10.78 | probable leucyl aminopeptidase | |
| CheC3 | OE2410R | OE4712F | C | 1 | 0 | N | 31.39 | CHP | |
| CheC3 | OE2410R | OE5234R | TM | 0 | 1 | N | 15.62 | CHP (nonfunctional, N-terminal part) | |
| CheC1 | OE2414R | OE1679R | MA(L) | 1 | 0 | N | 7.92 | ABC-type transport system periplasmic substrate-binding protein (probable substrate phosphate) | |
| CheC1 | OE2414R | OE2042F | TM | 1 | 0 | N | 14.09 | probable copper-transporting ATPase | |
| CheC1 | OE2414R | OE2190R | MA(L) | 0 | 1 | N | 23.51 | aldehyde dehydrogenase | |
| CheA | OE2415R | OE1319R | C | 1 | 0 | N | 10.56 | cell division protein ftsZ | |
| CheA | OE2415R | OE1536R | TM | 1 | 0 | N | 9.69 | htr14 (mpcT) | |
| CheA | OE2415R | OE1539F | TM | 1 | 0 | N | 9 | CHP | Might be indirect via Htr14 |
| CheA | OE2415R | OE1559R | C | 1 | 0 | N | 10.33 | cell division protein ftsZ | |
| CheA | OE2415R | OE1620R | C/MA(P) | 2 | 0 | Y | 40.77 | phosphoribosylglycinamide formyltransferase / phosphoribosylaminoimidazolecarboxamide formyltransferase | Trusted (reciprocally confirmed) |
| CheA | OE2415R | OE1929R | TM | 1 | 0 | N | 10.51 | htr16 | |
| CheA | OE2415R | OE2121F | TM | 1 | 0 | N | 18.44 | CHP | |
| CheA | OE2415R | OE2168R | TM | 2 | 0 | N | 24.55 | htr6 | |
| CheA | OE2415R | OE2170R | MA(L) | 1 | 0 | N | 22.73 | probable periplasmic substrate-binding protein | Probably indirect via htr6 |
| CheA | OE2415R | OE2189R | TM | 2 | 0 | N | 12.03 | htr4 | |
| CheA | OE2415R | OE2196F | MA(L) | 1 | 0 | N | 9.99 | periplasmic substrate-binding protein | Probably indirect via Htr18 |
| CheA | OE2415R | OE2374R | C | 1 | 0 | N | 15.09 | cheW2 | |
| CheA | OE2415R | OE2419R | C/MA(P) | 0 | 1 | Y | 40.81 | cheW1 | Trusted (reciprocally confirmed); exchange problem |
| CheA | OE2415R | OE3038F | MA(L) | 1 | 0 | N | 7.65 | methylenetetrahydrofolate dehydrogenase / methenyltetrahydrofolate cyclohydrolase | |
| CheA | OE2415R | OE3167F | TM | 2 | 0 | N | 18 | htr8 | |
| CheA | OE2415R | OE3190F | TM | 1 | 0 | N | 8.04 | hypothetical protein | |
| CheA | OE2415R | OE3227F | C | 0 | 1 | N | 10.43 | homolog to nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase | promiscuous prey |
| CheA | OE2415R | OE3347F | TM | 1 | 1 | N | 16.07 | htr1 | |
| CheA | OE2415R | OE3436R | TM | 1 | 0 | N | 23.78 | htr17 | |
| CheA | OE2415R | OE3474R | TM | 2 | 0 | N | 21.38 | htr5 (cosT) | |
| CheA | OE2415R | OE3476R | MA(L) | 1 | 0 | N | 8.2 | cosB | Probably indirect via CosT |

| | Bait | Prey | Loc. | #d | #i | R | MAS | Prey Protein Name | Comment |
|---|---|---|---|---|---|---|---|---|---|
| CheA | OE2415R | OE3480R | TM | 2 | 0 | N | 29.78 | sensory rhodopsin II | Probably indirect via Htr2 |
| CheA | OE2415R | OE3481R | TM | 1 | 0 | N | 16.6 | htr1I | |
| CheA | OE2415R | OE3611R | TM | 1 | 0 | N | 14.92 | htr3 (basT) | |
| CheA | OE2415R | OE3759R | C | 1 | 0 | N | 19.92 | CHP | |
| CheA | OE2415R | OE4643R | C | 2 | 0 | Y | 19.64 | CHP | Trusted (reciprocally confirmed) |
| CheB | OE2416R | OE1782F | C | 1 | 0 | N | 7.42 | sufB domain protein | |
| CheB | OE2416R | OE1783F | C | 2 | 0 | N | 16.4 | sufB domain protein | promiscuous prey |
| CheB | OE2416R | OE2274R | C | 1 | 0 | N | 11.15 | phosphoribosylformyl-glycinamidine synthase component II | |
| CheB | OE2416R | OE2397F | E | 1 | 0 | N | 7.59 | flagellin B1 | |
| CheB | OE2416R | OE2408R | C | 1 | 0 | N | 14.91 | cheD | |
| CheB | OE2416R | OE2414R | C | 0 | 1 | N | 19.27 | cheC1 | |
| CheB | OE2416R | OE2469F | E | 1 | 0 | N | 7.01 | flagellin A1 | |
| CheB | OE2416R | OE2470F | E | 1 | 0 | N | 15.83 | flagellin A2 | |
| CheB | OE2416R | OE3139R | C | 1 | 0 | N | 20.83 | amidophosphoribosyl-transferase | |
| CheB | OE2416R | OE3227F | C | 1 | 1 | N | 20.32 | homolog to nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase | promiscuous prey |
| CheB | OE2416R | OE3943R | C | 2 | 1 | N | 38.36 | CHP | promiscuous prey |
| CheB | OE2416R | OE4260R | C | 1 | 0 | N | 22.21 | probable N-acetyltransferase | promiscuous prey |
| CheY | OE2417R | OE1268F | C | 0 | 1 | N | 16.47 | probable transcription regulator boa1 | |
| CheY | OE2417R | OE1428F | C | 1 | 0 | N | 9.84 | CHP | |
| CheY | OE2417R | OE1536R | TM | 0 | 1 | N | 33.7 | htr14 (mpcT) | |
| CheY | OE2417R | OE1620R | C | 0 | 1 | N | 14.78 | phosphoribosylglycinamide formyltransferase / phosphoribosylaminoimidazolecarboxamide formyltransferase | |
| CheY | OE2417R | OE2168R | TM | 0 | 1 | N | 48 | htrVI | |
| CheY | OE2417R | OE2170R | MA(L) | 0 | 1 | N | 19.8 | probable periplasmic substrate-binding protein | Probably indirect via htr6 |
| CheY | OE2417R | OE2189R | TM | 0 | 1 | N | 18.6 | htr4 | |
| CheY | OE2417R | OE2195F | TM | 0 | 1 | N | 13.84 | htr18 | |
| CheY | OE2417R | OE2196F | MA(L) | 0 | 1 | N | 11.19 | periplasmic substrate-binding protein | Probably indirect via Htr18 |
| CheY | OE2417R | OE2392R | C | 0 | 1 | N | 7.06 | htr15 | |
| CheY | OE2417R | OE2402F | C | 1 | 0 | N | 10.84 | CHP | |
| CheY | OE2417R | OE2404R | C | 0 | 1 | N | 16.26 | CHP | |
| CheY | OE2417R | OE2415R | C/MA(P) | 0 | 1 | N | 48 | cheA | |
| CheY | OE2417R | OE3101R | C | 0 | 1 | N | 23.64 | bat | |
| CheY | OE2417R | OE3167F | TM | 0 | 1 | N | 48 | htr8 | |
| CheY | OE2417R | OE3227F | C | 0 | 1 | N | 13.04 | homolog to nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase | promiscuous prey |
| CheY | OE2417R | OE3347F | TM | 0 | 1 | N | 48 | htr1 | |
| CheY | OE2417R | OE3436R | TM | 0 | 1 | N | 48 | htr17 | |
| CheY | OE2417R | OE3474R | TM | 0 | 1 | N | 48 | htr5 (cosT) | |
| CheY | OE2417R | OE3480R | TM | 0 | 1 | N | 14.12 | sensory rhodopsin II | Probably indirect via Htr2 |
| CheY | OE2417R | OE3481R | TM | 0 | 1 | N | 34.49 | htr1I | |
| CheY | OE2417R | OE3611R | TM | 0 | 1 | N | 43.18 | htr3 (basT) | |
| CheY | OE2417R | OE3612R | TM, MA(L) | 0 | 1 | N | 48 | basB | |
| CheY | OE2417R | OE3943R | C | 0 | 1 | N | 27.01 | CHP | promiscuous prey |
| CheY | OE2417R | OE4651F | C | 0 | 1 | N | 31.1 | probable ribose-1,5-bisphosphate isomerase | |

Table S2: (continued)

| Bait | Prey | | Loc. | #d | #i | R | MAS | Prey Protein Name | Comment |
|------|------|------|------|----|----|---|-----|-------------------|---------|
| CheY | OE2417R | OE5234R | TM | 0 | 1 | N | 16.81 | CHP (nonfunctional, N-terminal part) | |
| CheW1 | OE2419R | OE1536R | TM | 0 | 1 | N | 12.84 | htr14 (mpcT) | |
| CheW1 | OE2419R | OE1539F | TM | 1 | 0 | N | 7.99 | CHP | Might be indirect via Htr14 |
| CheW1 | OE2419R | OE1560R | C | 0 | 1 | N | 33.4 | CHP | promiscuous prey |
| CheW1 | OE2419R | OE1613R | C | 0 | 1 | N | 33.14 | probable acylaminoacyl-peptidase | |
| CheW1 | OE2419R | OE1620R | C | 0 | 1 | N | 21.65 | phosphoribosylglycinamide formyltransferase / phosphoribosylaminoim-idazolecarboxamide formyltransferase | |
| CheW1 | OE2419R | OE1781F | C | 0 | 1 | N | 23.2 | ATP-binding sufC-like protein | |
| CheW1 | OE2419R | OE2149R | C | 0 | 1 | N | 8.71 | CHP | |
| CheW1 | OE2419R | OE2168R | TM | 1 | 1 | N | 18.91 | htrVI | |
| CheW1 | OE2419R | OE2170R | MA(L) | 1 | 0 | N | 12.04 | probable periplasmic substrate-binding protein | Probably indirect via htr6 |
| CheW1 | OE2419R | OE2189R | TM | 1 | 1 | N | 28.99 | htr4 | |
| CheW1 | OE2419R | OE2373F | C | 0 | 1 | N | 12.49 | probable phosphate acetyltransferase | |
| CheW1 | OE2419R | OE2392R | C | 0 | 1 | N | 31.52 | htr15 | |
| CheW1 | OE2419R | OE2415R | C/MA(P) | 0 | 1 | Y | 38.87 | cheA | Trusted (reciprocally confirmed); exchange problem |
| CheW1 | OE2419R | OE2443R | TM | 0 | 1 | N | 12.28 | hypothetical protein | |
| CheW1 | OE2419R | OE2474R | C | 0 | 1 | N | 9.97 | htr13 | |
| CheW1 | OE2419R | OE3167F | TM | 1 | 1 | N | 18.42 | htr8 | |
| CheW1 | OE2419R | OE3347F | TM | 1 | 1 | N | 18.48 | htr1 | |
| CheW1 | OE2419R | OE3474R | TM | 1 | 1 | N | 18.56 | htr5 (cosT) | |
| CheW1 | OE2419R | OE3476R | MA(L) | 1 | 0 | N | 18.24 | cosB | Probably indirect via htr5 (cosT) |
| CheW1 | OE2419R | OE3480R | TM | 1 | 0 | N | 7.42 | sensory rhodopsin II | Probably indirect via HtrII |
| CheW1 | OE2419R | OE3481R | TM | 1 | 1 | N | 20.68 | htr1I | |
| CheW1 | OE2419R | OE3611R | TM | 1 | 1 | N | 21.48 | htr3 (basT) | |
| CheW1 | OE2419R | OE3612R | TM, MA(L) | 1 | 1 | N | 29.6 | basB | Probably indirect via basT |
| CheW1 | OE2419R | OE3943R | C | 1 | 1 | N | 21.77 | CHP | promiscuous prey |
| CheW1 | OE2419R | OE4146F | C | 0 | 1 | N | 23.89 | TATA-binding transcription initiation factor | |
| CheW1 | OE2419R | OE4159F | C | 0 | 1 | N | 29.31 | adenosylhomocysteinase | |
| CheW1 | OE2419R | OE4339R | MA(P) | 0 | 1 | N | 10.04 | ABC-type transport system ATP-binding protein (probable substrate copper) | |
| CheW1 | OE2419R | OE4673F | C | 0 | 1 | N | 10.21 | probable carboxypeptidase | |
| CheW1 | OE2419R | OE5201F | C | 0 | 1 | N | 24.27 | aspartate carbamoyltransferase catalytic subunit | |
| CheW1 | OE2419R | OE5234R | TM | 0 | 1 | N | 48 | CHP (nonfunctional, N-terminal part) | |
| CheC2 | OE3280R | OE1079F | C | 0 | 1 | N | 9.39 | CHP | |
| CheC2 | OE3280R | OE1319R | C | 0 | 1 | N | 7.09 | cell division protein ftsZ | |
| CheC2 | OE3280R | OE1414R | C | 0 | 1 | N | 25.58 | cell division protein ftsZ | |
| CheC2 | OE3280R | OE1417F | C | 0 | 1 | N | 11.37 | CHP | |
| CheC2 | OE3280R | OE1451F | MA(L) | 0 | 1 | N | 8.7 | homolog to S-adenosyl-methionine-dependent methyltransferase | |
| CheC2 | OE3280R | OE1782F | C | 0 | 1 | N | 8.17 | sufB domain protein | |
| CheC2 | OE3280R | OE1783F | C | 0 | 1 | N | 24.45 | sufB domain protein | promiscuous prey |

Table S2: (continued)

| Bait | Prey | | Loc. | #d | #i | R | MAS | Prey Protein Name | Comment |
|------|------|------|------|----|----|---|-----|-------------------|---------|
| CheC2 | OE3280R | OE2402F | C | 0 | 1 | N | 22.42 | CHP | |
| CheC2 | OE3280R | OE2404R | C | 0 | 1 | N | 25.87 | CHP | |
| CheC2 | OE3280R | OE2408R | C | 0 | 1 | N | 10.28 | cheD | |
| CheC2 | OE3280R | OE2414R | C | 1 | 0 | N | 18.99 | cheC1 | |
| CheC2 | OE3280R | OE3328R | MA(P) | 0 | 1 | N | 9.87 | ribonucleoside-diphosphate reductase alpha subunit | |
| CheC2 | OE3280R | OE3356F | C | 0 | 1 | N | 8.91 | AAA-type ATPase (transitional ATPase homolog) | |
| CheC2 | OE3280R | OE4466R | C | 0 | 1 | N | 9.88 | DNA repair protein | |
| CheC2 | OE3280R | OE4677F | C | 0 | 1 | N | 8.44 | DNA helicase II | |
| CheC2 | OE3280R | OE4729R | C | 0 | 1 | N | 21.31 | translation elongation factor aEF-2 | |
| | OE4643R | OE1150R | C | 1 | 0 | N | 7.37 | CHP | |
| | OE4643R | OE1268F | C | 1 | 0 | N | 16.99 | probable transcription regulator boa1 | |
| | OE4643R | OE1428F | C | 1 | 0 | N | 16.84 | CHP | |
| | OE4643R | OE1500R | C | 0 | 1 | N | 8.69 | pyruvate, water dikinase | |
| | OE4643R | OE1560R | C | 0 | 1 | N | 50 | CHP | promiscuous prey |
| | OE4643R | OE1909F | C | 0 | 1 | N | 12.49 | CHP | |
| | OE4643R | OE2415R | C | 1 | 1 | Y | 40.12 | cheA | Trusted (reciprocally confirmed) |
| | OE4643R | OE3175F | C | 0 | 1 | N | 23.95 | propionyl-CoA carboxylase carboxyltransferase component | |
| | OE4643R | OE3177F | C | 0 | 1 | N | 20.7 | propionyl-CoA carboxylase biotin carboxylase component | |
| | OE4643R | OE4260R | C | 1 | 0 | N | 8.61 | probable N-acetyltransferase | promiscuous prey |
| | OE4643R | OE4346R | MA(L) | 0 | 1 | N | 17.24 | ribonucleoside-diphosphate reductase alpha subunit | |

