INAUGURAL – DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der

Ruprecht - Karls - Universität

Heidelberg

von

Rajagopala S.V., Masters of Sciences

aus

Seesandra, India

Tag der mündlichen Prüfung:

19 April 2006

Thema

The Protein-protein Interaction Map of the *Treponema pallidum* Flagellar Apparatus

Referees:

Prof. Dr. Uwe Strähle

Priv. Doz. Dr. Matthias Mayer

Diese Arbeit wurde in der Arbeitgruppe von Dr. Peter Uetz am Institut für Toxikologie und Genetik (ITG) des Forschungszentrum Karlsruhe (FZK), Mitglied der Helmholtz-Gemeinschaft, angefertigt.

Zusammenfassung

Diese Studie stellt den ersten umfassenden Versuch dar, die Protein-Protein-Interaktionen (PPIs) eines bakteriellen Flagellenapparats genomweit zu erfassen. Wir haben dafür *Treponema pallidum* ausgewählt, also Spirochäten, die auch die Syphilis auslösen. *T. pallidum* hat ein kleines Genom mit nur 1,041 Genen, die alle in klonierter Form vorliegen. Die Flagellen von Spirochäten besitzen ausserdem einige spezielle Eigenschaften wie die periplasmatische Lokalisation der zwei polaren Flagellenbündel, welche sie für vergleichende Analysen besonders interessant machen.

Im Lauf dieses Projekts habe ich eine *Yeast two-hybrid*-Bibliothek des *Treponema pallidum*-Proteoms konstruiert und damit umfassende genom-weite Screens mit allen 75 Flagellen- und Motilitäts-Proteinen durchgeführt. Insgesamt wurden dabei 268 PPIs zwischen 174 verschiedenen Proteinen identifiziert. Davon waren zuvor nur 15 homologe Interaktionen ("Interologe") aus der Literatur bekannt.

Um die funktionelle Bedeutung dieser neuen Interaktionen für die bakterielle Motilität zu untersuchen, habe ich zuerst 23 Orthologe aller *Treponema*-Interaktoren in *E. coli* und 30 in *B. subtilis* vorhergesagt. Die Gene von 23 Orthologen wurden dann in *E. coli* und/oder *B. subtilis* (7 Gene) mutiert und deren Phänotyp in Motilitätsassays überprüft. Ich fand dabei 10 neue Proteine in *E. coli* (amiA, bcsC, rfaH, rpe, rpmJ, ycfH, yciM, yggH, yggW, und yncE) und 5 Proteine in *B. subtilis* (yaaT , yhbE/YhbF, yllB, yloS und yviF), deren *Treponema*-Homologe mit Flagellenproteinen interagierten und deren Mutanten einen Motilitätsphänotyp zeigen. Aufgrund ihrer Interaktionen und Phänotypen konnte ich 8 bisher nicht charakterisierten Proteinen eine Rolle als Motilitätsgene zuordnen, und zwar TP0046 (yaaT), TP0048 (yhbE/yhbF), TP0383 (yllB), TP0421 (yncE), TP0561 (ydjH), TP0658 (yviF), TP0712 (HP1034) and TP0979 (tatD). Basierend auf den 268 two-hybrid-Interaktionen konnte ich ausserdem 1455 Interaktionen in 30 anderen Bakterien vorhersagen.
Mit 35 Flagellenproteinen von 30 Bakterienarten habe ich ausserdem einen phylogenetischen "Super-Stammbaum" erstellt, der die Flagellenapparate in einen evolutionären Kontext stellt und belegt, dass die Spirochäten einen ungewöhnlichen und abgeleiteten Flagellenapparat aufweisen.

Ein Protein unbekannter Funktion, TP0658 (yviF in *B. subtilis*) wurde näher untersucht. Die Y2H-Interaktionen von TP0658 mit Flagellinen (flaB1-3) wurden biochemisch verifiziert. TP0658 bindet an eine C-terminale Sequenz von FlaB1 zwischen L231 und D247 (dem so genannten Cα-Loop der Flagelline). Dieses epitop ist in *B. subtilis* und anderen Bakterien konserviert. Ein hochkonserviertes Asparagin (N237 in FlaB1 bzw. N255 in *B. subtilis* hag) ist entscheidend für die Bindung. Eine ΔyviF Mutante in B. subtilis hatte einen starken Motilitätsphänotyp und keine nachweisbaren Flagelline. Koexpression von TP0658 und FlaB1 in *E. coli* stabilisiert FlaB1. Interessanterweise binden sowohl TP0658 als auch FliS (ein bekanntes aber nicht verwandtes Chaperon) an das gleiche Epitop in Flagellin, welches bei dessen Polymerisierung beteiligt ist. Meine Daten sprechen daher dafür dass TP658 und seine Verwandten "Assembly-Chaperone" der bakteriellen Flagelle sind.

Summary

The study presented here is the first comprehensive attempt to identify protein-protein interactions (PPIs) of the bacterial flagellar apparatus on a genome-wide scale. We selected *Treponema pallidum,* the syphilis spirochete, as a model organism for flagellum interactome analysis because of its small genome (1,041 ORFs) which is available as a cloned ORFeome. Interestingly, the spirochete flagellum possesses unique features not found in other bacterial species such as a periplasmic localisation of two flagellum bundles fixed to the cell poles.

In the course of this project I constructed a *T. pallidum* yeast-two hybrid (Y2H) proteome library (prey array) and carried out comprehensive genome-wide Y2H screens for all 75 *T. pallidum* flagellum and chemotaxis proteins against the whole genome prey array. In total 268 protein-protein interactions, involving 174 unique proteins were identified. Only 15 interologous interactions were previously described by other studies.

To characterise the functional relevance of the proteins not previously known to be involved in motility but interacting with known flagellar proteins 23 orthologs were predicted in *E. coli* and another 30 in *B. subtilis*. These orthologs were tested for their functional involvement in motility using gene deletions of these orthologes protein. Motilty assays (swarming assays) were done for all 23 orthologous *E. coli* mutants and 7 *Bacillus subtilis* mutants. Five *B. subtilis* proteins (namely, yaaT, yhbE/YhbF, yllB, yloS and yviF) and 10 *E. coli* proteins (amiA, bcsC, rfaH, rpe, rpmJ, ycfH, yciM, yggH, yggW, and yncE) showed a motility phenotype, clearly supporting their functional relevance in bacterila motility. These included 8 uncharacterized proteins, for which I assigned a function in motility based on their motility defect in the orthologous gene mutant and their physical two-hybrid association with known flagellum proteins, namely, TP0046 (yaaT), TP0048 (yhbE/yhbF), TP0383 (yllB), TP0421 (yncE), TP0561 (ydjH), TP0658 (yviF), TP0712 (HP1034) and TP0979 (tatD).

