Gene expression noise and robustness in the *Escherichia coli* chemotaxis pathway

Dissertation submitted to the

Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

MSc. Linda Elisabeth Løvdok

born in: Drammen, Norway

December 2008

From the group of
Victor Sourjik
at the
ZMBH

First referee:	Prof. Dr. Bernd Bukau
Second referee:	Dr. Victor Sourjik
Oral-examination:	February 2009

Some of the results presented in this work were reported in published papers or during scientific meetings:

Publications

Markus Kollmann, Linda Løvdok, Kilian Bartholome, Jens Timmer, Victor Sourjik
"Design principles of a bacterial signalling network"
Nature, **438** (2005) 504-507

Linda Løvdok, Markus Kollmann, Victor Sourjik
"Co-expression of signaling proteins improves robustness of the bacterial chemotaxis pathway"
Journal of Biotechnology **129** (2007) 173-180

N. Vladimirov, L. Løvdok, D. Lebiedz, V. Sourjik
"Rapid Cell: a tool to investigate dependence of bacterial chemotaxis on attractant gradients and adaptation time"
PLoS Comput Biol., in press

L. Løvdok, N. Vladimirov, A. Müller, K. Bentele, D. Lebiedz, M. Kollmann, V. Sourjik
"Role of translational coupling in robustness of bacterial chemotaxis pathway"
submitted

Oral Presentations (presenting author underlined)

<u>L. Løvdok</u>, D. Kentner, S. Thiem, V. Sourjik
"Signaling and robustness in the *E. coli* chemotaxis pathway"
BLAST IX Meeting
2007.01.14 – 19., Laughlin, Nevada, USA

Poster Presentations

L. Løvdok, M. Kollmann, V.Sourjik
"Gene expression noise in the *E. coli* chemotaxis pathway"
International Summer School on Molecular Imaging
2006.09.04 – 08, Heidelberg, Germany

L. Løvdok, M. Kollmann, V.Sourjik

"Gene expression noise in the *E. coli* chemotaxis pathway"
BLAST IX Meeting
2007.01.14 – 19, Laughlin, Nevada, USA

L. Løvdok, M. Kollmann, V.Sourjik
"Gene expression noise and robustness in the *E. coli* chemotaxis pathway"
2nd FEBS Advanced lecture course on Systems Biology from Molecules to Life
2007.03.10 – 16, Gosau, Austria

L. Løvdok, N. Vladimirov, K. Bentele, M. Kollmann, V.Sourjik
"Gene expression noise in the *E. coli* chemotaxis pathway"
Sensory Transduction in Microorganisms
2008.01.13 – 18, Ventura, LA, USA

Gene expression noise and robustness in the *Escherichia coli* chemotaxis pathway

Dissertation

by

MSc. Linda Elisabeth Løvdok

Acknowledgements

First and foremost I must express my thanks to Prof. Dr. Bernd Bukau for kindly supervising this thesis.

Also, my warmest gratitude goes to Dr. Victor Sourjik for his guidance, mentoring and support during the past five years.

This work could not have been written without the fruitful collaboration with Dr. Markus Kollmann and Kajetan Bentele at the Institute for Theoretical Biology, Humboldt University Berlin, and Nikita Vladimirov at the Interdisciplinary Center for Scientific Computing, University of Heidelberg. Thanks to Julia Rausenberger at the University of Freiburg for the helpful discussions.

For help with cloning, the microscope, computer problems, data analysis and creating a wonderful working environment, cakes and shashlik, I would like to thank all the former and present colleagues in the lab; Anette Müller, Louisa Liberman, Sebastian Thiem, David Kentner, Gabriele Schwarz, Sonja Schulmeister, Silke Neumann, Hui Li, Olga Oleksiuk, Maren Emmerich, Susanne Wölfelschneider, Erik Sommer, Mohit Kumar, Tanja Erhard, Yiling Yang, Gabriele Malengo, Vladimir Jakovljevic, Barbara di Ventura, Francesca Chiovaro and Stefan Viel.

Special thanks go to Marianne Uteng, Katharina Paulsen , Aksana Labokha, Steffen Preissler and the bikers Jakob Tschäpe, Marco Pfeifer, Tomas Grübl and Florentine Wahl for help with flat tires, for their patience uphill and the fun outside of the lab.

Og så vil jeg takke jentene hjemme som ikke har glemt meg selv om jeg har vært borte så lenge.

Jeg vil takke min søster, Nina, og mine fantastiske foreldre for at de alltid har vært der for meg med oppmuntringer og støtte, og for at de har vært så tålmodige og forståelsesfulle gjennom mine lange studier.

Zum Schluss möchte ich mich bei meinem Liebsten, Mark, ganz herzlich bedanken. Ohne deine Unterstützung wäre ich ganz verloren. Ich habe so ein Glück gehabt, dass ich dich kennengelernt habe. Danke, dass du immer für mich da bist.

Contents

Zusammenfassung

Genetisch identische Zellen und Organismen zeigen eine große Vielfältigkeit, obwohl sie der selben Umwelt ausgesetzt worden sind. In Populationen genetisch identischer Zellen konnten große interzelluläre Variationen in der Konzentration von Proteinen festgestellt werden. Dieses so genannte Rauschen der Genexpression lässt sich auf stochastische Effekte bei der Transkription und Translation, sowie der ungleichen Verteilung von Proteinen während der Zellteilung, zurückführen. Insbesondere bei Proteinen, die an Signalsystemen beteiligt sind, können derartige Fluktuationen im Expressionsniveau zu Fehlern in der Signalprozessierung führen, was dramatische Auswirkungen auf die Zelle haben kann. Ziel dieser Arbeit war es, Eigenschaften und Auswirkungen des Rauschens der Genexpression besser zu verstehen. Beispielhaft wurde dafür das Chemotaxis-System von *E. coli* untersucht, das sich aufgrund seiner einfachen Topologie und guten Charakterisierung ideal als Modellsystem eignet. Um den Ursprung von Fluktuationen in der Genexpression zu identifizieren, erfolgte eine Quantifizierung des nativen Expressionsniveaus der jeweiligen Gene in den drei bekannten Hierarchieklassen der Flagellen- und Chemotaxisgene in einzelnen Zellen. Dabei konnte eine starke Korrelation der Expression der Flagellen- sowie Chemotaxisproteinen beobachtet werden, was darauf schliessen lässt, dass das Rauschen der Genexpression durch gemeinsame Faktoren aller Chemotaxisgene dominiert wird. Interessanterweise wurde eine Korrelation zwischen voneinander unabhängigen Genen beobachtet, wodurch globalen Faktoren der größte Anteil an den Fluktuationen zugeschrieben werden konnte. Verfolgt man eine Zelllinie über mehrere Generationen, so bleibt eine Korrelation der Proteinniveaus zwischen Mutter- und Tochterzelle für etwa eine Generation bestehen. Obwohl die absolute Proteinzahl wärend des Zellzyklus großen Schwankungen unterliegt, stellen sich die jeweiligen Konzentrationen in Abhängigkeit vom relativen Zellvolumen als stabiler dar.

Da das Rauschen der Genexpression eine Generation überdauert, wird erwartet, dass das Chemotaxissystem Mechanismen für die Robustheit des Signalwegs

bereit hält, die die auftretenden Fluktuationen kompensieren können. In der Tat ist das Signalnetzwerk, trotz seiner Sensitivität gegenüber einzelner Proteinniveaus, sehr robust gegenüber Co-Variationen aller Chemotaxisproteine, während augenscheinlich die Komplexizität des Netzwerks und der Aufwand für die Proteinexpression minimiert werden.
Erstaunlicherweise konnte gezeigt werden, dass die Robustheit des Signalwegs gegenüber unkorrelierten Variationen in den Proteinniveaus durch eine selektive paarweise Kopplung von individuellen Chemotaxis Genen erhöht werden kann. Ausserdem sind die Reihenfolge und Organisation dieser Gene im Chromosom, welche unter evolutiver Selektion stehen, eine der Besten, um das Rauschen auszugleichen.
Schlussendlich konnten zusätzliche Mechanismen, die zur Robustheit des Signalwegs beitragen, beschrieben werden. Die hier durchgeführten Experimente deuten darauf hin, dass die Aktivierung von CheB durch Phosphorylierung und die Konkurrenz zwischen CheY und CheB um die CheA P2 Bindedomäne, welche selbst nicht essentiell sind für die Chemotaxis, sich primär zur Kompensation der nicht-genetischen Individualität entwickelt haben.

Summary

Genetically identical cells and organisms exhibit remarkable diversity even when they have identical histories of environmental exposure. We have observed a great cell-to-cell variation in levels of proteins in an isogenic population of *E. coli* cells. Such gene expression noise arises from stochasticity of transcription and translation and from unequal partitioning of proteins by cell division. Variations in protein levels may lead to errors in a signal transduction system that can have a detrimental effect on the output of the system. Therefore it is important to study the characteristics and the effects of gene expression noise. The chemotaxis signal transduction pathway is well suited for noise studies due to its low complexity and the fact that it is well characterized.

One objective of this work was to identify the origin and time scale of gene expression noise in the chemotaxis pathway. To do so, we quantified native single cell expression levels of genes in all three classes in the hierarchy of flagella and chemotaxis genes. We observed strong correlation in expression levels of flagella and chemotaxis proteins, which suggests that gene expression noise is dominated by factors common to all chemotaxis genes. Interestingly, correlation between independent genes was also observed, revealing that global factors make the largest contribution to the observed fluctuations. It takes approximately one generation until the correlation in protein levels between mother and daughter cells is lost. Although, total protein levels vary greatly throughout the cell cycle, protein concentrations remain more stable.

Thus, noise persists on a generation time scale and consequently the pathway is expected to exhibit robustness mechanisms compensating for such fluctuations. Indeed, the signalling network, although sensitive to individual protein levels, was observed to be robust against co-variation of chemotaxis proteins while apparently minimizing network complexity and cost of protein expression.

Surprisingly, we observed that robustness of the pathway against the uncorrelated variations in protein levels can be enhanced by a selective pairwise coupling of individual chemotaxis genes on one mRNA, with the order of genes

in *E. coli*, which is subject to evolutionary selection, ranking among the best in terms of noise compensation.

Finally, additional topological features contributing to pathway robustness were discovered. Our experiments suggest that the activation of CheB by phosphorylation and competition between CheY and CheB for the CheA P2 binding domain, which are not essential for chemotaxis itself, have evolved primarily for compensation of non-genetic individuality.

1 Introduction

Networks of interacting biomolecules carry out many important and essential functions in living cells. The understanding of the underlying principles is poor, although quantitative analysis of a few simple systems, as well as construction of synthetic networks, has been performed to increase the knowledge of cellular systems [14, 160]. The chemotactic behaviour of *E.coli* bacteria has been studied carefully for several decades, which makes it one of the best understood physiological processes known. The pathway is thus often used as a paradigm for intracellular signal processing, and is a good model to study effects of noise and robustness on signal transduction.

1.1 Noise in gene expression

Non-genetic individuality, or noise, is defined as the diversity of phenotypes in a population of genetically identical cells. Observations of such diversity have been made in both prokaryotes and eukaryotes. Such individuality was first discovered to be a feature of the chemotactic response of coliform bacteria in the 1970's, with swimming behaviour and adaptation time varying from cell to cell in a population [83, 150].

Stochasticity and variations in networks of biomolecules can be exploited by some organisms to introduce diversity into a population. The variability can provide a mechanism for surviving exposure to stress, to repopulate the environment after stress or to migrate to and colonize new environments. Noise in genetic and metabolic networks can be detrimental to fitness of the cells by perturbing developmental pathways, disrupting cell cycle control or by forcing metabolites away from their optimal levels. The influence of noise has been recognized for decades, and numerous mathematical models describing noise in gene expression have been proposed. It is only recently however, that quantitative measurements have become feasible, and few experiments have been done

on the subject to verify these models [40, 124].
A better understanding of noise in gene expression could lead to an improved understanding of the biological function and evolutionary origin of naturally ocurring networks and engineering of new cellular circuits both in prokaryotes and eukaryotes [160].

1.1.1 Sources of noise

Many genes, mRNAs and proteins are present in a cell in low copy numbers. This leads to random fluctuations in the genetic networks. Hence, cell-to-cell variations observed in isogenic populations (as shown in figure 1.1), may be caused by stochastic effects in gene expression [40, 41]. Another source of noise

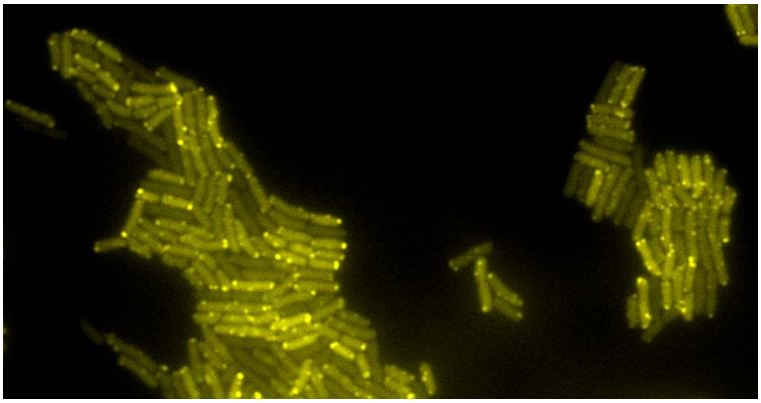

Figure 1.1: Variation in protein levels. Wild type *E.coli* cells expressing CheY as a fusion to YFP from the native promoter on the chromosome. The cells display a great cell-to-cell variation in fluorescence levels despite their identical genotype.

is spontaneous fluctuations in amounts or activities of regulatory proteins and polymerases that cause fluctuations in the output of a gene [107], hence transcription and translation are two main contributors to expression noise [124]. Moreover, previous studies in bacteria have shown that proteins are produced from an activated promoter in short bursts of variable numbers of proteins, and that these bursts occur at random time intervals [52].
For practical reasons, sources of noise are often divided into two categories: intrinsic noise sources create differences between two genes or proteins within one cell or within one specific pathway, organelle or sub-compartment of the cell. This means that stochastic events during the process of gene expression, from promoter binding via mRNA translation to protein degradation will manifest

as intrinsic noise. Differences between cells, eg. in local environment or in the concentration or activity of a factor that affects gene expression generally, or two genes or proteins equally in a cell or in a system, will result in extrinsic noise. Extrinsic noise can be divided into two subcategories: global noise which affects expression of all genes, and gene- or pathway specific extrinsic noise, such as fluctuations in the levels of a specific transcription factor or stochastic events in a specific signal transduction pathway.
Prokaryotes and eukaryotes are thought to have different sources of noise, however which sources are dominant has not been agreed on. Elowitz et al. showed, that the relative contributions of extrinsic and intrinsic variations to the total noise vary with expression level [41]. Experiments in *Saccharomyces cerevisiae* suggest that noise in eukaryotes is dominated by extrinsic factors such as global noise [22, 121, 128]. In prokaryotes transcription is suggested to be the dominant source of noise [68, 70, 106, 160]. Rosenfeld et al. quantified gene expression in *E. coli* cells over time, and found that extrinsic noise was the primary source of variability in gene expression, similar to the observation in budding yeast. The authors calculated autocorrelation times for noise, or the time scale over which the protein production rate fluctuates in any given cell. For intrinsic noise the autocorrelation time was $\geq$ 10 min, consistent with the hypothesis that rapid fluctuations in mRNA numbers are the source of intrinsic noise. Autocorrelation time for global noise factors in protein production rate was $\sim$ 40 min, similar to the observed cell cycle length, which suggested that whatever factors result in global noise persist on average for about one cell cycle [131].

1.1.2 Consequences of gene expression noise

Fast and slow fluctuations can affect the operation of genetic networks in cells in different ways. If small changes in protein levels persist long enough, they can have dramatic effects on fitness. On the other hand, large fluctuations may have no effect if they occur too frequently to affect a cellular process. The fact that the time scale of intrinsic noise is much shorter than that of extrinsic noise suggests that extrinsic noise has larger effect on the phenotype, at least in *E. coli* [131].
Fluctuations in protein levels may provide an advantage or a disadvantage for a cell or organism (fig. 1.2 A). Some examples of the consequences of heterogeneity in cells underline the importance of understanding gene expression noise. In eu-

karyotic cells, intrinsic noise can produce fluctuations in the relative expression of two alleles of the same gene in a heterozygote, potentially resulting in cells that express no allele, either individual allele, or both alleles. If the two alleles are functionally divergent, the population of cells could acquire heterogeneity (fig. 1.2 B). Such fluctuations may contribute to the still debated phenomenon of hybrid vigor (outbreeding enhancement). Alternatively, intrinsic noise in the case of haploinsufficiency may result in increased levels of noise or complete loss of function in a subset of cells. Such a mechanism has been proposed in the case of the human tumor suppressor gene NF1 [66] and prostate neoplasia formation in the mouse [104]. Another example of noise consequences is observed in the olfactory neurons. Each murine olfactory neuron expresses a single allele of one odorant receptor gene out of a choice of $\sim$ 1500 odorant receptor genes. A functional odorant receptor is required to prevent expression of other odorant receptors, which suggests that receptor choice occurs through a brief period of intrinsic noise with stochastic activation of a single promoter followed by a feedback with inhibitory signalling to the inactive odorant receptor promoters. The resulting heterogeneous population of olfactory neurons enables sensitive differentiation of odorant molecules; the stochastic nature of gene expression may create a functional sense of smell [31, 140].

Noise in gene expression in the context of positive feedback may be sufficient to create switching between two stable states such as the switch-like behaviour in fate decision seen in the λ phage [7, 94, 106]. The use of stochasticity to populate multiple steady states may play an important role in differentiation in multicellular organisms (fig. 1.2 C) or in survival in fluctuating environments for unicellular organisms.

Genes are organized into regulatory circuits where the expression of one gene can influence the expression of another. A consequence of this organization is that noise in the expression of one gene may propagate to affect noise in the expression of a downstream gene (fig. 1.2 D). Recent work in *E. coli* has demonstrated that a synthetic cascade of three transcription factors produces more noise in output than a linear cascade of two transcription factors or than one transcription factor alone [57]. Intrinsic noise in the expression of a transcription factor causes extrinsic noise in a downstream target gene. Additionally, global noise affecting expression of the transcription factor propagates to the downstream target. Thus, a low numbers of molecules it is not required for large fluctuations, because noise could be transmitted from upstream genes. Even in

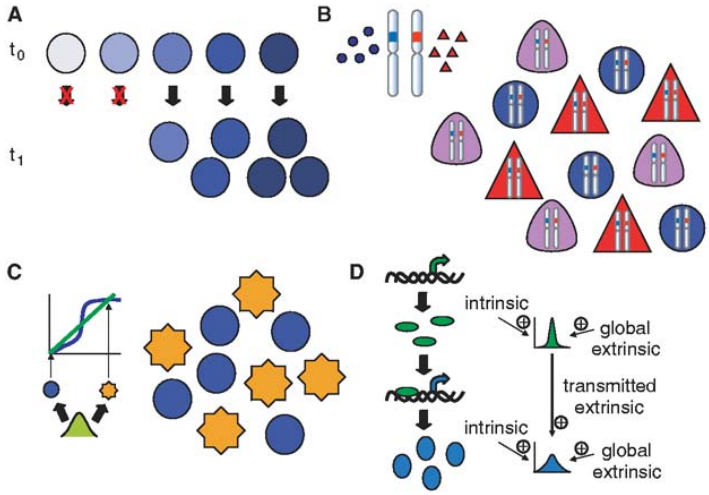

Figure 1.2: Consequences of noise. (A) Small differences in gene product abundance affect reproductive fitness. (B) In a heterozygous diploid population, cells display the phenotypes associated with each homozygote as well as the heterozygote. (C) Noise allows simultaneous achievement of multiple steady-state phenotypes in a population. (D) Noise can be transmitted from one gene, in this case a transcription factor, to a downstream target. The intrinsic and global extrinsic noise of the transcription factor can cause extrinsic noise in the downstream gene. Figure taken from [129].

a network where all components have low intrinsic noise, fluctuations can be substantial and the distributions of expression levels depend on the interactions between genes [125].

1.1.3 Robustness to perturbations

Various mechanisms have been suggested to reduce noise in biological networks. Becskei and Serrano demonstrated reduction of noise by means of negative feedback in a simple model in which a transcription factor negatively regulates its own synthesis [14]. For signal transduction systems, protein cascades have been theoretically shown to reduce the propagation of noise arising from a fluctuating cascade input [161]. Furthermore, it has been argued that the genetic networks underlying circadian rhythms have a structure that hinders biochemical stochasticity from disrupting the period of the circadian clock [10, 166]. Additionally, dimerization of transcription factors [29] and DNA looping [167] may also act to attenuate noise. The magnitude of the (intrinsic) noise for a particular protein is expected to be mostly set by levels of mRNA, rather than those of protein

itself [124, 158], and Swain demonstrated *in silico* that negative translational feedback has a greater efficiency at reducing stochasticity than negative transcriptional feedback [157]. Importantly, although noise reduction is a way to secure the output of a pathway, it should not be confused with robustness. Robustness does not lead to a reduction in noise, it reduces the consequence of noise and is thus the property that allows a system to maintain its functions despite external and internal perturbations [72]. A common misunderstanding is that robustness means staying unchanged regardless of stimuli or mutations. In other words, that the structure and components of the system, and therefore the mode of operation, is unaffected. In fact, robustness is the maintenance of specific functionalities of the system against perturbations, and it often requires the system to change its mode of operation in a flexible way. Thus, robustness allows changes in the structure and components of the system owing to perturbations, but specific functions are maintained.

Robustness against natural variation in protein levels in signal transduction systems primarily relies on the balance of opposing enzymatic activities, such as kinase and phosphatase. Translational coupling has been previously described in the tryptophan and galactose operons in *E. coli* [13, 120, 135], and could be an important mechanism to keep the ratio between proteins constant and thereby reducing negative effects of uncorrelated noise.

However, balance in enzyme activities is not always sufficient to cope with noise and several additional robustness mechanisms are known. Theoretical models [68, 70, 106, 160] and experimental evidence from *B. subtilis* [121] suggest that frequent transcription followed by inefficient translation results in lower intrinsic noise in protein levels than infrequent transcription followed by efficient translation. Here, the control of noise comes at the energetic cost of producing few proteins from numerous mRNA or the cost of repeated rounds of translation from few mRNA molecules. It has been noted that key regulatory proteins in *E. coli* display low translation rates, which could lower noise in protein levels [121]. Similarly, yeast genes that are essential or encode proteins involved in multi-subunit complexes tend to have higher rates of transcription and lower rates of translation [46].

In addition, a wide range of regulatory networks attain a robust response through control by positive and negative feedbacks as observed in cell cycle, the circadian clock and chemotaxis [3, 24, 115]. For instance, bacterial chemotaxis uses negative feedback to attain perfect adaptation which allows chemo-

taxis to occur in response to a wide range of stimuli [3, 9, 176]. Furthermore, in *Drosophila melanogaster* a bistability in positive feedback from gene products on their own expression is responsible for the robust segment polarity pattern formation [59, 168, 169].

Nevertheless, robust systems might be fragile or face a performance setback as a trade-off [72, 74]. The trade-offs between robustness, fragility and performance can be observed in biological systems at different levels. Chemotactic bacteria, for example, should be able to respond stronger without negative feedback, but this would sacrifice their precision in following a chemical gradient, hence the use of negative feedback improves the bacteria's ability to follow a chemical gradient, at the cost of reduced sensitivity.

Modularity is another mechanism of robustness for containing perturbation or damage locally to minimize the effects on the whole system. A cell in a multicellular organism can be seen as a module, but also a compartment, pathway or organelle within a cell are modules that could maintain the output of noise locally [108].

Moreover, numerous other cellular control mechanisms may exist to enable the switch between globally noisy or globally "quiet" states of gene expression. For instance, the Hsp90 chaperone activity is hypothesized to reduce the effect of stochastic molecular events that might otherwise result in developmental variability, due to the observation that reduction of heat-shock protein 90 (Hsp90) chaperone activity in *Arabidopsis thaliana* increased morphological diversity in inbred lines, in addition to revealing otherwise silent genetic variation among different lines [127].

Evolution often selects traits that might enhance robustness of the organism [33]. Hence, signalling systems have presumably employed fine tuning of reaction rates to evolve towards network architectures that are highly robust against variation in protein concentrations. Information processing cannot rely on network architecture alone and therefore cannot be made perfectly robust against both changes in reaction rates and changes in protein concentrations without loosing its function. However, kinetic constants can be expected to show in general small cell-to-cell variation. One reason for such robustness is that stochasticity of biochemical reactions have significantly smaller effect on the population-wide signalling precision than gene expression noise. Furthermore, kinetic constants are affected mostly in concert, e.g. by temperature or Mg^{2+} ions. As rate constants determine also protein synthesis and degradation, changes in expression

level of the pathway proteins can be used to partially compensate for changes in reaction rates.
Regardless of which noise reduction or robustness mechanisms are predominant in an organism, wasteful phenotypic variability in a population is supposedly suppressed when the population is well adapted to its environment [170]. However, if environmental conditions shift, phenotypic noise becomes advantageous because a noisy population will produce some members that are better adapted to the new environment. Recent work supports the idea that it is advantageous to increase variability in times of stress and decrease variability when organisms are well adapted to the environment [126]. Regulation of global noise factors could provide a molecular basis for such evolutionary flexibility.

1.2 Chemotaxis

1.2.1 The biased random walk

Chemotaxis is defined in the field of microbiology as a biased movement towards or away from a chemical stimulus via a metabolism-independent sensory transduction mechanism. Bacteria sense environmental signals via chemoreceptors located in the cytoplasmic membrane. In a spatial gradient of attractants cells that are swimming up the gradient sense a temporal gradient as changes in receptor occupancy and this decreases the probability of tumbling. Runs and tumbles occur in an alternating sequence where each run is constituting a step in a three-dimensional random walk (fig. 1.3). *E. coli* is propelled by about six flagellar filaments attached to the cell surface at random points (peritrichously). Each filament is powered by a rotary motor at its base. The motor is powered by the proton motive force, operating at close to 100 % efficiency, and it spins the flagellum with up to several hundred rotations per second [16]. When the motors turn counterclockwise (CCW), the filaments work together in a bundle to drive the cell forward in a run. Clockwise (CW) rotation leads to tumbling when the bundle flies apart as shown in fig. 1.3. Without attractant or repellent, the tumbling events occur about once per second [17]. Since the runs up a gradient are extended, a bias is imposed on the random walk that carries the cell in a favourable direction [5, 60, 100]. The bacterium is able to sense differences in concentrations over a time span long enough to overcome local fluctuations of

attractant. However, if exposure to constantly high concentrations of attractant is long enough, the cell starts to adapt to these concentration and returns to its original run and tumble behaviour [101].

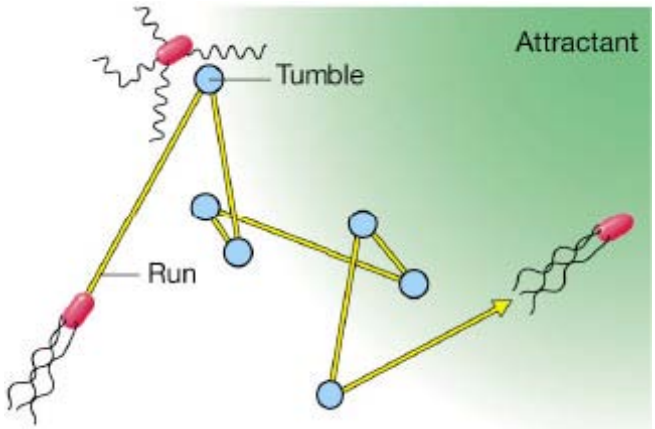

Figure 1.3: Chemotactic swimming behaviour of *E.coli*. The cells swim in straight runs interrupted by short tumbles in which the cell changes direction. Increasing attractant concentration prologs the run length so that the cell continues to move towards the attractant source. Repellents are avoided by increasing the tumbling frequency. Figure taken from [103].

1.2.2 The chemotaxis signalling pathway

In *Escherichia coli* the chemotaxis pathway is well understood and has been characterized on a genetic, structural and biochemical level. The structure of the chemotaxis pathway differs from the orthodox two component system scheme. In the prototypical two-component system (fig. 1.4), a sensor histidine kinase catalyzes its autophosphorylation and then subsequently transfers the phosphoryl group to a response regulator, which can then effect changes in cellular physiology, often by regulating gene expression. Many histidine kinases are bi-functional and also dephosphorylate their cognate response regulator. A histidine kinase autophosphorylates on a conserved histidine residue with subsequent transfer of the phosphoryl group to a cognate response regulator. Input domains on histidine kinases vary widely and typically do not share substantial homology to one another. The catalytic and ATPase (CA) domain of the histidine kinase is responsible for binding ATP and catalyzing autophosphorylation of a conserved histidine found within the dimerization- and histidine phosphotransferase (DHp) domain. The DHp domain mediates homodimerization and serves

as the phosphodonor for a cognate response regulator. Response regulators typically contain two domains, a receiver domain and an output domain. Receiver domains contain the phosphoacceptor aspartate and several other highly conserved amino acids that catalyze phosphotransfer from a histidine kinase. Output domains, which are activated by phosphorylation of the receiver domain, are varied, but often involved in binding DNA. Phosphorelays are a common variant of the two-component signalling paradigm. Receipt of a stimulus activates autophosphorylation of a hybrid histidine kinase. The phosphoryl group is then passed intramolecularly to a C-terminal receiver domain, similar to that found in response regulators. A histidine phosphotransferase (HPT) then shuttles the phosphoryl group from the hybrid kinase to a soluble response regulator containing an output domain [81]. The entire chemotaxis signal transduction pathway includes five attractant-specific receptors (Tsr, Tar, Trg, Tap, and Aer), six cytoplasmic chemotaxis proteins (CheA, CheW, CheR, CheB, CheY, and CheZ), and three proteins comprising a switch complex at the cytoplasmic face of the flagellar motor (FliM, FliN, and FliG) [146]. The sensor in the pathway is a receptor in the cytoplasmic membrane and the separate cytoplasmic histidine kinase is associated with the receptor. The response regulator interacts directly with the flagella motor. An overview of the chemotaxis signal transduction pathway is shown in fig. 1.5.

Information about changes in chemoeffector concentration is transmitted from specific chemoreceptors, methyl accepting chemotaxis proteins (MCPs), at the cytoplasmic membrane to the flagellar motors via the phosphorylated response regulator CheY. CheW is a SH3-like coupling protein, bound covalently to the receptors together with the histidine protein kinase CheA. When an attractant binds a receptor dimer, a conformational change involving the C-terminal domain of the MCP is passed on to CheA, resulting in reduced levels of its autophosphorylation rate and activity and thereby reduced levels of phospho-CheY (CheY-P). This prolongs the time the flagellar motors rotate CCW and the cell is smooth swimming. Binding of a repellent on the other hand, results in strong activation of the kinase, hence an increase in autophosphorylation activity. CheA transfers a phosphoryl group to CheY. Consequently, CheY-P levels increase leading to the formation of multiple CheY-P-FliM complexes which stabilize clockwise rotational rate and leads to extended bacterial tumbling. Dephosphorylation of CheY is greatly enhanced by the phosphatase CheZ [177].

CheR, a methyltransferase, and CheB, a methylesterase, mediate adaptation

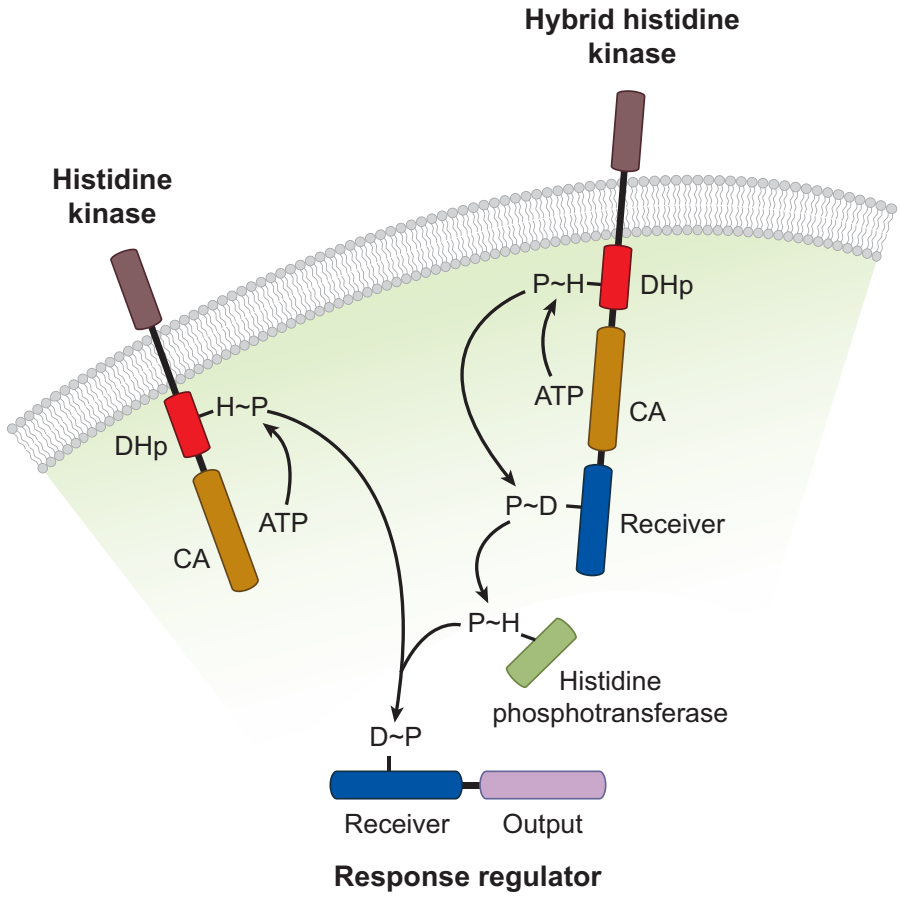

Figure 1.4: Schematic overview of a two-component signal transduction paradigm (left) and a phosphorelay (right), and the domain structure of each component. CA domain= catalytic and ATPase domain. DHp domain= dimerization and histidine phosphotransferase domain. HPT= histidine phosphotransferase. Figure taken from [81].

to a constant attractant concentration by adjusting the methylation level of receptors [109, 138, 154, 175]. An enhanced methylation level leads to more active receptors, and removal of methyl groups decreases the receptor activity. CheR is constitutively active, whereas CheB is a response regulator that also becomes activated through phosphorylation from CheA [6, 38] and thereby provides a negative feedback mechanism. Hence, binding of repellent, which leads to a kinase stimulation, will induce an opposite regulation by activating the methylesterase to remove methyl groups and thereby counteract the initial signal, until kinase activity returns to the pre-stimulus state. Cleverly, the adaptation system serves as a memory of past conditions because methylation works slower than the phosphorylation response, allowing temporal measurements of

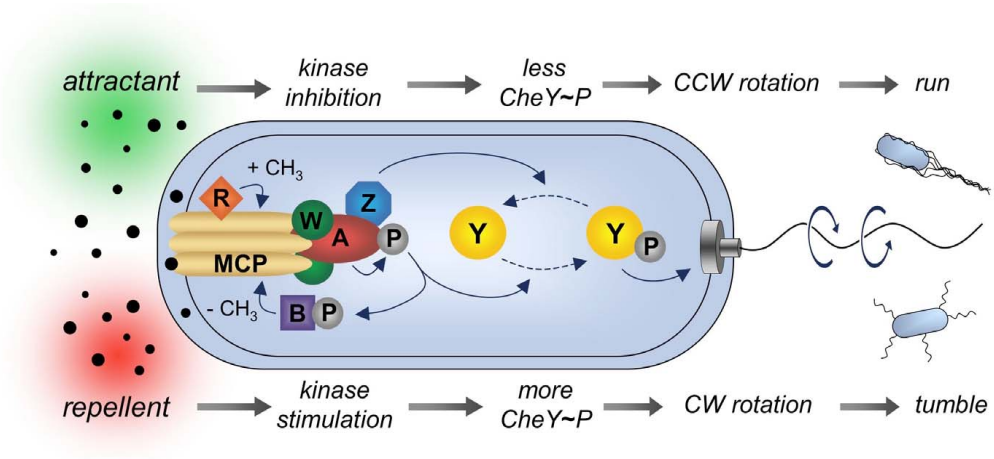

Figure 1.5: Schematic of the chemotaxis signalling pathway in *E.coli*. Figure kindly provided by D. Kentner.

concentration changes. The relatively simple chemotaxis system is nonetheless extremely sensitive, responding to ligand concentrations down to the nanomolar range, which corresponds to only a few molecules per cell, and up to 10^{-3} M, giving dynamic range up to five orders of magnitude. A change in receptor ligand occupancy of 0.2 % can cause a change in the motor bias of 23 % displaying a signal amplification, or gain of about 110 [137].

1.2.3 The core of the chemotaxis system: The chemoreceptors, CheW and CheA

The central processing unit in the chemotaxis machinery is the ternary complex formed by receptors, the assisting protein CheW and the histidine kinase CheA [49].

The chemoreceptors

The transmembrane receptors are used by *E.coli* to monitor the chemical composition of the environment. In *E. coli*, the high-abundance chemoreceptors Tsr (serine) and Tar (aspartate) are present in several thousand copies per cell, whereas the three low-abundance receptors, Trg, Tap, and Aer, are present in a few hundred copies [88].
Chemoreceptors are homodimers of $\sim$ 60 kDa subunits, which remain stable

both in absence and presence of ligands. They consist of an N-terminal periplasmic binding domain ($\sim$ 150 amino acids), a transmembrane region ($\sim$ 100 amino acids), a linker region (also called HAMP domain) and a C-terminal cytoplasmic signalling and adaptation domain ($\sim$ 300 amino acids). The periplasmic domain, responsible for the receptor specificity, consists of eight helices arranged in two symmetric four-helix bundles, one per subunit. Two helices from each subunit span the inner membrane bi-layer, where they form a transmembrane four-helix bundle. The cytoplasmic domain is a distinct four-helix bundle, formed by association of two helical hairpins, one per subunit. One helix in each subunit extends the entire length of the structure, connecting the ligand binding site at the membrane-distal end of the periplasmic domain with the kinase interaction region at the opposite end of the receptor.
Adaptation enzymes CheR and CheB are associated with major receptors at a C-terminal NWETF pentapeptide sequence, which is absent in minor receptors [8, 143]. Receptors have four conserved methylation sites, distributed in the membrane-proximal region of the cytoplasmic domain, which are subject to methylation by CheR and demethylation by CheB. Receptors are synthesized with two glutamate (E) and two glutamine (Q) residues, the latter being functionally similar to methylated glutamates. The two glutamines are promptly deamidated to glutamates by CheB, and can then be targeted by CheR.
The information transfer upon ligand binding is thought to take place by conformational changes of the receptor dimer in a pistonlike movement of the transmembrane helix α - 4 [44, 113] and by changes in supercoiling of the cytoplasmic four-helix bundle [173]. Methyl - glutamate or glutamine is neutral, and glutamate negatively charged, which lead to the assumption that methylation/demethylation also cause conformational changes to regulate receptor activity. Indeed, experiments showed that *in vivo* rates of disulphide formation between cysteine pairs spanning a signalling helix interface changed as a function of adaptational modification. Strikingly, those changes were opposite those caused by ligand occupancy. This suggests that adaptational modification resets the receptor complex to its null state by reversal of the conformational change generated by ligand binding [78, 151]. Figure 1.6 shows the structure of a dimeric bacterial receptor.