Preys were marked as promiscuous if they were found with at least three not directly interacting baits. The protein localisation (Loc.) is classified as follows: C cytoplasmic, TM transmembrane, MA(L) membrane associated with lipid anchor (as predicted by HaloLex), MA(P) membrane associated via PPI (when the protein is thought to be part of a complex with at least one TM or MA(L) protein; assigned after manual inspection of the protein's genomic context and annotation in HaloLex). The columns #d and #i give the number of direct and indirect experiments, where this interaction was found. R shows if the interaction was reciprocally confirmed. MAS is the highest association score with which the interaction was detected.
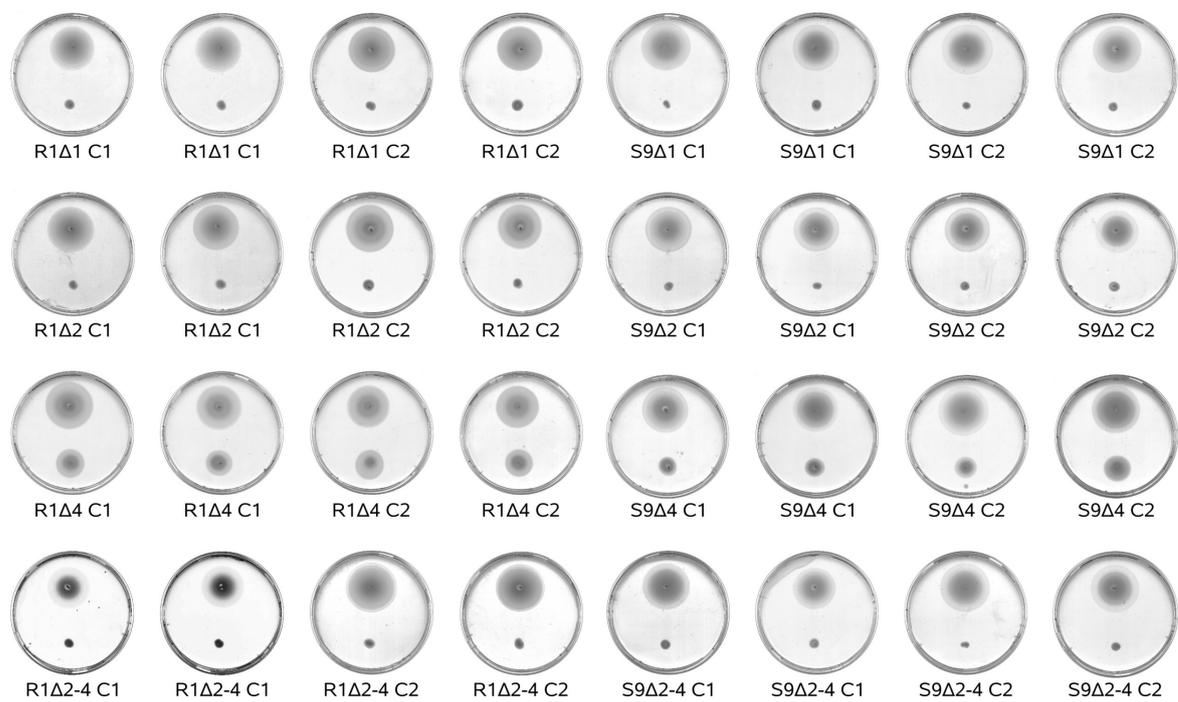
Figure S1: **Swarming ability of the deletion strains.** On each plate the deletion strain (bottom) is compared to the respective wildtype strain (top). For each deletion in both host strains, two clones were tested (C1 and C2). Each clone was examined on two plates.

Table S3: **Reversal frequencies as measured by computer-based cell-tracking.**

| Clone | | # Cells | # Reversals | % Reversals | $P_l$ (%) | $P_u$ (%) |
|---|---|---|---|---|---|---|
| | O | 256 | 183 | 71.48 | 65.46 | 76.85 |
| S9 | B | 162 | 111 | 68.52 | 60.69 | 75.45 |
| | S | 216 | 51 | 23.61 | 18.23 | 29.95 |
| | O | 127 | 100 | 78.74 | 70.41 | 85.29 |
| R1 | B | 100 | 81 | 81.00 | 71.67 | 87.89 |
| | S | 74 | 7 | 9.46 | 4.21 | 19.09 |
| | O | 166 | 11 | 6.63 | 3.52 | 11.84 |
| S9Δ1 C1 | B | 153 | 4 | 2.61 | 0.84 | 6.98 |
| | S | 201 | 3 | 1.49 | 0.39 | 4.65 |
| | O | 39 | 6 | 15.38 | 6.41 | 31.21 |
| S9Δ1 C2 | B | 77 | 11 | 14.29 | 7.69 | 24.55 |
| | S | 201 | 20 | 9.95 | 6.33 | 15.16 |
| | O | 58 | 0 | 0.00 | 0.16 | 7.74 |
| R1Δ1 C1 | B | 58 | 9 | 15.52 | 7.77 | 27.93 |
| | S | 157 | 23 | 14.65 | 9.70 | 21.38 |
| | O | 118 | 22 | 18.64 | 12.30 | 27.09 |
| R1Δ1 C2 | B | 36 | 5 | 13.89 | 5.23 | 30.29 |
| | S | 58 | 6 | 10.34 | 4.28 | 21.84 |
| | O | 252 | 6 | 2.38 | 0.97 | 5.36 |
| S9Δ2 C1 | B | 203 | 0 | 0.00 | 0.04 | 2.32 |
| | S | 125 | 1 | 0.80 | 0.04 | 5.03 |
| | O | 141 | 1 | 0.71 | 0.04 | 4.48 |
| S9Δ2 C2 | B | 211 | 3 | 1.42 | 0.37 | 4.44 |
| | S | 408 | 5 | 1.23 | 0.45 | 3.00 |
| | O | 97 | 2 | 2.06 | 0.36 | 7.97 |
| R1Δ2 C1 | B | 95 | 2 | 2.11 | 0.37 | 8.13 |
| | S | 247 | 7 | 2.83 | 1.25 | 6.00 |
| | O | 148 | 2 | 1.35 | 0.23 | 5.30 |
| R1Δ2 C2 | B | 216 | 2 | 0.93 | 0.16 | 3.66 |
| | S | 235 | 5 | 2.13 | 0.79 | 5.17 |
| | O | 179 | 126 | 70.39 | 63.04 | 76.85 |
| S9Δ4 C1 | B | 185 | 130 | 70.27 | 63.04 | 76.64 |
| | S | 280 | 58 | 20.71 | 16.22 | 26.03 |
| | O | 220 | 126 | 57.27 | 50.44 | 63.85 |
| S9Δ4 C2 | B | 142 | 97 | 68.31 | 59.90 | 75.71 |
| | S | 191 | 54 | 28.27 | 22.12 | 35.31 |
| | O | 35 | 23 | 65.71 | 47.74 | 80.32 |
| R1Δ4 C1 | B | 21 | 16 | 76.19 | 52.45 | 90.88 |
| | S | 58 | 17 | 29.31 | 18.46 | 42.91 |
| | O | 225 | 2 | 0.89 | 0.15 | 3.52 |
| S9Δ2-4 C1 | B | 199 | 9 | 4.52 | 2.22 | 8.69 |
| | S | 331 | 14 | 4.23 | 2.42 | 7.16 |
| | O | 236 | 3 | 1.27 | 0.33 | 3.98 |
| S9Δ2-4 C2 | B | 229 | 2 | 0.87 | 0.15 | 3.46 |
| | S | 393 | 12 | 3.05 | 1.66 | 5.42 |
| | O | 110 | 5 | 4.55 | 1.69 | 10.80 |
| R1Δ2-4 C1 | B | 92 | 1 | 1.09 | 0.06 | 6.76 |
| | S | 109 | 3 | 2.75 | 0.71 | 8.43 |

$P_l$ and $P_u$ are the limits of the 95 % confidence interval. The lines O, B, and S show the results of measurements with orange light step-down, with a blue light pulse, or without stimulus (spontaneous). C1 and C2 refer to different clones of the deletion strain.

Table S4: *che* and *fla* genes in archaeal genomes.

| | cheA | cheB | cheC | cheD | cheR | cheW | cheY | 439 | flaC | flaD/E | flaF | flaG | flaH | flaI | flaJ | flg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ape | | | | | | | | | | | 1901 | 1898 | 1896.1 | 1895.1 | | 1905, 1907 |
| Afu | 1040 | 1041 | 1039 | 1038 | 1037 | 1044 | 1042 | 1043 | | $1053^M$ | 1051 | 1052 | 1050 | 1049 | 1048 | 1054, 1055 |
| Mbo | 1336 | 1337, 0327*** | 1335 | 1334 | 1581, 1252, 0327*** | 1247, 1579, 1098 | 1338 | 1339 | | $1346^M$ | 1344 | 1345 | 1343 | 1342 | 1341 | 1347, 1348 |
| Hma | 2205 | 2204 | 2193, 0528, 1258, 2623 | 2192 | 2206 | 2203, 1484 | 2194 | 2209, 2213, 3221, 3231 | $2191^F$ | $2191^F$, 1482 | 2190 | 2187 | 2186 | 2184 | 2183 | 2198, pNG1026, rrnB0018* |
| Hsa | 2415R | 2416R | 2410R, 2414R, 3280R | 2408R | 2406R | 2374R, 2419R | 2417R | 2402F, 2404R | $2386R^F$ | 2390R, $2386R^F$ | 2385R | 2383R, 4607R | 2381R | 2380R | 2379R | 2397F, 2398F, 2399F, 2469F, 2470F, 2695F |
| Mse | | | | | | | | | | | 1327 | 1328 | 1326 | 1325 | 1324 | 1330 |
| Mbu | 0361 | 0360, 0399*** | 0362 | 0363 | 0364, 0399*** | 0357 | 0359 | 0358 | $1570^F$ | $0348^M$, $1570^F$, 1246 | 0350, 1571 | 0349, 1572 | 0351, 1573 | 0352, 1574 | 0353, 1575 | 0346, 0347, 0104 |
| Mja | | | | | | | | 1615* | 0894 | 0895, 0896 | 0897 | 0898 | 0899 | 0900 | 0901 | 0891, 0892, 0893 |
| MC5 | 0734 | 0733 | 0738, 0739 | 0735 | 0737 | 0732 | 0740 | 0741 | 1738 | 1737, 1736 | 1735, 1332 | 1734 | 1733 | 1732 | 1731 | 1739, 1740, 1741, 1742 |
| MC7 | 0174 | 0173 | 0178, 0179 | 0175 | 0177 | 0172 | 0180 | 0181 | 0942 | 0943, 0944 | 0945, 1343 | 0946 | 0947 | 0948 | 0949 | 0938, 0939, 0940, 0941 |
| MS2 | 0927 | 0926 | 0931, 0932 | 0928 | 0930 | 0925 | 0933 | 0934 | 1669 | 1670, 1671 | 1672, 0342 | 1673 | 1674 | 1675 | 1676 | 1666, 1667, 1668 |
| Mma | 0943, 0238 | 0944, 2125** | 0942 | 0941 | 1545 | 1544, 0240 | 0945 | 0946 | | $0953^M$ | 0951 | 0952 | 0950 | 0949 | 0948 | 0962, 0963, 1375, 1374 |

Table S4: (continued)

| | cheA | cheB | cheC | cheD | cheR | cheW | cheY | 439 | flaC | flaD/E | flaF | flaG | flaH | flaI | flaJ | flg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mac | 3066, 0014 | 3067, 0015, 1989***, 3542*** | 3065, 0012 | 3064, 0011 | 3063, 0013, 1989***, 3542*** | 3070, 0020 | 3068, 0016 | 3069 | | 3060[M], 3078[M] | 3058, 3080 | 3059, 3079 | 3057, 3081 | 3056, 3082 | 3055, 3083 | 3061, 3062, 3077 |
| Mba | 0984 | 0985, 2183*** | | 0982 | 0983, 2183*** | 0990 | 0986, 3321 | | | 1969[M] | 1967 | 1968 | 1966 | 1965 | 1964 | 1970 |
| Mmz | 0328, 1325 | 0329, 1326 | 0327, 1323 | 0326, 1322 | 0325, 1324 | 0332, 1330 | 0330, 1327 | 0331 | | 0321[M], 0417[M] | 0319, 0415 | 0320, 0416 | 0318, 0414 | 0317, 0413 | 0316, 0412 | 0322, 0323, 0418 |
| Mhu | 0110[F], 0494, 0989 | 0109, 0988, 0952**, 0887** | 0112, 0110[F], 2685, 1151, 2682 | 0111 | 0961, 0992, 0124 | 0007, 2533, 0991, 2550, 1642, 1423, 0496, 0960, 1925, 0898, 0003, 2399, 0993*, 2532* | 0108, 0126, 3040, 1439, 3041, 0315 | 0107 | | 0100[M] | 0102 | 0101 | 0103 | 0104 | 0105 | 3140, 3139, 1238 |
| Mva | 0220 | 0219 | 0257, 0258 | 0221 | 0259 | 0218, 0138 | 0256 | 0255 | 0969 | 0970, 0971 | 0972, 1352 | 0973 | 0974 | 0975 | 0976 | 0966, 0967, 0968 |
| RC-I | 571 | 570 | 572, X2603 | 573 | 584 | X655, X2603 | 569 | 568 | | 512[M], 510[M] | 501, 500 | 509, 507, 506, 505, 503 | 499 | 498 | 497 | 515, 514 |
| Mae | | | | | | | | | 0261 | 0262, 0263 | 0264 | 0265 | 0266 | 0267 | 0268 | 0256, 0257, 0258, 0259, 0260 |
| Nph | 2172A | 2174A | 2104A, 3118A | 2106A | 2170A | 4146A | 2102A | 2162A, 2166A | 2154A[F], 2686A | 2154A[F] | 2094A | 2096A, 2098A | 2156A | 2158A | 2160A | 2086A, 2088A, 2090A |
| Pab | 1332 | 1331 | 1334, 1333 | 1335 | 1329 | 1027 | 1330 | 1338 | 1381 | 1382 | 1383 | 1384 | 1385 | 1386 | 1387 | 1380, 1379, 1378*, |

Table S4: (continued)

| | cheA | cheB | cheC | cheD | cheR | cheW | cheY | 439 | flaC | flaD/E | flaF | flaG | flaH | flaI | flaJ | flg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pfu | | | | | | | | | 0336 | 0335 | 0334 | 0333 | 0332 | 0331 | 0330 | 0337, 0338 |
| Pho | 0484 | 0483 | 0487, 0488 | 0490 | 0481 | 0478 | 0482 | 0494 | 0552 | 0553 | 0553.1n | 0555 | 0556 | 0557 | 0559 | 0546, 0548, 0549, 0550, 0551 |
| Sac | | | | | | | | | | | 1175 | 1176 | 1174 | 1173 | 1172 | 1178 |
| Sso | | | | | | | | | | | 2319 | *is* | 2318 | 2316 | 2315 | 2323 |
| Sto | | | | | | | | | | | 2521 | 2520 | 2522 | 2523 | 2524 | 2518 |
| Tko | 0634, 0635 | 0633 | 0636, 0637 | 0639 | 0631 | 0629 | 0632 | 0641 | 0043 | 0044 | 0045 | 0046 | 0047 | 0048 | 0049 | 0038, 0039, 0040, 0041, 0042 |
| Tac | | | | | | | | | 0554 | 0555 | 0556 | 0557a | 0558 | 0559 | 0560 | 0553, 1407m |
| Tvo | | | | | | | | | 0608 | 0609 | 0610 | 0611 | 0612 | 0613 | 0614 | 0607, 1426 |