A flagellum phylogenetic supertree, using combined alignments of 35 orthologous flagellum proteins conserved across 30 species, indicates that spirochetes have a derived bacterial flagellum. Based on Y2H interactions in *T. pallidum* 1455 interologs were predicted for 30 selected bacterial species covering a wide range of bacterial diversity.

The Y2H interactions of TP0658/yviF, a protein of unknown function, with flagellins (FlaB1-3 in *T. pallidum* and FliC in *E. coli*), were biochemically verified. TP0658 binds in the C-terminal of FlaB1 comprising a peptide from L_{231} to D_{247} (the Cα loop of flagellin). This interaction epitope is conserved in *B. subtilis* and other bacteria. One well conserved Asparagine residue of flagellins (Asn_{237} of FlaB1 and Asn_{255} of its *B. subtilis* homolog *hag*) is crucial for binding of yviF. A *B. subtilis* ΔyviF mutant showed strongly reduced motility which was rescued by yviF overexpression and there are no detectable flagellins in the ΔyviF mutant. Co-expression of TP0658 and FlaB1 in *E. coli* suggests that TP0658 stabilizes flagellin. Strikingly, both TP0658/yviF and fliS, a known but unrelated flagellin-specific chaperone, bind to the C-terminal part of flagellin which is implicated in polymerization of flagellins. Hence I suggest that TP0658 is another conserved assembly factor for flagellin proteins, potentially working in concert with FliS or replacing it in certain subsets of bacterial species.

Acknowledgements

I owe my gratitude to Dr. Peter Uetz for showing confidence in me and providing free reigns to conduct research work independently. Further I would like to thank him for his encouragements and fruitful discussion.

I consider this as my greatest privilege and the best opportunity to express my heartfelt thanks to Prof. Dr. Uwe Strähle, Priv. Doz. Dr. Matthias Mayer and Prof. Dr. Bernd Bukau for accepting me as their graduate student and for their support.

My special thanks are to Mr. Bjoern Titz, a project coworker with whom I interacted the most during the project period, for is relentless support and concerted efforts in establishing the *Treponema* Y2H library and database. I am indebted to him for all his encouragement, help, and fruitful discussions. The co-ordination in the work established during the project period is much appreciated and it was my pleasure having worked with him in the project.
I express my sincere thanks to Mrs. Claudia Ester who assisted me with the peptide synthesis for the project, and also taking trouble in teaching me Deutsch.
I express my sincere thanks Mr. Johannes Goll for his help with excellent Bioinformatics inputs especially, data comparison and flagellum supertree construction.
Thanks also to Mrs. Sindhu Thomas, Mrs. Tanja Kuhn, and Ms. Cathrin Klumpp for their technical assistance in the project, Mrs. Carolina Muller especially for introducing me to the lab upon arrival and providing a nice working environment in the lab.

My special thanks to Mr. Roman Häuser and Ms. Katrin Wohlbold who did their diploma thesis on *Treponema* and also constructed a couple of *E. coli* mutants.

My sincere thanks to Prof. Timothy Palzkill and Mr. Matthew T. McKevitt for providing the *T. pallidum* ORFeome library and prey clones.
My sincere thanks also to Prof. Hirotada Mori for providing the *E. coli* knock-out library and complex purification data for comparative analysis. My thanks also to Prof. Russell L Finley for his *C. jejuni* Y2H data, Prof. Stephen Elledge for the UPS system, Prof. Schumann and Prof. Ehrlich for the *B. subtilis* strains and vectors.

I thank the Institute for Toxicology and Genetics, Forschungszentrums Karlsruhe for providing a conducive research environment and the DFG for financial support.

As for as my personal thanks: My salutations to my elder brother Dr. Jagadeesh without his support my higher studies would not have been possible.
My personal thanks are to my parents and all the family members for their continued and generous support
Finally, I thank my loving friends Deve, Shabana, Manju's, Ganesh, Joshi, Mayur, and Nagabhushana for their encouragements and moral support.

Table of Contents

Abbreviations

3-AT	3-Aminotriazole (3-AT)
AA	amino acid
AD	activation domain
APS	ammonium persulfate
bp	base pairs
BSA	albumin bovine serum
COG	cluster of orthologus groups
DBD	DNA-binding domain
DMSO	dimethyl sulfoxide
DNA	desoxy ribonucleic acid
dNTP	desoxy nucleosid triphosphate
EDTA	ethylenediaminetetraacetic acid
GAL4	yeast transcription factor of the galactose pathway
GST	glutathione S-transferase
HA	hemagglutinin
HRP	horse radish peroxidase
IPTG	isopropyl-beta-D-thiogalactopyranoside
KDa	kilo Dalton
Kan	kanamycine
kb	kilo base
MAT	mating type (locus)
µl	microliter
µM	micromolar
M	molar
mM	millimolar
OD	Optical density
ORF	open reading frame
ORFeome	all ORFs of a genome
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethyl sulphonyl fluoride

PPIs	protein-protein interactions
RT	room temperature
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyAcrylamide gel electrophoresis
TEMED	N,N,N',N'-tetramethylethylenediamine
TRIS	tris-(hydoxymethyl)-aminomethane
WT	wild type
YPD	yeast peptone dextrose
Y2H	yeast two-hybrid

1. Introduction

1.1 Motility in bacteria

In an ever-changing environment, it is essential that organisms are able to sense these changes, and to respond appropriately. The ability of bacterial cells for rapid and directed movement in a medium depends on the presence of the flagellum, the special motility organelle (Figure 1.1). The flagellum structural and functional proteins are encoded by ~50 genes (Jones and Aizawa 1991; Macnab 1992). Bacterial motility represents an important advantage for bacteria in moving towards favourable conditions or in avoiding detrimental environments and it allows flagellated bacteria to successfully compete with other microorganisms (Fenchel 2002). In addition to having locomotive properties, bacterial flagella play a crucial role in adhesion, biofilm formation and colonisation by micro-organisms, such as *Pseudomonas aeruginosa*, *Escherichia coli, Vibrio cholerae, Salmonella typhimurium and Helicobacter pylori* (Allen-Vercoe and Woodward 1999; Eaton et al., 1992; O'Toole and Kolter 1998; Pratt and Kolter 1998). Motility in pathogenic micro-organisms is usually considered a virulence factor, essential for colonisation of host organisms or target organs (Josenhans and Suerbaum 2002; Ottemann and Miller 1997). However, the flagellar filament bears strong antigenic properties with animal and plant hosts (Dangl and Jones 2001; Hayashi et al., 2001; Smith and Ozinsky 2002). Motility by means of flagella is very expensive for cellular economy in terms of the number of genes and the energy required for flagellar biosynthesis and functioning. Consequently, it is not surprising that the synthesis of flagella is highly regulated by external factors, including the interaction of bacterial cells with their host. With respect to this dual pathogenicity/antigenicity effect of flagella, several highly pathogenic bacteria, such as *Bordetella pertussis* (Akerley and Miller 1993; Akerley et al., 1992), *Shigella sp* (Tominaga et al., 1994), and *Yersinia pestis* (Kapatral et al., 1996), in contrast to their close relatives *Bordetella bronchiseptica, E. coli* and *Yersinia enterocolitica,* respectively, have lost the capacity to synthesise flagella, but yet possess the flagellar genes. Motility by means of flagella is widespread in the microbial world, and more than 80% of known bacterial species possess these organelles, including various flagellated archaea (Moens and Vanderleyden 1996; Thomas et al., 2001). The structure and arrangement of flagella on the cell differs from species to species and both seem to be related to the specific environments in which the cells reside (Joys 1988; Wilson and Beveridge 1993). Flagella can be arranged on the cell body in a variety of configurations, including single polar, multiple polar, and many peritrichous (or lateral) configurations.