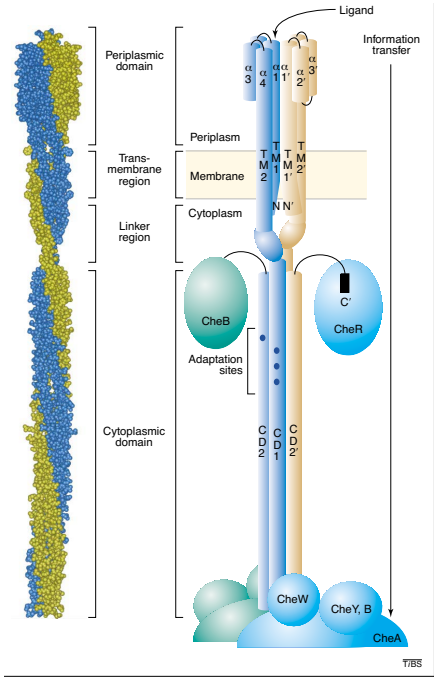

Figure 1.6: Structure of a dimeric bacterial chemoreceptor. (Left) Atomic structural model generated by combining crystal structures of the periplasmic and cytoplasmic domains of the aspartate and serine receptors, respectively, with modeled structures of the transmembrane and linker regions. The two symmetric subunits of the homodimer are in blue and gold, respectively. (Right) Schematic diagram showing structural and functional regions. For simplicity, helix supercoiling is omitted, and the pathway components that dock to the receptor in the assembled signalling complex are shown schematically (ellipsoids, spheres). Kinase docking, regulation and phosphotransfer events occur at the extreme cytoplasmic tip of the receptor. The adaptation enzymes interact with a conserved sequence at the C-terminus of certain receptors. Cytoplasmic sites of methylation and demethylation are shown as small ovals. Figure taken from [44].

Kinase CheA and adaptor protein CheW

The dimeric protein CheA consists of five domains per monomer (73 kDa); P1 - the phosphorylation domain carrying the phosphorylation site, P2 - CheY-and CheB-binding domain, P3 - dimerization domain, P4 - the catalytic domain, and P5 - the regulatory domain coupling CheA to receptor and CheW (see fig. 1.7). The response regulator binding domain P2 contains a docking site for CheY and CheB, and competition between the two has been shown biochemically [87]. However the P2 domain is not essential for chemotaxis and phosphorylation of CheY and CheB [61, 152]. Interaction with the P1 substrate domain from the other subunit in the same dimer is essential for trans-autophosphorylation and

dimer stability [21, 75]. The regulatory domain, P5, is homologous to and can bind CheW [20, 54, 84, 122]. Furthermore, P5 interacts with the receptor and is essential for receptor-mediated kinase regulation [26, 116].
CheA is expressed in two forms from two different start codons in the *cheA* gene. $CheA_L$ is the full length form, whereas the short version $CheA_S$ lacks the first 97 amino acids including the phosphorylation site. $CheA_S$ is only expressed in enterobacteria which have the phosphatase CheZ [110]. One third of the cells CheA pool is the short version, and both homo- and heterodimers of CheA short and long are thought to exist in the cell [75].
CheW is a 18 kDa helper protein consisting of two five-stranded β - barrels surrounding a hydrophobic core. Two putative binding sites for CheA and the receptors have been found. Recent work demonstrated a role of CheW in clustering of chemoreceptors at the pole [67]. Therefore CheW is believed to mediate interactions between the receptor and the kinase CheA, coupling chemoeffector binding events to kinase activity. However, it has been observed that a receptor:CheA complex can form in the absence of CheW [4, 84]. CheW has four known activities in vitro; binding to CheA; binding to receptors; promoting formation of CheA:CheW:receptor ternary complexes; and enabling receptors to stimulate and/or inhibit CheA autokinase activity [4, 25, 48, 49, 85, 95, 96, 102, 109, 136]. However, the excact functions *in vivo* remain unclear.

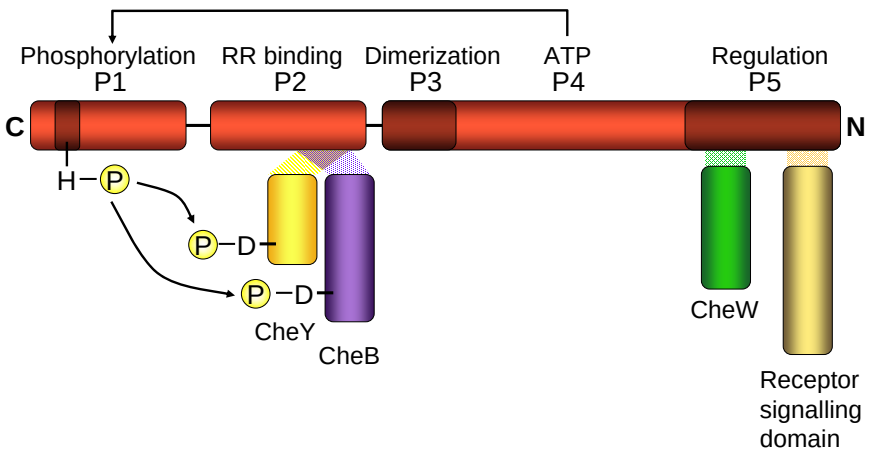

Figure 1.7: Schematic overview of the domain structure of CheA. CheY and CheB are competing for the docking site in the P2 domain. CheA can phosphorylate both CheY and CheB. P5 interacts with the receptor and CheW.

1.2.4 The response regulator CheY and its phosphatase CheZ

CheY is the response regulator and the messenger of the chemotaxis pathway. In its phosphorylated state it binds FliM, a part of the flagellar motor, to regulate the direction of rotation. With 14 kD, the protein folds as a compact, globular structure, in which a central fivestranded β - sheet is sandwiched between five α-helices. The phosphorylation site containing Asp57 is located at one end of the β - sheet [153]. In the adapted state, about half of the total CheY pool in the cell is phosphorylated, which is within the narrow range of phospho-CheY concentrations that allows the ultrasensitive motor to respond to both positive and negative stimuli [34]. The phospho-CheY level is balanced between CheY phosphorylation by CheA and dephosphorylation by CheZ. With its autodephosphorylation activity, the half-life of phospho-CheY is relatively long, about 10 seconds. Catalysis by CheZ lowers the half-life to only about 0.1 seconds, causing a rapid response to kinase inhibition by attractant [139].
The phosphatase CheZ is a dimer of two 24 kDa subunits, of which the central element is an elongated four-helix bundle. CheZ has two distinct interaction surfaces and the hinged CheZ molecule is thought to "clamp down" on the globular CheY molecule with a tight binding constant. The phosphatase is supposedly using the existing mechanism of CheY autodephosphorylation and render it more efficient by positioning the attacking water molecules in an appropriate geometry for attack, similar to the GTPase mechanism in the Ras/G_α families [177]. Many bacteria do not have CheZ but instead CheC, FliY and/or CheX proteins, which are homologous, but do not have any sequence similarity to CheZ [159]. Some bacteria have several CheY proteins which can be phosphorylated, but cannot control motility, and they are thought to act as a phosphate sink to draw off phosphoryl groups from the motility controlling response regulator. The benefit of having a separate phosphatase, instead of a higher CheY autodephosphorylation activity or several CheY molecules is not fully understood. CheZ localizes to the chemoreceptor cluster and thereby localizes the phosphate sink to the phosphate source (CheA). Cluster-associated phosphatase makes sure the gradient of CheY-P in the cytoplasm is flat, which could facilitate a concerted regulation of all flagellar motors distributed in the cell. The rapid dephosphorylation of phospho-CheY right at the site of production might also serve to buffer fluctuations in phosphorylation activity, so that only a substan-

tial up-regulation of the kinase creates enough phospho-CheY to escape the phosphatase and diffuse into the cytoplasm [30, 93, 164].

1.2.5 The adaptation system

A well established feature of chemotaxis is its property of adaptation: the steady-state tumbling frequency in a homogeneous ligand environment is insensitive to the value of ligand concentration. This property allows bacteria to maintain their sensitivity to chemical gradients over a wide range of attractant or repellent concentrations [9].
CheR, the methyl transferase, methylates the MCPs, whereas CheB, the methylesterase, catalyzes the reverse modification reaction. CheB catalyzes the deamidation of specific glutamine residues and the demethylation of methyl-glutamates introduced by CheR in the NWETF pentapeptide sequence of the cytoplasmic C terminal region of major chemoreceptors. Minor receptors lack the pentapeptide sequence. The adaptation enzymes show slower kinetic rates than CheY phosphorylation by CheA and CheY dephosphorylation by CheZ, which is why adaptation sets in after the initial movement response [8, 11, 143]. Both adaptation enzymes are two-domain proteins, with their active site in the C-terminal domain. They are present in *E.coli* in relatively low numbers of about 100-300 copies per cell [88].
Phosphorylation of CheB at Asp56 in the regulatory N-domain enhances methylesterase activity about 100 - fold [35–38]. In its unphosphorylated state, the regulatory domain inhibits methylesterase activity of the effector domain. Because the catalytic site is buried in an interaction of the two domains, access of substrate to the active site is sterically blocked. Phosphorylation is thought to disband the interdomain linkage, leading to an open conformation with increased enzymatic activity [35], thereby relieving the inhibition and stimulating the catalysis. The P2 domain of histidine kinase CheA has also been shown to inhibit the methylesterase activity of CheB. This inhibition is decreased upon phosphorylation of CheB [6].
Over a wide range of attractant concentrations, adaptation accurately returns kinase activity to the pre-stimulus state. Precise adaptation requires a feedback-mechanism, which couples kinase activity to receptor methylation. The Barkai-Leibler model of two-state receptors demonstrates that a dependence of CheR and CheB activity on receptor activity (not ligand binding or methylation level)

naturally leads to precise adaptation [9]. Interestingly, the control of CheB activity by phosphorylation is not essential for adaptation. Mutation of the Asp56 phosphorylation site in CheB can be compensated by overexpression of the protein. However, computer simulations argue that the CheB phosphorylation-feedback might serve to improve the robustness of the chemotaxis system to variations in the CheR and CheB expression [3, 74]. Experiments and computer simulations suggest that CheR and CheB not only act on the receptor they are bound to, but also on five to seven adjacent receptors, defining an "assistance neighbourhood" [55]. Such a mechanism would explain why minor receptors depend of the presence of major receptors for adaptation [89]. However, binding to receptors is not essential for adaptation, as deletion of the pentapeptide sequence can be compensated by overexpression of CheR [119]. Comparative analysis of prokaryotic genomes suggests that tethering the adaptation enzymes to receptors is a recently evolved mechanism, which is not essential, but only improves the efficiency of adaptation, so that lower CheR and CheB expression levels suffice [1].

1.2.6 Higher-order organization of chemotaxis receptors

Early biochemical studies suggested that the receptor-CheA interaction depends on CheW, which led to a model of one CheA dimer being linked to one receptor dimer by two CheW adaptors [49]. Individual receptor-kinase units were thought to regulate CheA activity independently, and it was suggested that the cell would optimally evenly distribute these units around the cell body to maximize detector coverage [18]. Immunogold electron microscopy [102] and fluorescence microscopy [148], surprisingly revealed that chemoreceptors form large clusters, where thousands of sensory complexes [88] are thought to come together. Figure 1.1 on page 2 shows cells expressing CheY as a fusion to YFP, making the clusters visible.
Clustering is conserved among all studied prokaryotic chemotaxis systems [51], which demonstrates its importance for signal processing. Different modes of interaction within the clusters have been suggested. Crystal structure of the cytoplasmic domain of the serine receptor Tsr, [71] and *in vivo* crosslinking studies suggest a trimeric structure [5, 56, 155, 156]. A trimer-of-dimers model has become widely accepted as the smallest signalling unit and the basic building block of a cluster [5, 123]. Receptor clustering seems to be strongly reduced in absence

of CheA and CheW [102, 148]. Hence, a model of a hexagonal cluster structure was suggested, in which trimers-of-dimers are connected by CheA dimers, coupled to receptors via CheW [84, 141].
Recently, a different model on receptor clustering was presented after solving the crystal structure of the cytoplasmic part of a *Thermotoga maritima* chemoreceptor [122]. Interestingly, these receptors did not crystallize as trimers-of-dimers, but rather as "hedgerows" of dimers with lateral interactions over a large surface. The same study analyzed the relative arrangement of CheA and CheW in complex, and suggested a model where receptors associate with CheA via CheW, with only little direct contact between receptors and CheA at the P3 domain. Interactions between the P5 domains of different CheA dimers were further proposed to connect multiple hedgerows into a two-dimensional lattice.
Receptor-CheW-CheA complexes are believed to be stable on the time scale of chemotactic signalling [49]. There is some evidence that attractant binding or demethylation of receptors decrease cluster stability [79, 144], but these effects appear to be minor *in vivo* [90, 97, 142, 164], although ligand binding might influence the relative arrangement and distances between receptors in the lattice [56, 165]. Association of the other chemotaxis proteins with the ternary complex is believed to be rather loose, but with the notable exception of CheY, no large changes in the degree of protein localization to the cluster have been observed upon stimulation [90, 164]. Assembly and disassembly of clusters thus does not appear to be directly involved in signal transduction. Because bacteria make temporal and not spatial comparisons of ligand concentration, the physical position of clusters in a cell is unimportant for signalling.
The main function of clustering appears to be in signal processing. Recent intriguing evidence for extensive functional interactions between receptors in clusters [5, 50, 77, 86, 146, 178] led to the understanding that signalling is not performed by independent receptor kinase units but an interconnected receptor network. Receptors of different types are intermixed in clusters [5], forming large allosteric protein complexes that integrate multiple stimuli and amplify weak signals [123, 146]. Computer models mathematically describing the behaviour of receptors explain the role of cooperativity in signal processing. It is assumed that single receptors behave as switches, which are, dependent on ligand occupancy and methylation level, either in an active (kinase stimulating) or inactive (kinase inhibiting) state (two-state model) [9]. Inside a cluster, receptors are believed to be coupled, stabilizing their respective state of activity in neighbouring

receptors, so that they tend to switch together from one state to the other, similar to allosteric multiprotein complexes. The Monod-Wyman-Changeux model suggests that a cluster might consist of many such cooperative units, with all receptors of one unit switching synchronously between inactive and active, due to their strong coupling (fig. 1.8) [69, 112, 114, 146]. Additionally, *in vivo* FRET

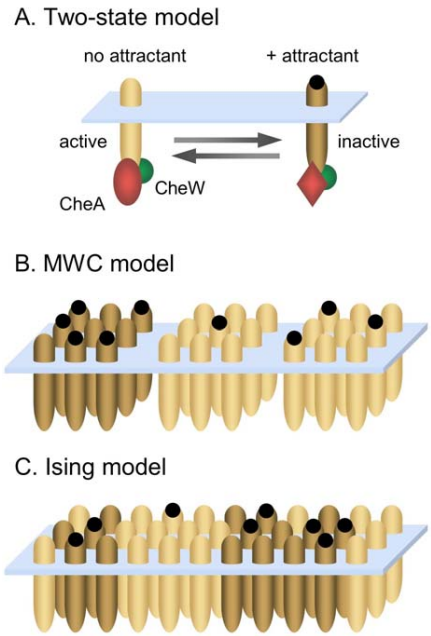

Figure 1.8: Models of receptor cooperativity. According to the Two-state model, single receptors, coupled to CheA and CheW, switch between an active and an inactive state. Attractant binding increases the probability of being in the inactive state, whereas methylation shifts the equilibrium towards the active state (A). The Monod-Wyman-Changeux (MWC) model claims the existence of cooperative receptor units. Depending on the ligand occupancy and methylation level of its receptors, the entire unit switches between active and inactive state (B). In the Ising model, all receptors together form a coherent network. Receptor coupling has a finite strength, causing local changes in activity (C). Taken from the PhD thesis of D.Kentner.

measurements showed that these cooperative units, or signalling teams, increase in size with receptor modification, suggesting an additional level of adaptation of the chemotaxis network [42]. Figure 1.9 shows how chemoreceptors cluster at different levels. In another model, described in a similar manner as ferromagnetism (hence; the Ising model), the entire receptor population might form a single coherent network, in which receptors, due to a finite coupling strength, only influence their adjacent neighbours, leading to local changes in activity (fig. 1.8) [39, 111].

Analyses of both models show that cooperative receptor units, or receptor neigh-

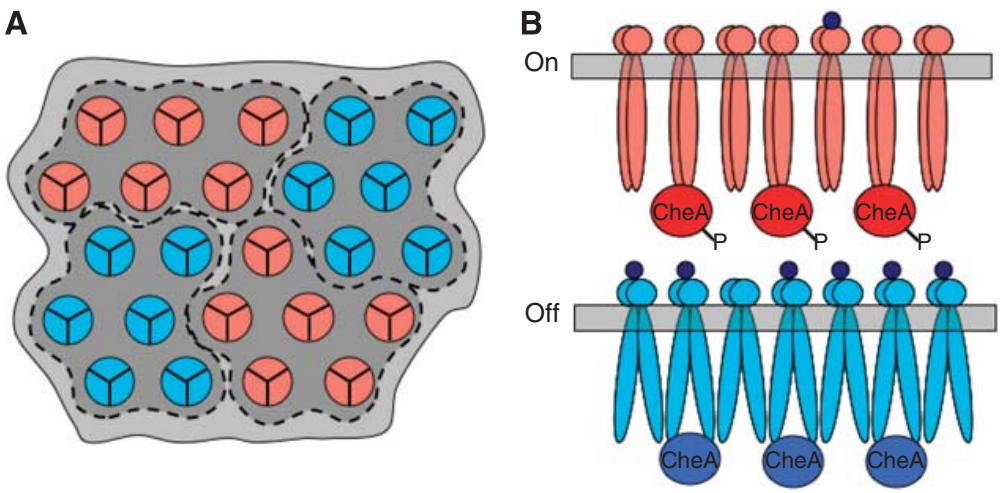

Figure 1.9: Model of the receptor cluster and signalling teams. Chemoreceptors cluster at different levels, from dimers, to trimers-of-dimers, to large polar and lateral clusters. Within a signalling team, receptors are assumed to be coupled strongly enough that the receptors are either all in the on state or all in the off state **(A)**. In the on state, CheA is active, autophosphorylating itself, and transferring the phospho-group to the response regulators CheY and CheB. In the off state, CheA is inactive, unable to autophosphorylate **(B)** [42].

bourhoods or signalling teams, function as highly sensitive antennae, the sensitivity of which is adjusted optimally by the adaptation system to react to both positive and negative simuli, so that the binding of a single ligand molecule can elicit a signal. In a mixed cluster, receptor cooperativity not only allows amplification, but also integration of various signals through coupling of different receptor species. Consistent with its importance, clustering appears to be essential for the functionality of the ternary complex. In contrast, abolishing localization of the other chemotaxis proteins by specific mutation only has a moderate effect on chemotaxis, and can be compensated by overexpression of the respective proteins [61, 119, 164]. It therefore seems likely that targeting these peripheral proteins to the cluster just enhances their enzymatic efficiency and specificity. In species with multiple chemotaxis systems, targeting them to separate clusters might additionally help to prevent unwanted cross-talk between the systems. The fact that many other sensory systems (B-cell, T-cell, synaptic) cluster indicates that receptor clustering is an important regulatory mechanism for the cell to adjust signalling properties or to recruit auxiliary proteins.

1.2.7 Flagella and motility

The bacterial flagellum is a motor organelle and a protein export and assembly apparatus. Flagella extend from the cytoplasm to the cell exterior and are assembled at the distal end. Hence all the protein subunits of the external elements have to be exported. Substrates diffuse down a narrow channel through the growing structure and assemble at the distal end often assisted by capping proteins [62, 99]. The flagellum consists of a long helical filament, a short curved structure called the hook and a basal body consisting of a central rod and several rings. Figure A.1 on page 117 shows a schematic of the assembly of the flagellum. The filament is connected to the cell by the hook which is built from a distinct subunit FlgE. Between the hook and the filament two junction proteins, FlgK and FlgL, known as hook-associated proteins, act as adapters. The hook is connected to the cell via the basal body. The distal rod, connects the hook to the L ring of the basal body. The L ring is followed by the P ring and then the proximal rod. These structures are mounted to the MS ring, which interacts with the C ring built up from FliG, FliM and FliN that in many copies form the switch complex. Two proteins MotA and MotB are integral to the cell membrane and necessary for motor rotation but not for motor switching. They are assumed to be stationary, whereas the FliG is assumed to be part of the rotor [98, 100]. The rotary motor that drives the flagellum is fuelled by the membrane gradient of protons. MotA and MotB form the stator and they function together to conduct ions across the membrane and to couple the ion flow to rotation. Each motor complex consists of 8-12 torque generating units consisting of four MotA and two MotB molecules [133].

1.2.8 Regulation of expression of the flagellar operons

Flagellar and related genes are expressed in 14 different operons. These are organized in a hierarchical fashion as a regulon divided into three promoter classes with expression of operons at one level affecting expression at lower levels. The system is turned on during exponential growth phase. The master operon *flhDC* lies at the top of the hierarchy in class 1 with both of the gene products being absolutely required for all other genes in the flagellar regulon. Not all proteins regulating the master operon have been found. However, cyclic AMP is known to be an activator [98]. LrhA is thought to be a key regulator controlling the

transcription of flagellar, motility and chemotaxis genes by regulating the synthesis of $FlhD_2C_2$ [82]. Additionally the global *hns* gene is a positive regulator of the flagella biosynthesis [80].
Class 2 consists of structural components of the basal body-hook structure such as the *fliFG* and *flgBC* genes. In addition, class 2 proteins also include the transcriptional activator for class 3 operons, FliA or σ^{28}, a sigma factor required for expression of the class 3 operons and some of the class 2 operons. Within the class 2 cluster the promoters are turned on sequentially with significant delays in the order *fliL, fliE, fliF, flgA, flgB, flhB, fliA*. DNA regulatory sites in the promoter regions of the operons are ranked in affinity. When the concentration of a transcription factor increases in the cell, it binds to and activates the operons with the highest affinity sites first [32, 63, 73]. This observed order corresponds to the spatial position of the gene products during flagellar motor assembly starting in the cytoplasm and extending to the extracellular side. The *fliL* operon genes form the cytoplasmic C ring and FliE and FliF form the MS ring in the inner membrane, thought to be the first assembled structure and *fliA* is the last class 2 gene to be switched on. When the transcription of *fliA* starts, the expression of FlhDC starts to decrease and the transcription of the class 3 operons can begin. The class 3 genes can be divided into two groups. The filament structural operons *flgK, fliD* and *fliC* are activated first, before *flgM* and the chemotaxis operons *meche* and *mocha*. The mocha operon consists of *motAB*, and the cytoplasmic signalling proteins CheA and CheW. The *meche* operon consists of *tar, tap, cheR, cheB, cheY and cheZ*. The three receptors Tsr, Trg and Aer are encoded elsewhere in the genome [88]. Expression of class 3 operons is dependent on FlhD and FlhC only indirectly via the expression of FliA. When FliA is expressed from a foreign promoter the class 3 operons are independent on the presence of FlhD and FlhC. The class 2 operons are absolutely dependent on presence of FlhD and FlhC, but the expression is enhanced by FliA. The class 3 FlgM is an anti-sigma factor and FliA's antagonist. FlgM binds to the sigma factor FliA and prevents it from binding to class 3 promoters. Assembly of the export apparatus through the expression of all the class 2 genes is needed for expression of class 3 operons. FlgM is removed from the cell by export at the point were expression of class 3 operons is needed. FliA remains in the cell and is then able to bind to the class 3 operons in absence of FlgM. Hence, which proteins can be exported at what stage of assembly is tightly regulated and the export of FlgM is a checkpoint mechanism to ensure

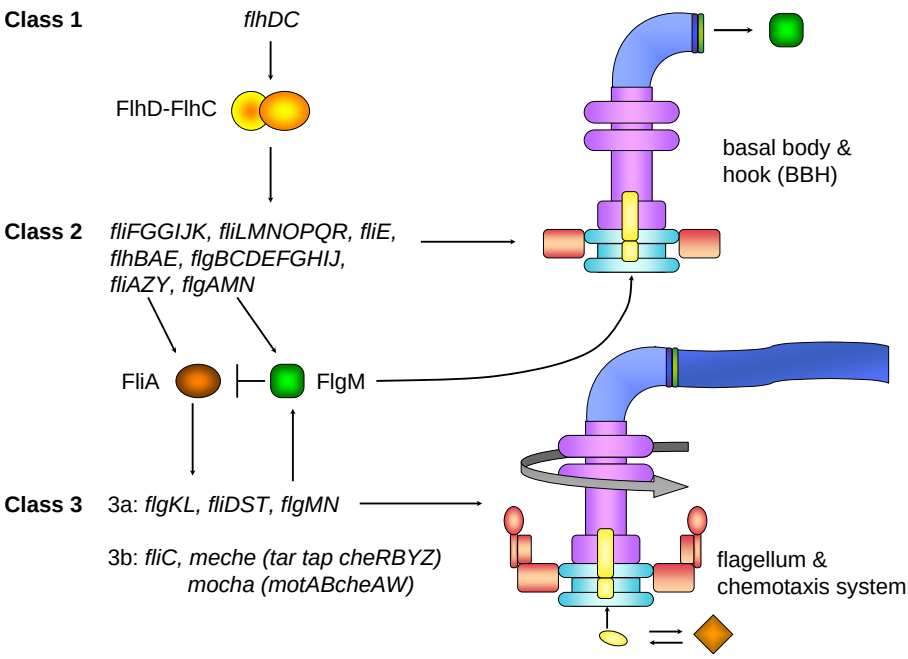

Figure 1.10: The genetically defined hierarchy of flagellar operons in *E.coli*. The master regulator FlhDC turns on class 2 genes, one of which, FliA, turns on class 3 genes. A checkpoint ensures that class 3 genes are not switched on until basal body and hook structures (BBH) are completed. This is implemented by FlgM, which binds and inhibits FliA. When BBH are completed, they export FlgM out of the cell, leaving FliA free to activate the class 3 operons. Note that *flgM* is transcribed from both a class 2 and a class 3 promoter.

that the filament - and chemotaxis proteins are not expressed before functional basal body-hook structures are completed [47, 58, 64, 65, 98, 118]. Several additional feedback mechanisms have been suggested. For instance FliZ, which is expressed from both class 2 and class 3 promoters is a positive activator of class 2 promoters [76, 117]. In addition, Aldridge et al. have recently observed that in *Salmonella enterica* serovar Typhimurium, FliZ regulates the concentration of $FlhD_4C_2$ posttranslationally [132]. The regulation hierachy is shown in figure 1.10.

1.3 Aims of the current work

One of the main goals of this work is to characterize gene expression noise in flagella and chemotaxis genes. Fundamental questions we want to answer are where noise originates and which noise sources are the most important in prokaryotic signal transduction. Furthermore, the time scale of fluctuations should be studied, as the impact of noise on the signal transduction pathway could be dependent on the duration of the variations. Hence, this work additionally focuses on the effects of variations in protein levels on the chemotaxis pathway. Our objectives are to evaluate any known mechanisms and possibly discover new pathway features compensating for the observed fluctuations in gene expression. Finally, assessing the optimal pathway topology to cope with noise and the evolutionary implications of noise on chemotaxis would be of great interest.
To address these questions we constructed fluorescent protein fusions to selected flagella and chemotaxis genes to quantify the variations in protein levels in the chemotaxis pathway on a single cell level by snap shot - and time lapse imaging. In addition, swarm assays and tethering experiments allowed us to assay the chemotactic ability of a population or of single cells respectively, while FRET measurements provide a readout of kinase activity. Under varying levels of different chemotaxis proteins, these methods were used to investigate the network design and its robustness in detail.

2 Materials and Methods

2.1 Chemicals and consumables

Chemicals and consumables used in this work are listed in table 2.1 on page 28.

2.2 Media and buffers

2.2.1 LB plates

10 g Tryptone
5 g Yeast extract
5 g NaCl
1 M NaOH
15 g Agar
Added H_2O to total volume of 1 l and adjusted pH to 7.

2.2.2 Tethering buffer

100 ml 0,1 M KPO_4
200 µl 0,5 M EDTA
13,4 ml 5 M NaCl
100 µl 10 mM Methionine
1 ml 10 M Lactic acid
Added H_2O to total volume of 1 l and adjusted pH to 7 with NaOH.

Table 2.1: Chemicals used in this work

Chemical	Company	Product no.
Alkaline phosphatase, shrimp	Roche	1758250
L-arigine	Sigma	A8094
Agarose ultra pure	Difco	77712
Agar Bacteriology	Fluka	40617
Agar Select	Difco	77699
Ampicillin	Applichem	A0839.0050
Chloramphenicol	Applichem	A1806.0100
Kanamycin sulphate	Applichem	A1493.0025
Bromophenol blue	Applichem	A3640.0025
Chloroform (Trochloromethane)	Roth	3313.4
Calcium chloride-dihydrate	Roth	5239.1
Di-sodium hydrogen phosphate	Applichem	A2943.0500
dNTP set	Invitrogen	1029718
1kb plus DNA ladder	Invitrogen	10787018
EDTA	Merck	74513
Ethanol	Applichem	A1613.2500
Ethidium bromide	Applichem	A1152.0025
Glycerol 99,5 %	Gerbu	
HCl,	Applichem	A0659.1000
IPTG	Roth	2316.3
Potassium Chloride	Applichem	A2939.1000
Potassium di-hydrogen phosphate	Applichem	A2946.0500
Di-potassium hydrogen phosphate	Applichem	A2945.0500
T4 DNA Ligase	Fermentas	EL 0011
T4 DNA Ligase buffer	Invitrogen	46300018
Lactic acid	Sigma	L6661
$MgSO_4$	ROTH	P027.1
Methanol	Applichem	A0688.2500
L-Methionine	Sigma	M9625-25G
Pwo DNA Polymerase	Roche	
Poly-L-Lysine	Sigma-Aldrich	P8920
Restriction enzymes	NEB	
Taq DNA Polymerase	Invitrogen	18038042
Yeast extract	Difco	Theoretikum 80705
X-gal	Roth	2315.2

2.2.3 TB

10 g Tryptone
5 g NaCl
1 ml NaOH (pH 7)
Added H_2O to 1 l total volume.

2.2.4 TAE buffer for gel electrophoresis (50x)

242 g Tris base
57,1 g Glacial acetic acid
100 ml 0,5 M EDTA, pH 8
Added H_2O to 1 l total volume

2.2.5 10x gel loading buffer for electrophoresis

0,25 % (w/v) Bromophenol blue
0,25 % (w/v) Xylene cyanol
Glycerol in 40 % H_2O

2.2.6 1 kb plus ladder for gel electrophoresis

10 µl ladder stock (table 2.1)
30 µl 10x loading buffer
60 µl H_2O

2.2.7 Buffers for Southern blot

SSC-buffer: 20 x stock solution

175.3 g 3,0 M NaCl
88.2 g 0,3 M Na-citrate
Added 800 ml H_2O to 1 l total volume and adjusted pH to 7.

Depurination buffer

1 L 0,25 M HCl
12 ml 12 N HCl in 988 ml H_2O

Denaturation buffer

1.5 M NaCl
0.5 M NaOH
300 ml 5 M NaCl
40 ml 12,5 N NaOH
660 ml H_2O

Washing buffer 1

500 ml 2x SSC
0.1 % SDS
50 ml 20x SSC
5 ml 10 % SDS
445 ml H_2O

Washing buffer 2

0,5x SSC
0,1 % SDS 12,5 ml 20x SSC
5 ml 10 % SDS
482,5 ml H_2O

2.2.8 Reaction kits

- QIAprep Spin Miniprep Kit 250, Qiagen, Hilden
- DNeasy Mini Kit Qiagen, Qiagen, Hilden
- Qiaprep Spin Miniprep Kit, Qiagen, Hilden
- Qiaquick Gel Extraction Kit, Qiagen, Hilden
- ECL Direct Nucleic Acid Labelling and Detection System kit, Amersham Biosciences (GE Healthcare), UK

2.3 Bacterial strains

Table 2.2 shows a list of the bacterial strains used in this work, with description of the genotype, source of the strain and references if the strain has been mentioned previously in a paper.

Table 2.2: Bacterial strains used in this work

Strain	Relevant genotype	Source or reference
DH5α	F^-	Hanahan et al. (1983), Invitrogen, Karlsruhe
RP437	Chemotaxis wild type	Parkinson and Houts (1982)
VS100	Δ*cheY*	Sourjik and Berg (2000)
VS102	Δ*flgM*	Sourjik and Berg (2002), Kollmann et al. (2005)
VS104	Δ*(cheY-cheZ)*	Sourjik and Berg (2002a)
VS116	Δ*flhC*	Kentner et al. (2006)
VS124	Δ*(cheB-cheZ)*	Sourjik and Berg (2002a)
VS126	Δ*cheR*	This work, Løvdok et al. (2007)
VS127	Δ*(cheY-cheZ)* Δ*cheR*	Sourjik and Berg (2002a)
VS137	Δ*(cheY-cheZ)* Δ*tsr*	This work
VS149	Δ*(cheR-cheZ)*	Vaknin and Berg (2004)
VS161	Δ*cheZ*	This work, Løvdok et al. JBT (2007)

Continued on next page

Table 2.2 – Continued from previous page

Strain	Relevant genotype	Reference or source
VS162	*cheY-yfp, cheZ-cfp (weak)*	Kollmann et al. (2005)
VS165	Δ*(cheY-cheZ)* Δ*tsr tar*Δ*pp*	This work
VS166	Δ*cheA*	This work
VS168	Δ*(cheY-cheZ)* Δ*cheA*	This work
VS177	Δ*(cheY-cheZ) cheA*Δ*P2*	A.Vaknin (gift), Jahreis et al. (2004)
VS181	Δ*(cheY-cheZ)* Δ*(tsr, tar, tap, trg, aer)*	Vaknin and Berg 2007
LL1	*cheY-yfp*, Δ*flgM*	This work
LL4	Δ*(cheY-cheZ)* Δ*flgM*	This work
LL5	Δ*(cheR-cheZ)* Δ*flgM*	This work
LL17	Δ*(cheY-cheZ)* Δ*flgM* Δ*tsr*	This work
LL18	Δ*(cheR-cheZ)* Δ*flgM* Δ*tsr*	This work
LL10	*cheY-yfp, fliM-cfp,* Δ*flgM*	This work
LL12	*cheY-yfp*	This work
LL13	*cheY-yfp, cheA-cfp*	This work
LL14	*cheY-yfp, fliM-cfp*	This work
LL15	*cheY-yfp, flgM-cfp*	This work
LL16	*cheY-yfp, flhC-cfp*	This work
LL19	*ptrc-yfp, flhC-cfp*	This work
LL22	*fliM-yfp*	This work
RP 2867	Δ*tap* Δ*(cheR-cheB)*	J.S. Parkinson (gift), Parkinson and Houts (1982)
RP4606	*cheW113*	J.S Parkinson (gift), Alexandre and Zhulin (2002)
RP4972	Δ*cheB*	J.S. Parkinson (gift), Stewart et al. (1990)
DHB 6521	λInCh1	J. Beckwith (gift), Boyd et al. (2000)
DHB 6501	host for λInCh1 system	J. Beckwith (gift), Boyd et al. (2000)
UU1250	Δ*(tsr, tar, tap, trg, aer)*	J.S. Parkinson (gift), Studdert and Parkinson (2004)

2.4 Plasmids

Ribosome binding sequences used for the fusion constructs in plasmids pVS142 to pVS520 and pLL33, pLL36 and pLL45 are listed in table 2.3. Table 2.4 on page 34 shows a list of the plasmids used in this work with relevant genotype, source of the plasmid and references.

Table 2.3: Upstream ribosome binding sequences of the fusion constructs made by A. Müller

	Upstream sequence
RBS^{CheR}	GAGCTCTTGAGAAGGCGCT**ATG**
RBS^{CheB}	GAGCTCAGTAAGGATTAACG**ATG**
RBS^{CheY}	GAGCTCCGTATTTAAATCAGGAGTGTGAA**ATG**
RBS^{CheZ}	GAGCTCCAGGGCATGTGAGGATGCGACT**ATG**
RBS^{CheYS}	ACTAGTGAAGGAGTGTGCC**ATG**
RBS$^{CheR\uparrow}$	GAGCTCGATAGGAAAGGCGCT**ATG**
RBS$^{CheB\uparrow}$	GAGCTCAAGAGGAAATTAACG**ATG**
RBS$^{CheY\uparrow}$	GAGCTCAATAGAGGAAATGTGAA**ATG**

2.5 Primers

Primers used in this work are listed in table 2.5 on page 38. Numbers denote the Linda-numbers of the primers. VIC-primers are made by Victor Sourjik. Primer ALD8 is a personal gift fromthe lab of Ady Vaknin. Forward and reverse primers are denoted f and r respectively.