The column 439 lists members of the protein family DUF439. The prefix of the gene identifiers was omitted if the rest is unambiguous (e.g. 2415F instead of OE2415F). [F]: Fusion protein, belongs to two groups; [M]: Different version of FlaD/E protein found in Methanomicrobia and Archaeoglobi; *: singleton, not included into ortholog cluster (see text); ** protein with CheB domain, but no response regulator domain; *** protein containing both a CheB and a CheR domain. *is*: gene present, but interrupted by an insertion element. The species are: Ape *Aeropyrum pernix* K1, Afu *Archaeoglobus fulgidus* DSM4304, Mbo *Candidatus Methanoregula boonei* 6A8, Hma *Haloarcula marismortui* ATCC43049, Hsa *Halobacterium salinarum* R1, Mse *Metallosphaera sedula* DSM5348, Mbu *Methanococcoides burtonii* DSM6242, Mja *Methanococcus jannaschii* DSM2661, MC5 *Methanococcus maripaludis* C5, MC7 *Methanococcus maripaludis* C7, MS2 *Methanococcus maripaludis* S2, Mma *Methanoculleus marisnigri* JR1, Mac *Methanosarcina acetivorans* C2A, Mba *Methanosarcina barkeri* fusaro, Mmz *Methanosarcina mazei* Goe1, Mhu *Methanospirillum hungatei* JF-1, Mva *Methanococcus vannielii* SB, RC-I uncultured methanogenic archaeon RC-I, Mae *Methanococcus aeolicus* Nankai-3, Nph *Natronomonas pharaonis* DSM2160, Pab *Pyrococcus abyssi* GE5, Pfu *Pyrococcus furiosus* DSM3638, Pho *Pyrococcus horikoshii* OT3, Sac *Sulfolobus acidocaldarius* DSM639, Sso *Sulfolobus solfataricus* P2, Sto *Sulfolobus tokodaii* 7, Tko *Thermococcus kodakaraensis* KOD1, Tac *Thermoplasma acidophilum* DSM1728, Tvo *Thermoplasma volcanium* GSS1. Also included in the analysis, but not listed in the table since no Che and Fla orthologs were detected, were: *Haloquadratum walsbyi* DSM16790, *Hyperthermus butylicus* DSM5456, *Ignicoccus hospitalis* KIN4 I, *Methanobacterium thermoautotrophicum* delta H, *Methanobrevibacter smithii* ATCC35061, *Methanocorpusculum labreanum* Z, *Methanopyrus kandleri* AV19, *Methanosaeta thermophila* PT, *Methanosphaera stadtmanae* DSM3091, *Nanoarchaeum equitans* Kin4-M, *Picrophilus torridus* DSM9790, *Pyrobaculum aerophilum* IM2, *Pyrobaculum arsenaticum* DSM13514, *Pyrobaculum calidifontis* JCM11548, *Pyrobaculum islandicum* DSM4184, *Staphylothermus marinus* F1, *Thermofilum pendens* Hrk 5.

# Bibliography

Alam, M. and Oesterhelt, D. (1984) Morphology, function and isolation of halobacterial flagella. *J Mol Biol* **176**: 459–475, URL http://www.ncbi.nlm.nih.gov/pubmed/6748081.

Albert, R., Jeong, H., and Barabasi, A.-L. (2000) Error and attack tolerance of complex networks. *Nature* **406**: 378–382, URL http://dx.doi.org/10.1038/35019019.

Alfarano, C., Andrade, C. E., Anthony, K., Bahroos, N., Bajec, M., Bantoft, K., Betel, D., Bobechko, B., Boutilier, K., Burgess, E., Buzadzija, K., Cavero, R., D'Abreo, C., Donaldson, I., Dorairajoo, D., Dumontier, M. J., Dumontier, M. R., Earles, V., Farrall, R., Feldman, H., Garderman, E., Gong, Y., Gonzaga, R., Grytsan, V., Gryz, E., Gu, V., Haldorsen, E., Halupa, A., Haw, R., Hrvojic, A., Hurrell, L., Isserlin, R., Jack, F., Juma, F., Khan, A., Kon, T., Konopinsky, S., Le, V., Lee, E., Ling, S., Magidin, M., Moniakis, J., Montojo, J., Moore, S., Muskat, B., Ng, I., Paraiso, J. P., Parker, B., Pintilie, G., Pirone, R., Salama, J. J., Sgro, S., Shan, T., Shu, Y., Siew, J., Skinner, D., Snyder, K., Stasiuk, R., Strumpf, D., Tuekam, B., Tao, S., Wang, Z., White, M., Willis, R., Wolting, C., Wong, S., Wrong, A., Xin, C., Yao, R., Yates, B., Zhang, S., Zheng, K., Pawson, T., Ouellette, B. F. F., and Hogue, C. W. V. (2005) The Biomolecular Interaction Network Database and related tools 2005 update. *Nucleic Acids Res* **33**: D418–D424, URL http://dx.doi.org/10.1093/nar/gki051.

Allers, T. and Mevarech, M. (2005) Archaeal genetics - the third way. *Nat Rev Genet* **6**: 58–73, URL http://dx.doi.org/10.1038/nrg1504.

Ames, P. and Parkinson, J. S. (1994) Constitutively signaling fragments of Tsr, the *Escherichia coli* serine chemoreceptor. *J Bacteriol* **176**: 6340–6348, URL http://www.ncbi.nlm.nih.gov/pubmed/7929006.

Aregger, M. (2003) Charakterisierung von cheW-Deletionsmutanten in *Halobacterium salinarum*. Ph.D. thesis, Ludwigs-Maximilians-Universität München.

Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.-C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., Kitagawa, M., Tomita, M., Kanaya, S., Wada, C., and Mori, H. (2006) Large-scale identification of protein-protein interaction of *Escherichia coli* K-12. *Genome Res* **16**: 686–691, URL http://dx.doi.org/10.1101/gr.4527806.

Armitage, J. P. (1999) Bacterial tactic responses. *Adv Microb Physiol* **41**: 229–289, URL http://www.ncbi.nlm.nih.gov/pubmed/10500847.

Bader, J. S., Chaudhuri, A., Rothberg, J. M., and Chant, J. (2004) Gaining confidence in high-throughput protein interaction networks. *Nat Biotechnol* **22**: 78–85, URL http://dx.doi.org/10.1038/nbt924.

Barabasi, A.-L. and Albert, R. (1999) Emergence of scaling in random networks. *Science* **286**: 509–512, URL http://www.ncbi.nlm.nih.gov/pubmed/10521342.

# Bibliography

Barak, R. and Eisenbach, M. (1992a) Correlation between phosphorylation of the chemotaxis protein CheY and its activity at the flagellar motor. *Biochemistry* **31**: 1821–1826, URL http://www.ncbi.nlm.nih.gov/pubmed/1737035.

Barak, R. and Eisenbach, M. (1992b) Fumarate or a fumarate metabolite restores switching ability to rotating flagella of bacterial envelopes. *J Bacteriol* **174**: 643–645, URL http://www.ncbi.nlm.nih.gov/pubmed/1729255.

Bardy, S. L., Ng, S. Y. M., and Jarrell, K. F. (2003) Prokaryotic motility structures. *Microbiology* **149**: 295–304, URL http://www.ncbi.nlm.nih.gov/pubmed/12624192.

Bardy, S. L., Ng, S. Y. M., and Jarrell, K. F. (2004) Recent advances in the structure and assembly of the archaeal flagellum. *J Mol Microbiol Biotechnol* **7**: 41–51, URL http://dx.doi.org/10.1159/000077868.

Barns, S. M., Delwiche, C. F., Palmer, J. D., and Pace, N. R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci U S A* **93**: 9188–9193, URL http://www.ncbi.nlm.nih.gov/pubmed/8799176.

Bartel, P., Chien, C. T., Sternglanz, R., and Fields, S. (1993) Elimination of false positives that arise in using the two-hybrid system. *Biotechniques* **14**: 920–924, URL http://www.ncbi.nlm.nih.gov/pubmed/8333960.

Berg, H. C. (2003) The rotary motor of bacterial flagella. *Annu Rev Biochem* **72**: 19–54, URL http://dx.doi.org/10.1146/annurev.biochem.72.121801.161737.

Berg, H. C. and Purcell, E. M. (1977) Physics of chemoreception. *Biophys J* **20**: 193–219, URL http://www.ncbi.nlm.nih.gov/pubmed/911982.

Besir, H. (2001) Untersuchung der lipidvermittelten Kristallisation der Ionenpumpen Bakteriorhodopsin und Halorhodopsin aus *Halobacterium salinarum*. Ph.D. thesis, Ludwig-Maximilians-Universität München, URL http://edoc.ub.uni-muenchen.de/196/.

Bieger, B., Essen, L.-O., and Oesterhelt, D. (2003) Crystal structure of halophilic dodecin: a novel, dodecameric flavin binding protein from *Halobacterium salinarum*. *Structure* **11**: 375–385, URL http://www.ncbi.nlm.nih.gov/pubmed/12679016.

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, URL http://www.ncbi.nlm.nih.gov/pubmed/9989504.

Bischoff, D. S., Bourret, R. B., Kirsch, M. L., and Ordal, G. W. (1993) Purification and characterization of *Bacillus subtilis* CheY. *Biochemistry* **32**: 9256–9261, URL http://www.ncbi.nlm.nih.gov/pubmed/8369293.

Bischoff, D. S. and Ordal, G. W. (1992) Identification and characterization of FliY, a novel component of the *Bacillus subtilis* flagellar switch complex. *Mol Microbiol* **6**: 2715–2723, URL http://www.ncbi.nlm.nih.gov/pubmed/1447979.

Blat, Y. and Eisenbach, M. (1996) Oligomerization of the phosphatase CheZ upon interaction with the phosphorylated form of CheY. The signal protein of bacterial chemotaxis. *J Biol Chem* **271**: 1226–1231, URL http://www.ncbi.nlm.nih.gov/pubmed/8557654.

162

Blum, H., Beier, H., and Gross, H. J. (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8**: 93–99, URL http://dx.doi.org/10.1002/elps.1150080203.

Bonneau, R., Facciotti, M. T., Reiss, D. J., Schmid, A. K., Pan, M., Kaur, A., Thorsson, V., Shannon, P., Johnson, M. H., Bare, J. C., Longabaugh, W., Vuthoori, M., Whitehead, K., Madar, A., Suzuki, L., Mori, T., Chang, D.-E., Diruggiero, J., Johnson, C. H., Hood, L., and Baliga, N. S. (2007) A predictive model for transcriptional control of physiology in a free living cell. *Cell* **131**: 1354–1365, URL http://dx.doi.org/10.1016/j.cell.2007.10.053.

Borkovich, K. A., Alex, L. A., and Simon, M. I. (1992) Attenuation of sensory receptor signaling by covalent modification. *Proc Natl Acad Sci U S A* **89**: 6756–6760, URL http://www.ncbi.nlm.nih.gov/pubmed/1495964.

Borkovich, K. A., Kaplan, N., Hess, J. F., and Simon, M. I. (1989) Transmembrane signal transduction in bacterial chemotaxis involves ligand-dependent activation of phosphate group transfer. *Proc Natl Acad Sci U S A* **86**: 1208–1212, URL http://www.ncbi.nlm.nih.gov/pubmed/2645576.

Boukhvalova, M. S., Dahlquist, F. W., and Stewart, R. C. (2002) CheW binding interactions with CheA and Tar. Importance for chemotaxis signaling in *Escherichia coli. J Biol Chem* **277**: 22251–22259, URL http://dx.doi.org/10.1074/jbc.M110908200.

Boulton, S. J., Vincent, S., and Vidal, M. (2001) Use of protein-interaction maps to formulate biological questions. *Curr Opin Chem Biol* **5**: 57–62, URL http://www.ncbi.nlm.nih.gov/pubmed/11166649.

Bren, A. and Eisenbach, M. (1998) The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. *J Mol Biol* **278**: 507–514, URL http://dx.doi.org/10.1006/jmbi.1998.1730.

Brooun, A., Bell, J., Freitas, T., Larsen, R. W., and Alam, M. (1998) An archaeal aerotaxis transducer combines subunit I core structures of eukaryotic cytochrome c oxidase and eubacterial methyl-accepting chemotaxis proteins. *J Bacteriol* **180**: 1642–1646, URL http://www.ncbi.nlm.nih.gov/pubmed/9537358.

Butland, G., Peregrín-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., and Emili, A. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli. Nature* **433**: 531–537, URL http://dx.doi.org/10.1038/nature03239.

Castagnoli, L., Costantini, A., Dall'Armi, C., Gonfloni, S., Montecchi-Palazzi, L., Panni, S., Paoluzi, S., Santonico, E., and Cesareni, G. (2004) Selectivity and promiscuity in the interaction network mediated by protein recognition modules. *FEBS Lett* **567**: 74–79, URL http://dx.doi.org/10.1016/j.febslet.2004.03.116.

Causier, B. (2004) Studying the interactome with the yeast two-hybrid system and mass spectrometry. *Mass Spectrom Rev* **23**: 350–367, URL http://dx.doi.org/10.1002/mas.10080.

Chaban, B., Ng, S. Y. M., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S.-I., and Jarrell, K. F. (2007) Systematic deletion analyses of the *fla* genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, *Methanococcus maripaludis. Mol Microbiol* **66**: 596–609, URL http://dx.doi.org/10.1111/j.1365-2958.2007.05913.x.

Chao, X., Muff, T. J., Park, S.-Y., Zhang, S., Pollard, A. M., Ordal, G. W., Bilwes, A. M., and Crane, B. R. (2006) A receptor-modifying deamidase in complex with a signaling phosphatase reveals reciprocal regulation. *Cell* **124**: 561–571, URL http://dx.doi.org/10.1016/j.cell.2005.11.046.

Chiu, W., Baker, M. L., and Almo, S. C. (2006) Structural biology of cellular machines. *Trends Cell Biol* **16**: 144–150, URL http://dx.doi.org/10.1016/j.tcb.2006.01.002.

Cline, S. W., Lam, W. L., Charlebois, R. L., Schalkwyk, L. C., and Doolittle, W. F. (1989) Transformation methods for halophilic archaebacteria. *Can J Microbiol* **35**: 148–152, URL http://www.ncbi.nlm.nih.gov/pubmed/2497937.

Cohen-Ben-Lulu, G. N., Francis, N. R., Shimoni, E., Noy, D., Davidov, Y., Prasad, K., Sagi, Y., Cecchini, G., Johnstone, R. M., and Eisenbach, M. (2008) The bacterial flagellar switch complex is getting more complex. *EMBO J* **27**: 1134–1144, URL http://dx.doi.org/10.1038/emboj.2008.48.

Cohen-Krausz, S. and Trachtenberg, S. (2002) The structure of the archeabacterial flagellar filament of the extreme halophile *Halobacterium salinarum* R1M1 and its relation to eubacterial flagellar filaments and type IV pili. *J Mol Biol* **321**: 383–395, URL http://www.ncbi.nlm.nih.gov/pubmed/12162953.

Cohen-Krausz, S. and Trachtenberg, S. (2008) The flagellar filament structure of the extreme acidothermophile *Sulfolobus shibatae* B12 suggests that archaeabacterial flagella have a unique and common symmetry and design. *J Mol Biol* **375**: 1113–1124, URL http://dx.doi.org/10.1016/j.jmb.2007.10.048.

Collura, V. and Boissy, G. (2007) From protein-protein complexes to interactomics. *Subcell Biochem* **43**: 135–183, URL http://www.ncbi.nlm.nih.gov/pubmed/17953394.

da Costa, M. S., Santos, H., and Galinski, E. A. (1998) An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv Biochem Eng Biotechnol* **61**: 117–153, URL http://www.ncbi.nlm.nih.gov/pubmed/9670799.

Dandekar, T., Snel, B., Huynen, M., and Bork, P. (1998) Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci* **23**: 324–328, URL http://www.ncbi.nlm.nih.gov/pubmed/9787636.