1.2 The bacterial flagellum

The bacterial flagellum proved to be a discrete system composed of three main components: the "*basal body"* localized within the cell wall, a prolonged "*external spiral protein filament"* composed of flagellin subunits, and flexible joint, the so-called *"hook"*, connecting the first two parts (DePamphilis and Adler 1971). The basal body is a complex structure, similar to type III secretion systems, begins inside the cell, spans the cell envelope including the cytoplasmic membrane, and extends well outside the cell. The basal body consists of an integral membrane ring called the MS ring (polymer of FliF protein), a cytosolic C-ring, a rod that traverses the periplasmic space (composed of FliE, FlgB and FlgC proteins), a periplasmic P ring (FlgI), and an outer membrane L ring (FlgH) (Figure 1.2).

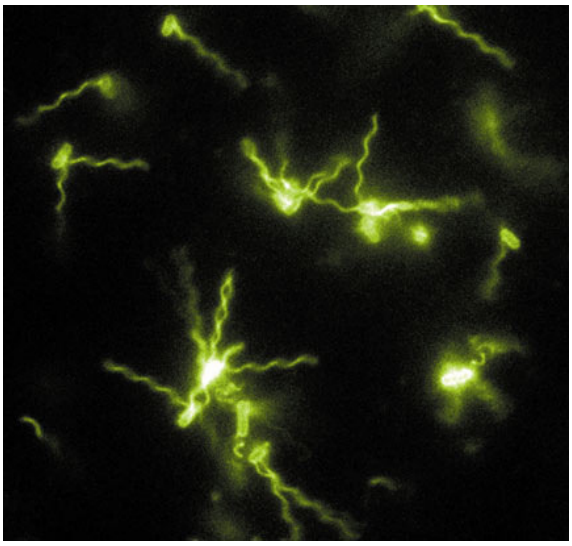


Figure 1.1 Cells of *Escherichia coli* labeled with an amino-specific Alexa Fluor dye, de-energized by exposure to light, and examined in a fluorescence microscope. Real-time imaging of fluorescent flagellar filaments, showing what flagella do as swimming cells run and tumble (Turner et al., 2000) (From www.mcb.harvard.edu/Faculty/Berg.html).

1.2.1 Flagellum Motor

Bacterial flagella are driven by a rotary motor that utilizes the free energy stored in the electrochemical proton gradient across the cytoplasmic membrane to do mechanical work (Meister et al., 1987). In some marine bacteria, sodium ions replace protons (Bakeeva et al., 1986; Hirota and Imae 1983). MotA forms the proton or sodium channel and consists of a MotA_ExbB domain (PFAM PF01618), which is shared with proteins from other protein export systems, such as the TonB complex. Full-length MotA orthologous are found in all flagellar systems. MotB is required for the rotation of the flagellar motor and is thought to be a linker that fastens the torque-generating machinery to the cell wall. It contains a C-terminal OmpA domain, shared with many outer-membrane proteins. MotB-like proteins could be identified in all flagellated bacteria. Curiously, multiple MotB-like proteins occur in several species: two in *Bacillus subtilis, Aquifex aeolicus, Pirellula sp*., and *Pseudomonas aeruginosa*, three in *Chromobacterium violaceum* and *Vibrio parahaemolyticus* (containing two flagellar systems), and four in *Leptospira interrogans* (Pallen et al., 2005b). The basal body receives torque from the motor and transmits it to the hook and then to the filament (Berg and Turner 1993; Bren and Eisenbach 1998).

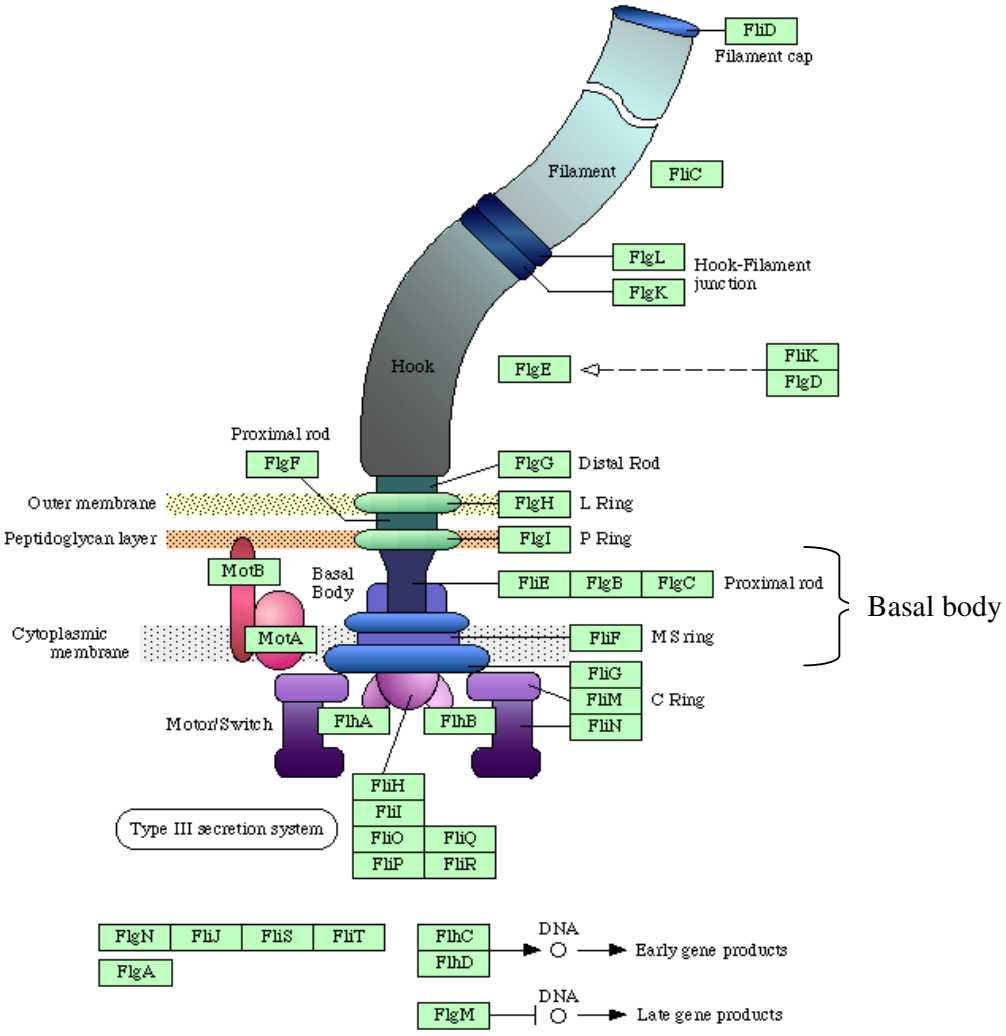


Figure 1.2 Schematic diagram of a bacterial flagellum: *E. coli* flagellum, the names of flagellum components are depicted in approximate stoichiometry (from www.kegg.org).