Table 2.4: Plasmids used in this work

Plasmid	Relevant genotype	Source or reference
pTrc99a	Expression vector; pBR ori, pTrc promoter, Amp^R	Amann et al. (1988)
pBAD33	Expression vector; pACYC ori, pBAD promoter, Cm^R	Guzman et al. (1995)
pBAD18K	Expression vector; pBR ori, pBAD promoter, KanR	Guzman et al. (1995)
pAMPTS	Cloning vector, ts origin of replication, Amp^R	Sourjik and Berg (2000)
pLL3	Cloning plasmid for integration of *cheY-yfp* into chromosome, pAMPTS derivative	This work
pLL6	FlgM-YFP expression plasmid; pTrc99a derivative	This work
pLL12	Cloning plasmid for integration of *fliM-yfp* into chromosome, pAMPTS derivative	This work
pLL13	Cloning plasmid for integration of *fliM-cfp* into chromosome, pAMPTS derivative	This work
pLL16	FlgM expression plasmid; pTrc99a derivative	This work, Kollmann et al. (2005)
pLL21	Cloning plasmid for integration of *flhC-cfp* into chromosome, pAMPTS derivative	This work
pLL22	Cloning plasmid for integration of *flgM-yfp* into chromosome, pAMPTS derivative	This work
pLL23	Cloning plasmid for integration of *flgM-cfp* into chromosome, pAMPTS derivative	This work

Continued on next page

Table 2.4 – Continued from previous page

Plasmid	Relevant genotype	Soure or reference
pLL24	Cloning plasmid for integration of *cheA-cfp* into chromosome, pAMPTS derivative	This work
pDK3	Expression vector for cloning of C-terminal eyfp fusions; pTrc99a derivative	Løvdok et al. (2007)
pVS55	YFP-CheA expression plasmid; pBAD33 derivative	V.S., This work
pVS18	CheY-YFP expression plasmid; pTrc99a derivative	Sourjik and Berg (2002a)
pVS54	CheZ-CFP expression plasmid; pBAD33 derivative	Sourjik and Berg (2002a)
pVS64	CheZ-YFP expression plasmid; pTrc99a derivative	Liberman et al.(2004)
pVS88	CheZ-CFP-CheY-YFP expression plasmid; pTrc99a derivative	Sourjik and Berg (2002a, 2004)
pVS137	CheR-YFP expression plasmid; pTrc99a derivative	Løvdok et al. (2007)
pVS138	CheB-YFP expression plasmid; pTrc99a derivative	Løvdok et al. (2007)
pDK19	CheR-YFP expression plasmid; pDK3 derivative	Løvdok et al. (2007)
pDK167	CheR^{D154A} expression plasmid; pBAD33 derivative	D. Kentner, This work
pVS142	RBS^{CheB}_CheB_CheY-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS145	RBS^{CheR}_CheR_CheB-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS261	RBS^{CheYS}_CheA-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS305	RBS^{CheY}_CheY_CheZ-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS319	-316_CheY-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS321	RBS$^{CheY\uparrow}$_CheY_CheZ-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS450	RBS$^{CheB\uparrow}$_CheB_CheY-YFP expression plasmid; pTrc99a derivate	A. Müller, This work

Continued on next page

Table 2.4 – Continued from previous page

Plasmid	Relevant genotype	Soure or reference
pVS451	RBS$^{CheR\uparrow}$_CheR_CheB-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS452	RBS$^{CheR\uparrow}$_CheR-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS487	RBS$^{CheB\uparrow}$_CheB-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS490	RBS^{CheYS2}_CheA_CheW-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS495	RBS$^{CheY\uparrow}$_CheY-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pVS520	RBS^{CheYS2}_CheA$_S$_CheW-YFP expression plasmid; pTrc99a derivate	A. Müller, This work
pLL34	FliA expression vector;pTrc99a derivative	This work
pLL33	-316_CheY-YFP expression plasmid; pBAD33 derivate	This work
pLL36	RBS^{CheB}_CheB_CheY-YFP expression plasmid; pBAD33 derivate	This work
pLL39	CheA-CFP expression plasmid; pBAD33 derivative	This work
pLL40	CheB_CheY_CheZ expression plasmid; pTrc99A derivative	This work
pLL42	CheY-CheZ expression plasmid;pTrc99a derivative	This work
pVS103	CFP-CheB expression plasmid; pBAD33 derivative	V.S., This work
pVS111	CheR-CFP expression plasmid; pBAD33 derivative	V.S., This work
pVS91	CheB expression plasmid; pBAD33 derivative	Liberman et al. (2004)
pVS97	CheB^{D56E} expression plasmid; pBAD33 derivative	V.S., This work
pVS112	CheBc (202 first aa deleted) expression plasmid; pBAD33 derivative	This work
pVS113	CheR expression plasmid; pBAD33 derivative	V.S., This work
pLL43	RBS$^{CheR\uparrow}$_CheR-YFP expression plasmid; pBAD33 derivative	This work
pLL44	RBS$^{CheR\uparrow}$_CheR_CheB expression plasmid; pBAD33 derivative	This work

Continued on next page

Table 2.4 – Continued from previous page

Plasmid	Relevant genotype	Soure or reference
pLL45	RBS$^{CheR\uparrow}$_CheR expression plasmid; pBAD33 derivative	This work
pAV8	CheZ^{F98S}-CFP-CheY-YFP expression plasmid; pTrc99A derivative	A. Vaknin (personal gift)
pVS571	CheR_CheB-YFP expresssion plasmid; pBAD33 derivative	A. Müller, This work

Table 2.5: Primers used in this work

Number	Sequence	Restriction site	Gene	Purpose
6a	gataaaccatggacacataaaagc	NcoI	*flgM* f	for pLL6
6b	catacgaggatccgttactctgcaagtcttg	BamHI	*flgM* r	for pLL6
7a	gctcgttctagatgtggaattgtgagcgga	XbaI	*cheW* f	downstr of *cheA* for pLL24
7b	gtcgaggagctccttgctcaccattccacc	BamHI	*cheW* r	downstr of *cheA* for pLL24
8a	gttacgtctagaacatttgattaacc	XbaI	*fliN* f	downstr of *fliM* for pLL12 and 13
8b	attcatggatcccggctcaggcggcgcattc	BamHI	*fliN* r	downstr of *fliM* for pLL12 and13
9a	tgccgatctagagatcaacgaagcgcagc	XbaI	*flgN* f	downstr of *flgM* for pLL22 and 23
9b	cggcctgcggatccggccggagataatct	BamHI	*flgN* r	downstr of *flgM* for pLL22 and 23
10a	cgcgcgtctagacctgacgactgaacatcct	XbaI	*motA* f	downstr of *flhC* for pLL21
10b	ggcgggggatccgatcggctgaacctaacg	BamHI	*motA* r	downstr of *flhC* for pLL21
27a	cggatagagctcgccgataactcatataacg	SacI	*fliA*	for pLL34
27b	caggtatctagattataacttacccag	XbaI	*fliA*	for pLL34
37a	ccggtctagaaatggtgagcaagggc	XbaI	*cfp* f	for pLL39
37b	cgcggaagctttacttgtacagctc	HindIII	*cfp* r	for pLL39
38b	tccacctctagatcaaatccaagactatc	XbaI	*cheZ* r	for pLL42 in combination with VIC131
39b	ccgcgctctagattaatccttacttagcgca	XbaI	*cheR* r	for pLL45 in combination with VIC131, recloning from pVS452

Continued on next page

Table 2.5 – Continued from previous page

Number	Sequence 5'-3'	Restriction site	Gene	Purpose
40b	gcgccgtctagattaaatacgtatcgcctgt	XbaI	*cheB* r	for pLL44 in combination with VIC131, recloning from pVS451
VIC23	cacaccgcgaattcagataacac	EcoRI	*cheY*	for pLL3
VIC66	ggtcagcttgccgtaggtggc	–	*yfp/cfp*	for sequencing ***-YFP/CFP fusions
VIC67	gcgcgatcacatggtcctgctggag	–	*yfp/cfp*	for sequencing YFP/CFP-*** fusions
VIC131	atgtgtggaattgtgagcgg	–	–	pTrc99a sequencing primer f
VIC132	ctgatttaatctgtatcagg	–	–	pTrc99a sequencing primer r
VIC121	tttatcgcaactctctactg	–	–	pBAD33 sequencing primer f
VIC122	atccgctctagaatactactcctggttc	–	–	pBAD33 sequencing primer r
VIC109	tccagctctagagcatgtgaggatg	XbaI	CheZ-CFP r	for pLL3
ALD8	ggtcgactcagaccctgattttgctc	XbaI	*cheR* r	for pLL45

2.6 Molecular cloning

2.6.1 PCR

1 µl template DNA (200ng/ µl)*
5 µl 10x PCR buffer
1 µl dNTP Stock(10mM)
0,5 µl primer 1 (100pmol/ µl)
0,5 µl primer 2 (100pmol/ µl)
0,5 µl polymerase
41,5 µl H_2O

PCR Cycle:
5 min 95 °C
30 sec 95 °C
45 sec 55 °C variable, depending on melting temperature of primers
1 min 72 °C variable, depending on length of fragment
25 cycles, or 35 with genomic DNA as template
10 min 72 °C

*For some experiments genomic DNA was used rather than template DNA in form of cells from a single colony on a freshly streaked LB-plate.
Fragments were separated by eletrophoresis on a 1 % agarose gel consisting of 0,4 g agarose in 40 ml TAE buffer (2.2.4) and 1 µl ethidium bromide.
Gel-extraction was carried out using QIAquick 250 gel-extraction kit from Qiagen, listed in 2.2.8.

2.6.2 Restriction digestion

Enzymes used were SacI, XbaI, PvuII, NotI, EcoRI, EcoRV, SmaI, KpnI, BamHI, HindIII, ClaI all from New England Biolabs or Fermentas. All (preparative) digestions were performed in 30-40 µl volume 2 hours in 37 °C, and analytical digestions to test clones were performed in 10-20 µl with half the amount of DNA.
10 µl DNA (1000-4000 ng)
0,5 µl enzyme 1

0,5 µl enzyme 2
10x buffer
10x BSA
ddH_2O to 30 or 40 µl

2.6.3 Dephosphorylation of cloning vector

To reduce the probability of vector religation, dephosphorylation of the cloning vector was performed.

4 µl vector DNA (200-1000 ng)
1 µl SAP (Shrimp Alkaline Phosphatase)
1 µl 10x SAP buffer
4 µl H_2O

The dephosphorylation was performed 15 minutes at 37 °C followed by 20 minutes at 65 °C.

2.6.4 Ligation

2 µl vector DNA (50 ng)
0,5 µl DNA ligase T4
1 µl 10 x ligation buffer
5,5 µl insert DNA

Ligations were performed in 10-20 µl volume at 16 °C over night.

2.6.5 Competent cells

The cells were grown in 5 ml LB at 37 °C over night and then diluted 100 fold in 100 ml fresh LB and grown at 37 °C to mid-late log phase at OD_{600} ∼0,7. Cells were then centrifuged 5 minutes at 2000 x g, 4 °C and resuspended in 50 ml 0,1 M chilled $MgCl_2$. After 45 minutes incubation on ice the cells were

centrifuged for 5 minutes at 1500 x g and 4 °C and resuspended in 50 ml 0,1M chilled $CaCl_2$. After 5 minutes of centrifugation under the same conditions the cells were resuspended in 3 ml chilled 0,1 M $CaCl_2$with 18 % glycerol and incubated on ice for two hours before aliquoted and frozen at -80°C.

2.6.6 Transformation

50 µl competent cells were thawed on ice. 0,2-1 µl plasmid DNA was added and the solution was gently mixed and left on ice. After 20-30 minutes, the cells were heat shocked at 42 °C water bath for 45 seconds and placed back on ice for 3 minutes. 5 ml LB medium was added and cells were incubated 45 minutes, gently shaking, at 37 °C or 30 °C depending on the plasmid. Finally, the cells were harvested by centrifugation at 6000 rpm, 4 minutes, resuspended in approximately 200 µl of the supernatant and plated on selective media to grow over night.
DNA was purified from the cells using QIAprep Spin Miniprep Kit 250.

2.6.7 Blue-white screening of recombinants

In the presence of β-galactosidase, Lac+ bacteria hydrolyse X-gal and a blue precipitate appears as blue colonies on the plate. Bacteria containing recombinant plasmids form white colonies because LacZ is no longer functional. Bacteria harbouring non-recombinant plasmids, plasmids that have been digested and ligated back together without an insert, form blue colonies. In-frame cloning of the fragment into the *lacZ* gene can however allow expression and lead to leaky phenotypes and variation in the intensity of the blue colonies. This phenomenon leads to false negative colonies. A problem with blunt end ligation is that some plasmids can religate after degradation of the sticky end without any inserted fragment. This leads to false positives in the form of white colonies. In the current experiments the recombinants were streaked on broth plates with ampicillin and 100 µl 50mg/ml X-gal and 10 µl 0,1 mM IPTG.

2.6.8 Frozen cell stocks

Cells were grown over night in 3 ml selective TB medium until stationary phase was reached. After centrifugation at 8000 rpm, pellets were resuspended in ∼1 ml TB with 20 % glycerol, transferred to cryotubes and frozen at -80 °C.

2.7 Cloning by recombination

2.7.1 Crossing in a gene using a plasmid with a temperature sensitive ori

The plasmid pAMPTs could be used to replace a native copy of a gene with a gene fusion. The fusion is crossed into the genome by homologous recombination with the gene of interest and the downstream region of the gene as regions of homology. A plasmid conferring temperature-sensitivity containing the fusion and downstream region of about 300 bp, was first constructed, and transformed into the recipient strain. Transformants were incubated overnight at 30 °C with resistance on LB-agar plates. A mixed population of the transformants was then restreaked and incubated with appropriate antibiotics at 42 °C to select for uptake of the desired gene and resistance . Large single colonies were picked and restreaked and incubated at 42 °C. This cleaning step was repeated twice to separate cells which had incorporated the resistance from those who were trying to survive on the gene from the plasmid. Next, single colonies were picked to grow in 50 ml LB shaking at 30 °C without antibiotics for 24 hours. This step was repeated by transferring 100 µl cells from the liquid culture to 50 ml fresh LB and grown for another 24 hours. Serial dilutions of the culture were made (optimal dilution 10^{-6}) and the cells were plated on LB without antibiotics and incubated at 30 °C. Colonies from the plate with the highest dilution were picked and replica plated onto LB with and without resistance and incubated at 30 °C. The cells that were not resistant to ampicillin were picked for a phenotypic screening since they would have lost the resistance gene along with the other plasmid genes and hopefully got the desired gene introduced into the chromosome. Phenotypic screening was performed for fusions by fluorescence microscopy, and for deletions by swarm plate experiments and PCR. Positive clones were finally confirmed by southern blot.

2.7.2 Crossing in a gene using λ InCh

pBR322 derived plasmids are widely used in studies with *E. coli* because of their high copy numbers in cells. For the production of large amounts of protein the high expression from these plasmids is desired, but in some studies low level

expression is necessary to obtain relevant measurements, for instance when assessing in vivo phenotypes. Overexpression of genes that lead to overproduction of proteins can affect the growth rate, stress responses and the properties of the protein itself, and the heterogeneity in copy number within a population results in variation in gene expression level between single cells and hence increase in noise.

One solution to this problem is to integrate the gene into the chromosome. Methods for integrating DNA from plasmids into the chromosome are numerous but often include cloning steps. Boyd et al. constructed a plasmid-chromosome shuttle system for *E. coli* named λ InCh that avoids time-consuming cloning steps [27]. The integration of the plasmid-borne genes into the chromosome is reversible and stable and only one specialized vector is required. Plasmids that confer ampicillin resistance and are derived from pBR322 can be used in the system. Transfer of the genetic material into the chromosome was accomplished in three in vivo steps involving homologous recombination and site specific recombination. In the first step, homology between sequences on a typical pBR322 derived plasmid and sequences on λ InCh permitted a double recombination, which conferred ampicillin resistance to the phage and resulted in pickup of the plasmid insert. The two homologous regions were a region near the origin of replication of pBR322 and fragment of the *bla* ampicillin resistance gene of pBR322. The *bla* fragment did not confer ampicillin resistance itself. A result of the double recombination in the first step was that the kanamycin resistance allele of the phage was replaced both by a complete bla gene, conferring ampicillin resistance to the lysogen, and by the segment between the two homologous regions which included the desired region. In the second step the ampicillin resistance gene in the phage and the linked plasmid insert were introduced into the chromosome of *E. coli* by site specific recombination of the phage into the λ *att* site. This was a recombination between the lambda attachment sites of the phage, *attP*, and *E. coli*, *attB*, that resulted in a lysogen. λ InCh phages that had picked up the region of interest formed ampicillin resistant, kanamycin-sensitive lysogens, in which the expression system was inserted into the chromosome at the lambda attachment site as part of the prophage. λ lysogens are less stable than ordinary chromosomal loci due to the activity of prophage genes that can lead to spontaneous partial induction and the loss or tandem duplication of the prophage DNA. Multiple lysogens are also often obtained when selection by antibiotic resistance is used. Therefore, most of the λ DNA was deleted in

the third step by a second homologous recombination event. To make this step possible, the phage carried a fragment of chromosomal DNA from one side of the *att* site, near *att* DNA, in a position such that most of the lambda DNA could be looped out and deleted by a single recombination event due to a direct repeat of DNA, one inside and one outside of the prophage. Direct selection for lysogenic strains that had deleted all the material between the two repeats was performed by heat induction. The Lambda phage carries the heat-inducible cI857 repressor, so that incubation at 42 °C kills all the cells containing the prophage. All the heat killing genes lie between the direct repeats, thus isolates that had lost the prophage genes in an ancestor of the surviving bacterium were selected. The temperature-independent derivatives were no longer lysogenic or unstable [27]. A detailed protocol for the transduction is to be found in section A.1.1.

2.8 Southern blot

To confirm that fusions were successfully inserted into the chromosome, southern blots were performed. The chromosomal DNA was extracted using the DNeasy kit from Qiagen and digested with EcoRV, PvuI or PstI enzymes, depending on the position and restriction sites of the gene of interest. The digested DNA was loaded on a 1 % agarose gel and separated by gel-electrophoresis at 60 volts for 4-5 hours until the sample had run two thirds of the gel and a smear of DNA was visible under UV light. The gel was photographed with a ruler on the side to know the position of the marker.
The gel was further processed by covering with depurination solution under gentle shaking for 10 minutes until the bromophenol blue had turned yellow. The depurination solution was replaced by distilled water for rinsing before the gel was covered in denaturation solution for 25 minutes until the bromophenol dye had returned to is blue colour. Following rinsing in distilled water, the gel was treated with neutralizing buffer for 45 minutes, with one buffer change after 30 minutes.
For capillary blotting a glass plate was put on a pyrex dish with 20x SSC buffer. A wick made from three sheets of filter paper (3 MM Whatman) was saturated in 20x SSC buffer and put on top of the plate. The gel was placed on the Whatman paper taking care to avoid air bubbles. Cling film was used to surround the gel to prevent SSC being absorbed directly by the paper towels. A sheet

of Hybond ECL nitrocellulose membrane of appropriate size was pre-wet and placed on the gel without trapping air bubbles, on top of which three sheets of pre-wetted Whatman were placed. Finally, a 7 cm stack of absorbent paper towels, a glass plate and a weight completed the stack, which was left over night for blotting.
For DNA fixation the membrane was rinsed with 6x SSC for 1 minute and then treated with UV. The fixed blots could be stored for several weeks wrapped in SaranWrap at 2-8 °C.
YFP was used as a probe to detect the fusions. The probe was diluted to a concentration of 10 ng/ µl denatured by heating for 5 minutes, cooled on ice for 5 minutes and mixed with an equal volume of labelling reagent (ECL Direct Nucleic Acid Labelling and Detection System kit, Amersham Biosciences). Glutaraldehyde solution (same kit) of an equivalent volume was added, and after 10 minutes incubation at 37 °C the mixture could be left on ice for 10-15 minutes. In the mean time, the blot was pre-hybridized for 15 minutes at 42 °C in a hybridization buffer containing 0.5 M NaCl and blocking agent to a final concentration of 5 % (w/v). The labelled probe was added and hybridization was performed over night.
Washing of the membrane was performed two times 20 minutes in 42 °C with the primary buffer, and two times 5 minutes at room temperature with the secondary wash buffer.
Signal generation and detection was performed with the ECL Direct Nucleic Acid Labelling and Detection System kit (Amersham Biosciences) and a blue light sensitive autoradiograaphy film (Amersham Hyperfilm ECL). Long exposure times was necessary to obtain the necessary signal strength.

2.9 Chemotaxis assays

2.9.1 Growth conditions

Overnight cultures were grown in tryptone broth containing suitable antibiotics until the cell density was saturated. After 100 times dilution in 10 ml fresh medium with antibiotics and inducer, the cells were allowed to grow 3,5 - 4 hours at 34 °C in a rotary shaker until exponential growth was reached at OD_{600}= 0,45-0,5. Cells were then harvested by centrifugation (8000 rpm, 1 min) and

washed twice in tethering buffer, resuspended and diluted 20 times in tethering buffer for FACScan measurements, not diluted for microscopy or concentrated 5 times for tethering experiments.

2.9.2 Tethering assay

Chemotactic ability can be quantified by recording the bias of the flagellar motor. This is achieved in tethering assays where single cells are attached to a potassium hydroxide (KOH) treated cover slip by one flagella with an anti-FliC antibody and the rotation of the cell in tethering buffer is observed under the microscope. Cells were grown as described in 2.9.1 and 5 ml cells were centrifuged at 8000 rpm, 1 minute and resuspended in 1 ml tethering buffer. The cells were sheared by passing the 1 ml volume back and forth between two 26G needles on 3cc syringes, connected by intramedic polyethylene tubing. When the fluid moves from one syringe into the other, this is counted as one stroke. Shearing of all flagella but one was usually accomplished after 100 strokes, as confirmed by observing 2 - 3 slow swimmers in a 20x field under the microscope. After shearing, cells were washed free of tethering buffer by centrifugation and resuspending two times. The final cell density for tethering was between 10^7 and 10^8 cells/ml.
Cover slips were cleaned by soaking in a saturated solution of KOH in 95 % ethanol for 5 minutes. After thourough rinsing under a running stream of distilled water, the surface of the glass was very hydrophilic. Tunnel slides were created by making two parallell grease lines across the short axis of a microscope slide about 1 cm apart. Two cover slips were placed on the slide using the grease to secure them about 1 cm apart. Two more grease lines were placed on the surface of the cover slips. The clean, dry KOH treated cover slip was placed on top over the 1 cm space, creating a tunnel of about 40 µl volume through which fluid can flow.
200 µl of cells were added to 100 µl anti-flagellin antibody, and 50 µl of this mixture was flowed onto each tunnel. The whole tunnel slide was inverted and placed in a humid chamber for 30 - 60 minutes. Good tethering gave about 5 spinning cells / 40x objective field. The rest of the cells were either not moving or only jiggling.
The recording of spinning cells was performed at 25 °C with a Zeiss Aiostar plus microscope eqipped with a CCD camera (Panasonic WV-BP330) and a

DV recorder (Panasonic AG-DV1DC). The analysis of the videos was performed using IPS image recognition software from the Visometrics Group.

2.9.3 Swarm assay

E. coli cells with a functional chemotactic system swarm on soft agar plates. The bacteria spread radially due to the attractant gradients formed by metabolizing the different nutrients in the agar. Plates were poured with 35 ml TB with 0,3 % agar containing the appropriate antibiotic and inducer. Cells from colonies on a freshly streaked plate were picked and carefully placed into the fresh swarm agar with a toothpick. The chemotaxis wild type strain (RP437) was alway used as a positive control for swarming efficiency. Swarm assays were used to screen clones by testing constructs for complementation in their corresponding null mutants, eg. a *flhC* clone was tested in a Δ*flhC* strain. For such screenings the cells were incubated at 34 °C for 6 - 8 hours or until the positive control had swarmed approximately 1 cm. Additionally, swarm assays were used to quantify chemotactic ability at different protein concentrations. Here the diameter of the swarm ring was quantified in pixels in ImageJ after photographing with a Canon EOS 300D (DS6041). The swarm diameters were always normalized to the swarm diameter of the positive control. Single-cell expression of fluorescent reporter proteins in swarm assays was quantified by taking cells from specified positions on the soft agar plate after 16 - 24 hours incubation in 30 °C. The cells were washed in tetheringbuffer and agar was removed by centrifuging it down at low speed (3500 rpm, 2 minutes). After attachment to poly-L-lysine treated cover slips the protein levels were quantified by microscopy and imaging as described in 2.10.2.

2.9.4 Stimulus-dependent FRET

Background

Fluorescence Resonance Energy Transfer describes an energy transfer mechanism between two chromophores. A donor chromophore in its excited state can transfer energy by a nonradiative, long-range dipole-dipole coupling mechanism to an acceptor chromophore in close proximity (typically <10nm). This energy transfer mechanism is termed Förster resonance energy transfer[45]. When both

molecules are fluorescent, the term "fluorescence resonance energy transfer" is often used, although the energy is not actually transferred by fluorescence. The FRET efficiency depends on the distance between the donor and acceptor, the spectral overlap of the donor emission spectrum and the acceptor absorption spectrum, and the relative orientation of the donor emission dipole moment and the acceptor absorption dipole moment. FRET depends on the donor-to-acceptor separation distance r with an inverse 6th power law due to the dipole-dipole coupling mechanism:

$$E = \frac{1}{1 + (\frac{r}{R_0})^6} \qquad \text{eq. (2.1)}$$

with R_0 being the Förster distance of this pair of donor and acceptor at which the FRET efficiency is 50 %. FRET is a useful tool to quantify molecular dynamics in biophysics and biochemistry, such as protein-protein interactions, protein-DNA interactions, and protein conformational changes. In this work CheY and CheZ has been used as previously described [147] to monitor FRET responses to chemotactic stimulation. The FRET pair was expressed from the same promoter on the plasmid pVS88, where *cheY* was fused to *yfp* and *cheZ* to *cfp*. The fusions to CheZ and CheY interact in a phosphorylation-dependent manner. Attractant inhibits the kinase, leading to a decrease in FRET. This is observed as a decrease in YFP/CFP ratio due to the reduced numbers of CheY-P-CheZ complexes. Conversely, FRET increases upon removal of attractant. Hence, measuring the YFP/CFP ratio gives a direct readout of kinase activity.

Preparation of cells

FRET experiments were carried out in different strains all containing the plasmid pVS88 which is encoding CheY-YFP and CheZ-CFP expressed together from the pTrc promoter. Cells were grown as described in 2.9.1, harvested by centrifugation at 4000 rpm for 10 minutes, and resuspended in 10 ml tethering buffer. To prevent further growth and protein production, the cells were incubated (4 °C) for 30 - 60 minutes prior to the measurements.
Coverslips (round, 12 mm diameter) were treated with 40 µl poly-L-lysine for 20 minutes. The poly-L-lysine was removed by pipetting, and shortly before the measurement, residues were washed off by pipetting 100 µl water up and down five times. 1 ml of the cooled cell suspension was centrifuged 1 min at 8000 rpm

and the pellet was resuspended in the residual 20 - 40 µl of tethering buffer. A liquid drop of this high density cell suspension was placed in the centre of a prepared cover slip and left for 20 min at room temperature for the cells to attach in a dense monolayer. Apiezon grease was used to attach the cover slip to the flow chamber. It was important to keep the chamber sealed and free of air, and as soon as the cover slip was attached, the pump was started to prevent cell death by oxygen depletion.

Data aquisition

A Zeiss Axiovert 200 microscope was used to perform FRET experiments. A 40x Objective (Zeiss Plan-Neofluar 40x/0.75) was used to focus a section of a dense monolayer of about 500 - 600 cells. Cells were excited at a CFP-specific wavelength (436/20 nM) from a 75W xenon lamp attenuated by factor 500 with neutral density filters. A 455 nm dichroic mirror was used to split the excitation and emission light in two spectral parts. CFP and YFP emission were detected through 480/40 nm band pass and 520 long-pass emission filters respectively and detected by photomultiplier tubes, PMTs (Hamamatsu photon counting head). A PCI-6034E counting board connected to a computer with LabView7/Template software was used for data aquisition as counts of detected photons per second.
During stimulus dependent FRET measurements, the flow chamber was kept under constant flow (0.3 ml/min, 1 ml/min for repellent measurements) of tethering buffer. A stable YFP/CFP ratio was established after approximately 20 minutes (base-line). For a change from buffer to attractant (the non- metabolizable α-methyl-aspartate (MeAsp) or serine) or repellent (nickel chloride) the pump was stopped and a clamp used to close the tube while changing to another vessel. Amplitudes of the ratio were measured in a range from 30 μM to 1 mM attractant or repellent for 1 - 2 min. For determination of adaptation time, the attractant was added for so long until base-line level had been reached, the time ranging from 1 min to 20 min depending on the strain measured. FRET is equivalent to kinase activity and was calculated from changes in ratio of YFP and CFP signals with the help of the following equation:

$$FRET = \frac{\Delta R_{max} - \Delta R}{\frac{\Delta Y}{\Delta C} + R_0 + \Delta R_{max} - \Delta R} \qquad \text{eq. (2.2)}$$

where Δ R_{max} represents the change of the $\frac{\Delta Y}{\Delta C}$ ratio under maximal stimulation with a saturating amount of α - methyl aspartate.
$\frac{\Delta Y}{\Delta C}$ represents the specific constant (= 1.1 or 2.4 depending on beamsplitter) defining the ratio of the increase of YFP emission compared to the increase of CFP emission after a stimulus.
R_0 is the YFP/CFP emission ratio in absence of a stimulus (= base-line) and Δ R is calculated as the difference of the ratio before and directly after addition of attractant.

2.10 Quantification of protein levels

2.10.1 FACScan

Flow cytometry is a powerful technique for analyzing large pools of single cells. The mechanism involves passing individual cells through the path of a laser beam by holding the cells in a thin stream of fluid. The light scattered from each cell can produce information related to cell size and density. The side scatter measured is dependent on the granularity of the cell, whereas the forward scatter varies with the size of the cell. In addition, cells can be labelled with fluorescent dyes specific for different structures (eg. DNA) or physiological conditions (eg. pH). The resulting fluorescence from the cells is collected by photomultiplier tubes (PMTs) that convert the light to a voltage or electrical output that is proportional to original fluorescence intensity and the voltage on the PMT so that the data can be recorded digitally and quantified.
Fluorescence activated cell sorting (FACS) allowed us to quantify the distribution of YFP fluorescence levels in a large cell population (fig. 2.1 A). A dilution series of cells induced with 1 mM IPTG was made to find the optimal cell density for the FACS measurements. Dilutions ranging from 5-fold to 100-fold were performed and gave similar results. The 20 - fold dilution was then selected for further measurements since it showed good expression curves and was already being used. Single-cell fluorescence measurements were carried out with a Beckton-Dickinson FACScan flow cytometer with a 488 nm argon excitation laser and a 525 nm emission filter, limiting the measurements to YFP - fusions only in our case. FACScan data were analyzed using the programme CellQuest-Pro 4.0.1. Also cells without a fluorescence marker were measured as negative

controls. The set up for the scanning was the same for all experiments. A million cells were measured in about two minutes while the apparatus was set to slow flow with almost 10000 events per second. Only fluorescence from cells in a defined area of the plot was saved so that the smallest possible gate in the side-scattering and forward scattering space were used to exclude noise due to different cell sizes and to subtract background fluorescence from the buffer. The laser had a power of 15,05 W and voltage of 6,4 V. Laser current was between 5,2 and 5,3 A. The voltages on the photo multiplier tubes, were 500 V for the tubes detecting side scatter and 800 V for the tubes detecting fluorescence. An amplifier gain of 6,23 was chosen for the tubes detecting forward scatter, and 8,35 for the tubes detecting the side scatter. Dot plots were observed for side scatter as a function of forward scatter and fluorescence as a function of side scatter.

2.10.2 Imaging

The advantage of microscopy and imaging is the possibility to quantify the levels of several fluorescence markers in single cells. Additionally, by measuring the total fluorescence and dividing by cell length or volume provides a measure for protein concentration. For protein concentration measurements the hemispherical cell poles are neglected, thus the cells are for simplicity assumed to be rectangular. Histograms of fluorescence values obtained from imaging resemble the ones obtained from FACS measurements, proving that both methods are suitable for quantification of protein levels (fig. 2.1).
Cells were grown and induced as described above. The cells were attached to poly-L-lysine treated cover slips. 100 µl poly-L-lysine was put on the cover slip and incubated for 10 minutes before it was rinsed off twice with water. 100 µl cells were incubated on the slide for 10 minutes and excess cells were then rinsed off 3-5 times with 1 ml tethering buffer. The cover slip was then placed onto the slide and any excess fluid was gently wiped off. For time-lapse imaging with growing cells, 10 µl of cell suspension was placed on thin agarose pads that contained 10 % tryptone broth. The agarose pad was sealed in a custom-made aluminium slide using coverslips on both sides; the cell suspension for these experiments was taken directly from the fresh growing culture, without washing in tethering buffer. The slide was placed into a custom-made temperature-controlled holder connected to a water bath (Lauda Ecoline Staredition RE104),

with the measured temperature at the stage set to about 35 °C. Florescence microscopy was performed on a Zeiss Axio Imager .Z1 microscope equipped with an ORCA AG CCD camera (Hamamatsu) and HE YFP (Excitation BP 500/25; Dichroic LP515; Emission BP 535/30) and HE CFP (Excitation BP 436/25; Dichroic LP455; Emission BP 480/40) filter sets. Images were made every 5, 20 or 30 minutes depending on which resolution was needed. A variety of image processing software is available for quantification of fluorescence as a measure of expression. The intensity in approximately 100 - 300 cells at each induction level was quantified in ImageJ version 1.37v (Wayne Rasband, NIH, USA, http://rsb.info.nih.gov/ij/). Integral density and mean intensity of the fluorescence in single cells as well as cell length were quantified using an automated custom-written ImageJ plug-in.

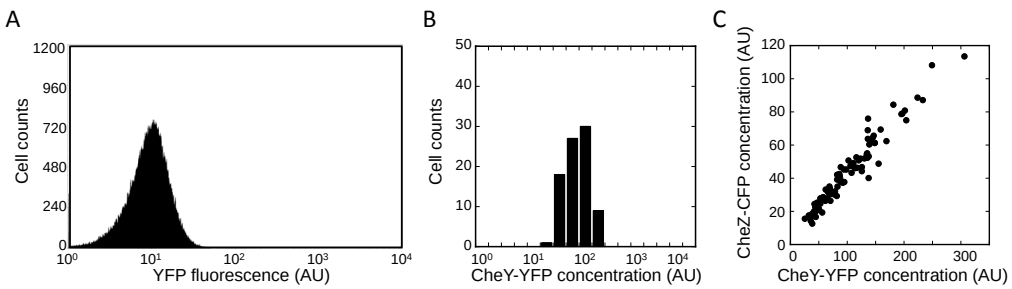

Figure 2.1: Methods for quantification of protein levels. (A) Distribution of YFP fluorescence in Δ*(cheY-cheZ)* cells expressing CheY-YFP and CheZ-CFP together from the pTrc promoter on plasmid pVS88 obtained by FACS. (B) Same as (A) however, distribution is obtained from imaging. Relative concentration values are acquired by dividing total fluorescence by cell length (C) Single cell concentrations of CheY-YFP and CheZ-CFP obtained from the same images as in (B) showing the cell-to-cell variations and correlations in protein levels. No inducer was added.

2.10.3 Fluorimetry

To calibrate fluorescence intensity in FACS and imaging data, a PerkinElmer LS55 luminescence spectrometer was used to determine the absolute number of reporter proteins in control cells. Cells were grown as described in 2.9.1 and washed in tethering buffer before dilution to adjust the cell density to be equal for all cell cultures. 2 ml of the cells were sonicated with a Branson Sonifier 450 three times ten seconds at constant intermediate power output, until complete lysis. YFP fluorescence was measured at 510 nm excitation and 560 nm emission

in a volume of 300 µl in a quarts cuvette. Sonicated cells without a fluorescence reporter were used as a negative control, and their autofluorescence was subtracted from all values as background. A solution of purified YFP with known concentration, determined by Bradford assay, and absorbance measurements by a Specord205 spectrophotometer from Analytik Jena, was used to produce a calibration curve, relating fluorescence to molecule number. Cell number in 1 ml culture was counted using a Neubauer counting chamber, and cell volume was determined by measuring cell width and length by imaging. These values for one culture were used to provide a conversion factor from FACS or imaging values to single-cell protein levels.

2.11 Data analysis

Data obtained from CellQuestPro after flow cytometry measurements and from ImageJ after microscopy was analyzed in the programme Kaleidagraph version 3.6 (http://www.synergy.com/). Statistics were obtained to achieve the mean values of expression. Noise values were calculated as the standard deviation in the distribution of expression values divided by the mean. Histograms were made with a fixed number of bins to compare the distribution of cells at different expression levels in FACS experiments and microscopy experiments.

2.12 Computer simulations and bioinformatics

2.12.1 Modelling the signal transduction pathway

The dynamical behaviour of the networks was described on the level of ordinary differential equations assuming Michaelis - Menten kinetics. The kinetic constants were taken from *in vivo* and *in vitro* measurements [www.pdn.cam.ac.uk/compcell], except for the methylation process and the CheY-CheZ feedback loop shown in fig. 3.27 D, where constants were determined from optimization of a population of 70 cells with respect to their least deviation from the midpoint of the motor response curve [34] (3.2 μM) under the experimentally determined gene expression noise [74]. Fully chemotactic cells were defined as cells for which the level of CheY-P was within the interval 2.2-4.3 μM [34] and for determination of their proportion in fig. 3.11 and

fig. 3.28 a population size of 10^4 cells was assumed. Expression levels for the chemotaxis strain RP437 where taken from [88]. Scatter plots were analyzed using a correlation coefficient defined by:

$$\rho = \frac{mean(x * y)}{std(x) * std(y)} \qquad \text{eq. (2.3)}$$

The mathematical model includes all known protein interactions among CheR, CheB, CheY, and CheZ. The adapted receptor activity is determined by the methylation level and consequently by the ratio between receptor-bound CheR and CheB, allowing us to omit all details of transient adaptation kinetics. The relation of phosphorylated CheY to the flagellar motor rotation bias follows the experimentally determined motor response curve. Our mathematical model reflects the experimentally observed robustness of the pathway output against concerted overexpression of all chemotaxis proteins but shows the expected sensitivity to independent variations in protein levels. Effects of translational noise on protein concentration has been simulated by Gaussian random variables with mean zero and standard deviation over mean of 0.05 to arrive at the experimentally observed cell-to-cell variations of the CW bias [74]. The translational coupling constant is set to 25 % of the mean translational efficiency to generate the rank list (fig. 3.30). A minimum of 10^5 individuals are needed in our computer simulations to assure that that each position of the rank list (fig. 3.30) does not change place beyond its nearest neighbours to 95 % confidence, as shown by data resampling using bootstrap. The influence of transcriptional noise or extrinsic noise on the gene order was not significant as both CheY-P level of our chemotaxis pathway model and measured clockwise rotation bias [74] are almost insensitive to increased transcriptional activity.