Danner, S. and Soppa, J. (1996) Characterization of the distal promoter element of halobacteria in vivo using saturation mutagenesis and selection. *Mol Microbiol* **19**: 1265–1276, URL http://www.ncbi.nlm.nih.gov/pubmed/8730868.

del Rosario, R. C. H., Staudinger, W. F., Streif, S., Pfeiffer, F., Mendoza, E., and Oesterhelt, D. (2007) Modelling the CheY(D10K,Yl00W) *Halobacterium salinarum* mutant: sensitivity analysis allows choice of parameter to be modified in the phototaxis model. *IET Syst Biol* **1**: 207–221, URL http://www.ncbi.nlm.nih.gov/pubmed/17708428.

DeLong, E. F. and Pace, N. R. (2001) Environmental diversity of bacteria and archaea. *Syst Biol* **50**: 470–478, URL http://www.ncbi.nlm.nih.gov/pubmed/12116647.

DeLong, E. F., Taylor, L. T., Marsh, T. L., and Preston, C. M. (1999) Visualization and enumeration of marine planktonic archaea and bacteria by using polyribonucleotide probes and fluorescent in situ hybridization. *Appl Environ Microbiol* **65**: 5554–5563, URL http://www.ncbi.nlm.nih.gov/pubmed/10584017.

Dennis, P. P. and Shimmin, L. C. (1997) Evolutionary divergence and salinity-mediated selection in halophilic archaea. *Microbiol Mol Biol Rev* **61**: 90–104, URL http://www.ncbi.nlm.nih.gov/pubmed/9106366.

Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., Henne, A., Wiezer, A., Bäumer, S., Jacobi, C., Brüggemann, H., Lienard, T., Christmann, A., Bömeke, M., Steckel, S., Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H.-P., Gunsalus, R. P., Fritz, H.-J., and Gottschalk, G. (2002) The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J Mol Microbiol Biotechnol* **4**: 453–461, URL http://www.ncbi.nlm.nih.gov/pubmed/12125824.

Desmond, E., Brochier-Armanet, C., and Gribaldo, S. (2007) Phylogenomics of the archaeal flagellum: rare horizontal gene transfer in a unique motility structure. *BMC Evol Biol* **7**: 106, URL http://dx.doi.org/10.1186/1471-2148-7-106.

Djordjevic, S. and Stock, A. M. (1998) Chemotaxis receptor recognition by protein methyltransferase CheR. *Nat Struct Biol* **5**: 446–450, URL http://www.ncbi.nlm.nih.gov/pubmed/9628482.

Dobreva, I., Fielding, A., Foster, L., and Dedhar, S. (2008) Mapping the Integrin-Linked Kinase Interactome Using SILAC. *J Proteome Res* **7**: 1740–1749, URL http://dx.doi.org/10.1021/pr700852r.

Eisenbach, M. (1990) Functions of the flagellar modes of rotation in bacterial motility and chemotaxis. *Mol Microbiol* **4**: 161–167, URL http://www.ncbi.nlm.nih.gov/pubmed/2187141.

Eisenberg, H., Mevarech, M., and Zaccai, G. (1992) Biochemical, structural, and molecular genetic aspects of halophilism. *Adv Protein Chem* **43**: 1–62, URL http://www.ncbi.nlm.nih.gov/pubmed/1442321.

Falke, J. J. and Hazelbauer, G. L. (2001) Transmembrane signaling in bacterial chemoreceptors. *Trends Biochem Sci* **26**: 257–265, URL http://www.ncbi.nlm.nih.gov/pubmed/11295559.

Felsenstein, J. (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author, URL http://evolution.genetics.washington.edu/phylip.html.

Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**: 245–246, URL http://dx.doi.org/10.1038/340245a0.

Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. L., and Bateman, A. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res* **34**: D247–D251, URL http://dx.doi.org/10.1093/nar/gkj149.

Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H.-R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. L., and Bateman, A. (2008) The Pfam protein families database. *Nucleic Acids Res* **36**: D281–D288, URL http://dx.doi.org/10.1093/nar/gkm960.

Formstecher, E., Aresta, S., Collura, V., Hamburger, A., Meil, A., Trehin, A., Reverdy, C., Betin, V., Maire, S., Brun, C., Jacq, B., Arpin, M., Bellaiche, Y., Bellusci, S., Benaroch, P., Bornens, M., Chanet, R., Chavrier, P., Delattre, O., Doye, V., Fehon, R., Faye, G., Galli, T.,

Girault, J.-A., Goud, B., de Gunzburg, J., Johannes, L., Junier, M.-P., Mirouse, V., Mukherjee, A., Papadopoulo, D., Perez, F., Plessis, A., Rossé, C., Saule, S., Stoppa-Lyonnet, D., Vincent, A., White, M., Legrain, P., Wojcik, J., Camonis, J., and Daviet, L. (2005) Protein interaction mapping: a *Drosophila* case study. *Genome Res* **15**: 376–384, URL http://dx.doi.org/10.1101/gr.2659105.

Fredrick, K. L. and Helmann, J. D. (1994) Dual chemotaxis signaling pathways in *Bacillus subtilis*: a sigma D-dependent gene encodes a novel protein with both CheW and CheY homologous domains. *J Bacteriol* **176**: 2727–2735, URL http://www.ncbi.nlm.nih.gov/pubmed/8169223.

Frickey, T. and Lupas, A. (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**: 3702–3704, URL http://dx.doi.org/10.1093/bioinformatics/bth444.

Frolow, F., Harel, M., Sussman, J. L., Mevarech, M., and Shoham, M. (1996) Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe-2S ferredoxin. *Nat Struct Biol* **3**: 452–458, URL http://www.ncbi.nlm.nih.gov/pubmed/8612076.

Fukuchi, S., Yoshimune, K., Wakayama, M., Moriguchi, M., and Nishikawa, K. (2003) Unique amino acid composition of proteins in halophilic bacteria. *J Mol Biol* **327**: 347–357, URL http://www.ncbi.nlm.nih.gov/pubmed/12628242.

Galagan, J. E., Nusbaum, C., Roy, A., Endrizzi, M. G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R. D., Cann, I., Graham, D. E., Grahame, D. A., Guss, A. M., Hedderich, R., Ingram-Smith, C., Kuettner, H. C., Krzycki, J. A., Leigh, J. A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J. N., Smith, K., Springer, T. A., Umayam, L. A., White, O., White, R. H., de Macario, E. C., Ferry, J. G., Jarrell, K. F., Jing, H., Macario, A. J. L., Paulsen, I., Pritchett, M., Sowers, K. R., Swanson, R. V., Zinder, S. H., Lander, E., Metcalf, W. W., and Birren, B. (2002) The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res* **12**: 532–542, URL http://dx.doi.org/10.1101/gr.223902.

Gambacorta, A., Trincone, A., Nicolaus, B., Lama, L., and Derosa, M. (1994) Unique Features Of Lipids Of Archaea. *Systematic And Applied Microbiology* **16**: 518–527.

Garrity, G. M., Boone, D. R., and Castenholz, R. W., editors (2001) Bergey's Manual of Systematic Bacteriology. Springer.

Garrity, L. F. and Ordal, G. W. (1997) Activation of the CheA kinase by asparagine in *Bacillus subtilis* chemotaxis. *Microbiology* **143**: 2945–2951, URL http://www.ncbi.nlm.nih.gov/pubmed/12094812.

Gavin, A.-C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dümpelfeld, B., Edelmann, A., Heurtier, M.-A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A.-M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J. M., Kuster, B., Bork, P., Russell, R. B., and Superti-Furga, G. (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**: 631–636, URL http://dx.doi.org/10.1038/nature04532.

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A.-M., Cruciat, C.-M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.-A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141–147, URL http://dx.doi.org/10.1038/415141a.

Gegner, J. A. and Dahlquist, F. W. (1991) Signal transduction in bacteria: CheW forms a reversible complex with the protein kinase CheA. *Proc Natl Acad Sci U S A* **88**: 750–754, URL http://www.ncbi.nlm.nih.gov/pubmed/1992467.

Gegner, J. A., Graham, D. R., Roth, A. F., and Dahlquist, F. W. (1992) Assembly of an MCP receptor, CheW, and kinase CheA complex in the bacterial chemotaxis signal transduction pathway. *Cell* **70**: 975–982, URL http://www.ncbi.nlm.nih.gov/pubmed/1326408.

Geiduschek, E. P. and Ouhammouch, M. (2005) Archaeal transcription and its regulators. *Mol Microbiol* **56**: 1397–1407, URL http://dx.doi.org/10.1111/j.1365-2958.2005.04627.x.

Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S., and Mische, S. M. (1999) Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis* **20**: 601–605, URL http://dx.doi.org/10.1002/(SICI)1522-2683(19990301)20:3<601::AID-ELPS601>3.0.CO;2-6.

Ghosh, M. and Sonawat, H. M. (1998) Kreb's TCA cycle in *Halobacterium salinarum* investigated by $^{13}$C nuclear magnetic resonance spectroscopy. *Extremophiles* **2**: 427–433, URL http://www.ncbi.nlm.nih.gov/pubmed/9827332.

Gingras, A.-C., Gstaiger, M., Raught, B., and Aebersold, R. (2007) Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* **8**: 645–654, URL http://dx.doi.org/10.1038/nrm2208.

Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chant, J., and Rothberg, J. M. (2003) A protein interaction map of *Drosophila melanogaster*. *Science* **302**: 1727–1736, URL http://dx.doi.org/10.1126/science.1090289.

Gloeckner, C. J., Boldt, K., Schumacher, A., Roepman, R., and Ueffing, M. (2007) A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics* **7**: 4228–4234, URL http://dx.doi.org/10.1002/pmic.200700038.

Goede, B., Naji, S., von Kampen, O., Ilg, K., and Thomm, M. (2006) Protein-protein interactions in the archaeal transcriptional machinery: binding studies of isolated RNA polymerase subunits and transcription factors. *J Biol Chem* **281**: 30581–30592, URL http://dx.doi.org/10.1074/jbc.M605209200.

# Bibliography

Goh, K.-I., Oh, E., Jeong, H., Kahng, B., and Kim, D. (2002) Classification of scale-free networks. *Proc Natl Acad Sci U S A* **99**: 12583–12588, URL http://dx.doi.org/10.1073/pnas.202301299.

Gonzalez, O., Gronau, S., Falb, M., Pfeiffer, F., Mendoza, E., Zimmer, R., and Oesterhelt, D. (2008) Reconstruction, modeling & analysis of *Halobacterium salinarum* R-1 metabolism. *Mol Biosyst* **4**: 148–159, URL http://dx.doi.org/10.1039/b715203e.

Gouet, P., Chinardet, N., Welch, M., Guillet, V., Cabantous, S., Birck, C., Mourey, L., and Samama, J. P. (2001) Further insights into the mechanism of function of the response regulator CheY from crystallographic studies of the CheY–CheA(124–257) complex. *Acta Crystallogr D Biol Crystallogr* **57**: 44–51, URL http://www.ncbi.nlm.nih.gov/pubmed/11134926.

Gropp, F. and Betlach, M. C. (1994) The bat gene of *Halobacterium halobium* encodes a trans-acting oxygen inducibility factor. *Proc Natl Acad Sci U S A* **91**: 5475–5479, URL http://www.ncbi.nlm.nih.gov/pubmed/8202511.

Guerrero, C., Tagwerker, C., Kaiser, P., and Huang, L. (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol Cell Proteomics* **5**: 366–378, URL http://dx.doi.org/10.1074/mcp.M500303-MCP200.

Hakes, L., Pinney, J. W., Robertson, D. L., and Lovell, S. C. (2008) Protein-protein interaction networks and biology–what's the connection? *Nat Biotechnol* **26**: 69–72, URL http://dx.doi.org/10.1038/nbt0108-69.

Hakes, L., Robertson, D. L., and Oliver, S. G. (2005) Effect of dataset selection on the topological interpretation of protein interaction networks. *BMC Genomics* **6**: 131, URL http://dx.doi.org/10.1186/1471-2164-6-131.

Hall, T. A. and Brown, J. W. (2004) Interactions between RNase P protein subunits in Archaea. *Archaea* **1**: 247–253, URL http://archaea.ws/archive/summaries/volume1/a1-247.html.

Han, J.-D. J., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J. M., Cusick, M. E., Roth, F. P., and Vidal, M. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**: 88–93, URL http://dx.doi.org/10.1038/nature02555.

Han, J.-D. J., Dupuy, D., Bertin, N., Cusick, M. E., and Vidal, M. (2005) Effect of sampling on topology predictions of protein-protein interaction networks. *Nat Biotechnol* **23**: 839–844, URL http://dx.doi.org/10.1038/nbt1116.

Hart, G. T., Ramani, A. K., and Marcotte, E. M. (2006) How complete are current yeast and human protein-interaction networks? *Genome Biol* **7**: 120, URL http://dx.doi.org/10.1186/gb-2006-7-11-120.

Hartley, J. L., Temple, G. F., and Brasch, M. A. (2000) DNA cloning using in vitro site-specific recombination. *Genome Res* **10**: 1788–1795, URL http://www.ncbi.nlm.nih.gov/pubmed/11076863.

Hartmann, R., Sickinger, H. D., and Oesterhelt, D. (1980) Anaerobic growth of halobacteria. *Proc Natl Acad Sci U S A* **77**: 3821–3825, URL http://www.ncbi.nlm.nih.gov/pubmed/6933439.

Haupts, U., Tittor, J., and Oesterhelt, D. (1999) Closing in on bacteriorhodopsin: progress in understanding the molecule. *Annu Rev Biophys Biomol Struct* **28**: 367–399, URL http://dx.doi.org/10.1146/annurev.biophys.28.1.367.

Hazelbauer, G. L. and Adler, J. (1971) Role of the galactose binding protein in chemotaxis of *Escherichia coli* toward galactose. *Nat New Biol* **230**: 101–104, URL http://www.ncbi.nlm.nih.gov/pubmed/4927373.

Hazelbauer, G. L., Falke, J. J., and Parkinson, J. S. (2008) Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* **33**: 9–19, URL http://dx.doi.org/10.1016/j.tibs.2007.09.014.

Hess, J. F., Oosawa, K., Kaplan, N., and Simon, M. I. (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* **53**: 79–87, URL http://www.ncbi.nlm.nih.gov/pubmed/3280143.

Hildebrand, E. and Schimz, A. (1990) The lifetime of photosensory signals in *Halobacterium halobium* and its dependence on protein methylation. *Biochim Biophys Acta* **1052**: 96–105, URL http://www.ncbi.nlm.nih.gov/pubmed/2322596.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S.-L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., ren Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A. R., Sassi, H., Nielsen, P. A., Rasmussen, K. J., Andersen, J. R., Johansen, L. E., Hansen, L. H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., rensen, B. D. S., Matthiesen, J., Hendrickson, R. C., Gleeson, F., Pawson, T., Moran, M. F., Durocher, D., Mann, M., Hogue, C. W. V., Figeys, D., and Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**: 180–183, URL http://dx.doi.org/10.1038/415180a.

Hoch, J. A. (2000) Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* **3**: 165–170, URL http://www.ncbi.nlm.nih.gov/pubmed/10745001.

Hope, I. A. and Struhl, K. (1986) Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**: 885–894, URL http://www.ncbi.nlm.nih.gov/pubmed/3530496.

Hou, S., Brooun, A., Yu, H. S., Freitas, T., and Alam, M. (1998) Sensory rhodopsin II transducer HtrII is also responsible for serine chemotaxis in the archaeon *Halobacterium salinarum*. *J Bacteriol* **180**: 1600–1602, URL http://www.ncbi.nlm.nih.gov/pubmed/9515936.

Hou, S., Larsen, R. W., Boudko, D., Riley, C. W., Karatan, E., Zimmer, M., Ordal, G. W., and Alam, M. (2000) Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature* **403**: 540–544, URL http://dx.doi.org/10.1038/35000570.

Huang, H., Jedynak, B. M., and Bader, J. S. (2007) Where have all the interactions gone? Estimating the coverage of two-hybrid protein interaction maps. *PLoS Comput Biol* **3**: e214, URL http://dx.doi.org/10.1371/journal.pcbi.0030214.

Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., and Stetter, K. O. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63–67, URL http://dx.doi.org/10.1038/417063a.

Imanaka, H., Yamatsu, A., Fukui, T., Atomi, H., and Imanaka, T. (2006) Phosphoenolpyruvate synthase plays an essential role for glycolysis in the modified Embden-Meyerhof pathway in *Thermococcus kodakarensis*. *Mol Microbiol* **61**: 898–909, URL http://dx.doi.org/10.1111/j.1365-2958.2006.05287.x.

Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**: 23–28, URL http://www.ncbi.nlm.nih.gov/pubmed/2265755.

Irihimovitch, V. and Eichler, J. (2003) Post-translational secretion of fusion proteins in the halophilic archaea *Haloferax volcanii*. *J Biol Chem* **278**: 12881–12887, URL http://dx.doi.org/10.1074/jbc.M210762200.

Irihimovitch, V., Ring, G., Elkayam, T., Konrad, Z., and Eichler, J. (2003) Isolation of fusion proteins containing SecY and SecE, components of the protein translocation complex from the halophilic archaeon *Haloferax volcanii*. *Extremophiles* **7**: 71–77, URL http://dx.doi.org/10.1007/s00792-002-0297-0.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**: 4569–4574, URL http://dx.doi.org/10.1073/pnas.061034498.

James, P., Halladay, J., and Craig, E. A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**: 1425–1436, URL http://www.ncbi.nlm.nih.gov/pubmed/8978031.

Jeong, H., Mason, S. P., Barabási, A. L., and Oltvai, Z. N. (2001) Lethality and centrality in protein networks. *Nature* **411**: 41–42, URL http://dx.doi.org/10.1038/35075138.

Kalmokoff, M. L. and Jarrell, K. F. (1991) Cloning and sequencing of a multigene family encoding the flagellins of *Methanococcus voltae*. *J Bacteriol* **173**: 7113–7125, URL http://www.ncbi.nlm.nih.gov/pubmed/1718944.

Kandler, O. (1994) Cell-Wall Biochemistry And 3-Domain Concept Of Life. *Systematic And Applied Microbiology* **16**: 501–509.

Karatan, E., Saulmon, M. M., Bunn, M. W., and Ordal, G. W. (2001) Phosphorylation of the response regulator CheV is required for adaptation to attractants during *Bacillus subtilis* chemotaxis. *J Biol Chem* **276**: 43618–43626, URL http://dx.doi.org/10.1074/jbc.M104955200.

Kastritis, P. L., Papandreou, N. C., and Hamodrakas, S. J. (2007) Haloadaptation: insights from comparative modeling studies of halophilic archaeal DHFRs. *Int J Biol Macromol* **41**: 447–453, URL http://dx.doi.org/10.1016/j.ijbiomac.2007.06.005.

Keegan, L., Gill, G., and Ptashne, M. (1986) Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231**: 699–704, URL http://www.ncbi.nlm.nih.gov/pubmed/3080805.

Kehry, M. R., Bond, M. W., Hunkapiller, M. W., and Dahlquist, F. W. (1983) Enzymatic deamidation of methyl-accepting chemotaxis proteins in *Escherichia coli* catalyzed by the cheB gene product. *Proc Natl Acad Sci U S A* **80**: 3599–3603, URL http://www.ncbi.nlm.nih.gov/pubmed/6304723.

Kehry, M. R. and Dahlquist, F. W. (1982) The methyl-accepting chemotaxis proteins of *Escherichia coli*. Identification of the multiple methylation sites on methyl-accepting chemotaxis protein I. *J Biol Chem* **257**: 10378–10386, URL http://www.ncbi.nlm.nih.gov/pubmed/6213619.

Keller, A., Eng, J., Zhang, N., jun Li, X., and Aebersold, R. (2005) A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol Syst Biol* **1**: 2005.0017, URL http://dx.doi.org/10.1038/msb4100024.

Kennelly, P. J. (2002) Protein kinases and protein phosphatases in prokaryotes: a genomic perspective. *FEMS Microbiol Lett* **206**: 1–8, URL http://www.ncbi.nlm.nih.gov/pubmed/11786249.

Kennelly, P. J. (2003) Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry. *Biochem J* **370**: 373–389, URL http://dx.doi.org/10.1042/BJ20021547.

Kentner, D. and Sourjik, V. (2006) Spatial organization of the bacterial chemotaxis system. *Curr Opin Microbiol* **9**: 619–624, URL http://dx.doi.org/10.1016/j.mib.2006.10.012.

Kerkar, S. (2005) Ecology of Hypersaline Microorganisms. In: Marine Microbiology: Facets & Opportunities, chap. 5. National Institute of Oceanography, Goa, 37–47, URL http://hdl.handle.net/2264/70.

Kim, C., Jackson, M., Lux, R., and Khan, S. (2001) Determinants of chemotactic signal amplification in *Escherichia coli*. *J Mol Biol* **307**: 119–135, URL http://dx.doi.org/10.1006/jmbi.2000.4389.

Kim, K. K., Yokota, H., and Kim, S. H. (1999) Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**: 787–792, URL http://dx.doi.org/10.1038/23512.

Kirby, J. R., Kristich, C. J., Saulmon, M. M., Zimmer, M. A., Garrity, L. F., Zhulin, I. B., and Ordal, G. W. (2001) CheC is related to the family of flagellar switch proteins and acts independently from CheD to control chemotaxis in *Bacillus subtilis*. *Mol Microbiol* **42**: 573–585, URL http://www.ncbi.nlm.nih.gov/pubmed/11722727.

Kirsch, M. L., Peters, P. D., Hanlon, D. W., Kirby, J. R., and Ordal, G. W. (1993a) Chemotactic methylesterase promotes adaptation to high concentrations of attractant in *Bacillus subtilis*. *J Biol Chem* **268**: 18610–18616, URL http://www.ncbi.nlm.nih.gov/pubmed/8395512.

Kirsch, M. L., Zuberi, A. R., Henner, D., Peters, P. D., Yazdi, M. A., and Ordal, G. W. (1993b) Chemotactic methyltransferase promotes adaptation to repellents in *Bacillus subtilis*. *J Biol Chem* **268**: 25350–25356, URL http://www.ncbi.nlm.nih.gov/pubmed/8244966.

Klein, A. (2005) Vergleichende Analyse halophiler Genome. Diploma thesis, Fachhochschule Bingen.

Klein, C., Aivaliotis, M., Olsen, J. V., Falb, M., Besir, H., Scheffer, B., Bisle, B., Tebbe, A., Konstantinidis, K., Siedler, F., Pfeiffer, F., Mann, M., and Oesterhelt, D. (2007) The low molecular weight proteome of *Halobacterium salinarum*. *J Proteome Res* **6**: 1510–1518, URL http://dx.doi.org/10.1021/pr060634q.

Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K., and Schiebel, E. (1999) Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast* **15**: 963–972, URL http://dx.doi.org/10.1002/(SICI)1097-0061(199907)15:10B<963::AID-YEA399>3.0.CO;2-W.

Koch, M. (2005) Investigations on halobacterial transducers with respect to membrane potential sensing and adaptive methylation. Ph.D. thesis, Ludwig-Maximilians-Universität München, URL http://edoc.ub.uni-muenchen.de/4954/.

Koch, M. K. and Oesterhelt, D. (2005) MpcT is the transducer for membrane potential changes in *Halobacterium salinarum*. *Mol Microbiol* **55**: 1681–1694, URL http://dx.doi.org/10.1111/j.1365-2958.2005.04516.x.

Koch, M. K., Staudinger, W. F., Siedler, F., and Oesterhelt, D. (2008) Physiological sites of deamidation and methyl esterification in sensory transducers of *Halobacterium salinarum*. *J Mol Biol* **380**: 285–302, URL http://dx.doi.org/10.1016/j.jmb.2008.04.063.

Kokoeva, M. V. and Oesterhelt, D. (2000) BasT, a membrane-bound transducer protein for amino acid detection in *Halobacterium salinarum*. *Mol Microbiol* **35**: 647–656, URL http://www.ncbi.nlm.nih.gov/pubmed/10672186.

Kokoeva, M. V., Storch, K.-F., Klein, C., and Oesterhelt, D. (2002) A novel mode of sensory transduction in archaea: binding protein-mediated chemotaxis towards osmoprotectants and amino acids. *EMBO J* **21**: 2312–2322, URL http://dx.doi.org/10.1093/emboj/21.10.2312.

Koshland, D. E. (1977) A response regulator model in a simple sensory system. *Science* **196**: 1055–1063, URL http://www.ncbi.nlm.nih.gov/pubmed/870969.

Kott, L., Braswell, E. H., Shrout, A. L., and Weis, R. M. (2004) Distributed subunit interactions in CheA contribute to dimer stability: a sedimentation equilibrium study. *Biochim Biophys Acta* **1696**: 131–140, URL http://www.ncbi.nlm.nih.gov/pubmed/14726213.

Kristich, C. J. and Ordal, G. W. (2002) *Bacillus subtilis* CheD is a chemoreceptor modification enzyme required for chemotaxis. *J Biol Chem* **277**: 25356–25362, URL http://dx.doi.org/10.1074/jbc.M201334200.

Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrín-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M. M., Vlasblom, J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rilstone, J. J., Gandi, K., Thompson, N. J., Musso, G., Onge, P. S., Ghanny, S., Lam, M. H. Y., Butland, G., Altaf-Ul, A. M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**: 637–643, URL http://dx.doi.org/10.1038/nature04670.

Kupper, J., Marwan, W., Typke, D., Grünberg, H., Uwer, U., Gluch, M., and Oesterhelt, D. (1994) The flagellar bundle of *Halobacterium salinarium* is inserted into a distinct polar cap structure. *J Bacteriol* **176**: 5184–5187, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=8051038.

172

Landy, A. (1989) Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem* **58**: 913–949, URL http://dx.doi.org/10.1146/annurev.bi.58.070189.004405.

Lanyi, J. K. (1974) Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol Rev* **38**: 272–290, URL http://www.ncbi.nlm.nih.gov/pubmed/4607500.

Lanyi, J. K. (2006) Proton transfers in the bacteriorhodopsin photocycle. *Biochim Biophys Acta* **1757**: 1012–1018, URL http://dx.doi.org/10.1016/j.bbabio.2005.11.003.

Lanyi, J. K. and Luecke, H. (2001) Bacteriorhodopsin. *Curr Opin Struct Biol* **11**: 415–419, URL http://www.ncbi.nlm.nih.gov/pubmed/11495732.

Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Henderson, R. K., King, D. S., Huang, L., Kustu, S., Berry, E. A., and Wemmer, D. E. (2001) Crystal structure of an activated response regulator bound to its target. *Nat Struct Biol* **8**: 52–56, URL http://dx.doi.org/10.1038/83053.

Lengeler, J. and Jahreis, K. (1996) Phosphotransferase systems or PTS as carbohydrate transport and as signal transduction systems., chap. Vol. 2. Elsevier Science, 573–598.

Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P.-O., Han, J.-D. J., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J.-F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., Heuvel, S. V. D., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E., and Vidal, M. (2004) A map of the interactome network of the metazoan *C. elegans*. *Science* **303**: 540–543, URL http://dx.doi.org/10.1126/science.1091403.

Li, X.-J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal Chem* **75**: 6648–6657, URL http://dx.doi.org/10.1021/ac034633i.

Lindley, E. V. and MacDonald, R. E. (1979) A second mechanism for sodium extrusion in *Halobacterium halobium*: a light-driven sodium pump. *Biochem Biophys Res Commun* **88**: 491–499, URL http://www.ncbi.nlm.nih.gov/pubmed/465050.

Lorenz, R. J. (1996) Grundbegriffe der Biometrie. Gustav Fischer Verlag.

Lukat, G. S., Stock, A. M., and Stock, J. B. (1990) Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis. *Biochemistry* **29**: 5436–5442, URL http://www.ncbi.nlm.nih.gov/pubmed/2201404.

Lupas, A. and Stock, J. (1989) Phosphorylation of an N-terminal regulatory domain activates the CheB methylesterase in bacterial chemotaxis. *J Biol Chem* **264**: 17337–17342, URL http://www.ncbi.nlm.nih.gov/pubmed/2677005.

Lux, R., Jahreis, K., Bettenbrock, K., Parkinson, J. S., and Lengeler, J. W. (1995) Coupling the phosphotransferase system and the methyl-accepting chemotaxis protein-dependent chemotaxis signaling pathways of *Escherichia coli*. *Proc Natl Acad Sci U S A* **92**: 11583–11587, URL http://www.ncbi.nlm.nih.gov/pubmed/8524808.

Lux, R., Munasinghe, V. R., Castellano, F., Lengeler, J. W., Corrie, J. E., and Khan, S. (1999) Elucidation of a PTS-carbohydrate chemotactic signal pathway in *Escherichia coli* using a time-resolved behavioral assay. *Mol Biol Cell* **10**: 1133–1146, URL http://www.ncbi.nlm.nih.gov/pubmed/10198062.

Ma, J. and Ptashne, M. (1988) Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* **55**: 443–446, URL http://www.ncbi.nlm.nih.gov/pubmed/3180218.

Macnab, R. M. and Koshland, D. E. (1972) The gradient-sensing mechanism in bacterial chemotaxis. *Proc Natl Acad Sci U S A* **69**: 2509–2512, URL http://www.ncbi.nlm.nih.gov/pubmed/4560688.

Maddock, J. R. and Shapiro, L. (1993) Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science* **259**: 1717–1723, URL http://www.ncbi.nlm.nih.gov/pubmed/8456299.

Maeder, D. L., Anderson, I., Brettin, T. S., Bruce, D. C., Gilna, P., Han, C. S., Lapidus, A., Metcalf, W. W., Saunders, E., Tapia, R., and Sowers, K. R. (2006) The *Methanosarcina barkeri* genome: comparative analysis with *Methanosarcina acetivorans* and *Methanosarcina mazei* reveals extensive rearrangement within methanosarcinal genomes. *J Bacteriol* **188**: 7922–7931, URL http://dx.doi.org/10.1128/JB.00810-06.

Marwan, W., Alam, M., and Oesterhelt, D. (1991) Rotation and switching of the flagellar motor assembly in *Halobacterium halobium*. *J Bacteriol* **173**: 1971–1977, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=2002000.

Marwan, W. and Oesterhelt, D. (1990) Quantitation of photochromism of sensory rhodopsin-I by computerized tracking of *Halobacterium halobium* cells. *J Mol Biol* **215**: 277–285, URL http://www.ncbi.nlm.nih.gov/pubmed/2213884.

Marwan, W. and Oesterhelt, D. (2000) Archaeal Vision and Bacterial Smelling. *ASM News* **66**: 83–89, URL http://newsarchive.asm.org/feb00/feature3.asp.

Marwan, W., Schäfer, W., and Oesterhelt, D. (1990) Signal transduction in *Halobacterium* depends on fumarate. *EMBO J* **9**: 355–362, URL http://www.ncbi.nlm.nih.gov/pubmed/2303030.

McCraith, S., Holtzman, T., Moss, B., and Fields, S. (2000) Genome-wide analysis of vaccinia virus protein-protein interactions. *Proc Natl Acad Sci U S A* **97**: 4879–4884, URL http://dx.doi.org/10.1073/pnas.080078197.

McEvoy, M. M., Bren, A., Eisenbach, M., and Dahlquist, F. W. (1999) Identification of the binding interfaces on CheY for two of its targets, the phosphatase CheZ and the flagellar switch protein FliM. *J Mol Biol* **289**: 1423–1433, URL http://dx.doi.org/10.1006/jmbi.1999.2830.

McEvoy, M. M., Hausrath, A. C., Randolph, G. B., Remington, S. J., and Dahlquist, F. W. (1998) Two binding modes reveal flexibility in kinase/response regulator interactions in the bacterial chemotaxis pathway. *Proc Natl Acad Sci U S A* **95**: 7333–7338, URL http://www.ncbi.nlm.nih.gov/pubmed/9636149.