1.2.2 Flagellum motor switch complex (C-ring)

The cytoplasmic C-ring forms the flagellum motor switch complex, it attaches to the MS ring, which is in the cytoplasmic membrane and is formed from a single protein, FliF (Figure 1.3). The C-ring consists of three proteins: FliG, FliM and FliN. FliG is conserved in all flagellar systems. Curiously, spirochaetes show marked variation in their repertoires of FliG homologues: two paralogues occur in *T. pallidum* and *B. burgdorferi*, whereas there are four in *L. interrogans*. Charon and Goldstein have suggested that the two FliG paralogues in *T. pallidum* and *B. burgdorferi* might play a role in flagellar asymmetrical rotation (Charon and Goldstein 2002), but the four in *L. interrogans* require further explanation. FliG is essential for assembly, rotation and clockwise/counter-clockwise (CW/CCW) switching of the bacterial flagellum. About 25 copies of FliG are present in a large rotor-mounted assembly termed the 'switch complex', which also contains the proteins FliM and FliN. Mutational data indicate that FliM binds to both domains of FliG-MC (middle and C-terminal domains), thereby suggesting that FliM might

control the relative orientation of the domains in the flagellum. FliM is the rotor protein most important for CW/CCW switching (Sockett et al., 1992; Welch et al., 1993). The crystal structure of the middle and C-terminal domains of FliG from *Thermotoga maritima* reveled that, the motor switching occurs by FliM-regulated movement of the C-terminal domain of FliG relative to other parts of the protein (Brown et al., 2002).

A key event in signal transduction during chemotaxis of bacterial species is the interaction between the phosphorylated form of the response regulator CheY (CheY-P) and the switch of the flagellar motor, located at its base. The consequence of this interaction is a shift in the direction of flagellar rotation from the default, counter clockwise, to clockwise.

FliM plays an important role in transmitting the chemotaxis signal to the flagellum motor switch protein FliG. CheY when phosphorylated (CheY-P), binds to FliM and produces a conformational change in FliM that is propagated to FliG. The N-terminal region of FliM binds to CheY-P, the middle region of FliM binds to FliG, and the C-terminal region binds to FliN (Figure 1.4).

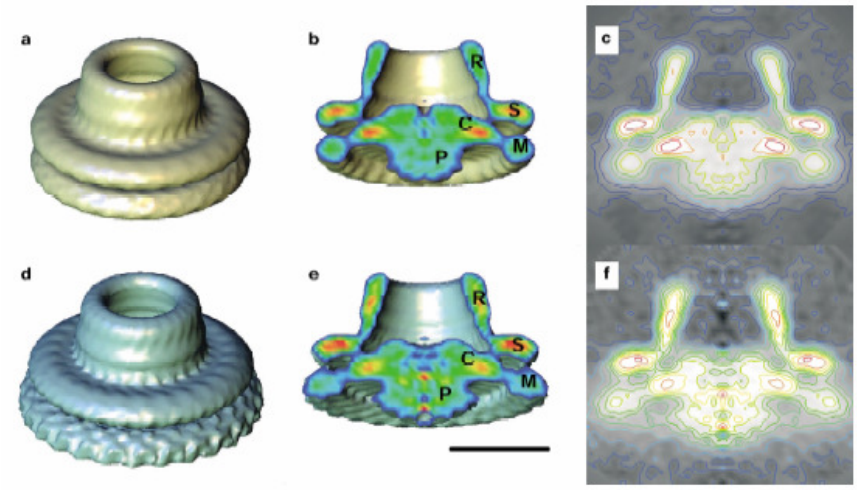


Figure 1.3 A Three-dimensional structures of the FliF and FliFG rings: a, b and c, The FliF ring; d, e and f, the FliFG ring. a and d, Solid surface representation. b and e, half-cut density through the axis with density levels on the section colorcoded from blue (low), green (medium) to red (high). c and f, Density distributions on the axial sections represented in gray-scale as well as by contour lines color coded in the same way as b and e. For solid surface representation, the volume of the FliF ring was estimated from its theoretical molecular mass, 1586 kDa. The contour level for the FliFG ring was chosen to make the volume of its proximal portion the same as that of the FliF ring. The scale bar represents 100 Å (from Suzuki et al., 2004).

The FliN protein is the third major component of the C-ring (Zhao et al., 1996), FliN interacts with FliG and FliM (Irikura et al., 1993). In addition to its functions in motor rotation and switching, FliN is thought to have a role in the export of proteins that form the exterior structures of the flagellum (i.e., rod, hook, and filament proteins) (Brown et al., 2005). The crystal structure of the FliN protein from *Thermotoga maritima* revealed that, FliN is a tightly intertwined dimer composed mostly of β sheets. Several well-conserved hydrophobic residues form a no polar

patch on the surface of the molecule. A mutation in the hydrophobic patch affected flagellar assembly and switching, showing that this surface feature is important for FliN function.. The *T. maritima* FliN is primarily a dimer in solution, and *T. maritima* FliN and FliM together form a stable $FliM_1$-$FliN_4$ complex while *Escherichia coli* FliN forms a stable tetramer in solution (Brown et al., 2005).

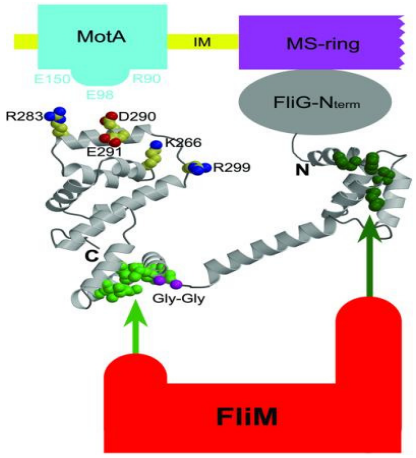


Figure 1.4 Proposed arrangement of FliG-MC (middle and C-terminal domains) relative to other proteins of the flagellum. The charge-bearing ridge on the C-terminal domain is oriented to allow interactions with conserved charged residues of MotA. FliM binds to both domains of FliG-MC, most likely through the EHPQ motif in domain I (dark green) and the surface hydrophobic patch on domain II (light green), (FliM–FliG binding interactions are indicated by the arrows). The flexible Gly–Gly linker (magenta) allows relative movement of the domains. IM = inner membrane (Brown et al., 2002).