2.12.2 Secondary structure predictions

Secondary structure of the *meche* mRNA sequence was predicted using the mfold programme [179], with pairing being restricted to bases closer than 600 to each other.

2.12.3 Analysis of gene order

Analysis of the consensus order of chemotaxis genes was performed using a custom written Perl programme. The programme scanned flat files of 392 microbial genomes from GenBank database using variable regular expressions to identify chemotaxis genes in the annotation. Upper and lower case letters, variable delimiters, order of chemotaxis terms and synonymous terms were taken into account. Due to incomplete annotation of many genomes, chemotaxis genes were selected in only 245 of them. Closest upstream and downstream neighbours of the selected gene were also recorded. In cases of ambiguous gene annotation, a BLAST analysis was performed to test whether the gene is homologous to one of the chemotaxis genes in other bacteria. The resulting triplets were used to calculate occurrences and relative order of chemotaxis genes.

2.12.4 Modelling movement of bacteria in gradient

RapidCell, a hybrid model of chemotactic *E. coli* that combines the MWC signal processing by mixed chemoreceptor clusters, the adaptation dynamics described by ODE, and a detailed model of cell tumbling. The model dramatically reduces computational costs and allows highly efficient simulation of *E. coli* chemotaxis. The MWC model was employed for mixed receptor clusters [43] with a mean - field approximation for adaptation kinetics [55]. An individual receptor homo dimer is described as a two-state receptor, being either 'on' or 'off'. Receptors form clusters with all receptors in a cluster either 'on' or 'off' together. The clusters are composed of mixtures of Tar and Tsr receptors. Both CheR and CheB are assumed to bind receptors independently on their activity. A bound CheR (CheB) can (de-)methylate any inactive (active) receptor within the assistance neighbourhood. To couple the bias of individual motors to the probability of tumbling, we applied a voting model for several independent motors, based on detailed experimental investigation

of tumbling mechanics [163], where the cell switches from 'Run' to 'Tumble', if at least 3 of its 5 motors rotate CW, and from 'Tumble' to 'Run', if at least 3 of 5 rotate CCW. In the model of the cells environment, different gradient forms were used (linear, Gaussian, nominal constant-activity) with different steepnesses.

These components were combined into a new simulator for *E. coli* chemotaxis RapidCell, which uses a hybrid pathway simulation instead of fully stochastic or ODE approaches, and is therefore computationally cheap. This allows to simulate populations of 10^4 - 10^5 cells on the time scale of hours using a desktop computer.

3 Results

3.1 Characterization of noise in expression of flagella and chemotaxis genes

For characterization of noise in the chemotaxis pathway we quantified protein expression levels by different methods as described in section 2.10.1 and 2.10.2.

3.1.1 Fluctuations in protein levels in the flagella assembly and chemotaxis pathway

As mentioned in section 1.2.8, flagella and chemotaxis genes are organized in a regulon consisting of 14 operons. These are arranged in a hierarchical fashion in three promoter classes where expression from one class can affect expression at a lower level in the hierarchy. To locate the origin of noise in the hierarchy, *yfp*- and *cfp* fusions to some of the genes (marked as red boxes in fig. 3.1) in all three classes were constructed. Each gene was replaced by its gene fusion by homologous recombination (2.7.1). The gene fusions were thus expressed from the native promoter on the chromosome. First, the distribution of single cell CFP concentrations in populations of cells was quantified by imaging, for CFP fusion to CheA in class 3, FlgM in class 2 and 3, FliM in class 2, and FlhC in class 1. Relative variations in protein levels in the three classes was evident from distributions of absolute concentrations in fig. 3.2 A. Normalized concentration distributions, (fig. 3.2), show that the width of the cell-to-cell distribution and hence the noise was larger for the proteins in class 3 than in classes 2 and 1. In all classes the distribution of CFP concentrations showed a tail towards higher protein concentrations (log-normal distribution) which corresponds well to the distribution of CheY-YFP seen in fig. 2.1. Log-normal distributions are frequently

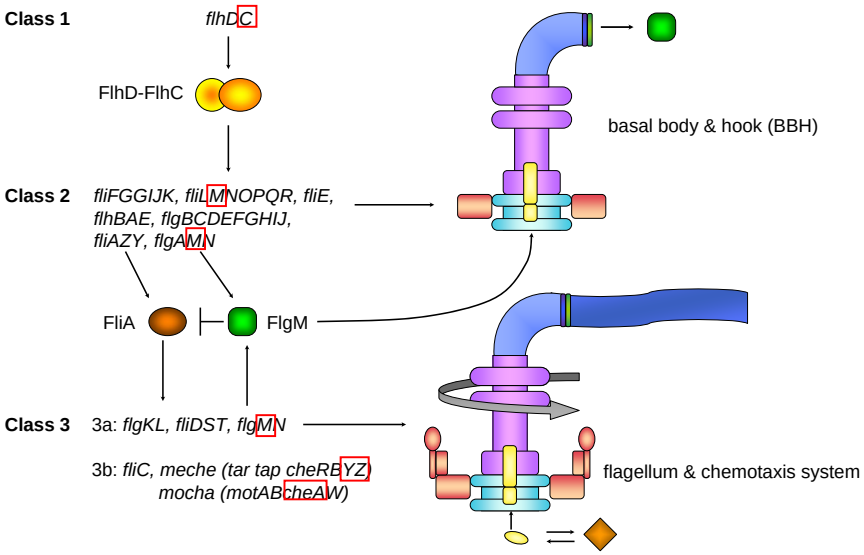

Figure 3.1: Fusions made in the hierarchy of flagella and chemotaxis genes. FlhC in class 1, FliM in class 2, FlgM in class 2 and 3 and CheA in class 3 were fused to CFP or YFP in order to quantify gene expression and localize the origin of noise in the hierarchy. Strains are described in table 2.2.

observed in nature and usually originate from stochasticity of promoter activation and transcription bursts. In other words log-normal distributions are typical for systems where variations in an upstream component affect downstream components such that random independent effects are multiplicative. Additive effects or a one-gene system would rather result in a normal distribution [15].

To investigate the origins of gene expression noise in the chemotaxis pathway further, single cell levels of the fluorescent reporters were compared between classes in the hierarchy by microscopy and imaging, using YFP as a reporter for CheY expression, and comparing its expression to that of other proteins in all three classes. CheA and CheY were chosen as representatives for class 3 due to their location in two different operons (*mocha* and *meche* respectively), which are both under the control of the same transcriptional activator, the sigma factor σ^{28}, also called FliA. Fig 3.3 shows scatter plots of normalized relative concentrations of YFP and CFP from four experiments, where each point represents a single cell. Pooled results from four experiments presented in different colours show that the day-to-day variations in expression levels is moderate. Strong correlations in expression levels were observed when comparing two genes in class 3, *cheY* and *cheA*. Expression of *cheY-yfp* (class 3) and *fliM-cfp* (class 2) or *flgM-cfp* (class 2/3) shows weaker correlation, similar to the correlation observed between the class 1 gene *flhC* and the class 3 gene *cheY* (fig. 3.3 A-C).

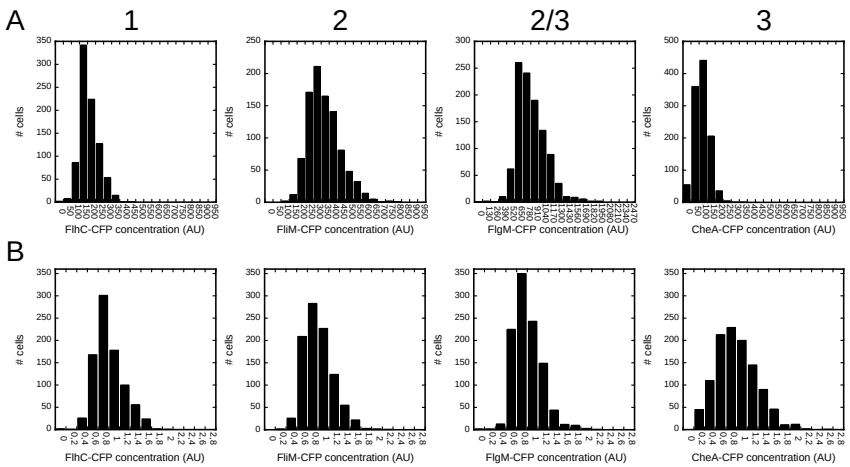

Figure 3.2: Distribution of single cell CFP levels at different levels in flagella gene hierarchy. FlhC in class 1, FliM in class 2, FlgM in class 2 and 3 and CheA and CheY in class 3 are expressed from the native promoter on the chromosome as fusions to CFP. Distributions of single cell CFP concentrations (A) and concentrations normalized to the mean (B). AU, arbitrary units. Pooled data from 4 experiments are shown. Concentrations were calculated as total fluorescence divided by cell length (2.10.2).

Additionally, we made a YFP fusion to FliM and compared the single cell levels of FliM-YFP to FlhC-CFP and FlgM-CFP in order to compare class 1 with class 2 and class 2 with class 2/3. Also here, correlation and the wider distribution in class 3 was observed. The strong co-variation in the expression levels of individual proteins, suggests that extrinsic noise is dominating the variations [41, 124]. This coupling is stronger for genes that belong to the same class, indicating contributions of factors specific for this class, for instance concentration of the sigma factor for class 3 genes. However, even unrelated genes seem to show strong correlation in *E. coli*, as evidenced by our negative control, where YFP inserted elsewhere in the chromosome and expressed from the pTrc promoter, is compared to FlhC-CFP (fig. 3.4 B). Moreover, also autofluorescence in YFP and CFP channel correlates (fig. 3.4 C). The reason for this correlation remains unclear. However, at such low fluorescence levels there might be a crosstalk between the two channels. Another possible reason might be that autofluorescence generally varies from cell-to-cell, in a normal distribution, and could be affected by growth or other factors independent on which compounds are causing the fluorescence leading to the same effect seen in both channels. The fact that correlation is observed between all the classes and even between non-reated genes, suggests that global factors affecting upstream genes contribute significantly to

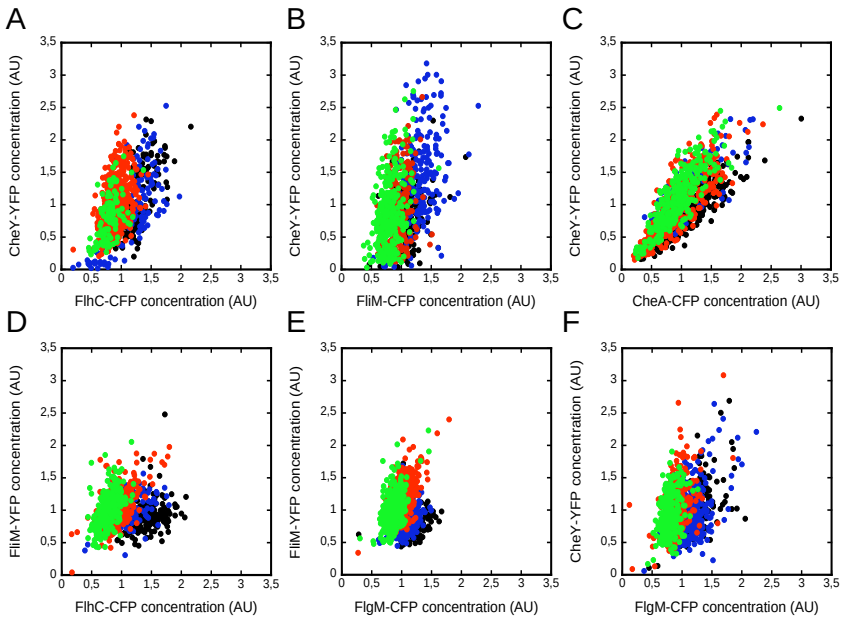

Figure 3.3: Scatter plots of single cell YFP and CFP levels in all three classes of the hierarchy. Normalized relative concentrations of CheY-YFP in class 3 compared to (A) FlhC-CFP in class 1, (B) FliM-CFP in class 2, (C) CheA-CFP in class 3 and (F) FlgM-CFP in class 2/3. Also FliM-YFP was compared to (D) FlhC-CFP and (E) FlgM-CFP. Results from four independent experiments are pooled, and each experiment is shown in a different colour.

the gene expression noise observed in the chemotaxis pathway. This is consistent with previous studies showing that much of extrinsic noise is of global nature [22, 128].

3.1.2 Testing possible additional noise sources

Another possible source of noise could be the sub-cellular localization of the proteins. Figure 3.5 shows YFP images of four of the compared double fusion strains. In LL13 CheY-YFP did not localize to the chemosensory complex but remained in the cytoplasm. Docking of CheY to CheA might be sterically hindered by the relatively large CFP molecule. We tested the effect of subcellular localization on the noise in protein levels by imaging. We expressed *cheY-yfp* from a plasmid in Δ*(Y-Z)* and Δ*(Y-Z)*ΔA strains. In absence of CheA, CheY remained in the cytoplasm, whereas in presence of CheA it localized to the receptor clusters. Both strains showed similar noise levels. Additionally, expression of YFP-CheA was tested in Δ*(Y-Z)*ΔA and Δ*(Y-Z)* Δ*(tsr tar tap trg aer)*

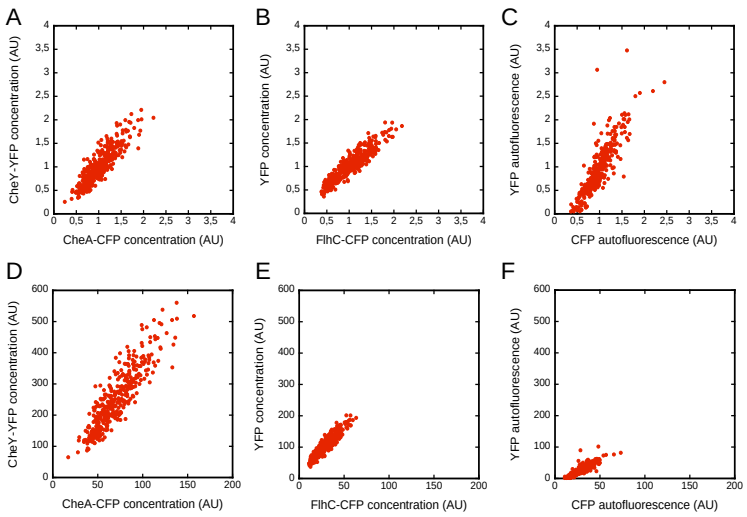

Figure 3.4: Correlation in YFP and CFP channel in negative controls. Concentrations in (A) LL13, (B) strain (LL19) expressing FlhC-CFP from the native promoter on the chromosome and YFP from a pTrc promoter elsewhere in the genome. (C) Chemotaxis wild type strain (RP437) without any fluorescence reporter, (D-F) same as (A-C) but absolute values.

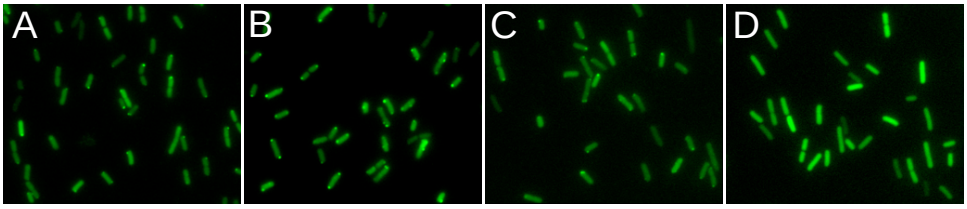

Figure 3.5: CheY-YFP expressed from the native promoter on the chromosome in the strains (A) LL16, (B) LL14, (C) LL13 and (D) LL15. CheY-YFP does not localize to the receptor clusters when CheA is expressed as a fusion in the same strain.

strains, where in the receptor deleted strain CheA remained in the cytoplasm. Also here, localization did not affect gene expression noise (fig. 3.6), hence noise is independent on sub-celluar localization of the proteins and usage of strain LL13 is appropriate for studies of noise in the hierarchy.

The assembly of flagella is a highly regulated process with several feedback mechanisms. σ^{28} has also been suggested to activate class 2 promoters [63], although in an unknown way. In addition, FliZ is suggested to activate class 2 promoters and regulate the concentration of the master regulator $FlhD_4C_2$ (see chapter 1.2.8). Such a positive feedback from class 2 to class 1 could amplify noise in downstream gene expression promoter providing an additional source for variations in class 3 protein levels. Hence, we tested how the levels of FlgM

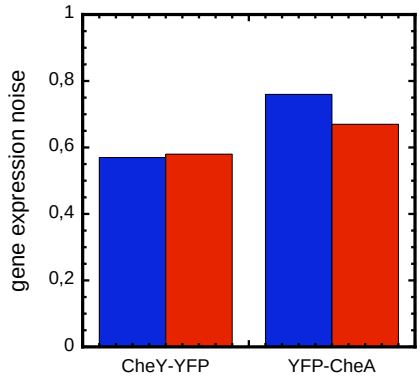

Figure 3.6: Clustering of CheY-YFP and YFP-CheA and its effect on noise. Noise in CheY-YFP (pVS18) and YFP-CheA (pVS55) levels when clustering, in in VS104 and VS168 respectively, (blue) and not clustering, VS168 and VS181 respectively, (red). Both fusions are expressed under full induction.

and FliA affect the expression in other genes in class 2 and 3. As shown in fig. 3.7, at increasing concentrations of the inhibitor FlgM expressed from an IPTG inducible promoter, CheY-YFP expression decreased, whereas FliM-CFP levels were unaffected. The same result was obtained by varying FliA levels, with overexpression of FliA increasing the CheY-YFP level, but not affecting FliM-CFP levels. Surprisingly, at very high levels of FliA, a part of the population had strongly increased CheY-YFP expression, whereas the other part had wild type levels of CheY-YFP. This phenomenon was not observed in the Δ*flgM* strain. We speculate that a possible cause might be bi-stability of transcription due to the high level of transcription factor. Nevertheless we can conclude that overexpression of FlgM or FliA has no effect on class 2 gene expression in the flagella gene hierarchy.

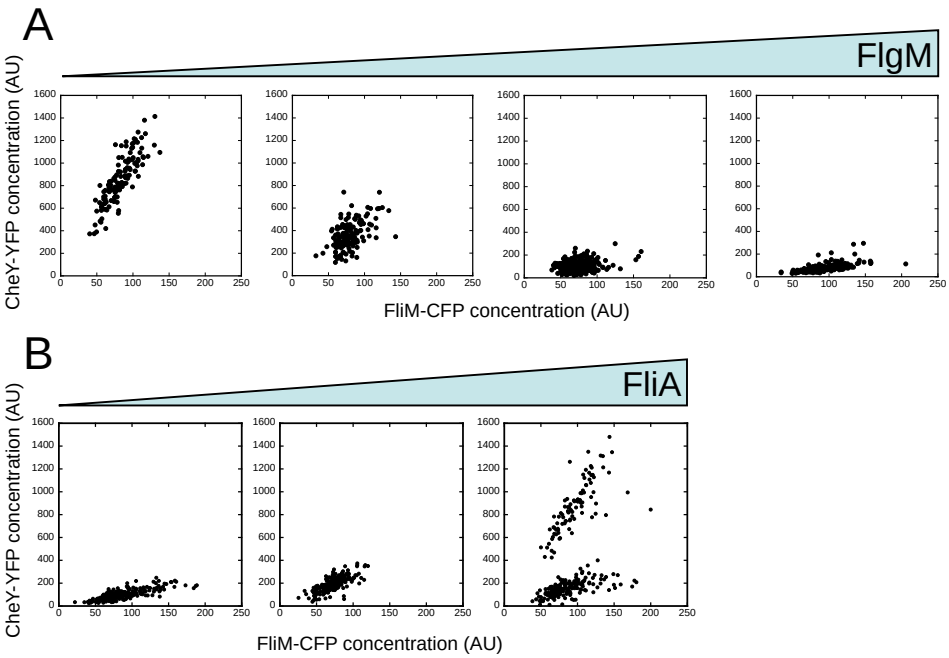

Figure 3.7: Expression of CheY-YFP and FliM-CFP at varying levels of the sigma factor FliA and its antagonist FlgM (A) At increasing concentrations of the inhibitor FlgM expressed from a pTrc promoter (pLL16) in a *cheY-yfp, fliM-cfp*Δ*flgM* strain (LL10), CheY-YFP expression decreases, whereas FliM-CFP levels remain unchanged. (B) overexpression of FliA (pLL34) in a *cheY-yfp, fliM-cfp* strain (LL14) increases the CheY-YFP level slightly. At very high levels of FliA a part of the population has strongly increased CheY-YFP expression. FliM-CFP levels are not affected.

3.1.3 Time scale of variations

Static snapshots to quantify steady state noise do not provide information on the temporal dynamics of gene expression. To assess how the protein levels in single cells vary with the cell cycle, high resolution time-lapse imaging (2.10.2) of a strain expressing *cheY-yfp* in the genome from the native promoter was performed. Total protein levels (integral density) were measured as well as concentrations (chapter 2.10.2 and fig. 2.1). Fig 3.8 shows the variation in total protein levels over time. At cell division (e.g. time points 5 min and 60 min) the total protein level in the cell drops dramatically. Fig 3.8 A shows that the two daughter cells contain about half the amount of CheY-YFP of their mother cell. As the two daughter cells grow, their protein levels increase until the next cell division, where the protein levels are halved again. On the other hand, protein concentrations remain relatively stable throughout the cell cycle (fig. 3.8 B). A

plot of cell length over time (fig. A.4) is almost identical with the total fluorescence over time, revealing that cell size decreases with time. Such cell shrinkage can be an effect of the growth conditions and nutrient depletion towards the end of the time lapse experiment. An average of the concentrations at each time point is plotted in fig. 3.8 C. The relatively small standard deviations signify how precisely the cells keep their protein concentrations stable. Taken together, these results suggest the presence of mechanisms keeping the level of CheY-P stable throughout the cell cycle to keep the cells optimally chemotactic, independent of growth stage.

Next, we studied the effect of growth on gene expression noise in CheY-YFP levels. For 5 cell lines, the cell-to-cell variation in concentration (integral density/length) of CheY-YFP between the daughter cells was quantified over time as SD/mean. As fig. 3.8 D shows, noise in CheY-YFP levels in these experiments seems to stabilize after 3 hours at about 0.15. A total noise of 0.15 was significantly lower than the expected 0.67 measured by FACS in a strain expressing CheY-YFP (fig. 3.11 A). Therefore, we also quantified the cell-to-cell variation in CheY-YFP levels among all the cells in the same time-lapse images. Here the noise decreased with time and stabilized on a level of 0.25 after 3 hours. Over time the noise seems to come together after 150 minutes. The change in noise over time might be an effect of the growth on agarose. Due to nutrient depletion the cell division process might slow down. To ascertain whether growth rate affects the variation in protein numbers, we cultured cells in different tryptone broth concentrations, from 20 % to 100 %, and quantified the cell-to-cell variations in their CheY-YFP levels. Figure 3.9 A shows the growth curves obtained and fig. 3.9 B shows the decrease in cell-to-cell variations in CheY-YFP concentrations with growth rate. Hence, the low noise values are not due to slow growth. This seemingly low level of noise compared to the total noise in CheY-YFP levels measured by FACS is due to the fact that background is not subtracted. Due to bleaching effects, short exposure times were used leading to low fluorescence which made data analysis impossible after background subtraction.

Furthermore, we compared the time scale of variation in protein levels over generation time. This time cells of the same strain were imaged every 30 minutes, and division occurred approximately once per hour, where the first division started after 30–60 minutes. Figure A.5 shows the relative CheY-YFP concentration in four representative cell lines. The descendants of the first two daughter

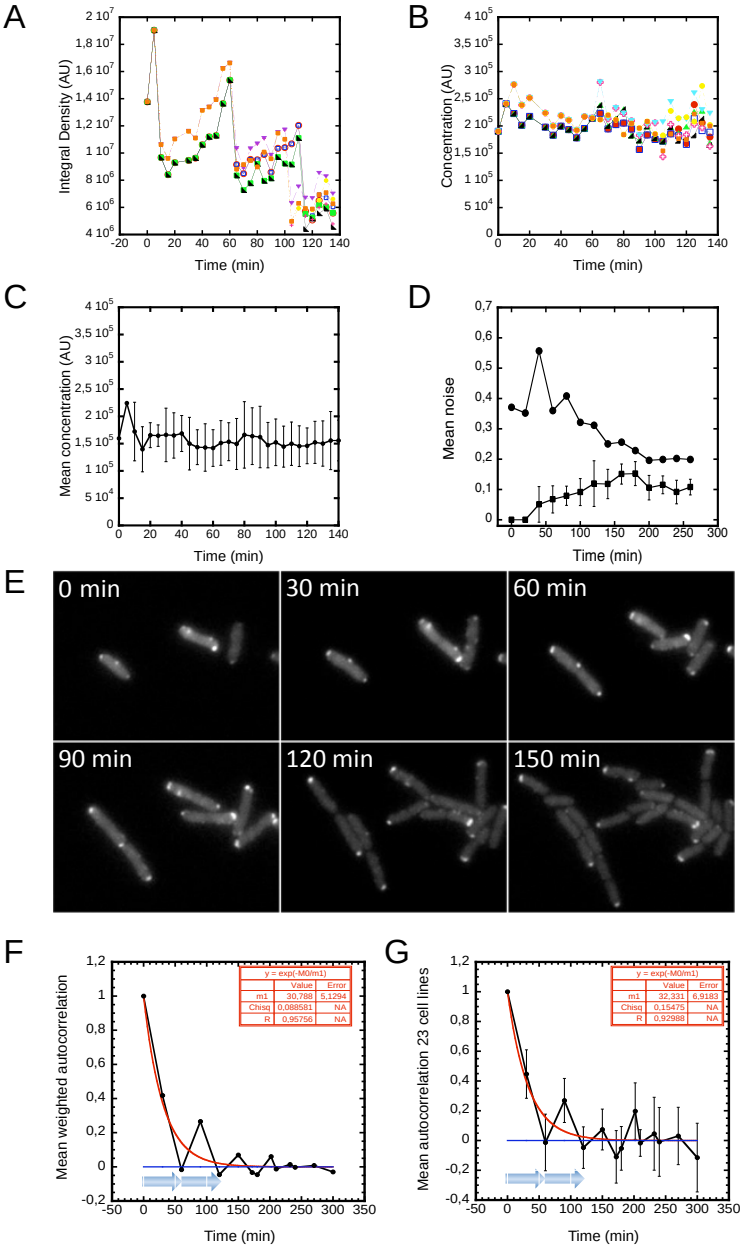

Figure 3.8: Variations in CheY-YFP levels over time. (A) Images of strain LL12 (CheY-YFP) were taken every 5 minutes. Integral density (total fluorescence) of YFP in cells descending from one mother cell. Different colours denote different daughter cells. (B) Concentrations calculated as integral density divided by cell length for the same cell line. (C) Mean concentration of all cells in the same cell line at each time point. Error bars represent standard deviations. (D) YFP images of strain LL14 were taken every 20 minutes. Mean noise in concentration of CheY-YFP between all the daughter cells in 5 cell lines (squares) and between all the cells in the image at each time point (circles). Error bars represent the standard deviations of the mean noise. (E) Images of three cell lines dividing three times over 150 minutes. (F) Autocorrelation function for concentration over time in dividing cells, simplified in the way that only signs were considered, not the size of the variation. Autocorrelation was averaged for 4 experiments where the values have been weighted according to number for cell lines in each experiment. (G) Autocorrelation function as mean of signs of all 23 cell lines with error bars showing the standard error. Both are fitted with the function $e^{-t/\tau}$.

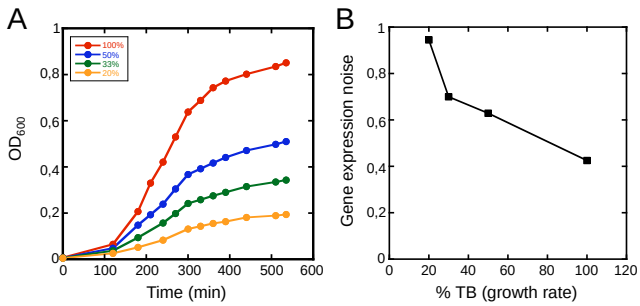

Figure 3.9: Dependency on growth rate on noise in CheY-YFP levels. (A) Growth curves for LL14 at four different TB concentrations. (B) Cell-to-cell variations in concentrations of CheY-YFP (in LL14), quantified from imaging, as a function of TB concentration and hence growth rate.

cells are shown in either red/orange or blue/green. Here, the cells divide on average 5 times within 300 minutes. Growth rate seems to slow down after 4 hours, probably due to lack of oxygen and/or nutrient depletion.
Images of three typical cells dividing three times within 150 minutes show that at timepoint 0, the mother cells have different CheY-YFP levels. The descendants of the brighter cell remain brighter than the others for a while. However, judging by eye, after 90 minutes all the daughter cells have the same fluorescence intensity. Hence, it seems to take approximately 60-90 minutes, or 1-2 generations until the protein levels in the cells do not correlate anymore (fig. 3.8 E). This was also quantified by measuring the intensities in the cells, concentrations, and using an autocorrelation function to describe how the protein levels in the cell at a time point correlate to the protein levels in the original mother cell. A data set of 23 cell lines from 4 different experiments was assessed. The concentration at each time point was subtracted the mean concentration of all cells at that timepoint. For the first, the size of the difference was not so interesting for us, only if it was positive or negative, thus we set positive values to +1 and negative to -1. All values were multiplied with the difference obtained at time point 0 and then divided by the number of cell lines as shown in eq. (3.1).

$$R(t) = \frac{\Sigma[sign(F_t - M_t)sign(F_0 - M_0)]}{N} \qquad \text{eq. (3.1)}$$

where R(t) is the autocorrelation function, t is the time in minutes, F is the fluorescence, M is the mean fluorescence and N the number of cell lines. When the autocorrelation function reaches zero, the protein levels in that cell are totally independent on the protein levels in the first generation. Figure 3.8

shows the resulting autocorrelation function. In fig. 3.8 F the average of the autocorrelation from 4 experiments is shown where each experiment has been weighted according to the number of cell lines measured in this experiment. Weighting errors in data sets that have already been averaged is not straightforward, hence error bars are not shown here. Pooling the signs in all the 23 cell lines, however, allowed us to calculate the standard error (fig. 3.8 G) and the result was similar to that in fig. 3.8 F. When data are fitted with an exponential decay (shown in red), the decay time, τ, was $\sim$30 which means it takes $\sim$30 minutes until the correlation has been reduced to a fraction 1/e. Hence, we could confirm that correlation in protein levels persists for about one generation.

3.2 Robustness of chemotactic signalling against gene expression noise

A key feature of the chemotaxis pathway is precise adaptation–its ability to return to the same level of pathway activity under conditions of continuous stimulation. Due to the steepness of the motor response curve (fig. 3.10), the CheY-P level in the cell has to be within a narrow concentration range. With CheY-P levels outside of this very narrow range, most of the sensitivity is lost, and the cell is unable to respond properly to a gradient (eg. spread in soft agar on swarm plates). A CheY-P concentration higher than 4.6 μM would lead to motors locked in CW rotation and continuos cell tumbling. A CheY-P level of below 2.3 μM, would result in smooth swimming with all motors rotating CCW [34]. Hence, *E. coli* has to keep CheY-P level and receptor- and kinase activites stable despite of great variations in protein levels, in order to navigate in a chemical gradient. Here we study the effects of gene expression noise on the chemotaxis pathway, and the means by which the cells compensate for such fluctuations.

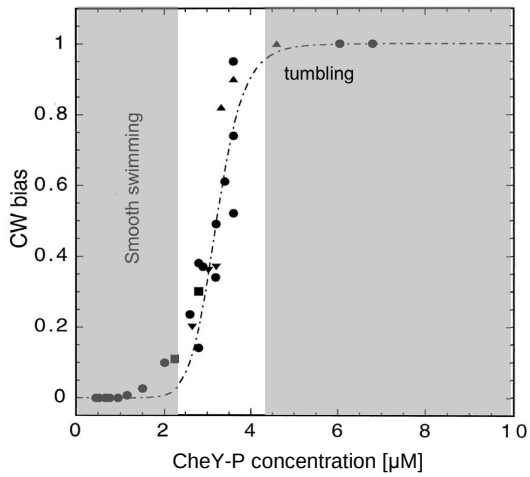

Figure 3.10: Characteristic response of individual motors as a function of CheY-P concentration. Each data point describes a simultaneous measurement of the motor bias and the CheY-P concentration in an individual bacterium. The dashed line shows the best fit obtained with a Hill function (Hill coefficient N_H= 10.3±1.1 and K_M= 3.1 μM). Motors were locked in (CW) state for tested CheY-P concentrations ranging from ∼4.6 to 25 μM. Data taken from [34].

3.2.1 Effect of co-variation of all chemotaxis proteins on chemotaxis

First, we experimentally determined the effect of co-variation in the levels of all signalling proteins on chemotactic behaviour (fig. 3.11). Chemotaxis efficiency was measured in Δ*flgM* cells expressing varying levels of FlgM from an IPTG-inducible plasmid (section 2.9.3, table 2.2). The relative mean expression level of chemotaxis proteins at each FlgM level was measured by FACS using strains with genomic *cheY-yfp* as a reporter (fig. 3.11 A). Concerted overexpression of all proteins up to 6.6-fold above the native level had little effect on chemotaxis efficiency (fig. 3.11 B). Thus, the CheY-P concentration in the overexpressing cells must be in the working range of the flagellar motor. This conclusion was further confirmed by the observation that the average time the motor spends rotating clockwise (clockwise bias) was nearly unaffected by protein overexpression (fig. 3.11 C). The motor bias was determined by tethering the cells by one flagellum to a cover slip, and watching the cells rotate in buffer (section 2.9.2). Tethering cell experiments provide a direct readout of the phenotypic differences in the adapted level of phosphorylated CheY-P. Notably, the corresponding standard deviation decreased with the level of expression, as expected from the decline in strength of gene expression noise (fig. 3.11 A, B). The distribution of the CW bias among a small population of cells for different expression levels is shown in fig. A.6, where A shows the bias in wild type cells. An overexpression of 6.6 fold wild type expression (fig. A.6 C), showed a slightly higher bias and a narrow distribution. Such reduced noise in bias at high expression levels of all chemotaxis proteins leads to an improved chemotaxis efficiency to above the wild type level (fig. 3.11 B).
Surprisingly, the wild type chemotaxis efficiency was observed at the lowest possible protein level. At expression levels lower than native levels, the chemotactic efficiency decreased rapidly, whereas increased protein levels would even benefit chemotaxis. To ascertain why wild type cells do not have higher levels of chemotaxis proteins to secure their chemotactic ability, we measured the growth rate at varying protein levels (fig. 3.11 B). Interestingly, the growth rate decreases at expression level above wild type level, whereas lower protein levels have no effect on growth, suggesting that there is a trade off between optimal chemotaxis ability and the metabolic burden of overproduction of proteins. Hence, *E. coli*

seems to have evolutionary optimized to be chemotactic at the lowest possible protein levels to keep the costs of protein production low.

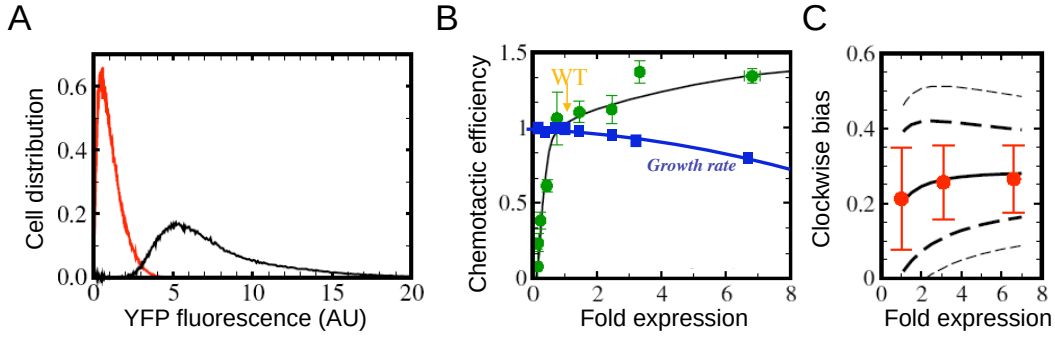

Figure 3.11: Effect of the total concentration of signalling proteins on chemotaxis. (A) Intercellular variation in the level of CheY-YFP, expressed from the native chromosomal position in wild type cells (LL12, red) and Δ*flgM* cells is (LL1, black) characterized by the means 1 and 6.6 and standard deviations of 0.67 and 3.17 respectively. (B) Chemotaxis efficiency of Δ*flgM* cells (VS102), expressing varying levels of FlgM from a plasmid (pLL16), determined in a swarm assay and normalized to the value of wild-type (RP437) cells (orange arrow). The relative mean expression of chemotaxis proteins at each FlgM level was measured by FACS, using the LL1 strain as a reporter for wild type CheY expression. The solid line is a guide to the eye, and error bars indicate standard errors. The blue line shows the corresponding growth curve, normalized to wild type growth. (C) Clockwise motor bias as a function of expression of chemotaxis proteins. Each point represents a mean of 20-30 cells. Error bars indicate standard deviations and illustrate intercellular variation. The solid line shows the clockwise bias calculated from the mathematical model of fig. 3.27 C under gene expression noise estimated from (A) (see chapter 3.3). The calculated standard deviation is depicted for the topology in fig. 3.27 C, assuming a steepness of the motor response curve with a Hill coefficient of five [133] (thick dashed lines) or ten [34] (thin dashed lines).