McNamara, B. P. and Wolfe, A. J. (1997) Coexpression of the long and short forms of CheA, the chemotaxis histidine kinase, by members of the family *Enterobacteriaceae*. *J Bacteriol* **179**: 1813–1818, URL http://www.ncbi.nlm.nih.gov/pubmed/9045846.

Metlina, A. L. (2004) Bacterial and archaeal flagella as prokaryotic motility organelles. *Biochemistry (Mosc)* **69**: 1203–1212, URL http://protein.bio.msu.ru/biokhimiya/contents/v69/full/69111477.html.

Mevarech, M., Frolow, F., and Gloss, L. M. (2000) Halophilic enzymes: proteins with a grain of salt. *Biophys Chem* **86**: 155–164, URL http://dx.doi.org/10.1016/S0301-4622(00)00126-5.

Miller, A. (2007) Untersuchung von putativen archaealen Chemotaxisproteinen durch Proteininteraktionsstudien und Charakterisierung von Deletionsmutanten in *Halobacterium salinarum*. Diploma thesis, Hochschule Darmstadt - University of Applied Sciences.

Milligan, D. L. and Koshland, D. E. (1988) Site-directed cross-linking. Establishing the dimeric structure of the aspartate receptor of bacterial chemotaxis. *J Biol Chem* **263**: 6268–6275, URL http://www.ncbi.nlm.nih.gov/pubmed/2834370.

Montrone, M., Eisenbach, M., Oesterhelt, D., and Marwan, W. (1998) Regulation of switching frequency and bias of the bacterial flagellar motor by CheY and fumarate. *J Bacteriol* **180**: 3375–3380, URL http://www.ncbi.nlm.nih.gov/pubmed/9642190.

Montrone, M., Marwan, W., Grünberg, H., Musseleck, S., Starostzik, C., and Oesterhelt, D. (1993) Sensory rhodopsin-controlled release of the switch factor fumarate in *Halobacterium salinarium*. *Mol Microbiol* **10**: 1077–1085, URL http://www.ncbi.nlm.nih.gov/pubmed/7934858.

Montrone, M., Oesterhelt, D., and Marwan, W. (1996) Phosphorylation-independent bacterial chemoresponses correlate with changes in the cytoplasmic level of fumarate. *J Bacteriol* **178**: 6882–6887, URL http://www.ncbi.nlm.nih.gov/pubmed/8955310.

Morag, E., Lapidot, A., Govorko, D., Lamed, R., Wilchek, M., Bayer, E. A., and Shoham, Y. (1995) Expression, purification, and characterization of the cellulose-binding domain of the scaffoldin subunit from the cellulosome of *Clostridium thermocellum*. *Appl Environ Microbiol* **61**: 1980–1986, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=7646033.

Motaleb, M. A., Miller, M. R., Li, C., Bakker, R. G., Goldstein, S. F., Silversmith, R. E., Bourret, R. B., and Charon, N. W. (2005) CheX is a phosphorylated CheY phosphatase essential for *Borrelia burgdorferi* chemotaxis. *J Bacteriol* **187**: 7963–7969, URL http://dx.doi.org/10.1128/JB.187.23.7963-7969.2005.

Muff, T. J. and Ordal, G. W. (2007) The CheC phosphatase regulates chemotactic adaptation through CheD. *J Biol Chem* **282**: 34120–34128, URL http://dx.doi.org/10.1074/jbc.M706432200.

Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Buillard, V., Cerutti, L., Copley, R., Courcelle, E., Das, U., Daugherty, L., Dibley, M., Finn, R., Fleischmann, W., Gough, J., Haft, D., Hulo, N., Hunter, S., Kahn, D., Kanapin, A., Kejariwal, A., Labarga, A., Langendijk-Genevaux, P. S., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Nikolskaya, A. N., Orchard, S., Orengo, C., Petryszak, R., Selengut, J. D., Sigrist, C. J. A., Thomas, P. D., Valentin, F., Wilson, D., Wu, C. H., and Yeats, C. (2007) New developments in the InterPro database. *Nucleic Acids Res* **35**: D224–D228, URL http://dx.doi.org/10.1093/nar/gkl841.

Ng, S. Y. M., Chaban, B., and Jarrell, K. F. (2006) Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J Mol Microbiol Biotechnol* **11**: 167–191, URL http://dx.doi.org/10.1159/000094053.

Ng, W. V., Kennedy, S. P., Mahairas, G. G., Berquist, B., Pan, M., Shukla, H. D., Lasky, S. R., Baliga, N. S., Thorsson, V., Sbrogna, J., Swartzell, S., Weir, D., Hall, J., Dahl, T. A., Welti, R., Goo, Y. A., Leithauser, B., Keller, K., Cruz, R., Danson, M. J., Hough, D. W., Maddocks, D. G., Jablonski, P. E., Krebs, M. P., Angevine, C. M., Dale, H., Isenbarger, T. A., Peck, R. F., Pohlschroder, M., Spudich, J. L., Jung, K. W., Alam, M., Freitas, T., Hou, S., Daniels, C. J., Dennis, P. P., Omer, A. D., Ebhardt, H., Lowe, T. M., Liang, P., Riley, M., Hood, L., and DasSarma, S. (2000) Genome sequence of *Halobacterium* species NRC-1. *Proc Natl Acad Sci U S A* **97**: 12176–12181, URL http://dx.doi.org/10.1073/pnas.190337797.

Ninfa, E. G., Stock, A., Mowbray, S., and Stock, J. (1991) Reconstitution of the bacterial chemotaxis signal transduction system from purified components. *J Biol Chem* **266**: 9764–9770, URL http://www.ncbi.nlm.nih.gov/pubmed/1851755.

Nowlin, D. M., Bollinger, J., and Hazelbauer, G. L. (1987) Sites of covalent modification in Trg, a sensory transducer of *Escherichia coli. J Biol Chem* **262**: 6039–6045, URL http://www.ncbi.nlm.nih.gov/pubmed/3032955.

Nutsch, T., Marwan, W., Oesterhelt, D., and Gilles, E. D. (2003) Signal processing and flagellar motor switching during phototaxis of *Halobacterium salinarum. Genome Res* **13**: 2406–2412, URL http://dx.doi.org/10.1101/gr.1241903.

Nutsch, T., Oesterhelt, D., Gilles, E. D., and Marwan, W. (2005) A quantitative model of the switch cycle of an archaeal flagellar motor and its sensory control. *Biophys J* **89**: 2307–2323, URL http://dx.doi.org/10.1529/biophysj.104.057570.

Oesterhelt, D. and Krippahl, G. (1983) Phototrophic growth of halobacteria and its use for isolation of photosynthetically-deficient mutants. *Ann Microbiol (Paris)* **134B**: 137–150, URL http://www.ncbi.nlm.nih.gov/pubmed/6638758.

Oesterhelt, D. and Stoeckenius, W. (1973) Functions of a new photoreceptor membrane. *Proc Natl Acad Sci U S A* **70**: 2853–2857, URL http://www.ncbi.nlm.nih.gov/pubmed/4517939.

Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**: 376–386, URL http://www.ncbi.nlm.nih.gov/pubmed/12118079.

Oren, A. (1991) Anaerobic growth of halophilic archaeobacteria by reduction of fumarate. *J Gen Microbiol* **137**: 1387–1390.

Oren, A. (1994) The Ecology Of The Extremely Halophilic Archaea. *Fems Microbiology Reviews* **13**: 415–439.

Oren, A. (1999) Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* **63**: 334–348, URL http://www.ncbi.nlm.nih.gov/pubmed/10357854.

Oren, A. (2002) Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications. *J Ind Microbiol Biotechnol* **28**: 56–63, URL http://dx.doi.org/10.1038/sj/jim/7000176.

Oren, A. and Trüper, H. G. (1990) Anaerobic growth of halophilic archaeobacteria by reduction of dimethysulfoxide and trimethylamine N-oxide. *FEMS Microbiology Letters* **70**: 33–36, URL http://www.blackwell-synergy.com/doi/abs/10.1111/j.1574-6968.1990.tb03772.x.

Ortenberg, R. and Mevarech, M. (2000) Evidence for post-translational membrane insertion of the integral membrane protein bacterioopsin expressed in the heterologous halophilic archaeon *Haloferax volcanii. J Biol Chem* **275**: 22839–22846, URL http://dx.doi.org/10.1074/jbc.M908916199.

O'Toole, P. W., Lane, M. C., and Porwollik, S. (2000) *Helicobacter pylori* motility. *Microbes Infect* **2**: 1207–1214, URL http://www.ncbi.nlm.nih.gov/pubmed/11008110.

Park, S.-Y., Beel, B. D., Simon, M. I., Bilwes, A. M., and Crane, B. R. (2004) In different organisms, the mode of interaction between two signaling proteins is not necessarily conserved. *Proc Natl Acad Sci U S A* **101**: 11646–11651, URL http://dx.doi.org/10.1073/pnas.0401038101.

Park, S.-Y., Borbat, P. P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A. M., Freed, J. H., Bilwes, A. M., and Crane, B. R. (2006) Reconstruction of the chemotaxis receptor-kinase assembly. *Nat Struct Mol Biol* **13**: 400–407, URL http://dx.doi.org/10.1038/nsmb1085.

Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M. J., Penn, C. W., Quail, M. A., Rajandream, M. A., Rutherford, K. M., van Vliet, A. H., Whitehead, S., and Barrell, B. G. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**: 665–668, URL http://dx.doi.org/10.1038/35001088.

Parkinson, J. S. and Kofoid, E. C. (1992) Communication modules in bacterial signaling proteins. *Annu Rev Genet* **26**: 71–112, URL http://dx.doi.org/10.1146/annurev.ge.26.120192.000443.

Parrish, J. R., Yu, J., Liu, G., Hines, J. A., Chan, J. E., Mangiola, B. A., Zhang, H., Pacifico, S., Fotouhi, F., DiRita, V. J., Ideker, T., Andrews, P., and Finley, R. L. (2007) A proteome-wide protein interaction map for *Campylobacter jejuni. Genome Biol* **8**: R130, URL http://dx.doi.org/10.1186/gb-2007-8-7-r130.

Patenge, N., Berendes, A., Engelhardt, H., Schuster, S. C., and Oesterhelt, D. (2001) The *fla* gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum. Mol Microbiol* **41**: 653–663, URL http://www.ncbi.nlm.nih.gov/pubmed/11532133.

Pfeiffer, F., Broicher, A., Gillich, T., Klee, K., Mejía, J., Rampp, M., and Oesterhelt, D. (2008a) Genome information management and integrated data analysis with HaloLex. *Arch Microbiol* **190**: 281–299, URL http://dx.doi.org/10.1007/s00203-008-0389-z.

Pfeiffer, F., Schuster, S. C., Broicher, A., Falb, M., Palm, P., Rodewald, K., Ruepp, A., Soppa, J., Tittor, J., and Oesterhelt, D. (2008b) Evolution in the laboratory: The genome of *Halobacterium salinarum* strain R1 compared to that of strain NRC-1. *Genomics* **91**: 335–346, URL http://dx.doi.org/10.1016/j.ygeno.2008.01.001.

Pitre, S., Alamgir, M., Green, J., Dumontier, M., Dehne, F., and Golshani, A. (2008) Computational Methods For Predicting Protein-Protein Interactions. *Adv Biochem Eng Biotechnol* URL http://dx.doi.org/10.1007/10_2007_089.

Porter, S. L. and Armitage, J. P. (2002) Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. *J Mol Biol* **324**: 35–45, URL http://www.ncbi.nlm.nih.gov/pubmed/12421557.

Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993) Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57**: 543–594, URL http://www.ncbi.nlm.nih.gov/pubmed/8246840.

Prasad, K., Caplan, S. R., and Eisenbach, M. (1998) Fumarate modulates bacterial flagellar rotation by lowering the free energy difference between the clockwise and counterclockwise states of the motor. *J Mol Biol* **280**: 821–828, URL http://dx.doi.org/10.1006/jmbi.1998.1922.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Séraphin, B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**: 218–229, URL http://dx.doi.org/10.1006/meth.2001.1183.

R Development Core Team (2008) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org, ISBN 3-900051-07-0.

Rain, J. C., Selig, L., Reuse, H. D., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schächter, V., Chemama, Y., Labigne, A., and Legrain, P. (2001) The protein-protein interaction map of *Helicobacter pylori. Nature* **409**: 211–215, URL http://dx.doi.org/10.1038/35051615.

Rajagopala, S. V., Titz, B., Goll, J., Parrish, J. R., Wohlbold, K., McKevitt, M. T., Palzkill, T., Mori, H., Finley, R. L., and Uetz, P. (2007) The protein network of bacterial motility. *Mol Syst Biol* **3**: 128, URL http://dx.doi.org/10.1038/msb4100166.

Ranish, J. A., Yi, E. C., Leslie, D. M., Purvine, S. O., Goodlett, D. R., Eng, J., and Aebersold, R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat Genet* **33**: 349–355, URL http://dx.doi.org/10.1038/ng1101.

Rao, C. V., Kirby, J. R., and Arkin, A. P. (2005) Phosphatase localization in bacterial chemotaxis: divergent mechanisms, convergent principles. *Phys Biol* **2**: 148–158, URL http://dx.doi.org/10.1088/1478-3975/2/3/002.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75**: 663–670, URL http://www.ncbi.nlm.nih.gov/pubmed/12585499.

Rollins, C. and Dahlquist, F. W. (1981) The methyl-accepting chemotaxis proteins of *E. coli*: a repellent-stimulated, covalent modification, distinct from methylation. *Cell* **25**: 333–340, URL http://www.ncbi.nlm.nih.gov/pubmed/7026041.

Rosario, M. M. and Ordal, G. W. (1996) CheC and CheD interact to regulate methylation of *Bacillus subtilis* methyl-accepting chemotaxis proteins. *Mol Microbiol* **21**: 511–518, URL http://www.ncbi.nlm.nih.gov/pubmed/8866475.

Rowsell, E. H., Smith, J. M., Wolfe, A., and Taylor, B. L. (1995) CheA, CheW, and CheY are required for chemotaxis to oxygen and sugars of the phosphotransferase system in *Escherichia coli. J Bacteriol* **177**: 6011–6014, URL http://www.ncbi.nlm.nih.gov/pubmed/7592359.

Ruden, D. M., Ma, J., Li, Y., Wood, K., and Ptashne, M. (1991) Generating yeast transcriptional activators containing no yeast protein sequences. *Nature* **350**: 250–252, URL http://dx.doi.org/10.1038/350250a0.

Rudolph, J. and Oesterhelt, D. (1995) Chemotaxis and phototaxis require a CheA histidine kinase in the archaeon *Halobacterium salinarium. EMBO J* **14**: 667–673, URL http://www.ncbi.nlm.nih.gov/pubmed/7882970.

Rudolph, J. and Oesterhelt, D. (1996) Deletion analysis of the *che* operon in the archaeon *Halobacterium salinarium. J Mol Biol* **258**: 548–554, URL http://dx.doi.org/10.1006/jmbi.1996.0267.

Rudolph, J., Tolliday, N., Schmitt, C., Schuster, S. C., and Oesterhelt, D. (1995) Phosphorylation in halobacterial signal transduction. *EMBO J* **14**: 4249–4257, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=7556066.

Ruepp, A. and Soppa, J. (1996) Fermentative arginine degradation in *Halobacterium salinarium* (formerly *Halobacterium halobium*): genes, gene products, and transcripts of the arcRACB gene cluster. *J Bacteriol* **178**: 4942–4947, URL http://www.ncbi.nlm.nih.gov/pubmed/8759859.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487–491, URL http://www.ncbi.nlm.nih.gov/pubmed/2448875.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463–5467, URL http://www.ncbi.nlm.nih.gov/pubmed/271968.

Saulmon, M. M., Karatan, E., and Ordal, G. W. (2004) Effect of loss of CheC and other adaptational proteins on chemotactic behaviour in *Bacillus subtilis. Microbiology* **150**: 581–589, URL http://www.ncbi.nlm.nih.gov/pubmed/14993307.

Schäfer, G., Engelhard, M., and Müller, V. (1999) Bioenergetics of the Archaea. *Microbiol Mol Biol Rev* **63**: 570–620, URL http://www.ncbi.nlm.nih.gov/pubmed/10477309.