1.2.3 Flagellar Rod

The bacterial flagellum rod is a short cylindrical structure of high mechanical stability which connects the basal body and hook of the flagellum. The proximal rod is formed from FlgF, FliE, FlgB and FlgC, the distal rod from FlgG. The flagellum rod has a diameter of about 180 Å for the distal portion (Kubori et al., 1992) and about 80 Å for the proximal portion (Francis et al., 1994). The amino acid sequences of FlgB, FlgC, FlgF and FlgG all have extensive heptad repeats of hydrophobic side-chains in their terminal regions, indicating that these regions form α-helical coiled coils. It has been proposed that a mutual interlocking of the α-helical coiled-coils is a common structural motif that makes the stable and continuous axial structure of the flagellum (Homma et al., 1990a; Homma et al., 1990b).

1.2.4 Flagellar Hook

The flagellar hook, a short, highly curved cylindrical tube, functions as an universal joint, it connects the basal body to the filament (Figures 1.2 and 1.5). In bacteria such as *Salmonella,* and *E. coli,* which have multiple flagella emerging from many positions on the cell, the hooks play the important role of enabling these flagella to function effectively as a bundle. The hook transmits the motor torque to the helical propeller over a wide range of its orientation for swimming and tumbling. A partial atomic model of the hook has been obtained by X-ray crystallography of

FlgE, a major proteolytic fragment of FlgE lacking the unfolded terminal regions, and by electron cryomicroscopy and three-dimensional helical image reconstruction of the hook. The model reveals the intricate molecular interactions and a plausible switching mechanism for the hook to be flexible in bending but rigid against twisting for its universal joint (Samatey et al., 2004a; Samatey et al., 2004b). There are two hook associated proteins, namely, hook-associated protein 1 (HAP1) and 3 (HAP3) [= FlgK and FlgL], forming a very short hook-filament junction,

which probably acts as a structural adaptor for a smooth transition between the two mechanically distinct structures: the hook is relatively flexible in bending, whereas the filament is much more rigid for its propeller function.

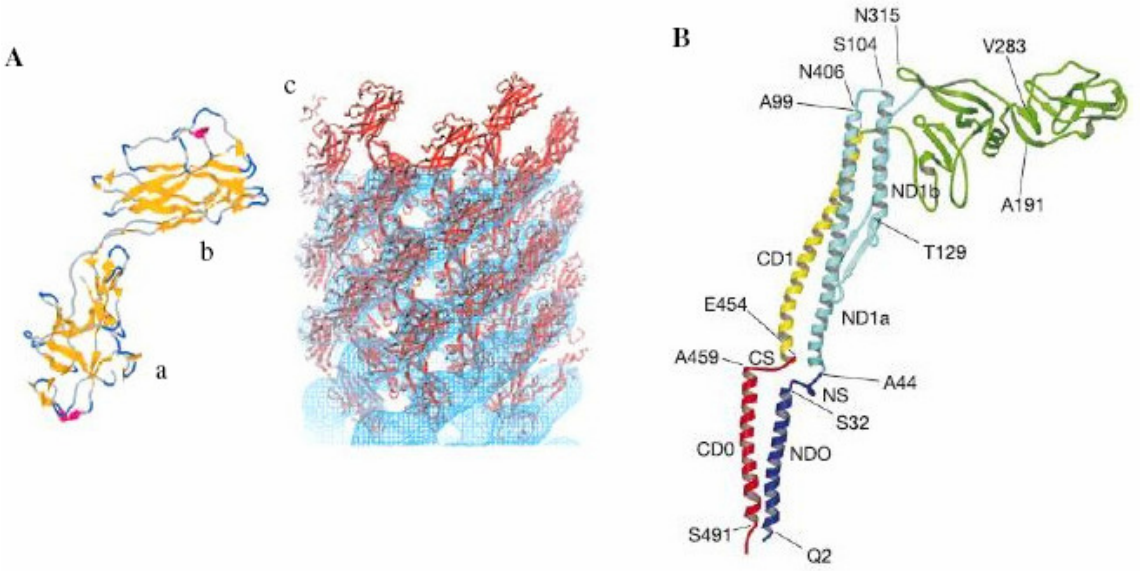


Figure 1.5 Flagellum hook and flagellin structures: A, The atomic model of the outer two domains of the hook. The structure of a major fragment of the hook subunit as determined by x-ray crystallography (Samatey et al., 2004a). β-Sheet are shown in yellow and α-helical segments in red. The lower domain (a) contains both N- and C-terminal regions, has the more evolutionarily conserved amino acid sequence, and corresponds to domain D1. The upper domain (b) corresponds to D2. (*c*) A stereo pair of the atomic structure docked into the *S. typhimurium* map (Shaikh et al., 2005). B, x-ray crystallography structure of flagellin (stereo diagram of the Cα- backbone). The chain is coloured as follows: residues 1–44, blue; 44–179, cyan; 179–406, green; 406–454, yellow; 454–494, red and the domains D0, D1, D2a, D2b and D3 (from Yonekura et al., 2003).

1.2.5 Flagellar Filament

The flagellar filament, is a long, thin cylindrical structure that is helical in shape and therefore when rotated functions like an propeller. The filament is a helical assembly of some tens of thousands of copies of a single protein, flagellin (= FliC). In other species, such as, *T. pallidum, B. burgdorferi* (spirochetes), three flagellin proteins, FlaB1-3 (homologous to FliC), are thought to form a heteropolymeric filament (Charon and Goldstein 2002). The polymorphic supercoiled forms of flagellin permit it to function as a helical propeller, which switches its helical pitch and handedness depending on the swimming mode (Macnab and Ornston 1977; Turner et al., 2000). The filament is a helical assembly of flagellin with roughly 11 subunits per two turns of the 1-start helix; it can also be described as a tubular structure comprising 11 protofilaments, which are nearly longitudinal helical arrays of subunits (O'Brien and Bennett 1972).

The crystal structure of FliC from *S. typhimurium* revealed three domains, D1, D2 and D3. Domain D1 forms the outer core of the filament and domains D2 and D3 form the projection on the filament surface. Atomic models of the two straight filaments have been built by combining all available data from X-ray crystallography, X-ray fibre diffraction, and electron

cryomicroscopy, and an insight into the twist-curvature coupling for the supercoiling was obtained from lateral interactions between domains D1 of neighbouring protofilaments (Yonekura et al., 2003).