3.2.2 Pathway performance is sensitive to the individual protein levels

We have thus shown that the chemotaxis pathway is surprisingly robust against concerted variations in protein expression. Some of this robustness naturally arises from the balance in enzymatic activities, which will be further discussed in section 3.3. Next, we experimentally investigated the effects of uncorrelated variations by analyzing the physiological performance of the pathway at a range of expression levels of four cytoplasmic chemotaxis proteins, CheR, CheB, CheY,

and CheZ (fig. 3.12). Proteins were expressed from plasmid-born constructs under control of an IPTG inducible pTrc promoter table 2.4, also as fusions to facilitate expression quantification, and soft agar swarm plates in the respective knock-out strains at different levels of induction was used to test the chemotactic efficiency 3.12 A. The YFP fusions did not appear to strongly affect signalling properties of the chemotaxis proteins, since they complemented well in the knock out strains on swarm plates, and could be used to analyze the robustness properties of the pathway topology. We further used FACScan to determine relative expression of every fusion protein at a given induction and plotted the chemotactic efficiency as a function of the protein level (fig. 3.12 B and C). Generally, the efficiency initially increased with the expression level of individual proteins until it reached a maximum, and then decreased again upon a further increase in expression. This was in a sharp contrast to a concerted variation in the levels of all proteins, where no decrease of the chemotactic efficiency was observed even at high protein expression (fig. 3.11 B). The optimal levels of complementation and the range of expression levels that supported chemotaxis were specific for each protein, with relative positions of the optima being in a good agreement with the native expression levels of corresponding proteins [88]. CheR and CheB fusions complemented optimally at the lowest expression levels (3.12 B), followed by CheZ and then by CheY (fig. 3.12 C).
Surprisingly, we found that an experimental co-expression of CheY-YFP with CheZ-CFP, led to a large increase in robustness against CheY overexpression (fig. 3.12 C). At low expression levels, the complementation of a Δ*(cheY-cheZ)* strain by this bi-cistronic construct was comparable to the complementation of a $\Delta cheY$ strain by CheY-YFP. The chemotactic efficiency dramatically improved, however, at the high expression levels. This result will be discussed further in section 3.2.3.

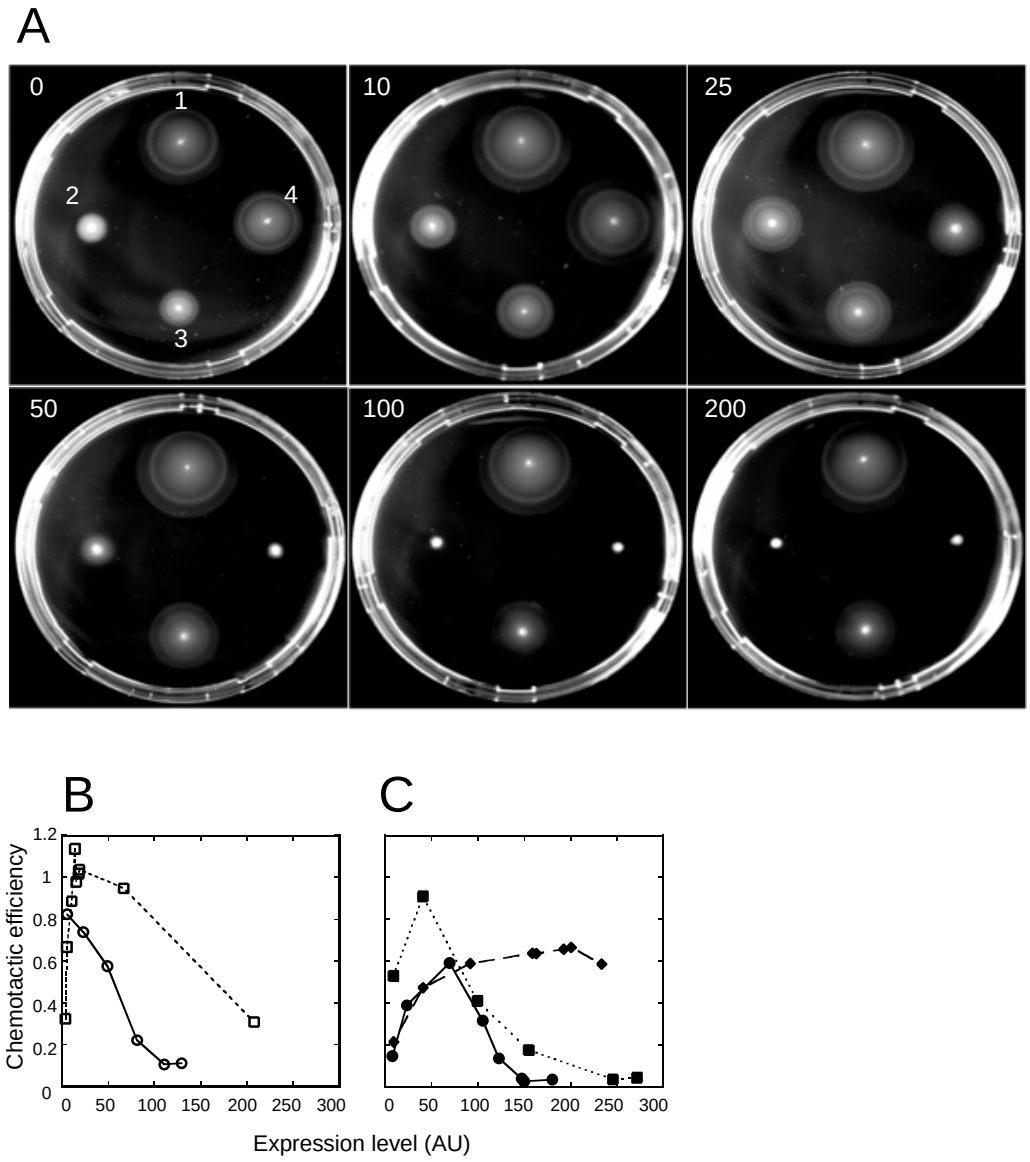

Figure 3.12: Physiological performance of the chemotaxis pathway at varying expression levels of signalling proteins. (A) Chemotaxis-driven spreading of wild-type cells with an empty vector pTrc99A (1), Δ*cheY* cells (VS100) expressing CheY-YFP from pVS18 (2), Δ*(cheY-cheZ)* cells (VS104) expressing CheY-YFP/CheZ-CFP from pVS88 (3), and Δ*cheZ* (VS161) cells expressing CheZ-YFP from pVS64 (4) on soft agar swarm plates at indicated IPTG concentrations, 0-200 μM. (B) Chemotactic efficiency, defined as the size of the outer swarm ring and normalized to the chemotactic efficiency of the wild type, as a function of the expression level of CheB-YFP expressed from pVS138 in a Δ*cheR* strain (RP4972, open circles, solid line) and CheR-YFP (open squares, dotted line). CheR-YFP was expressed either from pVS137 (low expression levels) or from pDK19 (high expression levels) in Δ*cheR* cells (VS126). Expression levels of the YFP fusion proteins were determined for liquid cultures using FACScan. (C) Chemotactic efficiency as a function of the expression level of CheY-YFP (closed circles, solid line), CheZ-YFP (closed squares, dotted line), and co-expressed CheY-YFP and CheZ-CFP (diamonds, dashed line).

To investigate this further, we measured FRET (as described in chapter 2.9.4) upon varying levels of CheB and CheR to ascertain whether kinase activity and adaptation is sensitive to variations in the individual protein levels. Figure 3.13 shows an example measurement of CheY/CheZ FRET in a wild type strain responding to a saturating level of attractant at 400 seconds, adapting back to pre-stimulus value and responding to the removal of attractant at 1050 seconds. The YFP/CFP ratio is proportional to the number of CheY-P-CheZ complexes and therefore to kinase activity. We observed that FRET efficiency

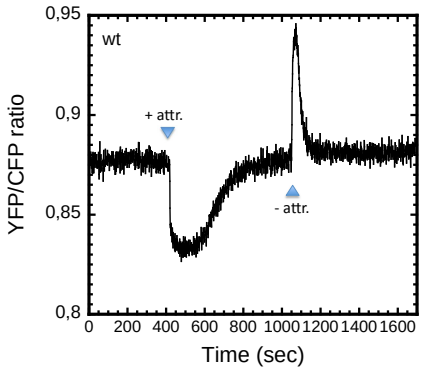

Figure 3.13: CheY/CheZ ratio upon addition and removal of attractant. CheY and CheZ are expressed as fusions from pVS88 in VS104 (Δ*(Y-Z)*). The CheY/CheZ ratio shows the responses to addition of saturating level of MeAsp, almost perfect adaptation back to pre-stimulus state and response to removal of the attractant. Note that the repellent response at attractant removal is higher than the attractant response.

(kinase activity) decreased with increasing CheB level. Additionally, we measured FRET in presence of the catalytically inactive mutant form of CheB, CheB^{D56E}, to assess whether the phosphorylation of CheB has an effect on FRET efficiency. Indeed, in absence of the phosphorylation feedback a higher expression level was needed to observe the same decrease as in presence of wild type CheB, consistent with the lower catalytic activity of CheB^{D56E} (fig. 3.14 A). The decrease in FRET upon overexpression of CheB can be explained by the lower kinase activity caused by an increased rate of de-methylation of the receptors. Adaptation to saturating levels of MeAsp was measured at varying levels of CheB^{D56E} by roughly estimating the time from addition of attractant until ratio reached half that of the steady state ratio without stimulation. For simplicity we also here excluded the phosphorylation feedback to CheB. Figure 3.14 C shows adaptation times to 1 mM MeAsp at different levels of CheB^{D56E}. The adaptation time increased slightly with increasing CheB^{D56E} levels until

4 fold native levels of CheB. At 14 fold native level of CheB, adaptation time had decreased again to similar adaptation time observed at native CheB levels. Comparison with the adaptation time to 150 μM MeAsp confirmed that this decrease was significant.

Similarly, we measured CheY/CheZ FRET upon overexpression of CheR. We observed an initial increase in FRET with increased CheR levels, however saturation was reached at 5 fold overexpression. An initial rise in FRET upon CheR overexpression is consistent with an increase in receptor activity due to methylation leading to higher kinase activity. Saturation in FRET follows from saturation of methylation sites counteracting inhibition of the kinase from attractant. YFP/CFP ratios to saturating levels of attractant and repellent at varying levels of CheR are shown in fig. A.7. Moreover, adaptation time to saturating levels of attractant decreases with CheR level (fig. 3.14 F) consistent with the fact that CheR increases the activity of the receptors, counteracting the effect of attractant binding. At high CheR levels, more attractant was needed for saturation, thus curves for adaptation to 150 and 300 μM are combined. Taken together, these results suggest that kinase activity and adaptation is sensitive to the levels of CheR or cheB.

It has previously been suggested that adaptation enzymes binding to receptor is not essential for adaptation and can be compensated for by overexpression of CheR [1, 119]. Therefore we measured FRET in a strain lacking Tsr and the pentapeptide sequence, NWETF, of Tar (see table 2.2). Without the pentapeptide sequence an attractant response was only visible at CheR levels higher than 10 fold that of wild type. At the highest induction levels we observed adaptation, however very slow, to 30 μM methyl aspartate (fig. 3.15).

Finally, we performed swarm experiments to exclude any sterical hindrance between CheR and CheB. A catalytical mutant of CheR, CheR^{D154A}, which can localize and bind to the receptors but is not catalytically active, was expressed in the wild type strain RP437 with native levels of CheB. Swarm efficiency did not vary with CheR levels, hence CheB is not sterically hindered by CheR (fig. A.8).

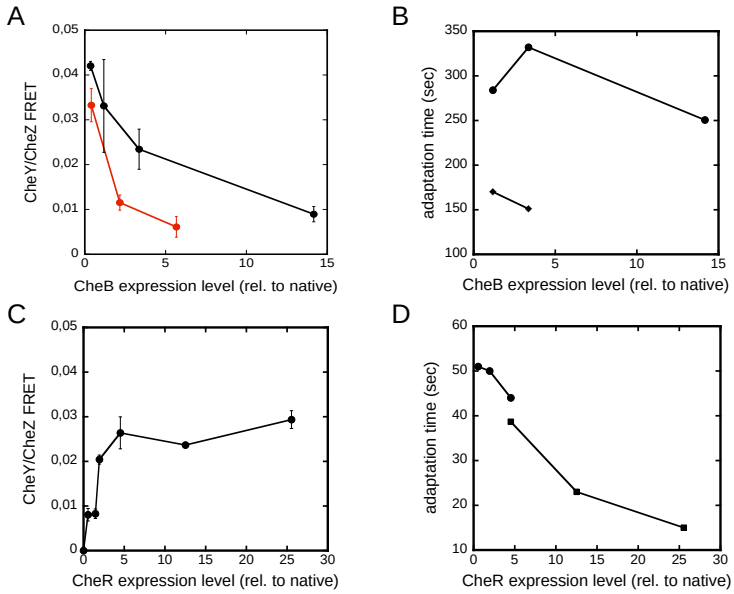

Figure 3.14: FRET efficiency at varying levels of wild type CheB, CheB^{D56E} and CheR. (A) CheY/CheZ FRET efficiency as function of CheB (red) or CheB^{D56E} (black) concentration (quantified by western blots, S. Schulmeister, unpublished data). CheB$_{wt}$ and CheB^{D56E} are expressed from plasmids pVS91 and pVS97 in a Δ*(cheB-cheZ)* strain (VS124). (B) Adaptation times to 1 mM (circles) and 150 μM (diamonds) MeAsp at varying CheB levels. (C) FRET efficiency as function of CheR level expressed from pVS113 (weak RBS) or pLL45 (strong RBS) in a Δ*cheR* Δ*(cheY-cheZ)* strain (VS127). CheR levels were determined by FACS as described in chapter 2.10. (D) Adaptation times to 150 μM (circels) and 300 μM (squares) MeAsp at varying CheR levels.

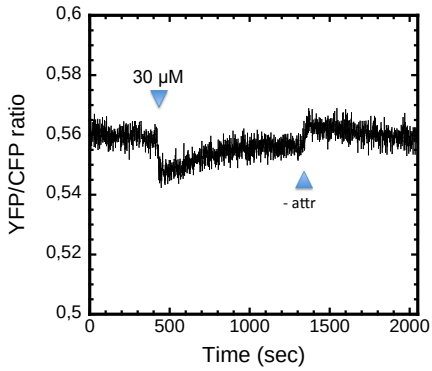

Figure 3.15: Adaptation time at overexpression of CheR in absence of the Tar pentapeptide sequence. YFP/CFP ratio at high CheR level showing adaptation to 30 and 100 μM MeAsp. CheR is expressed from pLL45 in a Δ*tsr*, *tar*Δ*pp* strain (VS165). Induced with 0.01 % arabinose CheR concentration is 25 fold higher than wild type level.

3.2.3 Role of translational coupling in robustness of the chemotaxis pathway

Maintaining a constant ratio between signalling proteins is important for a proper functioning of the pathway under varying protein levels, and we have shown that the chemotaxis system is much less sensitive to a concerted overexpression of CheY and CheZ than to the overexpression of each of these proteins individually (fig. 3.12 A,C). We thus assessed whether generally a co-expression of proteins from bi-cistronic constructs will improve performance of the pathway in a chemotaxis-driven spreading of bacteria in soft agar (fig. 3.16 A). Indeed, cells that express a YFP fusion to a particular gene as a mono-cistronic construct in the respective knock-out strain spread less efficiently than the cells that express this fusion as a downstream gene in bi-cistronic constructs at the same level (fig. 3.16 B). Co-expression of CheY/CheZ lead to an chemotaxis enhancement of 2.7 compared to CheZ alone, followed by CheB/CheY which enhanced chemotaxis by 2 fold compared to CheY alone. Co-expression of CheR/CheB leads to an enhancement of 1.5 fold that of CheB expressed alone.

To investigate the characteristics of such robustness to correlated levels of CheY and CheZ further, we measured the kinase activity upon overexpression of these two proteins. CheY/CheZ – FRET response amplitudes to saturating levels of MeAsp at varying levels of CheY-YFP and CheZ-CFP surprisingly scaled linearly with levels of CheY-YFP and CheZ-CFP in the cells (fig. 3.17). Hence, although kinase activity and the number of CheY-P-CheZ complexes in the cells increase with CheY/CheZ levels, the swarming efficiency and thereby the motor bias remains unaffected by overexpression.

The enhancement of chemotaxis of about 1.5 fold upon overexpression of CheR and CheB (fig. 3.16) was studied further. We measured kinase activity and adaptation time to stimuli using different expression levels of CheR and CheB (fig. 3.18). The adaptation enzymes were expressed from the pBAD promoter from the same operon. Amplitudes at saturating levels of MeAsp (100 μM) were recorded as a measure for kinase activity. We observed an increase in kinase activity upon increasing concentration of CheR and CheB until saturation was reached at 0.003 % arabinose. Further increase of the enzyme level did not affect the kinase activity as shown in fig. 3.18. In addition, adaptation times to 30 μM MeAsp decreased at low CheR/CheB concentrations, but reached a plateau at 0.002 % arabinose suggesting that adaptation is more robust against

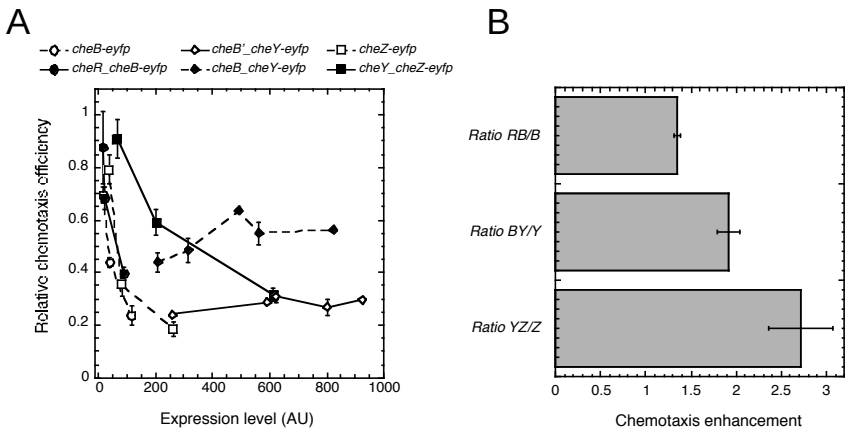

Figure 3.16: Improvement of the chemotactic efficiency by co-expression of the signalling proteins. (A) Dependence of the chemotactic spreading on the protein expression level for monocistronic (open symbols, dashed lines) or bicistronic (filled symbols, solid lines) constructs. Gene pairs were cloned as they appear in the genome, and the second gene was fused to an *yfp* reporter. Indicated constructs (plasmids described in table 2.4) were expressed in *E. coli* knock-out mutants that correspond to the YFP fusion gene. Expression was induced by 10, 25 or 100 μM IPTG or 0, 0.0005, 0.001, 0.003, 0.005 or 0.01 % arabinose, respectively. Expression levels were measured in liquid cultures grown under the same induction as described in section 2.9.1. (B) Enhancement of chemotactic efficiency by expression coupling. Enhancement was calculated as a ratio of chemotaxis efficiency at a given expression level of the monocistronic construct to the interpolated efficiency at the same expression level of the YFP fusion in the respective bicistronic construct, and values at different expression levels were averaged. Error bars indicate standard deviations.

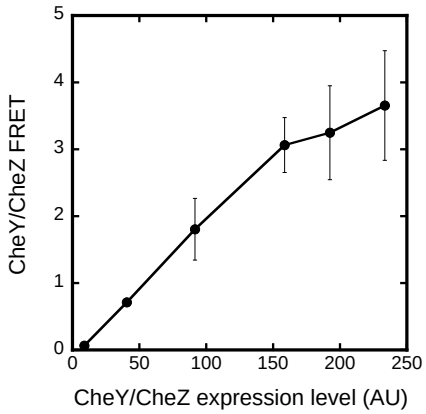

Figure 3.17: CheY-P-CheZ complex concentration upon overexpression of CheY and CheZ. CheY-YFP and CheZ-CFP were co-expressed from pVS88 at varying IPTG levels in Δ*(cheY-cheZ)* cells (VS104). CheY/CheZ-FRET was multiplied with expression level to obtain a measure for kinase activity (which is proportional to FRET). The expression levels are in arbitrary units of YFP expression and were measured by FACS.

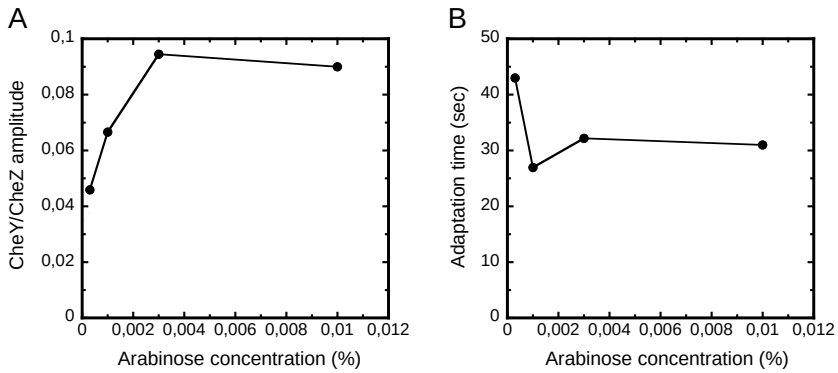

Figure 3.18: Kinase activity and adaptation time at co-varying levels of CheR and CheB. (A) Absolute amplitudes as a measure for kinase activity upon addition of 100 μM MeAsp (saturating, circles) at varying CheR/CheB levels expressed from pVS144 (table 2.4) in a Δ*(cheR-cheZ)* strain (VS149). 0.0004% arabinose corresponds to native levels of CheR and CheB. (B) Adaptation time to 30 μM MeAsp measured as time from addition of attractant to half the baseline is reached at co-varying levels of CheR/CheB.

co-expression of CheR and CheB than to overexpression of CheB or CheR alone (compare with fig. 3.14).

Our experiments have shown that coupled expression of chemotaxis proteins improves chemotaxis and increases the invariance of kinase activity and adaptation time, thereby improving the robustness of the pathway to perturbations. A high correlation of intercellular variation in levels of chemotaxis proteins could be partly attributed to the gene organization in poly-cistronic transcriptional units– operons – where multiple genes are transcribed as one mRNA. We observed a robustness of the chemotactic behaviour and of the motor bias against overexpression of all proteins (fig. 3.11). However, computer simulations predicted a larger variation of the motor bias in the population than observed experimentally (fig. 3.11 C), indicating presence of additional robustness mechanisms. This suggests that translational coupling [13, 91, 120, 135] could be a part of the robustness mechanisms in *E. coli* keeping a constant ratio of signalling proteins, thereby reducing negative effects of the uncorrelated noise.

3.2.4 Translational coupling between chemotaxis genes

To test whether expression levels of neighbouring chemotaxis genes is coupled on a translational level, we analyzed the neighbouring genes *cheA_ cheW* from the *mocha* operon in addition to the already mentioned pairs *cheR_cheB, cheB_*

cheY and *cheY_ cheZ* (fig. 3.16) from the *meche* operon. The level of translation of the first gene was then selectively varied by placing ribosome-binding sites (RBSs) of different strength in front of it. Translation levels were then quantified by FACS (2.10.1). As a control of the RBS strength, a *yfp* fusion to the first gene in the pair was placed under the same RBSs (3.19 A). Thus determined differences in the RBS strengths varied from 7 to 30 (fig. 3.19 B). For the *cheA_ cheW* pair, this strategy was complicated by the fact that CheA is expressed from two alternative translation initiation codons, yielding a long and a short version, CheA_L and CheA_S, respectively [145]. Consequently, changing the strength of the first RBS had only a moderate effect on the total expression level of CheA. Instead, we compared the constructs expressing CheA_L under the external RBS and CheA_S under the endogenous RBS with those expressing only CheA_S under the external RBS. The resulting net level of translation of CheA_L-YFP and CheA_S-YFP in the first construct was about four times higher than that of CheA_S-YFP in the second construct. For all pairs, stronger translation of the upstream gene resulted in an elevated expression of the downstream gene, implying the existence of a translational coupling (fig. 3.19 B). The coupling was quantified as a ratio of the indirect up-regulation seen in constructs that carry gene pairs to the direct up-regulation of the first gene. The strength of translational coupling varied among gene pairs, with *cheA_ cheW* showing the strongest and *cheR_ cheB* showing the weakest coupling (fig. 3.19 C).

3.2.5 Consensus order of chemotaxis genes in bacteria

Our analyses imply that the order of chemotaxis genes on the chromosome has to be subject to an evolutionary selection and is should thus be conserved among bacteria. Nikita Vladimirov from the Interdisciplinary Center for Scientific Computing (IWR) at the University of Heidelberg performed a comprehensive analysis of 245 sequenced bacterial genomes that contain annotated chemotaxis genes (section 2.12.3 on page 56). He could confirm that there is indeed a strong bias in the co-occurrence of these genes in the genome and in their order (fig. 3.20). The strength of the genetic coupling was roughly collinear with the experimentally observed strength of translational coupling. Notably, the resulting consensus order of chemotaxis genes (fig. 3.21 A) showed a nearly perfect match to their arrangement in *E. coli*, except for *cheZ* that is absent in most bacteria. Genetic

coupling and its order among chemotaxis proteins is shown in in the *E. coli* pathway (fig. 3.21) as an example.

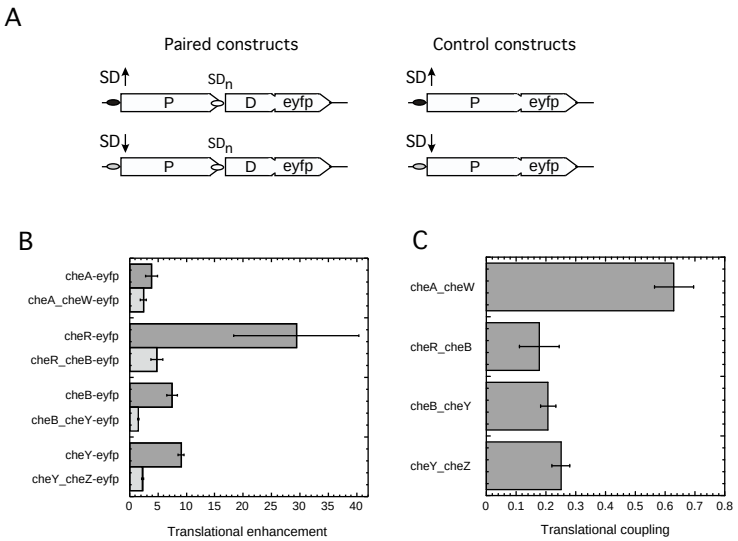

Figure 3.19: Translational coupling between neighbouring genes. (A) Experimental strategy. Bicistronic constructs that contained pairs of neighbouring chemotaxis genes in their chromosomal arrangement were cloned under RBS of different strength as indicated to create a C-terminal YFP fusion to a second gene. As a control of the strength of the RBS, the same sequence was placed in front of the monocistronic YFP fusion to the first gene. Expression of the constructs was analyzed using FACS as described in 2.10.1. (B) Direct (dark-grey) and indirect (light-grey) up-regulation of expression level of the fusion reporter by the stronger RBS. For the *cheA_cheW* pair, translation was regulated by using constructs that express either only short version of CheA or both long and short versions (see text for details). Error bars indicate standard deviations. (C) Translational coupling, defined as the ratio of indirect to direct up-regulation of expression levels by the stronger RBS.

	cheA (300)		*cheW* (440)		*cheR* (289)		*cheB* (219)		*cheY* (1828)		*cheZ* (81)		*mcp*[b] (2555)	
	left	right	left	right	left	right	left	right	left	right	left	right	left	right
cheA	0.7	1.3	**25.5**	5	10.4	4.2	17.4	10.5	**12.2**	**13.1**	-	**35.8**	0.6	1
cheW	5.3	**35.7**	8.4	6.4	**20.1**	6.6	4.1	1.4	1.3	1.1	-	2.5	3.3	2.8
cheR	1.7	4.7	3.9	13.6	0.3	0.3	**31.5**	6.8	1.8	0.4	-	-	0.2	1.8
cheB	5	12.3	0.9	2.3	6.2	**25.3**	0.5	0.5	2.6	1.2	1.2	-	0.2	0.2
cheY	**24.3**	4.3	4.8	5.2	2.1	7.3	6.4	**18.7**	3.7	4	**95.1**	-	0.9	1.3
cheZ	9.7	-	0.5	-	-	-	-	0.5	-	4.2	-	-	-	-
mcp[b]	8.6	6.7	17.9	**20**	16.3	2.1	1.8	3.2	1.9	1.5	-	-	**8.3**	**7.9**

Figure 3.20: Absolute frequencies of a pair wise occurence of chemotaxis genes calculated as a number of gene occurrences in -1 (left neighbour) or +1 (right neighbour) positions relative to a reference gene, normalized by the total number of analyzed reference gene counts (shown in brackets). Strongest genomic coupling on each side (highest co-occurrence frequency) is marked in bold. [b]Genes encoding chemoreceptors (methyl-accepting chemotaxis proteins).

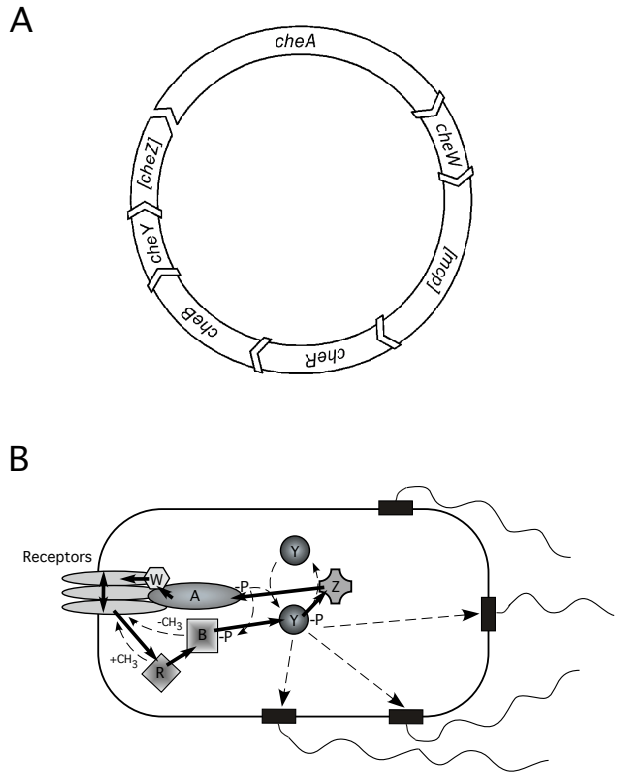

Figure 3.21: Genetic coupling of chemotaxis genes in bacteria. (A) Consensus order of chemotaxis gene arrangement among analyzed bacteria. Receptor (mcp) gene is taken in brackets because the number of receptor genes between *cheW* and cheR is variable; *cheZ* is taken in brackets because it is only present in a subset of bacteria. See fig. 3.20 for the frequencies of relative occurrence. (B) Genetic coupling and its order (solid arrows) among chemotaxis proteins, shown as an example in the *E. coli* pathway. Thin dashed arrows denote pathway reactions and CheY-P binding to flagellar motor.

3.2.6 Chemotactic selection experiments

We have shown that the chemotaxis pathway is robust against co-overexpression of some genes, by measuring the diameter of swarm rings at varying expression levels. To assess whether having correlated levels of specific proteins is indeed advantageous to the cells, we also picked cells at different parts of the swarm area, and measured the single cell protein levels by microscopy and imaging as described in 2.9.3 on page 48 to compare the correlation in cells with a good chemotactic ability with less- or non-chemotactic cells. CheY-YFP and CheZ-CFP were expressed from two different plasmids and under control of different promoters in the corresponding knock out strain. Means expression levels were similar in all samples, demonstrating that the expression itself is not affected by the position of cells in the swarm area. In contrast, relative expression levels in individual cells were strikingly different. In cells that remained back at the center (inoculation point), the levels of CheY and CheZ were clearly anti-correlated (fig. 3.22 B), whereas the best chemotactic cells at the outer edge of the swarm showed a strong positive correlation in these levels (fig. 3.22 D). These results clearly confirm that correlation in the single-cell levels of CheY and CheZ underlies an active selection during the process of spreading in soft agar and therefore has to benefit chemotaxis. In addition, our preliminary experiments also indicate selection for coupling between CheR/CheB and CheB/CheY protein pairs (data not shown).

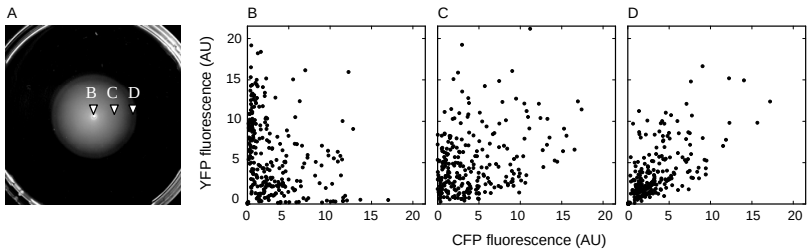

Figure 3.22: Chemotactic selection for the co-variation in the levels of CheY and CheZ. (A) Swarming of Δ(cheY-cheZ) cells (VS104) expressing *cheY-yfp* from pVS18 *cheZ-cfp* from pVS54 at 50μ M IPTG and 0.02% arabinose. (B-D) Expression levels of CheY-YFP and CheZ-CFP in individual cells, quantified by imaging as described in Section 2. Samples that were taken from the centre (B), 50% swarm diameter (C) and the outer edge (D) of the swarm ring as shown in (A). Each data point represents the total fluorescence intensity of YFP and CFP in a single cell; combined data from five experiments are shown. Respective mean values and correlation coefficients for the single-cell levels of CheZ-CFP and CheY-YFP, calculated as described in section 2.12, were 3.1, 5.6, and -0.26 (B); 4.2, 4.9, and 0.31 (C); and 3.2, 3.5, and 0.6(D).

Furthermore, we ascertained whether correlated levels of specific proteins provides an evolutionary advantage, by studying cells with improved chemotaxis through evolution. To promote a laboratory scale evolution, we inoculated Δ*(cheY-cheZ)* cells expressing CheY-YFP/CheZ-CFP from a plasmid on soft agar plates. After repeated re-inoculation of the best chemotactic cells (picked from the outer edge of the swarm ring) 5 times, chemotactic efficiency had improved by 30 % (fig. 3.23 A). In the evolved strain the single cell YFP and CFP levels were stronger correlated than in the initial strain (fig. 3.23 B). Back-transformation of the plasmid from the evolved strain into a fresh strain revealed that the mutation causing the improved chemotactic ability was not located in the plasmid. Epigenetic factors that could cause such improvement were ruled out given that repeated inoculation in over night cultures without any selection did not reduce the chemotaxis ability (data not shown). Hence, the improved chemotaxis obtained by evolution was caused by a stable, yet unknown, mutation in the chromosome.

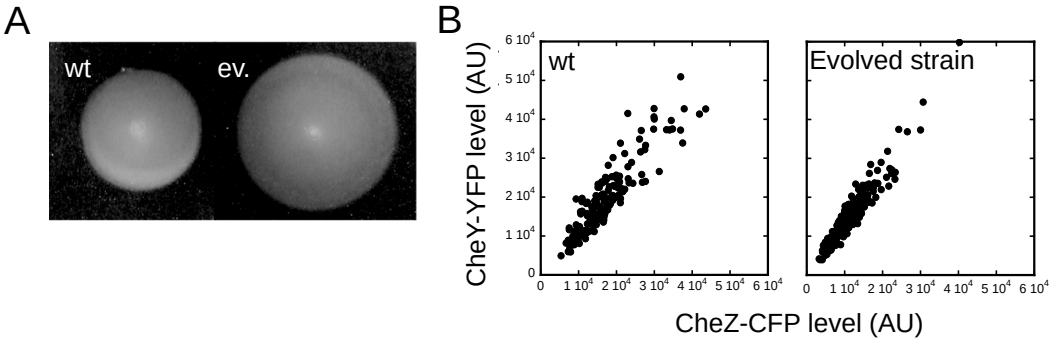

Figure 3.23: Evolved strain has increased correlation of CheY and CheZ. Cells expressing CheY-YFP and CheZ-CFP from a plasmid were re-inoculated 5 times to achieve a strain with improved chemotaxis. (A) The original strain before evolution (left) and an evolved strain with 30 % improved chemotaxis ability (right). (B) Single cell CheY-YFP and CheZ-CFP levels in the original (left) and evolved strain (right).

3.2.7 Effects of specific pathway features on robustness

The chemotaxis pathway has been shown to be extraordinarily robust against correlated levels of chemotaxis proteins and coupling of particular proteins seems

to improve chemotaxis. Such robustness cannot be explained entirely by counteracting enzymatic activities and transcriptional and translational coupling alone. To quest for further components of the chemotaxis pathway which could be responsible for the observed robustness and chemotactic selection for correlated protein levels, swarm selection experiments were performed upon removal of different features of the topology of the pathway.

Chemotactic selection in absence of the CheB phosphorylation feedback

The methyl-esterase CheB is activated by phosphotransfer from CheA. CheB phosphorylation has been previously shown to be non-essential for adaptation [3, 9]. Therefore, swarm selection experiments were performed to assess whether the selection for correlated levels of CheY and CheZ is affected by removal of the CheB phosphorylation feedback. Cells expressing CheY-YFP and CheZ-CFP together from a pTrc plasmid and $CheB^{D56E}$ from a pBAD plasmid were swarming in soft agar at native CheY/CheZ levels and varying levels of arabinose. As a positive control, CheY-YFP/CheZ-CFP was expressed in cells with wild type CheB showing a strong selection for correlated levels of CheY and CheZ. We observed that there is a selection for correlated levels of CheY and CheZ (fig. 3.24) also in absence of the phosphorylation feedback, however significantly weaker than in presence of wild type CheB. Thus, phosphorylation of CheB is important, however not essential, for the robustness of the chemotaxis pathway.