Schäfer, G., Purschke, W. G., Gleissner, M., and Schmidt, C. L. (1996) Respiratory chains of archaea and extremophiles. *Biochim Biophys Acta* **1275**: 16–20, URL http://www.ncbi.nlm.nih.gov/pubmed/8688447.

Scharf, B. E., Fahrner, K. A., Turner, L., and Berg, H. C. (1998) Control of direction of flagellar rotation in bacterial chemotaxis. *Proc Natl Acad Sci U S A* **95**: 201–206, URL http://www.ncbi.nlm.nih.gov/pubmed/9419353.

Schmitt, R. (2002) Sinorhizobial chemotaxis: a departure from the enterobacterial paradigm. *Microbiology* **148**: 627–631, URL http://www.ncbi.nlm.nih.gov/pubmed/11882696.

Schobert, B. and Lanyi, J. K. (1982) Halorhodopsin is a light-driven chloride pump. *J Biol Chem* **257**: 10306–10313, URL http://www.ncbi.nlm.nih.gov/pubmed/7107607.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**: 5857–5864, URL http://www.ncbi.nlm.nih.gov/pubmed/9600884.

Schulze, W. X. and Mann, M. (2004) A novel proteomic screen for peptide-protein interactions. *J Biol Chem* **279**: 10756–10764, URL http://dx.doi.org/10.1074/jbc.M309909200.

Schuster, S. C., Swanson, R. V., Alex, L. A., Bourret, R. B., and Simon, M. I. (1993) Assembly and function of a quaternary signal transduction complex monitored by surface plasmon resonance. *Nature* **365**: 343–347, URL http://dx.doi.org/10.1038/365343a0.

Selbach, M. and Mann, M. (2006) Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat Methods* **3**: 981–983, URL http://dx.doi.org/10.1038/nmeth972.

Semple, J. I., Sanderson, C. M., and Campbell, R. D. (2002) The jury is out on "guilt by association" trials. *Brief Funct Genomic Proteomic* **1**: 40–52, URL http://www.ncbi.nlm.nih.gov/pubmed/15251065.

Shah, D. S., Porter, S. L., Martin, A. C., Hamblin, P. A., and Armitage, J. P. (2000) Fine tuning bacterial chemotaxis: analysis of *Rhodobacter sphaeroides* behaviour under aerobic and anaerobic conditions by mutation of the major chemotaxis operons and cheY genes. *EMBO J* **19**: 4601–4613, URL http://dx.doi.org/10.1093/emboj/19.17.4601.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–2504, URL http://dx.doi.org/10.1101/gr.1239303.

Sherris, D. and Parkinson, J. S. (1981) Posttranslational processing of methyl-accepting chemotaxis proteins in *Escherichia coli*. *Proc Natl Acad Sci U S A* **78**: 6051–6055, URL http://www.ncbi.nlm.nih.gov/pubmed/6458812.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* **68**: 850–858, URL http://www.ncbi.nlm.nih.gov/pubmed/8779443.

Shimizu, T. S., Novère, N. L., Levin, M. D., Beavil, A. J., Sutton, B. J., and Bray, D. (2000) Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis. *Nat Cell Biol* **2**: 792–796, URL http://dx.doi.org/10.1038/35041030.

Shukla, D. and Matsumura, P. (1995) Mutations leading to altered CheA binding cluster on a face of CheY. *J Biol Chem* **270**: 24414–24419, URL http://www.ncbi.nlm.nih.gov/pubmed/7592655.

Shuman, S. (1991) Recombination mediated by vaccinia virus DNA topoisomerase I in *Escherichia coli* is sequence specific. *Proc Natl Acad Sci U S A* **88**: 10104–10108, URL http://www.ncbi.nlm.nih.gov/pubmed/1658796.

Silversmith, R. E., Appleby, J. L., and Bourret, R. B. (1997) Catalytic mechanism of phosphorylation and dephosphorylation of CheY: kinetic characterization of imidazole phosphates as phosphodonors and the role of acid catalysis. *Biochemistry* **36**: 14965–14974, URL http://dx.doi.org/10.1021/bi9715573.

Silversmith, R. E., Guanga, G. P., Betts, L., Chu, C., Zhao, R., and Bourret, R. B. (2003) CheZ-mediated dephosphorylation of the *Escherichia coli* chemotaxis response regulator CheY: role for CheY glutamate 89. *J Bacteriol* **185**: 1495–1502, URL http://www.ncbi.nlm.nih.gov/pubmed/12591865.

Sim, J.-H., Shi, W., and Lux, R. (2005) Protein-protein interactions in the chemotaxis signalling pathway of *Treponema denticola*. *Microbiology* **151**: 1801–1807, URL http://dx.doi.org/10.1099/mic.0.27622-0.

Simms, S. A., Stock, A. M., and Stock, J. B. (1987) Purification and characterization of the S-adenosylmethionine:glutamyl methyltransferase that modifies membrane chemoreceptor proteins in bacteria. *J Biol Chem* **262**: 8537–8543, URL http://www.ncbi.nlm.nih.gov/pubmed/3298235.

Simms, S. A. and Subbaramaiah, K. (1991) The kinetic mechanism of S-adenosyl-L-methionine: glutamylmethyltransferase from *Salmonella typhimurium*. *J Biol Chem* **266**: 12741–12746, URL http://www.ncbi.nlm.nih.gov/pubmed/2061337.

Soppa, J. (2006) From genomes to function: haloarchaea as model organisms. *Microbiology* **152**: 585–590, URL http://dx.doi.org/10.1099/mic.0.28504-0.

Sourjik, V. and Berg, H. C. (2000) Localization of components of the chemotaxis machinery of *Escherichia coli* using fluorescent protein fusions. *Mol Microbiol* **37**: 740–751, URL http://www.ncbi.nlm.nih.gov/pubmed/10972797.

Sourjik, V. and Berg, H. C. (2002a) Binding of the *Escherichia coli* response regulator CheY to its target measured in vivo by fluorescence resonance energy transfer. *Proc Natl Acad Sci U S A* **99**: 12669–12674, URL http://dx.doi.org/10.1073/pnas.192463199.

Sourjik, V. and Berg, H. C. (2002b) Receptor sensitivity in bacterial chemotaxis. *Proc Natl Acad Sci U S A* **99**: 123–127, URL http://dx.doi.org/10.1073/pnas.011589998.

Sourjik, V. and Berg, H. C. (2004) Functional interactions between receptors in bacterial chemotaxis. *Nature* **428**: 437–441, URL http://dx.doi.org/10.1038/nature02406.

Sourjik, V. and Schmitt, R. (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of *Rhizobium meliloti*. *Biochemistry* **37**: 2327–2335, URL http://dx.doi.org/10.1021/bi972330a.

Speranskii, V., Metlina, A., Novikova, T., and Bakeyeva, L. (1996) Disk-like lamellar structure as part of the archaeal flagellar apparatus. *Biophysics* **41**: 167–173.

Spraggon, G., Pantazatos, D., Klock, H. E., Wilson, I. A., Woods, V. L., and Lesley, S. A. (2004) On the use of DXMS to produce more crystallizable proteins: structures of the *T. maritima* proteins TM0160 and TM1171. *Protein Sci* **13**: 3187–3199, URL http://dx.doi.org/10.1110/ps.04939904.

Springer, W. R. and Koshland, D. E. (1977) Identification of a protein methyltransferase as the *cheR* gene product in the bacterial sensing system. *Proc Natl Acad Sci U S A* **74**: 533–537, URL http://www.ncbi.nlm.nih.gov/pubmed/322131.

Spudich, E. N. and Spudich, J. L. (1993) The photochemical reactions of sensory rhodopsin I are altered by its transducer. *J Biol Chem* **268**: 16095–16097.

Spudich, J. L. and Stoeckenius, W. (1979) Photosensory and chemosensory behavior of *Halobacterium halobium*. *Photobiochemistry and Photobiophysics* **1**: 43–53.

Stajich, J. E., Block, D., Boulez, K., Brenner, S. E., Chervitz, S. A., Dagdigian, C., Fuellen, G., Gilbert, J. G. R., Korf, I., Lapp, H., Lehväslaiho, H., Matsalla, C., Mungall, C. J., Osborne, B. I., Pocock, M. R., Schattner, P., Senger, M., Stein, L. D., Stupka, E., Wilkinson, M. D., and Birney, E. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* **12**: 1611–1618, URL http://dx.doi.org/10.1101/gr.361602.

Staudinger, W. (2001) Funktionsanalyse des Gens *flaK* in *Halobacterium salinarum* durch Deletion und Komplementation. Diploma thesis, Julius-Maximilians-Universität Würzburg.

Staudinger, W. (2007) Investigations on Flagellar Biogenesis, Motility and Signal Transduction of *Halobacterium salinarum*. Ph.D. thesis, Ludwig-Maximilians-Universität München.

Stewart, R. C., Roth, A. F., and Dahlquist, F. W. (1990) Mutations that affect control of the methylesterase activity of CheB, a component of the chemotaxis adaptation system in *Escherichia coli*. *J Bacteriol* **172**: 3388–3399, URL http://www.ncbi.nlm.nih.gov/pubmed/2188960.

Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) Two-component signal transduction. *Annu Rev Biochem* **69**: 183–215, URL http://dx.doi.org/10.1146/annurev.biochem.69.1.183.

Stock, A. M., Wylie, D. C., Mottonen, J. M., Lupas, A. N., Ninfa, E. G., Ninfa, A. J., Schutt, C. E., and Stock, J. B. (1988) Phosphoproteins involved in bacterial signal transduction. *Cold Spring Harb Symp Quant Biol* **53 Pt 1**: 49–57, URL http://www.ncbi.nlm.nih.gov/pubmed/3076087.

Stoeckenius, W., Lozier, R. H., and Bogomolni, R. A. (1979) Bacteriorhodopsin and the purple membrane of halobacteria. *Biochim Biophys Acta* **505**: 215–278, URL http://dx.doi.org/10.1016/0304-4173(79)90006-5.

Storch, K. F., Rudolph, J., and Oesterhelt, D. (1999) Car: a cytoplasmic sensor responsible for arginine chemotaxis in the archaeon *Halobacterium salinarum*. *EMBO J* **18**: 1146–1158, URL http://dx.doi.org/10.1093/emboj/18.5.1146.

Streif, S., Staudinger, W. F., Marwan, W., and Oesterhelt, D. (2008) Flagellar rotation in the archaeon *Halobacterium salinarum* depends on ATP. *Journal of Molecular Biology* URL http://dx.doi.org/10.1016/j.jmb.2008.08.057.

Stumpf, M. P. H., Wiuf, C., and May, R. M. (2005) Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proc Natl Acad Sci U S A* **102**: 4221–4224, URL http://dx.doi.org/10.1073/pnas.0501179102.

Szabó, Z., Sani, M., Groeneveld, M., Zolghadr, B., Schelert, J., Albers, S.-V., Blum, P., Boekema, E. J., and Driessen, A. J. M. (2007) Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol* **189**: 4305–4309, URL http://dx.doi.org/10.1128/JB.00042-07.

Szurmant, H., Bunn, M. W., Cannistraro, V. J., and Ordal, G. W. (2003) *Bacillus subtilis* hydrolyzes CheY-P at the location of its action, the flagellar switch. *J Biol Chem* **278**: 48611–48616, URL http://dx.doi.org/10.1074/jbc.M306180200.

Szurmant, H., Muff, T. J., and Ordal, G. W. (2004) *Bacillus subtilis* CheC and FliY are members of a novel class of CheY-P-hydrolyzing proteins in the chemotactic signal transduction cascade. *J Biol Chem* **279**: 21787–21792, URL http://dx.doi.org/10.1074/jbc.M311497200.

Szurmant, H. and Ordal, G. W. (2004) Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol Mol Biol Rev* **68**: 301–319, URL http://dx.doi.org/10.1128/MMBR.68.2.301-319.2004.

Tarasov, V. Y., Besir, H., Schwaiger, R., Klee, K., Furtwängler, K., Pfeiffer, F., and Oesterhelt, D. (2008) A small protein from the *bop-brp* intergenic region of *Halobacterium salinarum* contains a zinc finger motif and regulates *bop* and *crtB1* transcription. *Mol Microbiol* **67**: 772–780, URL http://www.ncbi.nlm.nih.gov/pubmed/18179416.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., and Natale, D. A. (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**: 41, URL http://dx.doi.org/10.1186/1471-2105-4-41.

Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997) A genomic perspective on protein families. *Science* **278**: 631–637, URL http://www.ncbi.nlm.nih.gov/pubmed/9381173.

Tebbe, A., Klein, C., Bisle, B., Siedler, F., Scheffer, B., Garcia-Rizo, C., Wolfertz, J., Hickmann, V., Pfeiffer, F., and Oesterhelt, D. (2005) Analysis of the cytosolic proteome of *Halobacterium salinarum* and its implication for genome annotation. *Proteomics* **5**: 168–179, URL http://dx.doi.org/10.1002/pmic.200400910.

Thar, R. and Kuhl, M. (2003) Bacteria are not too small for spatial sensing of chemical gradients: an experimental evidence. *Proc Natl Acad Sci U S A* **100**: 5748–5753, URL http://dx.doi.org/10.1073/pnas.1030795100.

Thomas, N. A., Bardy, S. L., and Jarrell, K. F. (2001a) The archaeal flagellum: a different kind of prokaryotic motility structure. *FEMS Microbiol Rev* **25**: 147–174, URL http://www.ncbi.nlm.nih.gov/pubmed/11250034.

Thomas, N. A. and Jarrell, K. F. (2001) Characterization of flagellum gene families of methanogenic archaea and localization of novel flagellum accessory proteins. *J Bacteriol* **183**: 7154–7164, URL http://dx.doi.org/10.1128/JB.183.24.7154-7164.2001.

Thomas, N. A., Mueller, S., Klein, A., and Jarrell, K. F. (2002) Mutants in *flaI* and *flaJ* of the archaeon *Methanococcus voltae* are deficient in flagellum assembly. *Mol Microbiol* **46**: 879–887, URL http://www.ncbi.nlm.nih.gov/pubmed/12410843.

Thomas, N. A., Pawson, C. T., and Jarrell, K. F. (2001b) Insertional inactivation of the *flaH* gene in the archaeon *Methanococcus voltae* results in non-flagellated cells. *Mol Genet Genomics* **265**: 596–603, URL http://www.ncbi.nlm.nih.gov/pubmed/11459179.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882, URL http://www.ncbi.nlm.nih.gov/pubmed/9396791.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680, URL http://www.ncbi.nlm.nih.gov/pubmed/7984417.

*Bibliography*

Thompson, S. R., Wadhams, G. H., and Armitage, J. P. (2006) The positioning of cytoplasmic protein clusters in bacteria. *Proc Natl Acad Sci U S A* **103**: 8209–8214, URL http://dx.doi.org/10.1073/pnas.0600919103.

Titz, B., Schlesner, M., and Uetz, P. (2004) What do we learn from high-throughput protein interaction data? *Expert Rev Proteomics* **1**: 111–121, URL http://dx.doi.org/10.1586/14789450.1.1.111.

Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., Hickey, E. K., Berg, D. E., Gocayne, J. D., Utterback, T. R., Peterson, J. D., Kelley, J. M., Cotton, M. D., Weidman, J. M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W. S., Borodovsky, M., Karp, P. D., Smith, H. O., Fraser, C. M., and Venter, J. C. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori. Nature* **388**: 539–547, URL http://dx.doi.org/10.1038/41483.

Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A. L., Sturch, K., Gilkes, N. R., Haynes, C. A., Warren, R. A., and Kilburn, D. G. (1998) Characterization and affinity applications of cellulose-binding domains. *J Chromatogr B Biomed Sci Appl* **715**: 283–296, URL http://www.ncbi.nlm.nih.gov/pubmed/9792516.

Triezenberg, S. J. (1995) Structure and function of transcriptional activation domains. *Curr Opin Genet Dev* **5**: 190–196, URL http://www.ncbi.nlm.nih.gov/pubmed/7613088.