1.3 The flagellum of the Spirochetes

Spirochetes are a medically important and ecologically significant group of motile bacteria with a distinct morphology including a unique periplasmic flagellum (Figure 1.6). The periplasmic flagella of several spirochete genera (*Leptospira*, *Treponema*, *Brachyspira*, and *Spirochaeta*) possess some features that distinguish them from the flagella of other bacteria (Charon et al., 1992; Chi et al., 2002). Because of a unique protein sheath surrounding the filament core, their flagellar filaments are thicker than those of other bacteria, e.g., 25 nm in diameter for *B. hyodysenteriae* compared to 20 nm for *E. coli* and *S. enterica* (Li et al., 2000b). This sheath is composed of a protein designated FlaA, and its sequence is well conserved among the spirochetes. The periplasmic flagellar core is composed of a family of FlaB proteins. FlaB proteins show 55% to 44% sequence identity to flagellin (fliC) proteins of other bacteria (For example, in *Leptospira* and *Bacillus*).

The FlaA protein is excreted into the periplasmic space by a SecA-dependent system, whereas FlaB is excreted by a Type III secretion system. The mechanism of assembly of the sheath around the core is not known, but the confining environment of the periplasmic space likely plays a critical role. The periplasmic flagellar composition of *B. burgdorferi* is somewhat different. It has only one FlaB protein, and there is considerably more FlaB than FlaA. The location of FlaA in *B. burgdorferi* is uncertain, but evidence suggests that it is closely associated with the periplasmic flagellar filament (Ge et al., 1998). Genetic analysis indicates that the interaction of FlaA with the FlaB subunits affects flagellar filament shape in a manner that has not been observed in any other flagellar system. The composition and genetics of periplasmic flagellar filament synthesis in *B. hyodysenteriae* is best understood, as its genes are more readily manipulated than other spirochetes (Rosey et al., 1995; Stanton et al., 2001). Because of the extensive homology of its FlaA and FlaB proteins with those of other spirochetes, conclusions drawn from the analysis of its periplasmic flagella are likely to be too relevant to other species such as *T. pallidum* and *T. denticola*. Wild-type cells of *B. hyodysenteriae* have periplasmic flagella that are left-handed with a defined shape, as measured by helix pitch and helix diameter (Li et al., 2000b).

Targeted mutagenesis of each of the genes encoding *flaA*, *flaB1*, *flaB2*, and *flaB3* resulted in cells that were still motile and possessed periplasmic flagella that retained their left-handed configurations. However, comparing mutant and wild-type purified periplasmic flagella shapes yielded surprising results (Li et al., 2000b). Each of the *flaB* mutants had periplasmic flagella that were slightly different from the wild-type. However, the *flaA* mutant's periplasmic flagella displayed major decreases in both helix pitch and helix diameter compared to the wild type. These results suggest that FlaA impacts the shape of the periplasmic flagella. Because bacterial flagella are often quasi-rigid and undergo helical transformations, perhaps the sheath helps stabilize the FlaB helical core into one of these configurations for optimal thrust as it rotates between the outer membrane sheath and cell cylinder.

In spirochetes the flagella are attached near each end of the cell within the periplasmic space. Translational motility requires the rotation of the two flagellar bundles in opposite directions,

which leads to swimming modalities that are quite different and more complex than the well-studied paradigms of *Escherichia coli* and *Salmonella*. The coordinated flagellar rotation, flexion, and direction reversals that occur in spirochete motility likely involve an efficient and novel signalling mechanism (Charon and Goldstein 2002). The coordinating signal would be rapidly transmitted from one end of the cell to the other, which in some cases is over 100- fold greater than the cell diameter. Therefore, the analysis of spirochete motility should lead to a better understanding of the nature of such long travelling signals in both spirochetes and other large bacteria. Finally, many spirochetes, including *Treponema*, *Borrelia*, and *Leptospira* are highly invasive pathogens. Motility is likely to play a major role in the disease process, and there is some recent evidence to support this conjecture (Kennedy et al., 1997; Lux et al., 2001; Lux et al., 2000; Rosey et al., 1996; Sadziene et al., 1991).

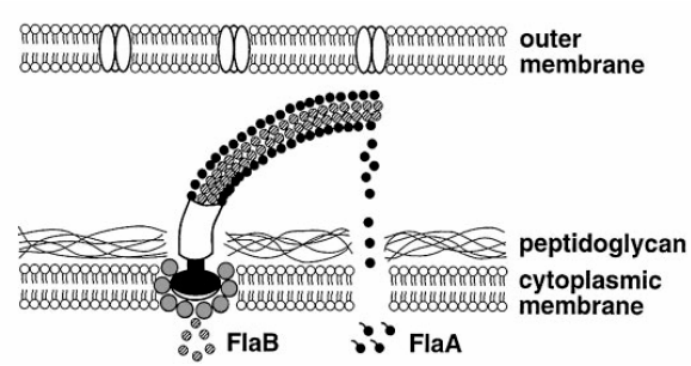


Figure 1.6 Structure of a spirochete periplasmic flagellum: In several spirochete species, the basal bodies lack the L and P rings found in *E. coli* and *S. enterica.* FlaA and FlaB are excreted to the periplasmic space by SecA and type III secretion systems, respectively. This complex periplasmic flagellar structure is found in *Treponema* sp., *S. aurantia*, *Leptospira* sp., and *B. hyodysenteriae* (from Charon and Goldstein 2002).

1.4 Self Assembly of the bacterial flagellum

Self-assembly is used in biological systems to construct large molecular complexes and cellular organelles. The assembly mechanism of the bacterial flagellum and type III secretion system involves dynamic conformational changes and partial folding. The bacterial flagellum is a classical example of such mechanisms (Namba and Vonderviszt 1997; Vonderviszt et al., 1989). Flagellum assembly vividly illustrates how the conformational flexibility and adaptability of biological macromolecules regulate the assembly processes. However, the structures of the assembly sites have never been visualized, even for these well-defined structures. The assembly process of several components of the bacterial flagellum was analyzed by cryo-electron microscopy and single-particle image analysis (Yonekura et al., 2001; Yonekura et al., 2000). The order of assembly of flagellar substructures is for the most part linear and sequential, proceeding from more proximal structures to more distal ones, i.e., from integral and peripheral membrane components to periplasmic components to outer membrane components and finally to components that lie in the cell exterior. The assembly process starts with FliF ring formation in the cytoplasmic membrane (Suzuki et al., 1978). A dedicated export apparatus, homologous to the type III protein export system (Kubori et al., 1998) is believed to be integrated at the cytoplasmic opening of the FliF ring channel and to export selectively a set of flagellar proteins into the channel of the flagellum by using the energy of ATP hydrolysis (Fan et al., 1997; Minamino and Macnab 1999). The flagellar proteins travel through the channel of the growing structure to the distal end, where the assembly occurs (Emerson et al., 1970).