Chemotactic selection when CheZ cannot localize to the chemosensory complex

A previously suggested additional feedback from CheA to CheZ, where localization and activation of CheZ is facilitated by $CheA_S$ in presence of CheY-P, should improve the robustness of the pathway [92]. Here, only CheA-bound phosphatase would be enzymatically active. Therefore, we tested if localization of CheZ affects the chemotactic ability or selection for correlated levels of CheY and CheZ. A mutated form of CheZ, $CheZ^{F98S}$, which does not localize to the chemosensory complex due to a disruption in the binding to $CheA_S$ [30], was expressed from an IPTG inducible plasmid as a CFP fusion [164] together with

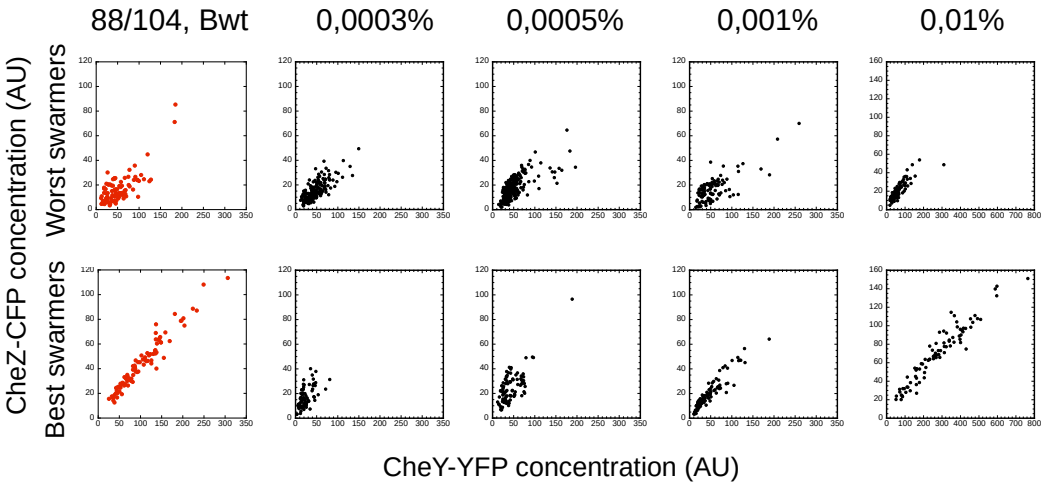

Figure 3.24: Correlation in levels of CheY-YFP and CheZ-CFP at varying levels of CheB^{D56E}. Single cell levels of CheY-YFP and CheZ-CFP (pVS88) expressed in Δ*(cheY-cheZ)* background (VS104) from different positions in the swarm ring (worst swarmers in the centre and best swarmer at the outer edge). In presence of wild type CheB (red), there is a strong selection for correlated levels of CheY-YFP and CheZ-CFP. A selection is also seen in presence of CheB^{D56E} (black), however only at the two highest induction levels (0,001 % and 0,01 % arabinose), where CheB^{D56E} is highly overexpressed.

CheY-YFP in Δ*(cheY-cheZ)* background. Swarm plates at varying IPTG concentrations revealed that localization of CheZ to the chemosensory complex does not have an effect on the chemotaxis ability and pathway robustness (fig. 3.25). Furthermore, FRET between CheY-YFP and CheZ^{F98S}-CFP was indistinguishable from FRET with wild type CheZ (fig. 3.25 B). This was also confirmed by picking cells at three different positions in the swarm ring to quantify the single cell levels of CheY-YFP and CheZ^{F98S}-CFP. Cells with a high correlation in the levels of the two proteins were selected for in the process of spreading in soft agar (fig. 3.26 A). Hence, the suggested feedback with activation of CheZ seems not to exist in the *E. coli* chemotaxis pathway.

Competition between CheY and CheB for the P2 domain of CheA and its effect on robustness

We investigated further whether the enhanced robustness could result from phosphorylation of CheY or CheB. The P2 domain of CheA contains a docking site

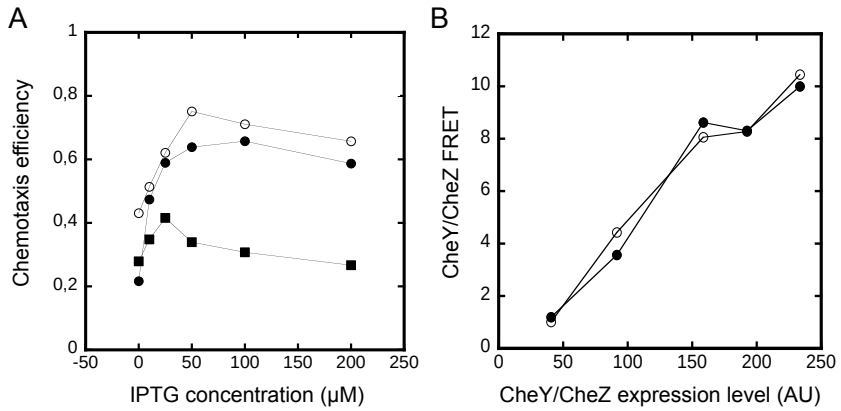

Figure 3.25: Chemotactic efficiency and FRET at varying levels of CheY and CheZ in absence of P2 or clustering of CheZ (A) Chemotaxis efficiency. Closed circles show the chemotactic efficiency in cells expressing CheY-YFP and CheZ-CFP (pVS88) in Δ*(cheY-cheZ)* background (VS104) at varying IPTG levels. CheZ is replaced by a mutant form preventing it from localizing to the chemosensory complex (pAV8, open circles). Squares show swarming in absence of the P2 domain of CheA (VS177) at varying levels of CheY and CheZ (pVS88). (B) CheY/CheZ FRET (pVS88, closed circles) and CheY/CheZ^{F98S} FRET (pAV8, open circles) in Δ*(cheR-cheZ)* background (VS149) at varying CheYCheZ levels.

for CheY and CheB and competition between the two has been shown biochemically [87]. Although the P2 domain is not essential for chemotaxis and phosphorylation [61, 152], it might be important for robustness. We performed swarm plate assays with CheY-YFP/CheZ-CFP expressed in a strain lacking the P2 domain of CheA in addition to CheY and CheZ. The swarming efficiency in the strain lacking P2 was generally reduced to about 25 % of that in presence of P2. At optimal CheY and CheZ levels the swarming ability and increased to about 40 % before at high IPTG levels the efficiency decreased towards 25 % again, hence the the cells are less robust against overexpression of CheY and CheZ in the absence of the P2 domain of CheA (fig. 3.25). In addition, the selection for cells with correlated levels of CheY and CheZ was not as strong in the absence of P2 (fig. 3.26 B) as observed for pAV8 (fig. 3.26 A) and for wild type (fig. 3.24). Thus, if CheY or CheB cannot bind the kinase due to the absence of the P2 domain of CheA, the chemotaxis pathway is less robust against co-overexpression of CheY and CheZ (fig. 3.25). Hence, the P2 domain of CheA seems to be important to robustness of the chemotaxis pathway and will be further discussed in chapter 3.3.

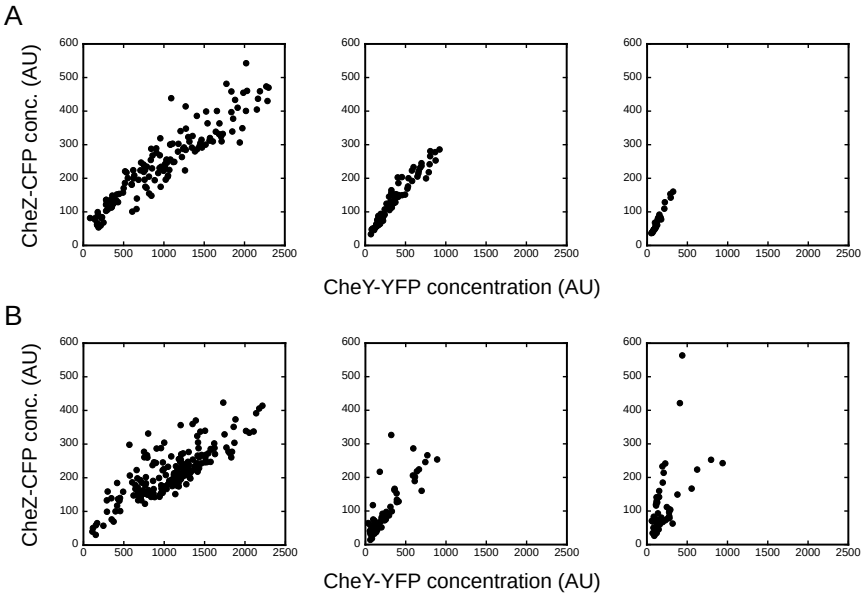

Figure 3.26: Selection for correlated levels of CheY and CheZ in absence of P2 or clustering of CheZ. Correlation between CheY-YFP and CheZ-CFP at three positions on the swarm ring where (A) CheZ cannot localise or (B) in absence of the CheA P2 domain. Same strains and plasmids are used as in fig. 3.25.

3.3 Mathematical analysis of robustness of the pathway topology

In collaboration with Markus Kollmann and his group at the center for theoretical biology at the University of Berlin, we studied the robustness of the chemotaxis pathway by simulations based on different network topologies. The most simple topology of a chemotactic signalling network that allows for precise adaptation is a two-state model proposed in [9] (BL model). This model is schematically drawn in figure 3.27 A. The BL model can readily explain the robustness of precise adaptation against a wide range of variations in kinetic parameters and protein concentrations as a consequence of integral feedback control at the methylation level [9, 111]. However, precise adaptation would be only physiologically relevant when the adapted level of CheY-P falls within the working range of the flagellar motor [34]. It is thus the stationary level of CheY-P rather than precise adaptation per se that should be under evolutionary selection, but in the BL model this quantity is not robust to large perturbations [3]. By extending the BL model, it is possible to construct larger adaptive topologies having equal input-output characteristics (fig 3.27 B-D) and analyse their

robustness against intercellular variations in protein concentrations arising from protein synthesis. The latter represents the dominating source of variations in a bacterial cell population [41, 129].

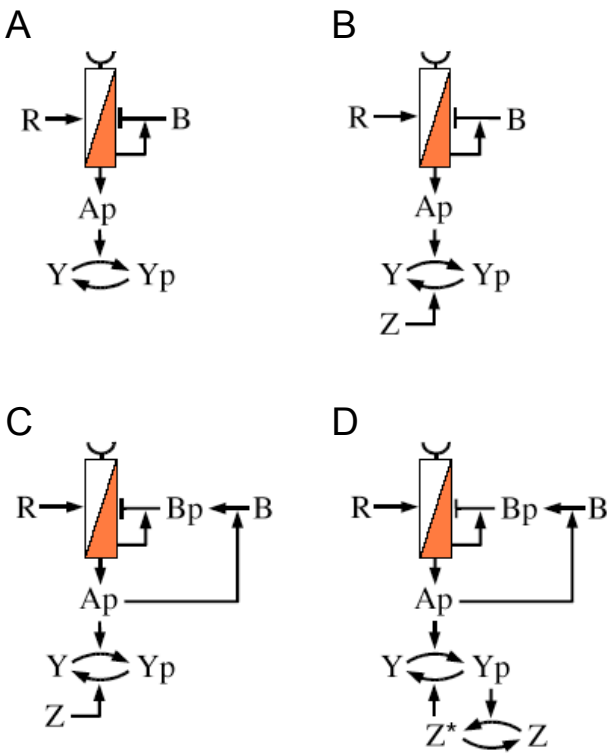

Figure 3.27: Four possible network topologies of bacterial chemotaxis showing precise adaptation. Links between proteins indicate activations (arrows) or repressions (bar ends). The receptor can either be in an active state (red) or an inactive state (white). The proteins involved are denoted, for example, by A = CheA, and their phosphorylated forms by Ap = CheA-P. (A) Minimal model as proposed in ref. [9]. (B) Same as model a but with a phosphatase CheZ substituting auto-dephosphorylation of CheY-P in topology A. (C) Same as model B but only the phosphorylated form of CheB can form a complex with active receptors. This topology represents the experimentally established network of *E. coli*. (D) Same as topology C but with an active form (Z*) of the CheY phosphatase [2, 23].

3.3.1 Co-expression of signalling proteins improves robustness of the pathway

We analysed mathematically the robustness of the four network topologies (fig. 3.27) under conditions of physiological perturbations. The mathematical description used, relied on five assumptions. First the numbers of protein copies are sufficiently large, such that stochastic effects on the protein level can be neglected. Each stable receptor complex includes one kinase protein CheA. A receptor complex can exist only in two functional states, active or inactive (two state model) [9]. The rate of CheA phosphorylation is assumed to be proportional to the average number of active receptor complexes in the cell. Finally, the protein-protein interactions can be described by Michaelis-Menten kinetics. The chemotactic efficiency of a population depends on the fraction of bacteria for which the CheY-P levels are within certain limits [34]. We tested the different hypothetical network topologies, shown in fig. 3.27, for their ability to reproduce qualitatively the experimental data for concerted variations of mean expression levels (fig. 3.11). In the computer simulations, we used the experimentally determined gene expression noise (Fig. 3.11 A) as an estimate for the minimum intercellular variations. The established topology for *E. coli* (Fig. 3.27 C) reproduces accurately the experimental data, whereas the simpler topology (Fig. 3.27 A) fails to match the data (fig. 3.28 A).
To quantify further the robustness of the different network topologies, we calculated the fraction of fully chemotactic bacteria for different strengths of gene expression noise (fig. 3.28). The topologies in fig. 3.27 C, D allow for the highest fraction of cells in the population to respond accurately to changes in ligand concentration. Furthermore, the topologies in fig. 3.27 B-D are sufficiently robust to compensate for co-variations in expression levels (fig. 3.28 A) but tolerate only a moderate increase in independent fluctuations in protein concentrations (fig. 3.28 B). Thus, the experimentally established design of the chemotaxis network in *E. coli* represents a minimal topology providing high robustness to perturbations in protein levels. This network design can compensate for the observed strong co-variations in gene expression but the negative effect of uncorrelated variations on the efficiency of chemotaxis can only be attenuated. Similar correlations in expression have been found recently in eukaryotes for genes under identical control [128]. We can therefore expect that analogous design principles, compensating for intercellular variations, will apply to all signalling networks

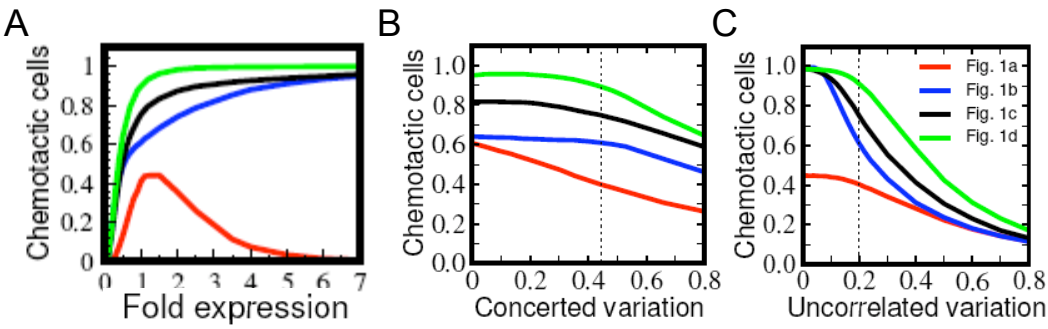

Figure 3.28: Simulated fraction of chemotactic cells as a function of gene expression noise.. The native level of gene expression noise is indicated by dotted lines. (A) Simulated fraction of fully chemotactic cells for topologies in fig. 3.27 A (red), B (blue), C (black) and D (green) as function of expression level. (B) Concerted variations ranging up to twofold of the wild-type strength with uncorrelated variations kept at the wild-type value, $\eta_{in} = 0.2$. (C) Same as (B) but with varying strength of uncorrelated variations and with concerted variations fixed to the wildtype value, $\eta_{ex} = 0.44$.

and gene regulation systems whenever precise regulation of an output signal is demanded.

3.3.2 CheB phosphorylation and competitive binding to P2 contributes to compensation of non-genetic individuality

Our model has proven inadequate in describing the robustness of the pathway to uncorrelated protein levels. For instance, the fact that chemotaxis seems to be more robust to variations in CheR and CheB levels than to CheY or CheZ, cannot be explained theoretically. The model is based on several assumptions; kinase activity is determined by the amount of receptor bound CheA. The probability for an allosteric receptor complex to be active is determined by its corresponding ligand occupation and methylation level. Phosphatase activity results from the concentration of free CheZ. Although synthesis and degradation of chemotaxis proteins is highly correlated and the pathway uses design concepts to compensate for the resulting concerted fluctuations in protein levels, this is by no means sufficient to keep the CheY-P level within the optimal range. The reason is that functional chemotaxis receptors are organized in large clusters. Cells have several clusters which vary greatly in size. As a consequence receptors will be unevenly distributed among daughter cells after cell division. Chemotaxis proteins lack

active degradation and therefore discrepancy between abundance of functional receptors, and phosphatase, remains for at least one generation. An optimal network design would require CheZ phosphatase activity to be proportional to the amount of kinases that can be activated. As a result, our simulations would require that only bound CheZ has phosphatase activity. However, we and others [164] have shown that chemotaxis ability is independent on localization of CheZ (fig. 3.25). This result suggests that kinase activity has to be adjusted to keep the CheY-P level optimal, which requires a linear relation between the amount of kinases associated with functional receptors and the amount of CheZ. Three specific biochemical modifications, absent in other two-component-systems, can fulfill this requirement.

First, in the adapted state, the probability for a receptor to be active has to be small. This in turn requires a high maximum kinase activity to compensate for phosphatase activity as given by the cellular amount of CheZ. Indeed, CheA shows about a 100-fold higher activity than other bacterial histidine kinases, e.g. EnvZ.

Second, receptor bound CheB, has to scale inversely with the amount of cellular CheY. This is achieved by activation of CheB by phosphorylation predominantly at the P2 binding domain of CheA and by CheY competing with CheB in binding to P2. As CheY concentration $10 - 20\mu$M exceeds that of CheB by almost two orders of magnitude the free P2 domains scale approximately inversely with CheY concentration.

Third, CheB once bound to P2 has to be phosphorylated with high probability. This requires a dissociation rate of CheB from the P2 domain that is significantly slower than the corresponding rate of CheY (consistent with previous observations of exchange rates at the cluster measured by FRAP, [134]).

Our differential equations based on these three modifications suggest that if the cellular concentrations of CheY and CheZ are linearly related, kinase activity is always balanced by phosphatase activity for fixed concentrations of receptor localized CheR and CheB. This theoretical finding is confirmed by the fact that CheY/CheZ FRET is nearly independent on the number of receptors, but scales linearly with CheY/CheZ level as shown in fig. 3.17 and fig. 3.29 B where a sevenfold increase of all chemotaxis proteins - except CheY and CheZ - does not lead to a similar elevation of kinase activity, but to a maximum 4 fold increase depending on CheY/CheZ expression level.

Elimination of cellular receptor concentration in our equations implies per-

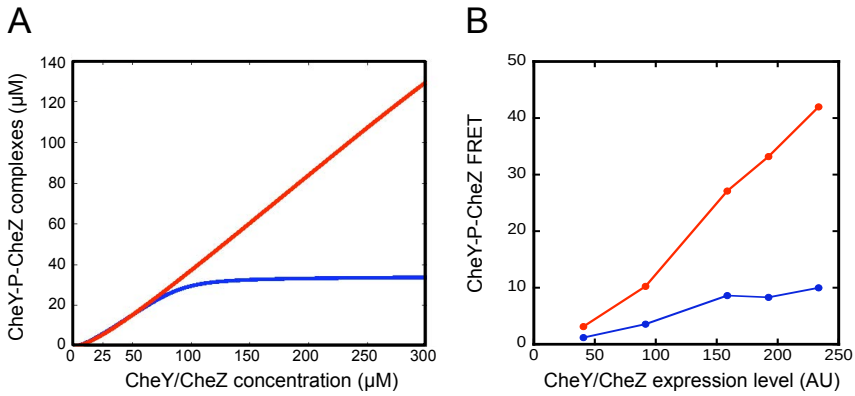

Figure 3.29: Simulated and experimentally measured CheY-P-CheZ complexes at varying levels of CheY/CheZ. (A) CheY-P-CheZ complexes at varying CheY/CheZ levels in Δ*(cheR-cheZ)* strain (VS149, blue) and Δ*(cheR-cheZ)* Δ*flgM* strain (LL5, red) comparable to fig. 3.17. (B) Experimentally measured FRET (proportional to CheY-P-CheZ complexes) in Δ*(cheR-cheB)* background with (blue) and without (red) FlgM

fect relation between CheY and CheZ concentrations to arrive at the optimal adapted CheY-P level. Indeed this was observed after extracting cells from different radii fig. 3.24 of the swarm ring (section 3.2.7 on page 85), where the bacteria with the highest chemotaxis performance showed strong correlation of these two proteins.

Taken together with the results shown in section 3.2.7, which exclude CheZ localization as important robustness and show that CheB phosphorylation is not necessary for adaptation, the conclusion from our analysis is that CheB phosphorylation and competitive binding to the P2 domain have been introduced mainly for compensation of non-genetic individuality and are therefore non-essential for primary pathway function. This is confirmed by the fact that CheA P2 domain improves selection for cells with correlated levels of CheY and CheZ (fig. 3.26).

3.3.3 Development of model to explain coupled expression of chemotaxis proteins

The topology alone and our early simulations were unable to explain robustness to pairwise overexpression of proteins. Hence we used our new and more complex model to describe the experimentally observed translational coupling and to ascertain why some proteins are coupled through sequential arrangement

in one operon and others are not. As mentioned above, enhancing robustness against gene expression noise is the most likely mechanism by which translational coupling could benefit chemotaxis. We thus used computer simulations to test whether chemotaxis gene order on the chromosome can improve robustness of the pathway output – adapted clockwise rotation (CW) bias of flagellar motor – against translational noise when translational coupling is taken into account. We simulated the effect of pair-wise coupling of individual chemotaxis proteins *cheR, cheB, cheY* and *cheZ* on the fraction of fully chemotactic cells, that is, cells with the steady-state level of CheY-P in the working range of flagellar motor, 2.2-4.3 μM [34]. Our *in silico* chemotaxis network model indeed confirmed that positive correlations between expression of adjacent genes via translational coupling affect deviations from the optimal adapted CW bias within a population (fig. 3.30). When levels of translational noise were assumed to be equal for all four genes, 16 gene orders out of possible 24 permutations were predicted to reduce variation of the bias in the population compared to the simulation in absence of coupling, whereas eight gene orders increased that variation (fig. 3.30 A). For eight gene orders, including the native gene order in *E. coli*, the noise reduction was particularly pronounced. The native gene order ranked first when a more physiological situation of 1.5 fold higher noise in expression of the weakly translated genes CheR and CheB was assumed (fig. 3.30 B). The gene order can be understood by simulating a 100 % pairwise translational coupling between different genes in the background of uncorrelated fluctuations of all other genes (fig. 3.30 C). Four adjacent gene pairs – *cheY_cheZ, cheR_cheZ, cheY_cheB*, and *cheR_cheB* – show favorable reduction in the standard deviation of CW bias. In all these cases, a positive effect is observed whenever a gene that enhances CheY-P level upon overexpression is followed by a downstream gene that reduces CheY-P level upon overexpression or vice versa (see Discussion). A negative effect is observed by coupling genes with similar effects on the CheY-P level – *cheY_cheR* and *cheB_cheZ* – which leads to an increased variation in CW bias. Notably, the degree of improvement in robustness due to coupling of individual genes apparently correlated with the experimentally observed strength of translational coupling (fig. 3.19 C).

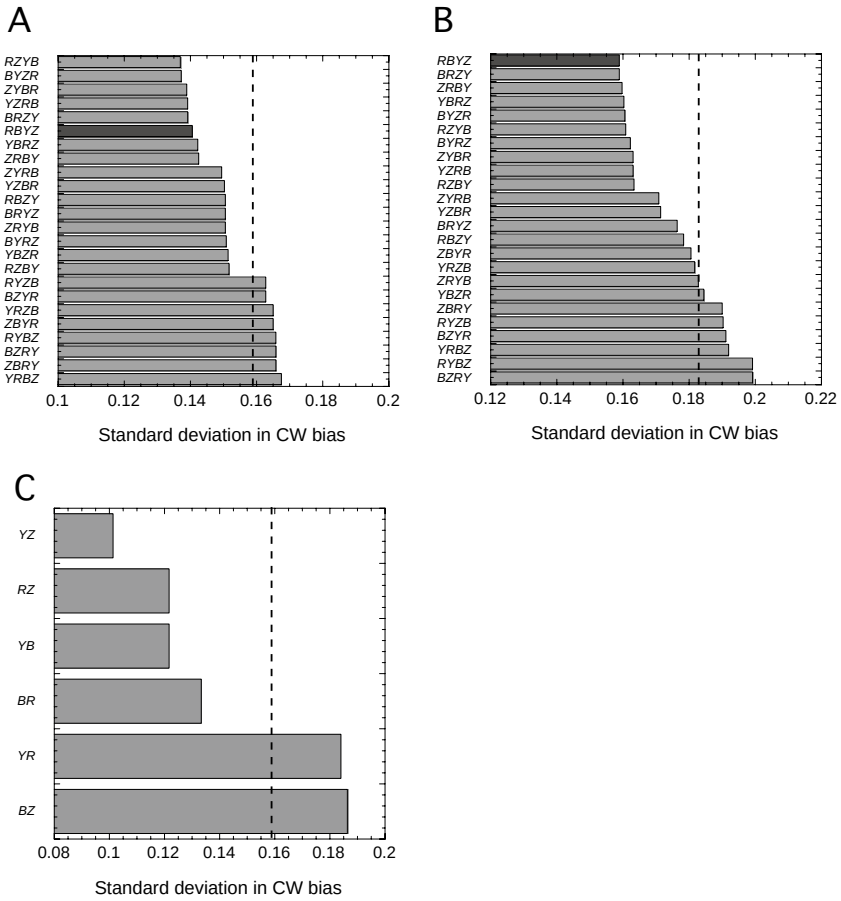

Figure 3.30: Simulated effects of translational coupling on robustness of the signalling output. Standard deviation of the CW motor bias in a population of 10^5 cells was simulated in presence of gene expression noise as described in Materials and Methods. (A,B) Simulations for different arrangements of translationally coupled chemotaxis genes, performed at equal noise levels for all genes (A) or at 1.5-fold higher noise for the weakly expressed *cheR* and *cheB* genes (B). Dark grey bars indicate gene order in *E. coli.* (C) Simulations for 100 % pair-wise coupling of chemotaxis genes, with remaining genes being uncoupled. Standard deviation of CW bias in absence of coupling is indicated by dashed lines. Genes are indicated by single letters, i.e. Y=CheY etc.

3.4 Dependence of chemotaxis on gradient shape and adaptation rate

Also in collaboration with Nikita Vladimirov at IWR , the dependence of chemotaxis on the shape of attractant gradient and adaptation rate was studied. Nikita wrote a computer program that simulates the behaviour of populations of *E. coli* in a computationally cheap way. RapidCell is based on the MWC model of mixed receptor clusters, the dynamics described by ODE and a detailed model of cell tumbling. The model was used to investigate chemotaxis in different gradients in liquid media and porous media (soft agar plates).
To study the effect of different adaptation rates on swimming in liquid medium and agar, we simulated cell populations with log-normal distribution of CheR/CheB concentration values. The mean (1.6) and standard deviation (0.48) are fitted to reproduce the variability of adaptation times observed for wild-type cells [19]: $T_{ad} = 311 \pm 150$ sec in response to a 0–10^{-3} M MeAsp step. We assume that the intensity of the moving gradient remains constant, and used constant-activity gradients as a simple model for the swarm assay simulations. The nominal circular constant-activity gradient provides constant cellular drift velocity at any distance from the centre of the plate. This property allows us to use it as a stationary model of the real moving gradient of attractants. Gradients with different steepnesses are denoted N1, N2 and N3. We plotted the distances travelled by single cells along the gradient N2 (steepness 4.56×10^{-3}) in liquid medium as a function of relative adaptation rate, and observed that a subpopulation with optimal CheR/CheB levels drifts more rapidly than other cells (fig. 3.31 A). Simulations in the N3 gradient (steepness 9.11×10^{-3}) in agar show that cells with low CheR/CheB levels are hindered by agar traps, while other cells drift successfully (fig. 3.31 B). In fig. 3.31 C,D the same cells are coloured from deep blue to red, according to their CheR/CheB levels. The outer edge of a bacterial ring in liquid medium contains many blue cells with CheR/CheB concentration between 0.5 and 2. In contrast, the outer edge in agar contains a uniform mixture of cells with different CheR/CheB levels, while deep blue cells with low CheR/CheB tend to stay behind. In order to study chemotaxis efficiency at different adaptation rates in agar, we measured experimentally the chemotactic efficiency on soft agar plates of cells with different CheR and CheB levels in the respective knock out strain. Expression of *cheR_cheB-yfp* as parts of one operon under control of the pBAD promoter and

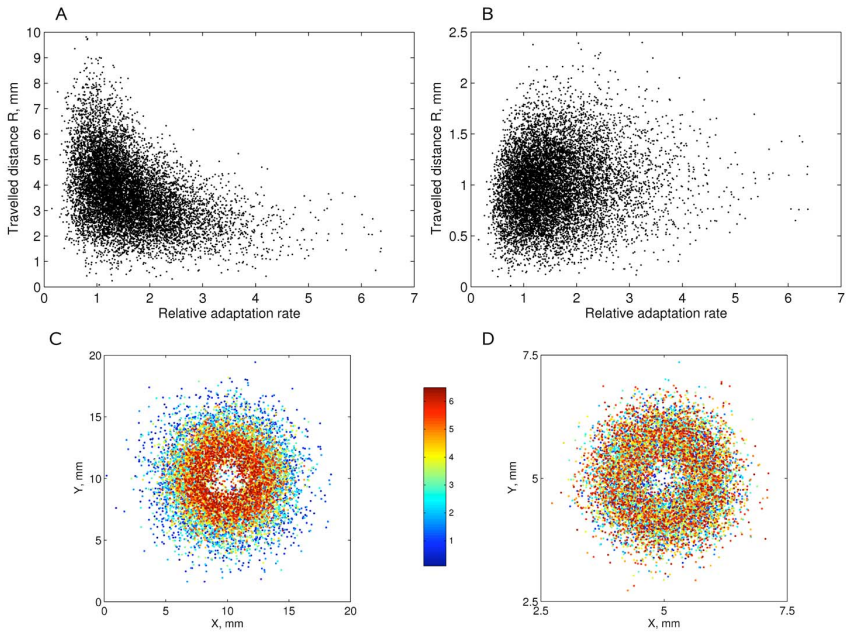

Figure 3.31: Simulation of motility in liquid medium and agar with physiological CheR/CheB concentration distribution. The distances R travelled by 10^4 cells after 1000 s of simulation in (A) liquid medium N2 gradient; (B) agar N3 gradient. The (x,y)-positions of cells colured from deep blue to red, according to their CheB/CheB concentration are shown in (C) for liquid medium, (D) for agar. The smallest CheR/CheB concentration values crrespond to deep blue, the highest to red. Note different scales of figures.

native ribosome binding sites was induced by varying arabinose concentrations which produced CheR/CheB levels in a wide range below or above the wild-type value. Furthermore, additional experiments with *cheR_cheB-yfp* expressed from the pTrc promoter were performed. This plasmid gave a much higher expression level and showed that only very high CheR/CheB levels, over 50 fold that of the wild type, can significantly decrease chemotactic efficiency on soft agar plates. Figure 3.32 A-B, E shows the simulated and experimentally determined relative chemotaxis efficiency as a function of CheB molecules per cell. The chemotaxis efficiency in the simulations was estimated as the average distance travelled by the cells, divided by the distance at the optimal CheR/CheB concentration.
The experiment could confirm the simulations which show that cells with CheR and CheB above a certain threshold perform chemotaxis equally efficient as in wild type, whereas lower CheR/CheB levels would severely impair chemotactic behaviour. Cells with low CheR/CheB would according to the simulations tend to run without tumbling and stay trapped in the pores in the agar most of the time. On the other hand, cells with extremely high CheR/CheB levels

loose their sensitivity to the gradient and also have poor chemotaxis efficiency. This suggests that in liquid medium there can be positive selection for cells with optimal CheR/CheB levels, they can reach the nutrient source faster and have more available substrates for growth. In contrast, swimming in agar poses mainly negative selection, where cells with low CheR/CheB levels are filtered out from the chemotactic population. The motor bias for optimal chemotaxis is different in liquid medium and soft agar. Cells with higher CCW bias would drift faster in liquid medium, but not in agar where the time of staying trapped increases with CCW bias.

To confirm that there is a selection for cells by their CheR/CheB levels in the swarm plates, cells were picked from two positions, at the center and at the outer edge of the swarm ring. By microscopy and imaging the relative single cell levels of CheB-YFP were quantified as described in 2.9.3 on page 48 and 2.10.2 on page 52. For the CheR/CheB expressed from the pBAD promoter, the cells collected in or near the center at normal agar concentration (0.27 %) had on average relatively low copy numbers of adaptation enzymes while cells in the outer edge of the swarm ring had relatively high copy numbers of CheR/CheB (fig. 3.32 C), confirming the prediction that there is a selection against low copy numbers. Two different agar concentrations were used for the swarm assays to obtain different steepnesses of the gradient. An agar concentration below 0.20 % does not produce stable gel structure, and therefore it is probably the most liquid agar that can be used for swarm plate experiments. As expected, in the swarm plates with reducedagar concentration (0.20 %) the difference between center and outer edge was much smaller (fig. 3.32 C, D), suggesting that there is no strong selection against low copy numbers in liquid medium. Our simulations and additional swarm plate selection experiments with pTrc promoter, which gives much higher basal expression level of CheR/CheB, show that very high levels of the adaptation enzymes, over 20 fold, can again decrease chemotaxis efficiency in agar and there is a selection against such high levels (fig. 3.32 E-F).

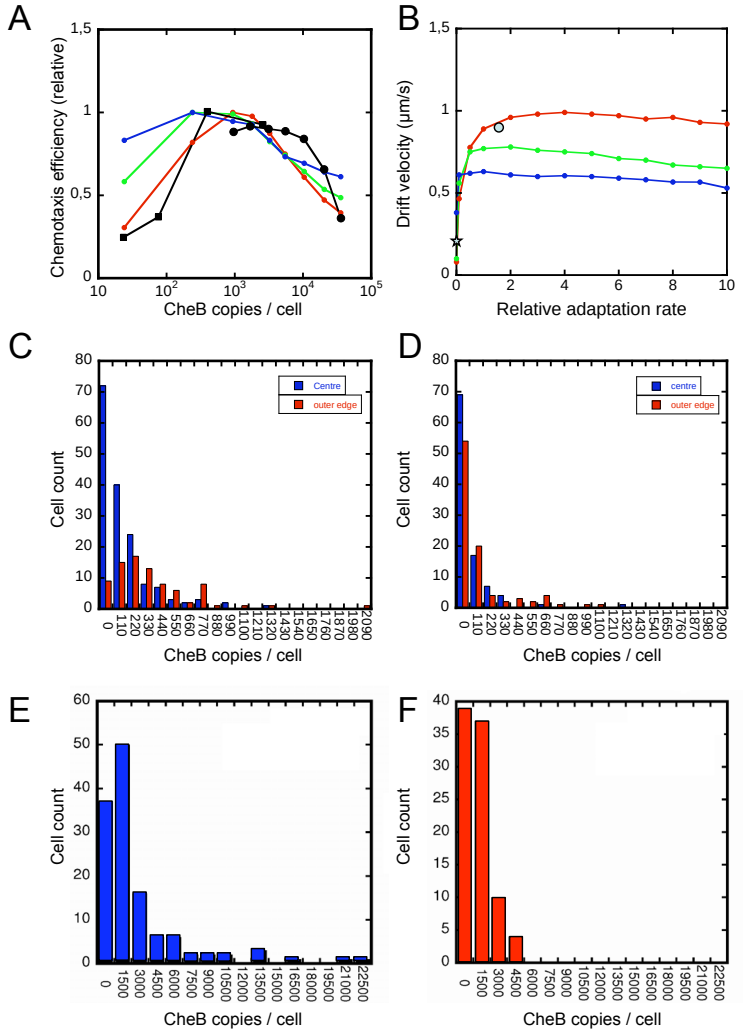

Figure 3.32: Dependency of chemotaxis efficiency on adaptation rate (A) Experimentally measured chemotaxis efficiency (black) at different expression levels of *cheRcheB-yfp* operon under control of the pBAD promoter (pVS571, squares) and the pTrc promoter (pVS145, circles) in a Δ*(tap-cheB)* strain (RP2867). The CheB-YFP levels reflect concerted CheRCheB-YFP concentration due to strong translational coupling. Wild type level of CheB is 240 copies/cell [88]. Simulated chemotaxis efficiency as a function of CheR/CheB levels are shown in colours. Cells are simulated in the constant activity gradients N1 (blue), N2 (green), N3 (red). The chemotaxis efficiency was estimated as the average distance travelled by cells, divided by the distance at the optimal CheR/CheB level. (B) Simulated drift velocity in the same gradients as in (A) as a function of relative adaptation rate. The light blue circle shows experimentally observed drift velocity of wild type cells, estimated from fig. 4 of [174]. The star shows the drift velocity from non-adapting cells, from fig. 6 of [174]. CheRCheB-YFP concentration in individual cells at different points of the swarm ring for plates with (C) normal soft agar (0.27 %);(D) liquid agar (0.20 %). Blue columns show the least swarming cells in the center of the swarm plate, and the red ones the best swarming cells from the outer edge. The expression of *cheR_cheB-yfp* was under control of the pBAD promoter and induced with 0,0003 % arabinose, which gave a basal expression level close to the wild-type value. The bin size is 110 copies/cell. (E) Same as (C) but CheRCheB-YFP are under the control of the pTrc promoter (pVS145) and overexpressed. The least swarming cells are shown in blue (E) and (F) shows the best swarming cells in red.

4 Discussion

4.1 Characterization of gene expression noise

Intercellular variation in protein levels in a genetically homogeneous cell population is the major source of perturbations that affect performance of all cellular pathways. In prokaryotes as in eukaryotes, the largest part of this noise appears to originate from fluctuations of global factors that affect expression of all genes in a cell, and from stochastic variations in promoter activity [121, 128, 129, 131]. Noise in biological systems could be expected to influence many processes in the cell, particularly the ones where a defined output has to be produced upon a certain stimulus, as in signal transduction. Hence, in this work we quantify noise in the chemotaxis system in order to discover its origins and ascertain its architecture.

Fluorescent protein fusions to several chemotaxis proteins were constructed in order to quantify their expression levels *in vivo*. Copy number variations might affect the result and lead to overestimated noise values. Therefore, genes were replaced by their fusion on the chromosome such that expression could take place from the native promoter. Relative fluorescence values were sufficient to quantify noise, and total fluorescence in a cell was assumed to be proportional to total protein level. Relative concentration values were obtained by dividing the total fluorescence by cell length measured from the images. The rod shaped *E. coli* grows by elongation with no change in cell width, thus the volume of a cell is mainly dependent on its length [105, 162].