Uetz, P., Dong, Y.-A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S. V., Roupelieva, M., Rose, D., Fossum, E., and Haas, J. (2006) Herpesviral protein networks and their interaction with the human proteome. *Science* **311**: 239–242, URL http://dx.doi.org/10.1126/science.1116804.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae. Nature* **403**: 623–627, URL http://dx.doi.org/10.1038/35001009.

Uetz, P. and Vollert, C. S. (2006) Protein-Protein Interactions., chap. 16. Springer, 1548–1552.

Usui, K., Katayama, S., Kanamori-Katayama, M., Ogawa, C., Kai, C., Okada, M., Kawai, J., Arakawa, T., Carninci, P., Itoh, M., Takio, K., Miyano, M., Kidoaki, S., Matsuda, T., Hayashizaki, Y., and Suzuki, H. (2005) Protein-protein interactions of the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Genome Biol* **6**: R98, URL http://dx.doi.org/10.1186/gb-2005-6-12-r98.

Valencia, A. and Pazos, F. (2002) Computational methods for the prediction of protein interactions. *Curr Opin Struct Biol* **12**: 368–373, URL http://www.ncbi.nlm.nih.gov/pubmed/12127457.

Vojtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993) Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* **74**: 205–214, URL http://www.ncbi.nlm.nih.gov/pubmed/8334704.

von Mering, C., Jensen, L. J., Kuhn, M., Chaffron, S., Doerks, T., Krüger, B., Snel, B., and Bork, P. (2007) STRING 7–recent developments in the integration and prediction of protein interactions. *Nucleic Acids Res* **35**: D358–D362, URL http://dx.doi.org/10.1093/nar/gkl825.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., and Bork, P. (2002) Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**: 399–403, URL http://dx.doi.org/10.1038/nature750.

Wadhams, G. H. and Armitage, J. P. (2004) Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol* **5**: 1024–1037, URL http://dx.doi.org/10.1038/nrm1524.

Wagner, G., Hartmann, R., and Oesterhelt, D. (1978) Potassium uniport and ATP synthesis in *Halobacterium halobium*. *Eur J Biochem* **89**: 169–179, URL http://www.ncbi.nlm.nih.gov/pubmed/29755.

Wang, H. and Matsumura, P. (1996) Characterization of the CheAS/CheZ complex: a specific interaction resulting in enhanced dephosphorylating activity on CheY-phosphate. *Mol Microbiol* **19**: 695–703, URL http://www.ncbi.nlm.nih.gov/pubmed/8820640.

Weidinger, E. (2007) Herstellung und Charakterisierung von Chemotaxiemutanten von *Halobacterium salinarum*. Diploma thesis, Fachhochschule München.

Welch, M., Chinardet, N., Mourey, L., Birck, C., and Samama, J. P. (1998) Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. *Nat Struct Biol* **5**: 25–29, URL http://www.ncbi.nlm.nih.gov/pubmed/9437425.

Welch, M., Oosawa, K., Aizawa, S., and Eisenbach, M. (1993) Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria. *Proc Natl Acad Sci U S A* **90**: 8787–8791, URL http://www.ncbi.nlm.nih.gov/pubmed/8415608.

Woese, C. R. and Fox, G. E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **74**: 5088–5090, URL http://www.ncbi.nlm.nih.gov/pubmed/270744.

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**: 4576–4579, URL http://www.ncbi.nlm.nih.gov/pubmed/2112744.

Woese, C. R., Magrum, L. J., and Fox, G. E. (1978) Archaebacteria. *J Mol Evol* **11**: 245–251, URL http://www.ncbi.nlm.nih.gov/pubmed/691075.

Wolanin, P. M., Thomason, P. A., and Stock, J. B. (2002) Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol* **3**: REVIEWS3013, URL http://www.ncbi.nlm.nih.gov/pubmed/12372152.

Wolfe, A. J. and Berg, H. C. (1989) Migration of bacteria in semisolid agar. *Proc Natl Acad Sci U S A* **86**: 6973–6977, URL http://www.ncbi.nlm.nih.gov/pubmed/2674941.

Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., and Ishihama, A. (2005) Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *J Biol Chem* **280**: 1448–1456, URL http://dx.doi.org/10.1074/jbc.M410104200.

Yan, J., Barak, R., Liarzi, O., Shainskaya, A., and Eisenbach, M. (2008) In vivo acetylation of CheY, a response regulator in chemotaxis of *Escherichia coli*. *J Mol Biol* **376**: 1260–1271, URL http://dx.doi.org/10.1016/j.jmb.2007.12.070.

Yonekura, K., Maki-Yonekura, S., and Namba, K. (2002) Growth mechanism of the bacterial flagellar filament. *Res Microbiol* **153**: 191–197, URL http://sfx.mpg.de/sfx_local?sid=Entrez:PubMed&id=pmid:12066889.

Zhang, W., Brooun, A., Mueller, M. M., and Alam, M. (1996) The primary structures of the Archaeon *Halobacterium salinarium* blue light receptor sensory rhodopsin II and its transducer, a methyl-accepting protein. *Proc Natl Acad Sci U S A* **93**: 8230–8235, URL http://www.ncbi.nlm.nih.gov/pubmed/8710852.

Zhao, R., Collins, E. J., Bourret, R. B., and Silversmith, R. E. (2002) Structure and catalytic mechanism of the *E. coli* chemotaxis phosphatase CheZ. *Nat Struct Biol* **9**: 570–575, URL http://dx.doi.org/10.1038/nsb816.

# Appendix

## List of abbreviations

| | |
|---|---|
| 3-AT | 3-aminotriazole |
| aa | amino acid(s) |
| ACN | acetonitrile |
| AD | activation domain |
| Ade | adenine |
| ADI | arginine deiminase |
| ADP | adenosine diphosphate |
| Amp | ampicillin |
| AP | affinity purification |
| ATP | adenosine triphosphate |
| bp | base pair(s) |
| BR | bacteriorhodopsin |
| CCW | counterclockwise |
| che | chemotaxis |
| CFE | cell-free extract buffer |
| CW | clockwise |
| dATP | deoxyadenosine triphosphate |
| DBD | DNA-binding domain |
| dCTP | deoxycytosine triphosphate |
| dGTP | deoxyguanosine triphosphate |
| DHFR | dihydrofolate reductase |
| DIG | digoxygenin |
| DMSO | dimethylsulfoxide |
| dNTPs | mixture dATP, dTTP, dCTP, dGTP |
| DTT | dithiothreitol |
| dTTP | deoxythymidine triphosphate |
| dUTP | deoxyuridine triphosphate |
| EDTA | ethylenediaminetetraacetate |
| e. g. | for example |
| ESI | electrospray ionisation |
| FA | formic acid |
| fla | flagella accessory |
| HAP | histidine-aspartate phosphorelay |
| His | histidine |
| HM | Halomedium |
| HPK | histidine protein kinase |
| HR | halorhodopsin |
| Htr | halobacterial transducer |
| i. e. | that is |
| IPTG | isopropyl-$\beta$-D-thiogalactopyranoside |

| | |
|---|---|
| kDa | kilodalton |
| LB | lysogeny broth (also known as Luria broth or Luria-Bertani broth) |
| LC | liquid chromatography |
| Leu | leucine |
| M | molar |
| MALDI | matrix-assisted laser desorption/ionization |
| MCP | methyl-acceepting chemotaxis protein |
| MS | mass spectrometry |
| $OD_{600}$ | optical density at 600 nm |
| PAGE | polyacrylamide gel electrophoresis |
| PDB | protein data bank |
| PEG | polyethylene glycol |
| $P_i$ | inorganic phosphate |
| PMF | proton-motive force or peptide mass fingerprinting |
| PMSF | phenylmethanesulphonylfluoride |
| PPI | protein-protein interaction |
| RB | resuspension buffer |
| rpm | revolutions per minute |
| RR | response regulator |
| RT | room temperature (20‑23 °C) |
| SC | synthetic complete (yeast medium) |
| SDS | sodium dodecyl sulfate |
| S-layer | surface layer |
| SM | synthetic medium |
| TCA | trichloroacetic acid (for protein precipitation) or tricarboxylic acid (Krebs cycle) |
| TOF | time-of-flight |
| Tris | tris(hydroxymethyl)aminomethane |
| Trp | tryptophane |
| vs. | versus |
| v/v | volume per volume |
| wt | wild type |
| w/v | weight per volume |
| X-gal | 5-bromo-4-chloro-3-indolyl-$\beta$-D-galacto pyranoside |
| YPD | yeast extract, peptone, dextrose (complex medium for yeast) |

# Publications

Schlesner M., Miller A., Streif S., Staudinger W., Müller J., Scheffer B., Siedler F. and Oesterhelt D. (2008) Identification of Archaea-specific Chemotaxis Proteins Which Interact with the Flagellar Apparatus, *submitted*

Schlesner M., Miller A., Besir H., Müller J., Scheffer B., Siedler F. and Oesterhelt D. (2008) The Protein Interaction Network of the *Halobacterium salinarum* Taxis Signal Transduction System, *in preparation*

Aivaliotis M., Oberwinkler T., Klee K., Damsma G., Schlesner M., Siedler F., Pfeiffer F., Lottspeich F. and Oesterhelt D. (2008) Arginine Fermentation – a Unique Feature to Gain Energy in an Extremophilic Archaeon, *in preparation*

Konstantinidis K., Schwarz C., Tebbe A., Schmidt A., Schlesner M., Klein C., Aivaliotis M., Reichelt P., Schwaiger R., Pfeiffer F., Siedler F., Lottspeich F. and Oesterhelt D. (2008) Quantitative analyses of the heat shock response in *Halobacterium salinarum* and the regulatory effects of a TFB protein, *in preparation*

Titz B., Schlesner M. and Uetz P. (2004) What do we learn from high-throughput protein interaction data? *Expert Rev Proteomics* 1, 111-121

# Danksagung

Mein Dank gilt...

Meinem Doktorvater Prof. Dr. Dieter Oesterhelt für die Möglichkeit, meine Doktorarbeit unter seiner Betreuung anfertigen zu können, für die fortwährende Unterstützung, und für das Vertrauen, das er mir entgegengebracht hat. Die große wissenschaftliche Freiheit und die vielen Möglichkeiten, die ich bei meiner Arbeit hier genossen habe und genieße, haben mir wesentlich mehr gebracht, als diese Arbeit widerzuspiegeln vermag.

Prof. Dr. Wolfgang Marwan für die Bereitschaft, das Zweitgutachten für die vorliegende Arbeit zu übernehmen.

Judith Müller, die einen erheblichen Teil der Experimente, die für diese Arbeit notwendig waren, und auch der Experimente, die dann doch nicht nötig waren, durchgeführt hat. Ich danke Judith auch dafür, dass sie mit unerschütterlicher Geduld Gele zerschnitten, Chromatogramme kontrolliert und Proben entsalzt hat.

Hüseyin Besir für die vielen Hilfestellungen, Tipps und richtungsweisenden Diskussionen am Anfang meiner Doktorarbeit, und für das Korrekturlesen, obwohl er Projekt, Abteilung und Institut schon lange verlassen hat.

Arthur Miller, der erst in seiner Diplomarbeit und später als wissenschaftlicher Mitarbeiter wesentlich zum Kapitel „Neue Chemotaxieproteine" beigetragen hat.

Bea Scheffer für die Durchführung von über tausen MS-Läufen, die Produktion einiger Terabyte an Daten, das Packen ungezählter Vor- und Hauptsäulen und den andauernden Kampf mit LC und QTOF. Michalis Aivaliotis für die Unterstützung bei der Massenspektrometrie, wenn Bea mal nicht da war, und für so einige hilfreiche Diskussionen. Birgit Bisle für die geduldige Hilfestellung bei meinen ersten Gehversuchen in Sachen Massenspektrometrie und Proteomik. Und Frank Siedler für die Bereitstellung der Massenspektrometrie-Infrastruktur und die andauernde Unterstützung bei den Messungen.

Friedhelm Pfeiffer für die Starthilfe bei der Programmierung innerhalb des HaloLex-Systems, und für zahlreiche Diskussionen und Erklärungen.

Valery Tarasov für viele Tipps und Tricks im Umgang mit Halos, und für viele wichtige und weniger wichtige Diskussionen.

Wilfried Staudinger und Stefan Streif für viele hilfreiche Diskussion über Taxie, Motilität, und die Tücken der Halos.

Yue Noeth-Liu und Mandy Handtke, die im Rahmen ihrer Master- bzw- Diplomarbeiten einige Aspekte des Interaktionsnetzwerks untersucht haben.

Vielen anderen Mitarbeitern der Abteilung Oesterhelt, die das Arbeiten hier so angenehm gemacht haben. Besonders hervorheben möchte ich Silvia Haslbeck, die uns immer mit sauberem Labormaterial versorgt, unsere Kolbenberge abtransportiert und literweise Medium bereitgestellt hat, und Frau Haack für die Hilfe in allen möglichen organisatorischen Belangen.

Meinen Eltern für die fortwährende Unterstützung während des Studiums und der Promotion, ohne die diese Arbeit niemals zustande gekommen wäre. Meinem Vater danke ich außerdem für das Korrekturlesen der schriftlichen Arbeit.

Meiner Frau Simone, die mir bei so vielem den Rücken freigehalten hat, Verständnis für so manchen langen Arbeitstag hatte, und mich nie vergessen lässt, dass es noch wichtigere Dinge gibt als Halobakterien, Massenspektrometrie und Proteininteraktionen.

Und natürlich meinem Sohn Lennart für die viele Zeit, in der er mich nicht hat arbeiten lassen.

# Lebenslauf

| | | |
|---|---|---|
| **Name** | | Schlesner |
| **Vorname** | | Matthias |
| **Anschrift** | | Gernstraße 16 |
| | | 82223 Eichenau |

| | | |
|---|---|---|
| **Geburtsdatum** | | 8. August 1978 |
| **Geburtsort** | | Kiel |
| **Familienstand** | | verheiratet, ein Kind |
| **Nationalität** | | deutsch |

| | | |
|---|---|---|
| **Schulbildung** | 1984 - 1988 | Grundschule Groß Vollstedt |
| | 1988 - 1997 | Gymnasium, Max-Planck-Schule in Kiel |
| | 05/1997 | Abitur (Note: 1,2) |

| | | |
|---|---|---|
| **Zivildienst** | 07/1997 - 07/1998 | Ergotherapie, Altenzentrum St. Nicolai, Kiel |

| | | |
|---|---|---|
| **Studium** | 10/1998 - 06/2003 | Humanbiologie, Philipps-Universität Marburg |
| | 09/2000 | Vordiplom (Note: sehr gut) |
| | 10/2001 - 09/2002 | Hauptfach Infektionsbiologie |
| | | Nebenfach Bioinformatik |
| | 10/2002 - 06/2003 | Diplomarbeit |
| | | "Erstellung cDNA-gestützter Systeme zur Untersuchung der Pathogenitätsdeterminanten von Influenzaviren" |
| | | Institut für Virologie, Marburg |
| | | Betreuer: Prof. Dr. Hans-Dieter Klenk |
| | 06/2003 | Abschluss der Diplomprüfung (Note 1,0) |

| | | |
|---|---|---|
| **Praktikum** | 07/2001 - 10/2001 | Bayer AG, Forschungszentrum Wuppertal |
| | | „Identifikation und Untersuchung neuer antibakterieller Wirkstoffe" |

| | | |
|---|---|---|
| **Promotion** | 10/2003 - 03/2004 | Doktorand |
| | | Arbeitsgruppe Peter Uetz |
| | | Institut für Toxikologie und Genetik |
| | | Forschungszentrum Karlsruhe |
| | | „Erstellung einer genomweiten Protein-Protein-Interaktionskarte von *Treponema pallidum*" |
| | seit 04/2004 | Doktorand |
| | | Abteilung Membranbiochemie |
| | | MPI für Biochemie, Martinsried |
| | | „The *Halobacterium salinarum* Taxis Signal Transduction Network: a Protein-Protein Interaction Study" |
| | | Betreuer: Prof. Dr. Dieter Oesterhelt |

Martinsried, 07.10.2008