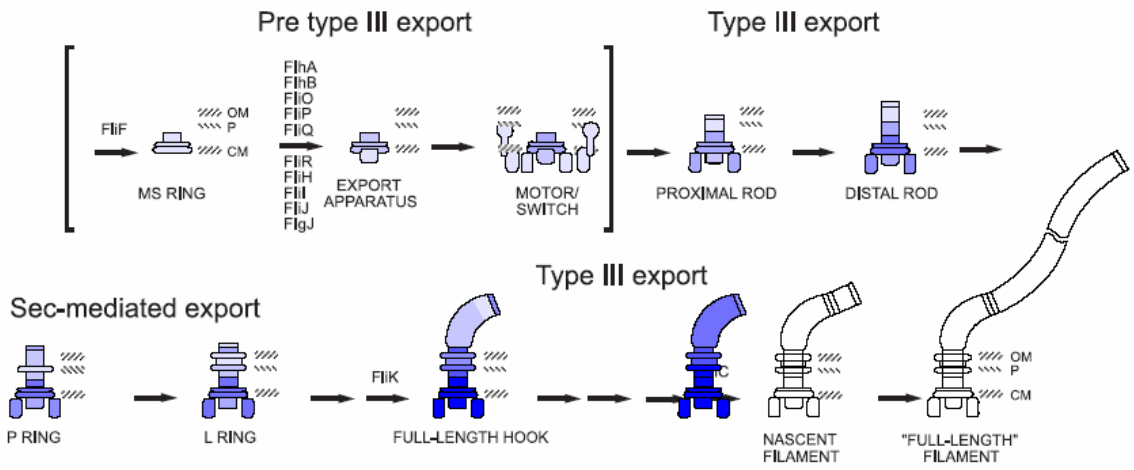


Figure 1.7 Type III secretion: Flagellar morphogenesis in Salmonella, emphasizing the type of pathway used at each stage. Substructures are shaded from dark blue (early) to light blue (late). Pre-type III assembly presumably involves Sec-mediated insertion for integral membrane structures, spontaneous assembly for peripheral structures, and transient association for soluble components; the order of assembly for this stage is not fully understood. Type III export is used for rod assembly, followed by Sec-mediated export with signal peptide cleavage for the P and L rings. Type III export resumes for the remainder of flagellar morphogenesis (from Macnab 2004).

1.4.1 Type III secretion and flagellar protein export system

Bacteria employ a number of different systems for transferring proteins from their cytoplasm to external structures, to the external medium, or into host cells. Flagellar protein export occurs by a type III secretion system (TTSS) which also contains essential virulence factors of many bacteria pathogenic to humans, animals and plants (Galan and Collmer 1999). These systems mediate the transfer of bacterial virulence proteins directly into the host cell cytoplasm. Proteins are thought to travel this pathway in a largely unfolded manner, and families of customized cytoplasmic chaperones, which specifically bind cognate secreted proteins, are essential for secretion.

The physical apparatus of a typical flagellar export system consists of both integral membrane and soluble components (Table 1.1). There are nine proteins that are truly central to the flagellar export apparatus, in the sense that they participate in the export of flagellum rod, hook, and filament proteins (Table 1.3) (Minamino and Macnab 1999; Minamino and MacNab 2000). Three of the proteins—FliH, FliI and FliJ—are soluble, the remaining six are integral membrane proteins and are believed to be located in a patch of specialized membrane within a central pore in the basal-body MS ring (Figure 1.7) (Macnab 1999; Suzuki et al., 1998).

Minamino et al (2000) used an affinity blot assay to detect the interactions among components of the type III flagellar export apparatus and between the export components and their substrates. Most substrates bound (to varying degrees) to $FlhA_C$, FliH, FliI and FliJ. With respect to binding to $FlhB_C$, there was a clear difference between rod / hook-type substrates, which bound strongly and filament-type substrates, which bound weakly. This may reflect the role of FlhB in substrate specificity switching (Table 1.3) (Minamino and MacNab 2000).

Minamino et al propose a model for the flagellar export apparatus in which FlhA and FlhB and the other four integral membrane proteins (FliO, FliP, FliQ and FliR) of the apparatus form a complex at the base of the flagellar motor (Figure 1.8). A soluble complex of at least three proteins (FliH, FliI and FliJ) bind the protein to be exported and then interact with the complex at the motor to deliver the protein, which is then exported in an ATP-dependent process mediated by FliI (Minamino and MacNab 2000). Recently Akeda et al. showed that InvC, an ATPase associated with a Salmonella enterica type III secretion system, has a critical function in substrate recognition. Furthermore, InvC induces chaperone release from and unfolding of the cognate secreted protein in an ATP-dependent manner which indicates a similarity between the mechanisms of substrate recognition by type III protein secretion systems and AAA + ATPase disassembly machines (Akeda and Galan 2005).

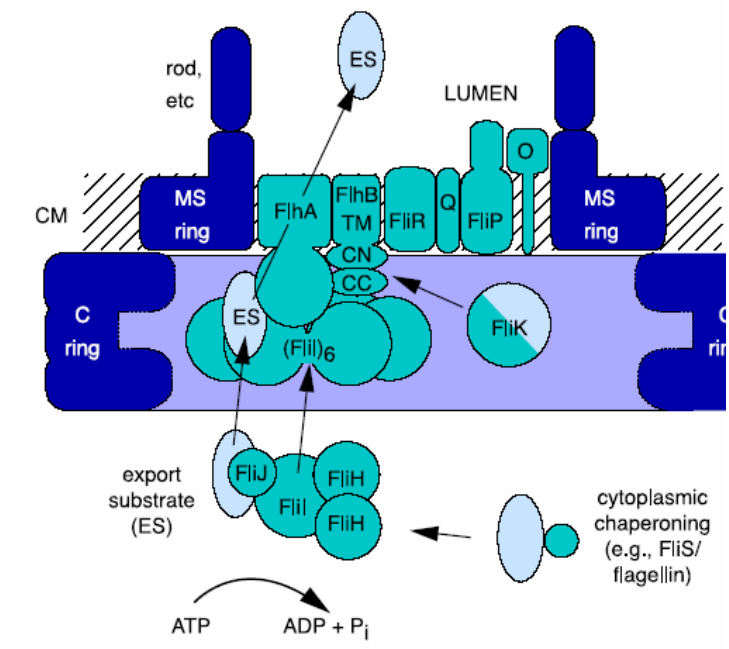


Figure 1.8 Cartoon of a transverse section of the flagellar protein export apparatus. Its components (turquoise) are either membrane proteins housed within the cellmembrane (CM) basal-body MS ring (dark blue) or soluble proteins that dock with the cytosolic domains of those membrane components (FlhA and FlhB), which project into the C ring (dark blue and lilac), (from Macnab 2004).