With the purpose of characterizing the noise in the chemotaxis pathway, 5 proteins that were considered interesting were chosen. CheY is the response regulator of the pathway, CheA the kinase, FliM stabilizes the CW bias upon binding of CheY-P, FlgM is the inhibitor of σ^{28} and thus of expression of class 3 genes, and FlhC is part of the master regulator of the whole pathway. Fluctuations in the levels of these proteins in the three different promoter

classes in the hierarchy of flagella and chemotaxis genes were shown to be wider in the lowest class, class 3, than in the other classes. The distributions of protein levels were asymmetrical with a positive skew. A lognormal distribution indicates the presence of multiplicative random independent fluctuations which suggests cascade effects. This is typical for protein production. However, the width of the distributions in class 2 and 1 were indistinguishable, thus a propagation of noise through the hierarchy, which could have amplified noise at the lowest level, was not observed. CheA-CFP and CheY-YFP, both in class 3, showed strong correlation in single cell levels, even though they are expressed on two different chemotaxis operons. The large co-variations in the levels of both proteins should arise from extrinsic factors affecting both genes equally such as fluctuations in transcriptional activity, since the proteins are under the same transcriptional regulation. The fact that we observed correlations between all classes, and even among unrelated genes, suggests that the extrinsic noise causing the observed cell-to-cell variations mainly originates from global factors (fig. 3.3 and fig. 3.4, p. 62 - 63).

Snap shot imaging only provides information about the noise at one particular time point. However, we were also interested in the time scale over which fluctuations persist. Although fluctuations may average out fast, slow fluctuations could introduce errors into the pathway such that the accuracy of the system is reduced. In order to assess the temporal dynamics of gene expression we applied time-lapse imaging to cells expressing *cheY-yfp* from its native promoter on the chromosome (fig. 3.8, p. 67). We observed that cells keep their protein concentrations relatively stable over several generations. To achieve this, the cells seemed to double their protein levels in the time of one cycle, such that the daughter cells had exactly the same concentration of CheY as the mother cell. The increase in total protein levels was identical to the increase in cell length through the cell cycle. This coupling of growth with gene expression might be important to allow the daughter cells to be chemotactic immediately after division. It is known that the CheY-P level has to be within a very narrow range in order for the cell to be chemotactic, hence the importance of keeping the CheY concentration stable throughout generations seems reasonable.

Next we compared the time scale of fluctuations with generation time. A simplified autocorrelation function, which considered only the sign of a variation (positive or negative) and not the absolute value, was calculated to assess how long the protein levels in the daughter cells are dependent on the levels in

the first generation. The autocorrelation function decayed exponentially with a characteristic time of ∼30 minutes. These results are in good agreement with previous studies. For instance, Rosenfeld et al. suggested that variations in protein levels can persist on the generation time scale because there is no evidence for an active degradation of chemotaxis proteins under standard growth and assay conditions, and a decrease in protein levels mainly results from dilution during cell division [131].

4.2 Robustness of the bacterial chemotaxis pathway

Given that fluctuations persist on a generation time scale, the chemotaxis pathway is expected to exhibit robustness features compensating such noise in order to keep the CheY-P level stable. We studied the effects of the observed gene expression noise on the robustness of the pathway by first overexpressing all the chemotaxis proteins to observe how this affected the chemotactic ability of the cells in soft agar on swarm plates (fig. 3.11, p. 72). Astonishingly, an overexpression of all chemotaxis proteins by 6.6 fold above the native level hat little effect on chemotaxis ability. Expression levels below native level however, had a detrimental effect. We also measured the motor bias as a direct readout of the pathway output in tethering experiments. Clockwise bias of the flagellar motor was nearly unaffected by protein overexpression, confirming pathway robustness. Below the wild type level of chemotaxis proteins, only a small fraction of switching cells was observed, which hampered a reliable determination of the mean and standard deviation. The standard deviation decreased with expression level as expected.
Interestingly, plotting the growth rate (normalized to wild type growth rate) on the same graph as the swarming efficiency revealed the reason why wild type cells have such low levels of chemotaxis proteins. Overproducing proteins is a great metabolic burden to the cells, as seen from the decreasing growth rate with increasing protein levels, and *E. coli* seems to have evolutionary optimized to be chemotactic at the lowest possible protein levels.
Next, we analyzed the robustness of different hypothetical pathway topologies (fig. 3.27, p. 90) theoretically by simulating the number of chemotactic cells

under co-variation of chemotaxis protein levels (fig. 3.28 A, p. 92). We demonstrated that the established pathway topology reproduced the experimental data well. Simpler topologies failed to match the data. However, the topology with an additional feedback with phosphatase activation also matched the data rather well. Additionally, we simulated the fraction of chemotactic cells as a function of strength of correlated and uncorrelated variations (fig. 3.28 B, C, p. 92), and discovered that topologies B-D are sufficiently robust to compensate for co-variations, however tolerated only a low level of independent variations. Several features are primarily responsible for such robustness to correlated protein levels, the main being a balance of enzymatic activities in the pathway. Such compensation mechanism will be particularly efficient at high protein levels, whereas at the low levels chemotaxis becomes limited by the level of CheY independent of the assumed topology. There are two key features accounting for the higher robustness of the topologies shown in fig. 3.27 C, D. First, robustness against concerted variations requires a balance of phosphatase and kinase reactions and similar requirements for the methylation process. The condition for a robust adaptive chemotaxis pathway is that CheY-P demands a phosphatase (CheZ) whereas CheB-P must not have one. But this kind of robustness is only valid for concentrations larger than the wild-type level (fig. 3.11A). The strong decrease in the number of chemotactic bacteria at lower expression levels (fig. 3.28 A) arises because the total CheY concentration drops below the working range of the flagellar motor. The increase in chemotatic performance to above the wild type of the topologies in fig. 3.27 B-D with the mean expression level can be explained by an accompanying decline of the gene expression noise [41]. The second topological feature leading to higher swarming efficiency is CheB phosphorylation resulting in an additional negative feedback loop (fig. 3.27 C, D). This second feedback loop compensates partially for deviations from the optimal CheY-P level without changing the input–output characteristics. CheB phosphorylation has been previously shown to be non-essential for adaptation [3, 9] and our analysis suggests that its main function might be in noise reduction. The essential features for the two design principles described above seem to be present among most established pathway topologies of bacterial chemotaxis [159]. In particular, the CheB phosphorylation feedback is almost universal, and although many bacteria lack CheZ, the function of CheY phosphatase is taken over by another protein or by the kinase itself. The even higher robustness of the topology in fig. 3.27 D arises from the activation

of CheZ by CheY-P that follows the same noise compensatory mechanism as the CheB phosphorylation feedback. Reflecting strong selection, the chemotaxis pathway in *E. coli* seems to be optimized for high sensitivity, fast response and perfect adaptation [28, 146, 171, 172]. The experimentally established design of the chemotaxis network in *E. coli* represents a minimal topology providing high robustness to physiological perturbations. Taken together, our mathematical analysis suggests that the experimentally established design of the chemotaxis pathway (fig. 3.27 C) is robust to concerted variation in protein levels and variation in the level of phosphorylated CheY should be much smaller than the variation in the levels of chemotaxis proteins. Simulations however, predicted that the pathway is not robust against uncorrelated fluctuations.

Therefore, we investigated the effects of uncorrelated variations further experimentally by perturbing individual levels. In general, the chemotactic efficiency was indeed much more sensitive to the overexpression of individual proteins than to the correlated overexpression, thereby confirming our theoretical analysis. For the individual variation of CheY or CheZ levels, model predictions were rather exact and the experimental chemotactic efficiency dropped below 50 % of the maximum already upon a two-fold protein overexpression. However, the pathway was more robust than expected to the overexpression of the adaptation enzymes CheR and CheB, indicating that an additional robustness mechanism might be present at the level of the sensory complex. To investigate this further, the response to attractants and the adaptation time was measured by FRET. As expected, a decrease in kinase activity was observed upon overexpression of CheB and increase upon overexpression of CheR (fig. 3.14 A, B, p. 77), but these changes were more gradual than predicted. Both in FRET and swarm experiments, the pathway seemed more robust against CheR than CheB overexpression. Furthermore, adaptation time to addition of attractants was sensitive to individual levels of the adaptation enzymes (fig. 3.18 B, p. 80) where adaptation seemed more robust against CheB levels than CheR levels. Additionally, adaptation time to addition of attractant was similar for wild type CheB and CheB_{D56E}, confirming that the phosphorylation feedback activating CheB is not required for adaptation. Consistent with previous work which suggested that receptor binding of adaptation enzymes is not required for adaptation and can be compensated for by overexpression of CheR [1, 119], we observed a slow adaptation at very high levels of CheR in absence of the C-terminal pentapeptide binding sequence of the receptor Tar.

One of our most striking observations was that a concerted expression of just two proteins, CheY/CheZ, CheB/CheY or CheR/CheB enabled efficient chemotaxis in a much larger range of protein levels (fig. 3.16, p. 79). The positive effect of such concerted expression on the chemotactic performance was most clear at high protein levels. Additionally, cells with correlated levels of CheY and CheZ were evolutionary selected for in the process of spreading in soft agar. Making the levels of CheY and CheZ balance each other is physiologically meaningful, because the native expression of these proteins is strongly coupled. Such coupling between expression CheY and CheZ serves to reduce the level of CheY-P when CheY is up-regulated, by increasing the level of phosphatase and thereby returning the pathway to homeostasis. Inversely, coupling could increase the rate of CheY phosphorylation when CheZ is up-regulated. As expected due to the opposing activities of CheR and CheB, kinase activity was also robust against co-variation of the two as was also the adaptation time (fig. 3.18, p. 80). From the point of robustness, coupling between CheB and CheY is not surprising either. On one hand, these two proteins compete for CheA-dependent phosphorylation. On the other hand, higher CheB activity reduces the level of receptor methylation and thereby the rate of CheY phosphorylation. A co-elevated level of CheY would thus counteract an increase in the level of CheB both directly, by reducing CheB phosphorylation, and indirectly, by increasing the level of phospho-CheY. Similarly, an up-regulation of CheB should counterbalance an increased level of CheY. In theoretical simulations we found that the CheB phosphorylation feedback can have either positive or negative effects on chemotaxis efficiency, depending on how it shifts the average CheY-P level relatively to the region of linear motor response. In the case of non-perfect ratio of CheR to CheB, the CheB phosphorylation mechanism can partially repair the negative effect of unbalanced CheR/CheB levels, by shifting the average CheY-P towards the optimal region. This confirms that CheB phosphorylation can improve chemotactic properties of cells with deviations in the ratio of CheR/CheB, as well as in the ratios of other proteins, from the optimum [74].

Furthermore, our results suggest that - along with robust pathway topology and transcriptional coupling, translational coupling appears to specifically compensate the output level of CheY-P and thereby CW motor bias against stochastic variations in the translation efficiencies of individual genes. All the pairs mentioned above in addition to CheA and CheW were shown to be

coupled on a translational level (fig. 3.19, p. 82). The uncorrelated translational component of gene expression noise, which results from stochastic variations in the translation efficiencies of individual genes, is not negligible and results in significant uncorrelated variation in the levels of two proteins produced from one polycistronic mRNA [74]. Such "translation noise" could originate, for example, from stochastic fluctuations in the number of ribosomes that bind upstream of a particular gene. It is thus not surprising that bacteria evolved a mechanism to reduce the variation in the stoichiometry of functionally coupled proteins on the level of translation.

Translational coupling between bacterial genes that are encoded as one polycistronic mRNA has been described before, primarily in metabolic operons [53, 94, 120, 130, 135], but also between genes encoding ribosomal proteins [13], and a two-component sensor [91]. Such coupling mostly happens when the stop codon of a downstream gene is close to or overlaps with the start codon or with the Shine-Dalgarno sequence of the upstream gene. The molecular basis for translational coupling may be a reinitiation of translation, unwinding of the Shine-Dalgarno (SD) sequence of the distal gene, or a combination of both. The translation of the proximal gene will locally increase the levels of ribosomes close to the initiation codon of the distal gene, which could then efficiently reinitiate translation of the distal gene even in absence of a strong SD sequence [53]. Ribosomes translating the proximal gene will also unwind any secondary structure that might form around the SD sequence of the distal gene, as long as this sequence belongs to the translated region of the proximal gene. Such opening of the SD sequence will facilitate both reinitiation of translation by already bound ribosomes and the entry of new ribosomes [130]. Whatever is the mechanism of coupling, it has been proposed to enable a tighter control of the stoichiometry of protein complexes [120].

Another relevant observation was that an order of chemotaxis genes in bacterial genomes is not random, with a strong bias towards the same gene sequence that is present in *E. coli* (fig. 3.20, p. 83). In addition to the coupling between aforementioned genes, our analysis revealed a strong coupling between receptors *(mcp)* and *cheW*, in agreement with these gene products being parts of the same stable signalling complex. This coupling was stronger than that between receptors and *cheA*, apparently consistent with a role of CheW as an adapter between receptors and CheA [49]. Coupling between *cheZ* and *cheA*, which was also statistically significant, could serve a similar function

to the coupling between *cheY* and *cheZ* and compensate for an increase in the level of phosphatase by an increase in the kinase activity. Interestingly, there was a strong coupling between *cheY* and *cheA*, which agrees with an observation that additional CheYs can function as a "phosphate sink" and thus substitute for a phosphatase function in bacterial species that do not have a homologue of CheZ [149]. The reason for coupling between receptor genes (or *cheW*) and *cheR* is less obvious, but keeping a proper ratio between receptors and methyltransferase activity might be important for maintaining a constant steady-state level of receptor methylation. These findings are supported by computer simulations of the intercellular variation in CW motor bias in presence and in absence of translational coupling (fig. 3.30, p. 96). *In silico* analysis with an improved model clearly indicated higher robustness of particular arrangements of chemotaxis genes against translational noise, namely those that couple genes with opposing effects on the CheY-P level. The detailed ranking among several arrangements that fulfill this requirement depends only little on the reaction rates in the pathway but strongly on the translational noise strength of the individual genes. Although for equal noise strength for all genes the *E.coli* gene oder is on sixth place, it takes the first place if we assume that the weakly translated genes *cheR* and *cheB* have slightly higher (1.5 fold) noise level than the more efficiently translated genes *cheY* and *cheZ*.

An observed dramatic improvement of the robustness of the population upon coupling the expression of just two chemotaxis proteins and the observed strong bias towards a particular order of chemotaxis genes on the chromosome, stresses the significance of gene order and the organization into operons and regulons. It is widely believed that such organization mostly serves to facilitate the transcriptional regulation of the genes that share a common function. Consistent with theoretical predictions [157], such organization might also be used to reduce the effects of the gene expression noise on the pathway output. Placing related genes under the same control couples the variation in their expression, and the pathway can subsequently evolve to be robust against such concerted noise. Such organization is evolutionary beneficial because it improves robustness of the signalling output without adding a cost of the increased complexity and is thus expected to be ubiquitous in bacterial networks. This might be true even for eukaryotic pathways, because – although polycistronic organization is absent in most eukaryotes – their genes are organized in regulons, and the expression levels of neighbouring genes are

frequently coupled on the level of chromatin remodelling [12].
With our modified theoretical model of the chemotaxis we looked for additional robustness features explaining the increased robustness to the individual protein levels. We found experimentally that chemotaxis ability is independent of CheZ localization (fig. 3.25, p. 88). Additionally, the phosphorylation feedback to CheB was not absolutely required for robustness (fig. 3.24, p. 87). The only plausible robustness feature left is the competition of CheY and CheB for the P2 domain of CheA, and our experiments demonstrated that robustness of chemotaxis was indeed reduced in absence of the P2 domain (fig. 3.25) and the P2 domain improved selection for correlated levels of CheY and CheZ (fig. 3.26, p. 89), confirming our hypothesis that competition for the P2 domain of the kinase is an important robustness factor in the chemotaxis pathway (fig. 4.1).
Finally, to ascertain whether gene expression noise can have positive effects

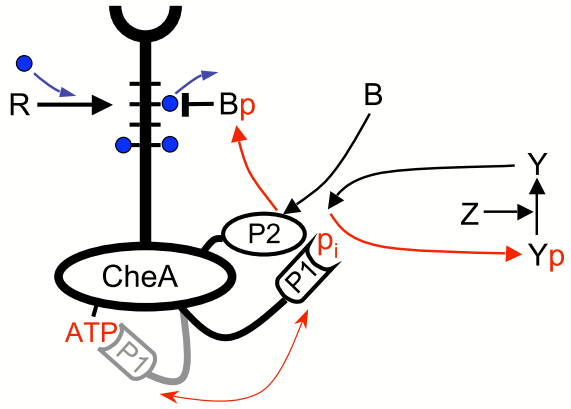

Figure 4.1: Topology of the pathway with competition for the CheA P2 domain increasing the robustness A receptor dimer has 8 methylation sites. The histidine kinase CheA binds with high affinity to these clusters to form functional CheA-CheW–Receptor complexes. Upon receptor activation, CheA autophosphorylates through interaction of its phosphorylation site domain, P1, with the ATP binding domain. Phosphotransfer from P1 to the response regulator CheY can occur in two ways, rapid transfer to CheY located at the P2 binding domain of CheA with rate k_2, and direct transfer at a significantly lower rate k_1. Phosphorylation dynamics of CheB works similar to that of CheY but with a much higher auto-dephosphorylation rate and absence of a designated phosphatase.

on the chemotaxis pathway, we measured the chemotactic ability in different nutrition gradients under varying concentrations of the adaptation enzymes (fig. 3.32, p. 100). We observed that cells with CheR and CheB above a certain threshold perform chemotaxis equally efficient as in the wild type, whereas lower CheR/CheB levels would severely impair chemotactic behaviour. Additionally,

very high CheR/CheB levels significantly decreased chemotactic efficiency. Cells with low CheR/CheB would according to the simulations tend to run without tumbling and stay trapped in the pores in the agar most of the time. On the other hand, cells with extremely high CheR/CheB levels loose their sensitivity to the gradient and also have poor chemotaxis efficiency. Indeed, swarm selection experiments revealed that there is a selection against low copy numbers of CheR/CheB and against extremely high levels. In contrast, in liquid culture, no selection against low copy numbers was observed and simulations revealed that cells with optimal CheR/CheB levels drifted more rapidly than other cells, with the optimal levels depending on the gradient (fig. 3.31, p. 98). These results suggest some evolutional implications. In the liquid medium with variable food sources and gradient intensities, variability of adaptation times (protein levels) among cells can help the whole population to respond to different gradients more readily, due to positive selection of cells with optimal CheR/CheB levels. In other words, for any given gradient steepness there will be a subpopulation which has the best CheR/CheB levels to follow this gradient. In contrast, agar poses mainly negative selection on cell population - cells with low CheR/CheB levels are altered out from competition, while all other cells travel with approximately equal efficiency. Taken together, these results suggest that gene expression noise can indeed also be advantageous to the cells.

5 Conclusions and outlook

Chemotaxis allows bacteria to colonize their environment more efficiently and to find optimal growth conditions, and is consequently under strong evolutionary selection in a noisy environment. In this work, we were able to characterize the observed fluctuations in the chemotaxis- and flagella protein levels. Cell-to-cell variations in protein levels seemed to be dominated by global variations and to a lesser degree variations in transcription. Such fluctuations remain in the cell for one generation, underlining the importance of robustness mechanisms to keep the pathway output, the CheY-P concentration, within the optimal range for chemotaxis.

The experimentally established design of the chemotaxis network in *E. coli* appears to represent a minimal topology providing high robustness to physiological perturbations. This network design can compensate for the observed strong co-variations in gene expression, consistent with the experimentally observed transcriptional coupling, but the negative effect of uncorrelated variations on the efficiency of chemotaxis can only be attenuated. Much of this robustness arises from balance of counteracting enzymes, and in addition we revealed a pair-wise translational coupling of adjacent genes enhancing the pathway robustness to uncorrelated fluctuations. In collaboration with theorists we found that the order of genes in *E. coli* ranks among the best in terms of noise compensation. Bioinformatics analysis further revealed that *E. coli* gene order corresponds to a consensus in sequenced bacterial genomes, confirming evolutionary selection for noise reduction.

By measuring chemotaxis efficiency and kinase activity, we observed that the phosphorylation feedback from CheA to CheB is not necessary for functional chemotaxis, but may improve robustness. Furthermore, cells lacking the P2 domain of CheA were less robust against co-variations in CheY and CheZ than wild type cells. Thus, we discovered that activation of CheB through phosphorylation, and particularly competition of CheB and CheY for the binding to CheA are features of the chemotaxis pathway with the function to increase its

robustness to fluctuations in protein levels.
To increase the understanding of the robustness of the chemotaxis pathway, we started some preliminary experiments to test swarming in cells with different topologies of the signalling pathway. Strains, unable to swarm in soft agar due to an in-frame deletion of an individual chemotaxis gene, were used for evolution experiments. After several generations in soft agar, suppressor mutants appeared. These mutants were selected and re-inoculated on a fresh plate, and over several re-inoculations (generations), the swarming ability of a population improved, as shown in fig. 5.1. Our future perspective is to purify cells with improved chemotactic ability and perform targeted sequencing to obtain the mutation causing the improvement and to exclude epigenetic factors. A question that arises is if one would always find the same mutation, or if other suppressor mutants could also be achieved. Our goal would be to systematically screen several suppressor mutants in all deletion strains, and finally test the robustness against variations in protein levels in these strains. This approach should increase the knowledge about the robustness mechanisms of the chemotaxis pathway, and provide a greater understanding of robustness in similar signalling pathways in other organisms.

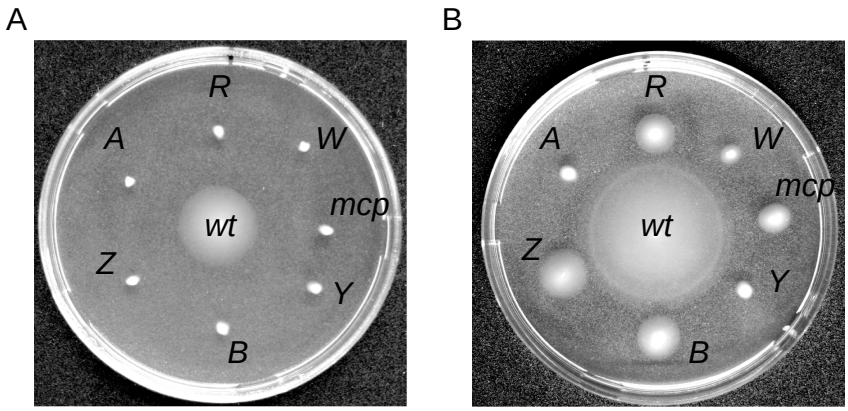

Figure 5.1: Swarming of suppressor mutant cells after evolution. (A) Cells with deletions as marked with eg. *R* for $\Delta cheR$, which are not swarming. *mcp* denotes cells without any methyl-accepting chemotaxis proteins (receptors). Wild type (RP437) is shown as a positive control. (B) Same strains after two rounds of evolution. Strains lacking CheZ or CheB have the best suppressors and have a swarming efficiency about 30 % of the wild type. Also cells lacking CheW and the chemoreceptors seem to have started swarming.

A Appendix

A.1 Supplementary methods

A.1.1 Crossing in a gene using LambdaInCh

([27] (http://rcc.med.harvard.edu/ dboyd.html) First a λInCh lysate was grown. The strain Joe59 or DHB6521 is lysogenic for λInCh1. λInCh2 can be used in the same way.
a. Streak Joe59 or DHB6521 on a broth plate and incubate overnight at 30 °C.
b. Pick a single colony and grow a 5 ml overnight culture in broth at 30 °C.
c. Dilute 100X in 5 ml broth with 2 mM MgSO4.
d. Incubate at 42 °C for 15 minutes.
e. Incubate at 37 °C with good aeration. Lysis should be apparent after about 1 hour.
f. Add 1/100 th volume of chloroform and vortex.
g. Continue incubation with aeration at 37 °C for 15 min.
h. Spin 10,000 RPM 10 min at 4 °C.
i. Decant the supernatant into a sterile tube with 1/100th volume of chloroform and vortex. This is the primary λInCh1 lysate. Store it at 4 °C.
Second, titer the lysate on a permissive strain.
a. Grow a 5ml overnight culture of SM551 or DHB6501 from a single colony in broth at 37 °C.
b. Dilute 100X in 5 ml broth with 2 mM $MgSO_4$ and 0.2 % maltose.
c. Incubate at 37 °C with aeration until OD_{600} is 0.2.
d. Put the culture in ice. This is the plating culture.
e. Make serial 10X dilutions of the lysate in 5ml broth with 2mM MgSO4. Make dilutions of 10^{-6} to 10^{-12}.
f. Mix 0.1 ml of each lysate dilution with 0.1ml of plating culture. g. Incubate 5 minutes at 37 °C.

h. Add 3.3 ml of melted 0.7 % top agar at 47 °C to the mixture of lysate and plating culture. Vortex and immediately pour onto a broth agar plate containing 2 mM $MgSO_4$ and 0.2 % maltose. Distribute the melted agar uniformly before it hardens.
i. Incubate at 37 °C overnight.
j. Determine the phage titer by counting plaques on a plate with 30 to 300 colonies.
Third, make a lysogen in your plasmid containing strain. Be sure that your plasmid confers ampicillin resistance and is capable of pickup by λInCh1. (The Homology Detector applications, InCh for pBR and InCh for pUC, which run on Macintosh computers are useful for determining whether your plasmid insert can be picked up.)
a. Grow an overnight culture from a single colony of your strain in broth with 2 mM $MgSO_4$ 0.2 % maltose and 0.2 mg/ml of ampicillin to select for maintenance of your plasmid.
b. Mix 0.1 ml of overnight with 10^8 λInCh1 and incubate at 30 °C for 15 minutes. Dilute to 1 ml in broth with 2 mM $MgSO_4$ and incubate with aeration for 1 hour at 30 °C.
c. Plate serial dilutions on media with 0.2 mg/ml ampicillin and 0.04 mg/ml kanamycin and incubate overnight at 30 °C.
d. Pick and purify by successive restreaking several colonies from the highest dilution plate that has any colonies. These should be λInCh lysogens of your strain.
e. Test these strains by streaking on a pair of plates. Incubate one at 42 °C and the other at 30 °C overnight. Lysogens should have no or few colonies at 42 °C but should grow normally at 30 °C. Choose a your primary lysogen for the next step.
Fourth, grow a lambda inch lysate from your primary lysogen. This is similar to the first step above except that lysis must be induced with chloroform (unless your strain is a supF strain, which is permissive for the growth of λInCh1). During growth of the phage a double recombination event can result in transfer of ampicillin resistance and your plasmid insert onto the phage DNA. This is a reasonably common event even though the amount of homology is not great because recombination frequencies are high during lambda growth.
a. Pick a single colony of your lysogen and grow a 5 ml overnight culture in broth at 30 °C selecting for maintenance of your plasmid with ampicillin.

b. Dilute 100X in 5 ml broth with 2 mM $MgSO_4$.
c. Incubate at 42 °C for 15 minutes.
d. Incubate at 37 °C with good aeration for 3 hours.
e. Add 1/100 th volume of chloroform and vortex.
f. Continue incubation with aeration at 37 °C for 15 min. Lysis may or may not be apparent.
g. Spin 10,000 RPM 10 min at 4 °C.
h. Decant the supernatant into a sterile tube with 1/100 th volume of chloroform and vortex.
This is the low frequency transducing (LFT) lysate. Most of the phage particles should carry the KanR allele of λInCh1 but some should have recombined with your plasmid and carry the ampR gene (and your insert) instead.
Fifth, make an ampicillin resistant secondary lysogen. This step involves insertion of the recombinant lambda onto the chromosome by site specific recombination, the second of the three recombination events mentioned above. This is an efficient process.
a. Grow an overnight culture of DHB6501 in 5 ml broth with 2 mM $MgSO_4$ and 0.2 %maltose.
b. Mix 0.1ml serial 10-fold dilutions of the LFT lysate with 0.1 ml aliquots of the culture.
c. Incubate 15 min at 30 °C, dilute to 1 ml in broth with 2 mM $MgSO_4$ and incubate 1 hour at 30 °C.
d. Spread 0.1 ml of each culture on broth plates with 0.025 mg/ml ampicillin. Concentrate the cells in the culture with the highest number of phage by centrifugation and plate on the same medium. The low concentration of ampicillin used in this step is appropriate for use with pBAD and pDHB60 derived plasmids. The Ampicillin resistance gene on these plasmids confers only low level resistance when in single copy.
e. Incubate overnight at 30 °C
f. Pick colonies from the lowest dilution plate that has colonies. These should be lysogenic for lInCh1 with your insert recombined into the phage.
g. Purify several colonies on broth agar plates without antibiotics and test for AmpR, KanR, and growth at 42 °C and 30 °C. Chose your secondary lysogen from among the colonies with the correct phenotype, AmpR, KanS and TS.
Note that it is possible to transduce the plasmid in this step, presumably as a cointegrate which is resolved upon lysogenization. Use of high concentrations

of ampicillin selects for this.
Sixth, grow a high frequency transducing (HFT) lysate. Make and titer a lysate from the strain obtained in step 5 above, the secondary lysogen, using the procedures of steps 1 or 4 and 2 above.
Seventh, select lysogens in your target strain(s) using the HFT lysate and the procedures of step 5 above but with your strain(s) as recipient and the HFT lysate as the phage stock. These tertiary lysogens are the precursors of the final cured strains you will construct. This step again involves site specific recombination at the att site. It actually repeats the second of the three recombination events that are involved in the construction. Eighth, cure the tertiary lysogens to stabilize the chromosomal construct. This step involves homology dependent recombination between the chromosomal region just before the att site and a region in the prophage. This is the third recombination involved in the construction. This deletion results in loss of all functional lambda genes and almost all the lambda DNA. Cells with the deletion are no longer temperature sensitive so they can grow at 42 °C.
a. Grow overnight cultures in broth at 30 °C
b. Plate dilutions on broth plates and incubate overnight at 42 °C
c. Pick well isolated colonies from the plates with the fewest cells plated.
d. Purify by streaking on broth plates and incubating at room temperature.
e. Test the cured strains for antibiotic resistance and temperature sensitivity as in step 5. The cured strains should be resistant to .025 mg/ml ampicillin and no longer temperature sensitive.
Ninth, transducing the ampicillin resistance. The ampicillin resistance allele should be 100% linked to your plasmid insert at the att site on the chromosome in the cured strains. You can transduce it to other strains with P1 selecting resistance to .025 mg/ml ampicillin. To do this either cure the secondary lysogen obtained in step 5 using the method outlined in step 8 or use one of the cured strains obtained in step 8. Grow P1 on the cured strain and use the lysate to transduce ampicillin resistance.

A.2 Supplementary figures

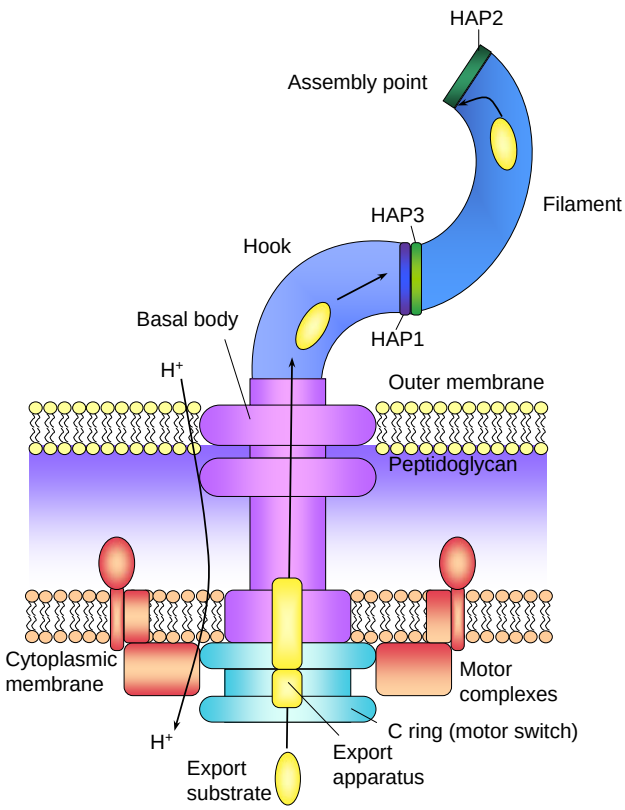

Figure A.1: Model of the bacterial flagellum-structure and assembly. The bacterial flagellum consists of three parts: the filament, the hook and the complex basal body. The proton gradient across the cytoplasmic membrane drives rotation of the rotor and the attached filament. CCW, counterclockwise; CW, clockwise; HAP, hook-associated protein.

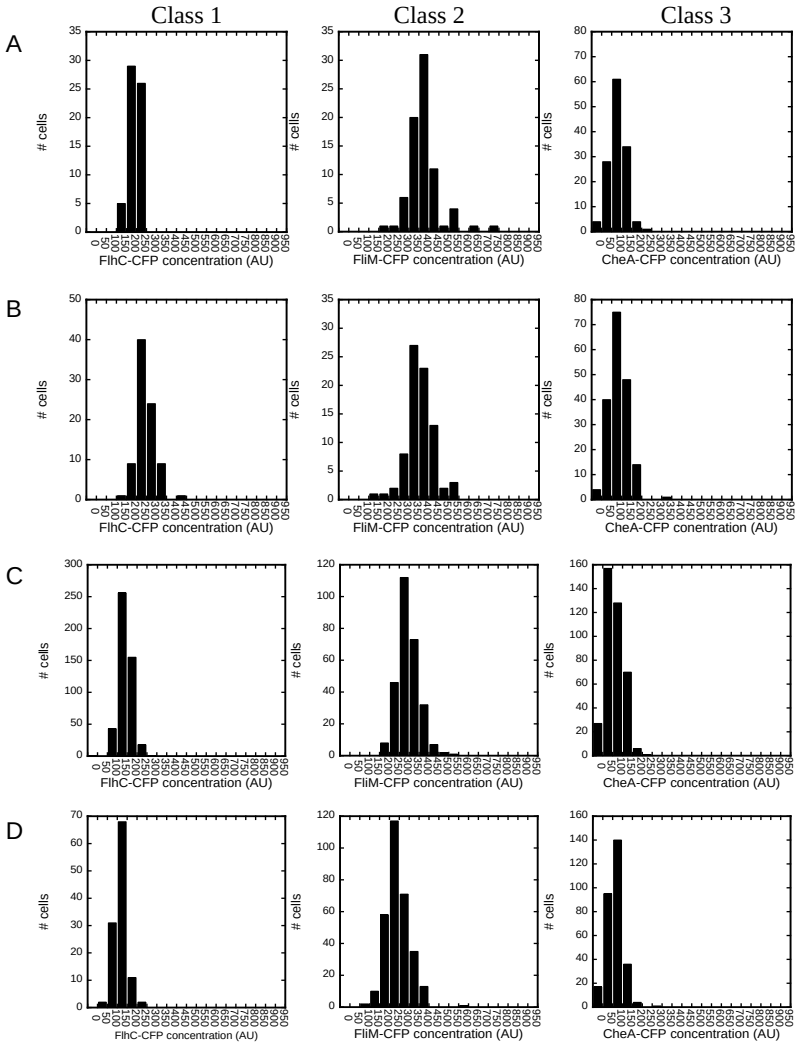

Figure A.2: Distribution of single cell CFP levels in different classes in flagella gene hierarchy. FlhC in class 1, FliM in class 2 and CheA in class 3 are expressed from the native promoter on the chromosome as fusions to CFP. Distributions of single cell absolute CFP concentrations from 4 different experiments (A-D). AU, arbitrary units.

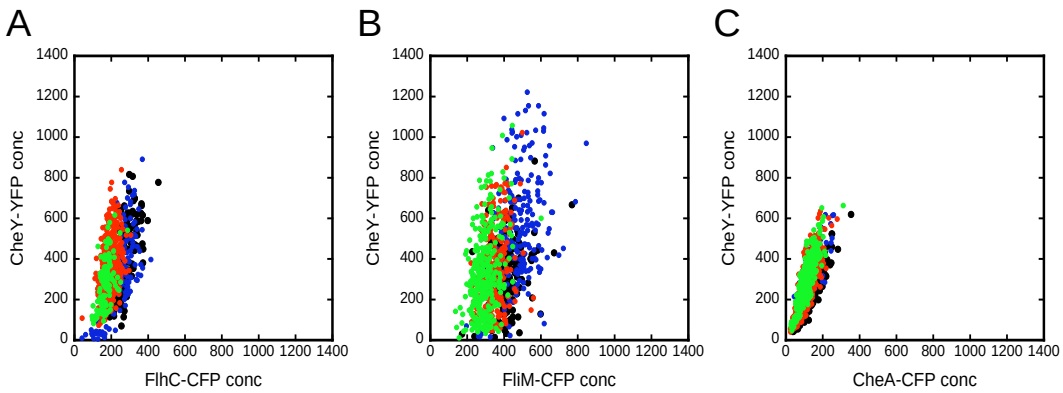

Figure A.3: Scatterplots of single cell YFP and CFP levels in all three classes of the hierarchy. Relative concentrations of CheY-YFP in class 3 compared to (A) FlhC-CFP in class 1, (B) FliM-CFP in class 2 and (C) CheA-CFP in class 3. Results from four independent experiments are pooled, and each experiment is shown in a different colour.

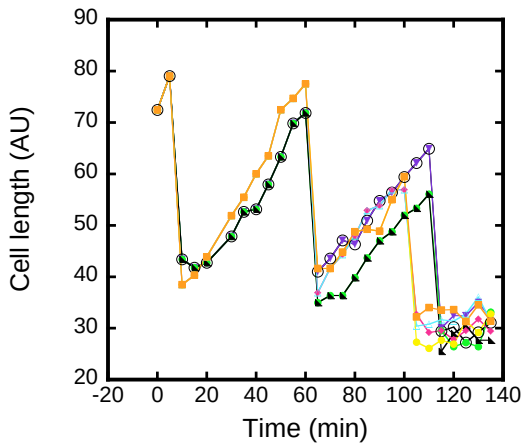

Figure A.4: Cell length over time for one cell line, corresponing to fig. 3.8 B. Cell size decreases over time under our experimental conditions.

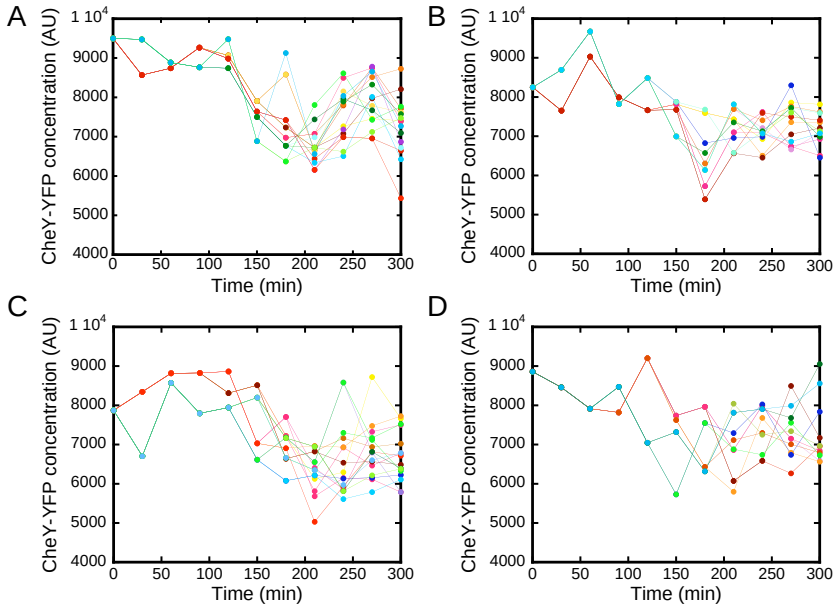

Figure A.5: CheY-YFP concentration over time in four representative cell lines The descendants of the first two daughter cells are shown in either red/orange or blue/green. Light blue arrows show the generation time for each generation, which varies for every cell line, but on average is about 60 minutes.

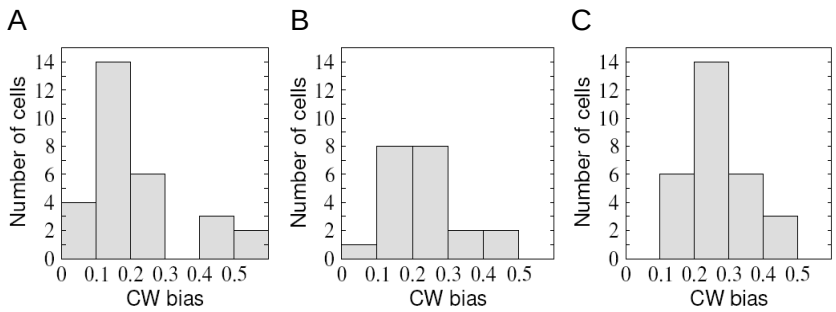

Figure A.6: Histograms of CW bias distributions at varying levels of chemotaxis proteins in (A) wild type cells, (B) IPTG inducible cells with 3.2 fold wild type expression and (C) FlgM minus strain with 6.6 fold wild type expression.

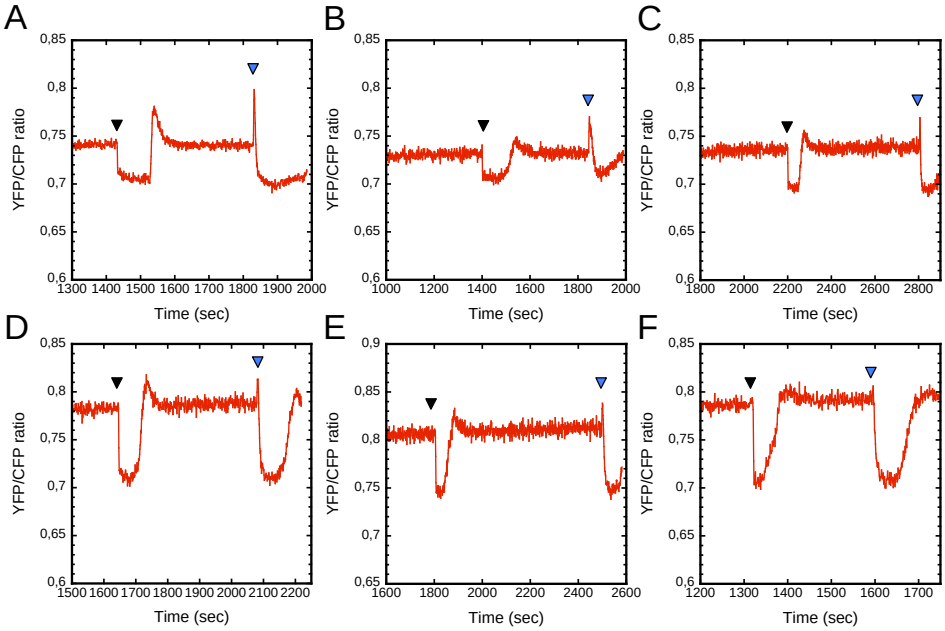

Figure A.7: Response amplitudes to saturating levels of attractant and repellent upon overexpression of CheR. Responses to 1 mM methyl-aspartate (black triangle) and 1 mM $NiCl_2$ (blue triangle). (A) Wild type cells, (B-D) CheR expressed from pVS113 at 0.0004, 0.001 and 0.01 % arabinose respectively, (E-F) CheR expressed from pLL45 at 0.001 and 0.01 % arabinose.

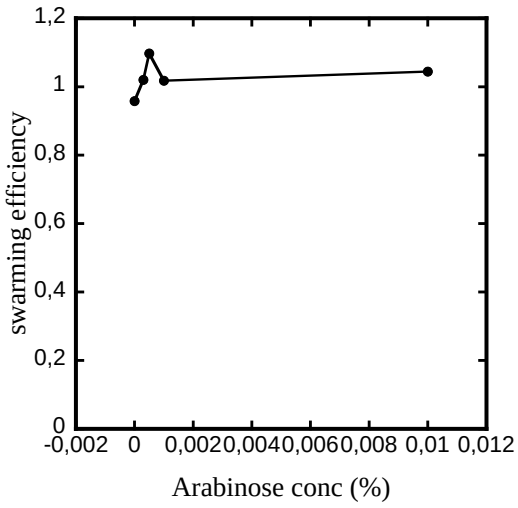

Figure A.8: Swarming efficiency upon overexpression of a catalytical mutant of CheR A catalytical mutant of CheR, $CheR^{D154A}$, which can localize and bind to the receptors but is not catalytically active, was expressed from pDK167 in the wild type strain RP437 with native levels of CheB. Swarm efficiency was the same at all CheR levels, hence there is no sterical hindrance for CheB.

Bibliography

[1] Alexander, R. P. and Zhulin, I. B. (2007). Evolutionary genomics reveals conserved structural determinants of signaling and adaptation in microbial chemoreceptors. *Proc Natl Acad Sci U S A*, 104(8):2885–90. 18, 76, 105

[2] Almogy, G., Stone, L., and Ben-Tal, N. (2001). Multi-stage regulation, a key to reliable adaptive biochemical pathways. *Biophys J*, 81(6):3016–28. 90

[3] Alon, U., Surette, M. G., Barkai, N., and Leibler, S. (1999). Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–71. 6, 7, 18, 86, 89, 104

[4] Ames, P. and Parkinson, J. S. (1994). Constitutively signaling fragments of Tsr, the *Escherichia coli* serine chemoreceptor. *J Bacteriol*, 176(20):6340–8. 15

[5] Ames, P., Studdert, C. A., Reiser, R. H., and Parkinson, J. S. (2002). Collaborative signaling by mixed chemoreceptor teams in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 99(10):7060–5. 8, 18, 19

[6] Anand, G. S. and Stock, A. M. (2002). Kinetic basis for the stimulatory effect of phosphorylation on the methylesterase activity of CheB. *Biochemistry*, 41(21):6752–60. 11, 17

[7] Arkin, A., Ross, J., and McAdams, H. H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics*, 149(4):1633–48. 4

[8] Banno, S., Shiomi, D., Homma, M., and Kawagishi, I. (2004). Targeting of the chemotaxis methylesterase/deamidase CheB to the polar receptor-kinase cluster in an *Escherichia coli* cell. *Mol Microbiol*, 53(4):1051–63. 13, 17

[9] Barkai, N. and Leibler, S. (1997). Robustness in simple biochemical networks. *Nature*, 387(6636):913–7. 7, 17, 18, 19, 86, 89, 90, 91, 104

[10] Barkai, N. and Leibler, S. (2000). Circadian clocks limited by noise. *Nature*, 403(6767):267–8. 5

[11] Barnakov, A. N., Barnakova, L. A., and Hazelbauer, G. L. (2001). Location of the receptor-interaction site on CheB, the methylesterase response regulator of bacterial chemotaxis. *J Biol Chem*, 276(35):32984–9. 17

[12] Batada, N. N., Urrutia, A. O., and Hurst, L. D. (2007). Chromatin remodelling is a major source of coexpression of linked genes in yeast. *Trends Genet*, 23(10):480–4. 109

[13] Baughman, G. and Nomura, M. (1983). Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in *Escherichia coli*. *Cell*, 34(3):979–88. 6, 80, 107

[14] Becskei, A. and Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–3. 1, 5

[15] Bengtsson, M., Stahlberg, A., Rorsman, P., and Kubista, M. (2005). Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Res*, 15(10):1388–92. 60

[16] Berg, H. C. (2003). The rotary motor of bacterial flagella. *Annu Rev Biochem*, 72:19–54. 8

[17] Berg, H. C. and Brown, D. A. (1972). Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*, 239(5374):500–4. 8

[18] Berg, H. C. and Purcell, E. M. (1977). Physics of chemoreception. *Biophys J*, 20(2):193–219. 18

[19] Berg, H. C. and Tedesco, P. M. (1975). Transient response to chemotactic stimuli in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 72(8):3235–9. 97

[20] Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. (1999). Structure of CheA, a signal-transducing histidine kinase. *Cell*, 96(1):131–41. 15

[21] Bilwes, A. M., Quezada, C. M., Croal, L. R., Crane, B. R., and Simon, M. I. (2001). Nucleotide binding by the histidine kinase CheA. *Nat Struct Biol*, 8(4):353–60. 15

[22] Blake, W. J., M, K. A., Cantor, C. R., and Collins, J. J. (2003). Noise in eukaryotic gene expression. *Nature*, 422(6932):633–7. 3, 62

[23] Blat, Y., Gillespie, B., Bren, A., Dahlquist, F. W., and Eisenbach, M. (1998). Regulation of phosphatase activity in bacterial chemotaxis. *J Mol Biol*, 284(4):1191–9. 90

[24] Borisuk, M. T. and Tyson, J. J. (1998). Bifurcation analysis of a model of mitotic control in frog eggs. *J Theor Biol*, 195(1):69–85. 6

[25] 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(4):1208–12. 15

[26] Bourret, R. B., Davagnino, J., and Simon, M. I. (1993). The carboxy-terminal portion of the CheA kinase mediates regulation of autophosphorylation by transducer and CheW. *J Bacteriol*, 175(7):2097–101. 15

[27] Boyd, D., Weiss, D. S., Chen, J. C., and Beckwith, J. (2000). Towards single-copy gene expression systems making gene cloning physiologically relevant: lambda InCh, a simple *Escherichia coli* plasmid-chromosome shuttle system. *J Bacteriol*, 182(3):842–7. 44, 45, 113

[28] Bray, D., Levin, M. D., and Morton-Firth, C. J. (1998). Receptor clustering as a cellular mechanism to control sensitivity. *Nature*, 393(6680):85–8. 105

[29] Bundschuh, R., Hayot, F., and Jayaprakash, C. (2003). The role of dimerization in noise reduction of simple genetic networks. *J Theor Biol*, 220(2):261–9. 5

[30] Cantwell, B. J., Draheim, R. R., Weart, R. B., Nguyen, C., Stewart, R. C., and Manson, M. D. (2003). CheZ phosphatase localizes to chemoreceptor patches via CheA-short. *J Bacteriol*, 185(7):2354–61. 17, 86

[31] Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. *Cell*, 78(5):823–34. 4

[32] Chilcott, G. S. and Hughes, K. T. (2000). Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica serovar typhimurium* and *Escherichia coli*. *Microbiol Mol Biol Rev*, 64(4):694–708. 23

[33] Ciliberti, S., Martin, O. C., and Wagner, A. (2007). Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput Biol*, 3(2):e15. 7

[34] Cluzel, P., Surette, M., and Leibler, S. (2000). An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science*, 287(5458):1652–5. 16, 54, 70, 72, 89, 91, 95

[35] Djordjevic, S., Goudreau, P. N., Xu, Q., Stock, A. M., and West, A. H. (1998). Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. *Proc Natl Acad Sci U S A*, 95(4):1381–6. 17

[36] Djordjevic, S. and Stock, A. M. (1997). Crystal structure of the chemotaxis receptor methyltransferase CheR suggests a conserved structural motif for binding S-adenosylmethionine. *Structure*, 5(4):545–58. 17

[37] Djordjevic, S. and Stock, A. M. (1998a). Chemotaxis receptor recognition by protein methyltransferase CheR. *Nat Struct Biol*, 5(6):446–50. 17

[38] Djordjevic, S. and Stock, A. M. (1998b). Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system. *J Struct Biol*, 124(2-3):189–200. 11, 17

[39] Duke, T. A. and Bray, D. (1999). Heightened sensitivity of a lattice of membrane receptors. *Proc Natl Acad Sci U S A*, 96(18):10104–8. 20

[40] Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–8. 2

[41] Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297(5584):1183–6. 2, 3, 61, 90, 104

[42] Endres, R. G., Oleksiuk, O., Hansen, C. H., Meir, Y., Sourjik, V., and Wingreen, N. S. (2008). Variable sizes of *Escherichia coli* chemoreceptor signaling teams. *Mol Syst Biol*, 4:211. 20, 21

[43] Endres, R. G. and Wingreen, N. S. (2006). Precise adaptation in bacterial chemotaxis through "assistance neighborhoods". *Proc Natl Acad Sci U S A*, 103(35):13040–4. 56

[44] Falke, J. J. and Hazelbauer, G. L. (2001). Transmembrane signaling in bacterial chemoreceptors. *Trends Biochem Sci*, 26(4):257–65. 13, 14

[45] Foerster, T. (1948). Zwischen Energiewanderung und Fluoreszenz. *Annalen der Physik*, 6(2). 48

[46] Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J., and Eisen, M. B. (2004). Noise minimization in eukaryotic gene expression. *PLoS Biol*, 2(6):e137. 6

[47] Frisk, A., Jyot, J., Arora, S. K., and Ramphal, R. (2002). Identification and functional characterization of textitflgM, a gene encoding the anti-sigma 28 factor in *Pseudomonas aeruginosa*. *J Bacteriol*, 184(6):1514–21. 24

[48] 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(3):750–4. 15

[49] 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(6):975–82. 12, 15, 18, 19, 107

[50] Gestwicki, J. E. and Kiessling, L. L. (2002). Inter-receptor communication through arrays of bacterial chemoreceptors. *Nature*, 415(6867):81–4. 19

[51] Gestwicki, J. E., Strong, L. E., and Kiessling, L. L. (2000). Tuning chemotactic responses with synthetic multivalent ligands. *Chem Biol*, 7(8):583–91. 18

[52] Golding, I., Paulsson, J., Zawilski, S. M., and Cox, E. C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–36. 2

[53] Govantes, F., Andujar, E., and Santero, E. (1998). Mechanism of translational coupling in the *nifLA* operon of *Klebsiella pneumoniae*. *Embo J*, 17(8):2368–77. 107

[54] Griswold, I. J. and Dahlquist, F. W. (2002). The dynamic behavior of CheW from *Thermotoga maritima* in solution, as determined by nuclear magnetic resonance: implications for potential protein-protein interaction sites. *Biophys Chem*, 101-102:359–73. 15

[55] Hansen, C. H., Endres, R. G., and Wingreen, N. S. (2008). Chemotaxis in *Escherichia coli*: a molecular model for robust precise adaptation. *PLoS Comput Biol*, 4(1):e1. 18, 56

[56] Homma, M., Shiomi, D., Homma, M., and Kawagishi, I. (2004). Attractant binding alters arrangement of chemoreceptor dimers within its cluster at a cell pole. *Proc Natl Acad Sci U S A*, 101(10):3462–7. 18, 19

[57] Hooshangi, S., Thiberge, S., and Weiss, R. (2005). Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci U S A*, 102(10):3581–6. 4

[58] Hughes, K. T., Gillen, K. L., Semon, M. J., and Karlinsey, J. E. (1993). Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science*, 262(5137):1277–80. 24

[59] Ingolia, N. T. (2004). Topology and robustness in the drosophila segment polarity network. *PLoS Biol*, 2(6):e123. 7

[60] Ishihara, A., Segall, J. E., Block, S. M., and Berg, H. C. (1983). Coordination of flagella on filamentous cells of *Escherichia coli*. *J Bacteriol*, 155(1):228–37. 8

[61] Jahreis, K., Morrison, T. B., Garzon, A., and Parkinson, J. S. (2004). Chemotactic signaling by an *Escherichia coli* CheA mutant that lacks the binding domain for phosphoacceptor partners. *J Bacteriol*, 186(9):2664–72. 14, 21, 88

[62] Jarrell, K. F. and McBride, M. J. (2008). The surprisingly diverse ways that prokaryotes move. *Nat Rev Microbiol*, 6(6):466–76. 22

[63] Kalir, S. and Alon, U. (2004). Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. *Cell*, 117(6):713–20. 23, 63

[64] Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M. G., and Alon, U. (2001). Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science*, 292(5524):2080–3. 24

[65] Karlinsey, J. E., Tanaka, S., Bettenworth, V., Yamaguchi, S., Boos, W., Aizawa, S. I., and Hughes, K. T. (2000). Completion of the hook-basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and fliC transcription. *Mol Microbiol*, 37(5):1220–31. 24

[66] Kemkemer, R., Schrank, S., Vogel, W., Gruler, H., and Kaufmann, D. (2002). Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 in vitro. *Proc Natl Acad Sci U S A*, 99(21):13783–8. 4

[67] Kentner, D. and Sourjik, V. (2006). Spatial organization of the bacterial chemotaxis system. *Curr Opin Microbiol*, 9(6):619–24. 15

[68] Kepler, T. B. and Elston, T. C. (2001). Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys J*, 81(6):3116–36. 3, 6

[69] Keymer, J. E., Endres, R. G., Skoge, M., Meir, Y., and Wingreen, N. S. (2006). Chemosensing in *Escherichia coli*: two regimes of two-state receptors. *Proc Natl Acad Sci U S A*, 103(6):1786–91. 20

[70] Kierzek, A. M., Zaim, J., and Zielenkiewicz, P. (2001). The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J Biol Chem*, 276(11):8165–72. 3, 6

[71] 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(6746):787–92. 18

[72] Kitano, H. (2004). Biological robustness. *Nat Rev Genet*, 5(11):826–37. 6, 7

[73] Kojima, S. and Blair, D. F. (2004). Solubilization and purification of the MotA/MotB complex of *Escherichia coli*. *Biochemistry*, 43(1):26–34. 23

[74] Kollmann, M., Lovdok, L., Bartholome, K., Timmer, J., and Sourjik, V. (2005). Design principles of a bacterial signalling network. *Nature*, 438(7067):504–7. 7, 18, 54, 55, 106, 107

[75] 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(1):131–40. 15

[76] Kutsukake, K., Ikebe, T., and Yamamoto, S. (1999). Two novel regulatory genes, *fliT* and /textitfliZ, in the flagellar regulon of Salmonella. *Genes Genet Syst*, 74(6):287–92. 24

[77] Lai, R. Z., Manson, J. M., Bormans, A. F., Draheim, R. R., Nguyen, N. T., and Manson, M. D. (2005). Cooperative signaling among bacterial chemoreceptors. *Biochemistry*, 44(43):14298–307. 19

[78] Lai, W. C., Beel, B. D., and Hazelbauer, G. L. (2006). Adaptational modification and ligand occupancy have opposite effects on positioning of the transmembrane signalling helix of a chemoreceptor. *Mol Microbiol*, 61(4):1081–90. 13

[79] Lamanna, A. C., Ordal, G. W., and Kiessling, L. L. (2005). Large increases in attractant concentration disrupt the polar localization of bacterial chemoreceptors. *Mol Microbiol*, 57(3):774–85. 19

[80] Landini, P. and Zehnder, A. J. (2002). The global regulatory hns gene negatively affects adhesion to solid surfaces by anaerobically grown escherichia coli by modulating expression of flagellar genes and lipopolysaccharide production. *J Bacteriol*, 184(6):1522–9. 23

[81] Laub, M. T. and Goulian, M. (2007). Specificity in two-component signal transduction pathways. *Annu Rev Genet*, 41:121–45. 10, 11

[82] Lehnen, D., Blumer, C., Polen, T., Wackwitz, B., Wendisch, V. F., and Unden, G. (2002). Lrha as a new transcriptional key regulator of flagella, motility and chemotaxis genes in escherichia coli. *Mol Microbiol*, 45(2):521–32. 23

[83] Levin, M. D. (2003). Noise in gene expression as the source of non-genetic individuality in the chemotactic response of escherichia coli. *FEBS Lett*, 550(1-3):135–8. 1

[84] Levit, M. N., Grebe, T. W., and Stock, J. B. (2002). Organization of the receptor-kinase signaling array that regulates *Escherichia coli* chemotaxis. *J Biol Chem*, 277(39):36748–54. 15, 19

[85] Levit, M. N., Liu, Y., and Stock, J. B. (1999). Mechanism of CheA protein kinase activation in receptor signaling complexes. *Biochemistry*, 38(20):6651–8. 15

[86] Li, G. and Weis, R. M. (2000). Covalent modification regulates ligand binding to receptor complexes in the chemosensory system of escherichia coli. *Cell*, 100(3):357–65. 19

[87] Li, J., Swanson, R. V., Simon, M. I., and Weis, R. M. (1995). The response regulators CheB and CheY exhibit competitive binding to the kinase CheA. *Biochemistry*, 34(45):14626–36. 14, 88

[88] Li, M. and Hazelbauer, G. L. (2004). Cellular stoichiometry of the components of the chemotaxis signaling complex. *J Bacteriol*, 186(12):3687–94. 12, 17, 18, 23, 55, 73, 100

[89] Li, M. and Hazelbauer, G. L. (2005). Adaptational assistance in clusters of bacterial chemoreceptors. *Mol Microbiol*, 56(6):1617–26. 18

[90] Liberman, L., Berg, H. C., and Sourjik, V. (2004). Effect of chemoreceptor modification on assembly and activity of the receptor-kinase complex in escherichia coli. *J Bacteriol*, 186(19):6643–6. 19

[91] Liljestrom, P., Laamanen, I., and Palva, E. T. (1988). Structure and expression of the *ompB* operon, the regulatory locus for the outer membrane porin regulon in *Salmonella typhimurium* LT-2. *J Mol Biol*, 201(4):663–73. 80, 107

[92] Lipkow, K. (2006). Changing cellular location of CheZ predicted by molecular simulations. *PLoS Comput Biol*, 2(4):e39. 86

[93] Lipkow, K., Andrews, S. S., and Bray, D. (2005). Simulated diffusion of phosphorylated CheY through the cytoplasm of *Escherichia coli*. *J Bacteriol*, 187(1):45–53. 17

[94] Little, J. W., Shepley, D. P., and Wert, D. W. (1999). Robustness of a gene regulatory circuit. *Embo J*, 18(15):4299–307. 4, 107

[95] Liu, J. D. and Parkinson, J. S. (1991). Genetic evidence for interaction between the CheW and Tsr proteins during chemoreceptor signaling by *Escherichia coli*. *J Bacteriol*, 173(16):4941–51. 15

[96] Liu, Y., Levit, M., Lurz, R., Surette, M. G., and Stock, J. B. (1997). Receptor-mediated protein kinase activation and the mechanism of transmembrane signaling in bacterial chemotaxis. *Embo J*, 16(24):7231–40. 15

[97] Lybarger, S. R., Nair, U., Lilly, A. A., Hazelbauer, G. L., and Maddock, J. R. (2005). Clustering requires modified methyl-accepting sites in low-abundance but not high-abundance chemoreceptors of escherichia coli. *Mol Microbiol*, 56(4):1078–86. 19

[98] Macnab, R. (1996). *Escherichia coli* and *Salmonella*: Cellular and molecular biology. In F., C., N., editor, *Escherichia coli and Salmonella*, volume 1, pages 123–142. American Society for Microbiology, Washington DC. 22, 24

[99] Macnab, R. M. (1999). The bacterial flagellum: reversible rotary propellor and type III export apparatus. *J Bacteriol*, 181(23):7149–53. 22

[100] Macnab, R. M. (2003). How bacteria assemble flagella. *Annu Rev Microbiol*, 57:77–100. 8, 22

[101] Macnab, R. M. and Koshland, D. E., J. (1972). The gradient-sensing mechanism in bacterial chemotaxis. *Proc Natl Acad Sci U S A*, 69(9):2509–12. 9

[102] Maddock, J. R., Alley, M. R., and Shapiro, L. (1993). Polarized cells, polar actions. *J Bacteriol*, 175(22):7125–9. 15, 18, 19

[103] Madigan, M. T. (2000). *Brock biology of microorganisms*. Prentice Hall, 9th edition. 9

[104] Magee, J. A., Abdulkadir, S. A., and Milbrandt, J. (2003). Haploinsufficiency at the Nkx3.1 locus. A paradigm for stochastic, dosage-sensitive gene regulation during tumor initiation. *Cancer Cell*, 3(3):273–83. 4

[105] Marr, A. G., Harvey, R. J., and Trentini, W. C. (1966). Growth and division of *Escherichia coli*. *J Bacteriol*, 91(6):2388–9. 101

[106] McAdams, H. H. and Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A*, 94(3):814–9. 3, 4, 6

[107] McAdams, H. H. and Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet*, 15(2):65–9. 2

[108] McAdams, H. H., Srinivasan, B., and Arkin, A. P. (2004). The evolution of genetic regulatory systems in bacteria. *Nat Rev Genet*, 5(3):169–78. 7

[109] McNally, D. F. and Matsumura, P. (1991). Bacterial chemotaxis signaling complexes: formation of a CheA/CheW complex enhances autophosphorylation and affinity for CheY. *Proc Natl Acad Sci U S A*, 88(14):6269–73. 11, 15

[110] 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(5):1813–8. 15

[111] Mello, B. A. and Tu, Y. (2003). Perfect and near-perfect adaptation in a model of bacterial chemotaxis. *Biophys J*, 84(5):2943–56. 20, 89

[112] Mello, B. A. and Tu, Y. (2005). An allosteric model for heterogeneous receptor complexes: understanding bacterial chemotaxis responses to multiple stimuli. *Proc Natl Acad Sci U S A*, 102(48):17354–9. 20

[113] Miller, A. F. and Falke, J. J. (2004). Chemotaxis receptors and signaling. *Adv Protein Chem*, 68:393–444. 13

[114] Monod J., Wyman J., C. J. (1965). On the Nature of Allosteric Transitions: a Plausible Model. *Journal of molecular biology*, 12:88–118. 20

[115] Morohashi, M., Winn, A. E., Borisuk, M. T., Bolouri, H., Doyle, J., and Kitano, H. (2002). Robustness as a measure of plausibility in models of biochemical networks. *J Theor Biol*, 216(1):19–30. 6

[116] Morrison, T. B. and Parkinson, J. S. (1997). A fragment liberated from the *Escherichia coli* CheA kinase that blocks stimulatory, but not inhibitory, chemoreceptor signaling. *J Bacteriol*, 179(17):5543–50. 15

[117] Mytelka, D. S. and Chamberlin, M. J. (1996). *Escherichia coli fliAZY* operon. *J Bacteriol*, 178(1):24–34. 24

[118] Ohnishi, K., Kutsukake, K., Suzuki, H., and Lino, T. (1992). A novel transcriptional regulation mechanism in the flagellar regulon of salmonella typhimurium: an antisigma factor inhibits the activity of the flagellum-specific sigma factor, sigma f. *Mol Microbiol*, 6(21):3149–57. 24

[119] Okumura, H., Nishiyama, S., Sasaki, A., Homma, M., and Kawagishi, I. (1998). Chemotactic adaptation is altered by changes in the carboxy-terminal sequence conserved among the major methyl-accepting chemoreceptors. *J Bacteriol*, 180(7):1862–8. 18, 21, 76, 105

[120] Oppenheim, D. S. and Yanofsky, C. (1980). Translational coupling during expression of the tryptophan operon of *Escherichia coli*. *Genetics*, 95(4):785–95. 6, 80, 107

[121] Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nat Genet*, 31(1):69–73. 3, 6, 101

[122] 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(5):400–7. 15, 19

[123] Parkinson, J. S., Ames, P., and Studdert, C. A. (2005). Collaborative signaling by bacterial chemoreceptors. *Curr Opin Microbiol*, 8(2):116–21. 18, 19

[124] Paulsson, J. (2004). Summing up the noise in gene networks. *Nature*, 427(6973):415–8. 2, 6, 61

[125] Pedraza, J. M. and van Oudenaarden, A. (2005). Noise propagation in gene networks. *Science*, 307(5717):1965–9. 5

[126] Pigliucci, M. and Murren, C. J. (2003). Perspective: Genetic assimilation and a possible evolutionary paradox: can macroevolution sometimes be so fast as to pass us by? *Evolution*, 57(7):1455–64. 8

[127] Queitsch, C., Sangster, T. A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature*, 417(6889):618–24. 7

[128] Raser, J. M. and O'Shea, E. K. (2004). Control of stochasticity in eukaryotic gene expression. *Science*, 304(5678):1811–4. 3, 62, 91, 101

[129] Raser, J. M. and O'Shea, E. K. (2005). Noise in gene expression: origins, consequences, and control. *Science*, 309(5743):2010–3. 5, 90, 101

[130] Rex, G., Surin, B., Besse, G., Schneppe, B., and McCarthy, J. E. (1994). The mechanism of translational coupling in *Escherichia coli*. Higher order structure in the atpHA mRNA acts as a conformational switch regulating the access of de novo initiating ribosomes. *J Biol Chem*, 269(27):18118–27. 107

[131] Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., and Elowitz, M. B. (2005). Gene regulation at the single-cell level. *Science*, 307(5717):1962–5. 3, 101, 103

[132] Saini, S., Brown, J. D., Aldridge, P. D., and Rao, C. V. (2008). FliZ Is a posttranslational activator of $FlhD_4C_2$-dependent flagellar gene expression. *J Bacteriol*, 190(14):4979–88. 24

[133] 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(1):201–6. 22, 72

[134] Schulmeister, S., Ruttorf, M., Thiem, S., Kentner, D., Lebiedz, D., and Sourjik, V. (2008). Protein exchange dynamics at chemoreceptor clusters in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 105(17):6403–8. 93

[135] Schumperli, D., McKenney, K., Sobieski, D. A., and Rosenberg, M. (1982). Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. *Cell*, 30(3):865–71. 6, 80, 107

[136] 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(6444):343–7. 15

[137] Segall, J. E., Block, S. M., and Berg, H. C. (1986). Temporal comparisons in bacterial chemotaxis. *Proc Natl Acad Sci U S A*, 83(23):8987–91. 12

[138] Segall, J. E., Ishihara, A., and Berg, H. C. (1985). Chemotactic signaling in filamentous cells of *Escherichia coli*. *J Bacteriol*, 161(1):51–9. 11

[139] Segall, J. E., Manson, M. D., and Berg, H. C. (1982). Signal processing times in bacterial chemotaxis. *Nature*, 296(5860):855–7. 16

[140] Serizawa, S., Miyamichi, K., and Sakano, H. (2004). One neuron-one receptor rule in the mouse olfactory system. *Trends Genet*, 20(12):648–53. 4

[141] Shimizu, T. S., Le Novere, N., 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(11):792–6. 19

[142] Shiomi, D., Banno, S., Homma, M., and Kawagishi, I. (2005). Stabilization of polar localization of a chemoreceptor via its covalent modifications and its communication with a different chemoreceptor. *J Bacteriol*, 187(22):7647–54. 19

[143] Shiomi, D., Homma, M., and Kawagishi, I. (2002). Intragenic suppressors of a mutation in the aspartate chemoreceptor gene that abolishes binding of the receptor to methyltransferase. *Microbiology*, 148(Pt 10):3265–75. 13, 17

[144] Shrout, A. L., Montefusco, D. J., and Weis, R. M. (2003). Template-directed assembly of receptor signaling complexes. *Biochemistry*, 42(46):13379–85. 19

[145] Smith, R. A. and Parkinson, J. S. (1980). Overlapping genes at the cheA locus of *Escherichia coli*. *Proc Natl Acad Sci U S A*, 77(9):5370–4. 81

[146] Sourjik, V. (2004). Receptor clustering and signal processing *in E. coli* chemotaxis. *Trends Microbiol*, 12(12):569–76. 10, 19, 20, 105

[147] Sourjik, V. and Berg, H. C. (2002). 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(20):12669–74. 49

[148] Sourjik, V., Muschler, P., Scharf, B., and Schmitt, R. (2000). VisN and VisR are global regulators of chemotaxis, flagellar, and motility genes in Sinorhizobium (Rhizobium) meliloti. *J Bacteriol*, 182(3):782–8. 18, 19

[149] Sourjik, V. and Schmitt, R. (1998). Phosphotransfer between CheA, CheY1, and CheY2 in the chemotaxis signal transduction chain of Rhizobium meliloti. *Biochemistry*, 37(8):2327–35. 108

[150] Spudich, J. L. and Koshland, D. E., J. (1976). Non-genetic individuality: chance in the single cell. *Nature*, 262(5568):467–71. 1

[151] Starrett, D. J. and Falke, J. J. (2005). Adaptation mechanism of the aspartate receptor: electrostatics of the adaptation subdomain play a key role in modulating kinase activity. *Biochemistry*, 44(5):1550–60. 13

[152] Stewart, R. C., Jahreis, K., and Parkinson, J. S. (2000). Rapid phosphotransfer to CheY from a CheA protein lacking the CheY-binding domain. *Biochemistry*, 39(43):13157–65. 14, 88

[153] Stock, A. M., Mottonen, J. M., Stock, J. B., and Schutt, C. E. (1989). Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. *Nature*, 337(6209):745–9. 16

[154] Stock, J. S. M. (1996). Cellular and molecular biology. In F.C.N., editor, *Escherichia coli and Salmonella*, pages 1103–1124. American Society for Microbiology, Washington DC. 11

[155] Studdert, C. A. and Parkinson, J. S. (2004). Crosslinking snapshots of bacterial chemoreceptor squads. *Proc Natl Acad Sci U S A*, 101(7):2117–22. 18

[156] Studdert, C. A. and Parkinson, J. S. (2005). Insights into the organization and dynamics of bacterial chemoreceptor clusters through in vivo crosslinking studies. *Proc Natl Acad Sci U S A*, 102(43):15623–8. 18

[157] Swain, P. S. (2004). Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J Mol Biol*, 344(4):965–76. 6, 108

[158] Swain, P. S., Elowitz, M. B., and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A*, 99(20):12795–800. 6

[159] Szurmant, H. and Ordal, G. W. (2004). Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol Mol Biol Rev*, 68(2):301–19. 16, 104

[160] Thattai, M. and van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci U S A*, 98(15):8614–9. 1, 2, 3, 6

[161] Thattai, M. and van Oudenaarden, A. (2002). Attenuation of noise in ultrasensitive signaling cascades. *Biophys J*, 82(6):2943–50. 5

[162] Trueba, F. J. and Woldringh, C. L. (1980). Changes in cell diameter during the division cycle of *Escherichia coli*. *J Bacteriol*, 142(3):869–78. 101

[163] Turner, L., Ryu, W. S., and Berg, H. C. (2000). Real-time imaging of fluorescent flagellar filaments. *J Bacteriol*, 182(10):2793–801. 57

[164] Vaknin, A. and Berg, H. C. (2004). Single-cell FRET imaging of phosphatase activity in the *Escherichia coli* chemotaxis system. *Proc Natl Acad Sci U S A*, 101(49):17072–7. 17, 19, 21, 86, 93

[165] Vaknin, A. and Berg, H. C. (2006). Osmotic stress mechanically perturbs chemoreceptors in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 103(3):592–6. 19

[166] Vilar, J. M., Kueh, H. Y., Barkai, N., and Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. *Proc Natl Acad Sci U S A*, 99(9):5988–92. 5

[167] Vilar, J. M. and Leibler, S. (2003). DNA looping and physical constraints on transcription regulation. *J Mol Biol*, 331(5):981–9. 5

[168] von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. (2000). The segment polarity network is a robust developmental module. *Nature*, 406(6792):188–92. 7

[169] Von Dassow, G. and Odell, G. M. (2002). Design and constraints of the Drosophila segment polarity module: robust spatial patterning emerges from intertwined cell state switches. *J Exp Zool*, 294(3):179–215. 7

[170] Waddington, C. H. (1959). Canalization of development and genetic assimilation of acquired characters. *Nature*, 183(4676):1654–5. 8

[171] Wadhams, G. H. and Armitage, J. P. (2004). Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol*, 5(12):1024–37. 105

[172] Webre, D. J., Wolanin, P. M., and Stock, J. B. (2003). Bacterial chemotaxis. *Curr Biol*, 13(2):R47–9. 105

[173] Winston, S. E., Mehan, R., and Falke, J. J. (2005). Evidence that the adaptation region of the aspartate receptor is a dynamic four-helix bundle: cysteine and disulfide scanning studies. *Biochemistry*, 44(38):12655–66. 13

[174] Wolfe, A. J. and Berg, H. C. (1989). Migration of bacteria in semisolid agar. *Proc Natl Acad Sci U S A*, 86(18):6973–7. 100

[175] Wolfe, A. J., Conley, M. P., Kramer, T. J., and Berg, H. C. (1987). Reconstitution of signaling in bacterial chemotaxis. *J Bacteriol*, 169(5):1878–85. 11

[176] Yi, T. M., Huang, Y., Simon, M. I., and Doyle, J. (2000). Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc Natl Acad Sci U S A*, 97(9):4649–53. 7

[177] 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(8):570–5. 10, 16

[178] Zimmer, M. A., Szurmant, H., Saulmon, M. M., Collins, M. A., Bant, J. S., and Ordal, G. W. (2002). The role of heterologous receptors in McpB-mediated signalling in Bacillus subtilis chemotaxis. *Mol Microbiol*, 45(2):555–68. 19

[179] Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, 31(13):3406–15. 56

List of Figures