Table 1.1 Proteins of the flagellum export apparatus

Protein Name	Function	Cellular location	Stoichiometry (approx.)[a]	Molecular mass (kDa)
FliI	ATPase; drives type III flagellar export	Cytoplasm		49.3
FliH	Negative regulator of FliI	Cytoplasm		25.8
FliJ	General chaperone	Cytoplasm		17.3
FlhA	Export component; target for soluble export complex	Membrane component; Center of MS ring*b*	≥2	74.8
FlhB	Export component; substrate specificity	Membrane component; Center of MS ring*b*	≥2	42.4
FliO	Export component	Membrane component; Center of MS ring*b*	≥1	13.1
FliP	Export component	Membrane component; Center of MS ring*b*	~4	26.8
FliQ	Export component	Membrane component; Center of MS ring*b*	≥1	9.6
FliR	Export component	Membrane component; Center of MS ring*b*	≥1	28.9

*a*Stoichiometry per flagellum, *b*These proteins are believed to reside in a specialized patch of the membrane within a pore in the MS ring (from Macnab 2003; Macnab 2004).

Table 1.2. Protein-protein interactions between export components assayed by affinity blotting

	$FlhA_C$	$FlhB_C$	FliH	FliI	FliJ	FliO	FliP	FliQ
$FlhA_C$	+	++	++	+/-	+/-	++	++	++
$FlhB_C$	+/-	+/-	+/-	+/-	+/-	-	-	-
FliH	++	+++	+++	+++	+++	-	-	-
FliI	++	++	+++	++	++	-	-	-
FliJ	+++	+	+++	++	+	-	-	-

Table 1.2 and table 1.3: Semiquantitative scale to reflect the strength of the interaction, as follows: no evidence for interaction (-); ambiguous interaction (+/-); weak interaction (+); moderate interaction (++); and strong interaction (+++) (Minamino and MacNab 2000). $FlhA_C$: C-terminal fragment of FhlA. $FlhB_C$: C-terminal fragment of FhlB.

Table 1.3 Protein-protein interactions between type III export substrates and export components assayed by affinity blotting.

	Rod/hook-type substrates				Filament-type substrates		
	FliE	FlgB	FlgE	FlgD	FlgK	FlgL	FliC
$FlhA_C$.	-	+/-	++	+++	++	++	+++
$FlhB_C$.	+++	+++	+++	+++	+	+/-	-
FliH.	-	++	+++	+++	+++	+++	+++
FliI.	+++	+++	+++	+++	+++	+++	+++
FliJ.	+/-	+/-	++	++	++	++	++

1.4.2 Flagellum export chaperones

Assembly of the bacterial flagellum and type III secretion of pathogenic bacteria require cytosolic export chaperones that interact with mobile components to facilitate their secretion. There are at least three to four specific cytoplasmic chaperones depending on species, FlgN, FliS, FliJ and FliT (Auvray et al., 2001; Bennett and Hughes 2000; Bennett et al., 2001; Fraser et al., 1999; Yokoseki et al., 1995).

Table 1.4 flagellum export chaperones

Protein Name	Function	Cellular location
FliJ	General chaperone	Cytoplasm
FlgN	FlgK-, FlgL-specific chaperone	Cytoplasm
FliS	FliC-specific chaperone	Cytoplasm
FliT	FliD-specific chaperone	Cytoplasm
FlgA	FlgI-specific chaperone	Periplasmic

FlgN: is a type-III secretion chaperone, which is proposed to regulate flagellar assembly via a negative-feedback loop containing its chaperone substrates (FlgK and FlgL), and to act as a translational regulator of FlgM (Aldridge et al., 2003).

FliJ: is a component of the flagellar export apparatus and has a chaperone-like activity for both rod-type, and hook-type substrates (Auvray et al., 2001; Fraser et al., 2003).

FliS: is a flagellin-specific export chaperone; it binds to flagellin (FliC) and prevents polymerization of flagellins in cytosol. The crystal structure for the FliS homologue from *A. aeolicus* has revealed that, FliS adopts a novel fold that is clearly distinct from those of the type III secretion chaperones, indicating that they do not share a common evolutionary origin. However, the structure of FliS in complex with a fragment of FliC (flagellin) revealed that, like the type III secretion chaperones, flagellar export chaperones bind their target proteins in extended conformation (Evdokimov et al., 2003). Mysteriously, *C. crescentus* lacks a FliS homologue but possesses six flagellins (Ely et al., 2000).

FlgA: a specific periplasmic chaperone, FlgA, for the P-ring protein (flgI) (Nambu and Kutsukake 2000). Whether or not they all play a direct role in presenting their substrates to the export apparatus is not clear, but they do prevent substrate degradation (Ozin et al., 2003).

Other assembly proteins that play important roles in various aspects of the assembly process, such as the hook length control, muramidase/rod-capping, hook-capping, and the filament-capping (Table 1.5). With the exception of FliD, these proteins are not present in the final flagellar structure**.**

Table 1.5 Other assembly proteins

Protein Name	Function	Reference
FliK	Hook lenght control protein	(Hirano et al., 1994; Patterson-Delafield et al., 1973; Williams et al., 1996)
FlgD	Hook-capping protein	(Ohnishi et al., 1994)
FliD	Filament-capping protein	(Yonekura et al., 2000)

1.5 Regulation cascade of flagellar gene expression

Bacterial flagellum synthesis genes form an ordered cascade in which the expression of one gene at a given level requires the transcription of another gene at a higher level. At the top of the hierarchy is the flhDC master operon in enterobacteria (Komeda 1982; Komeda 1986; Kutsukake et al., 1990), and fleQ or flrA master genes in *P. aeruginosa*, and *V. cholerae* (Klose and Mekalanos 1998b; Prouty et al., 2001), respectively. The organisation of the flagellar system has been extensively studied in enterobacteria and multiple levels of flhDC regulation were observed, such as transcriptional and posttranscriptional control in *E. coli* and protein stability control in *Proteus mirabilis* (Claret and Hughes 2000). The expression of the σ^{70} dependent flhDC operon is controlled by numerous environmental signals, e.g. temperature, osmolarity and pH (Li et al., 1993; Shi et al., 1993; Soutourina et al., 2002) and by global regulatory proteins, such as H-NS and the cAMP-CAP complex (catabolite gene activator protein) (Figure 1.9) (Bertin et al., 1994; Soutourina et al., 1999). Moreover, the stability of its mRNA is controlled by the RNA binding regulator CsrA (Romeo 1998; Wei et al., 2001). In contrast, the regulation of master genes governing the synthesis of polar flagella remains largely unknown.

1.5.1 Gram-negative bacteria

In Gram-negative bacteria, the hierarchy of the flagellar regulatory system was first well characterised in micro-organisms with peritrichous flagella, such as *E. coli* and *S. typhimurium* (Komeda 1982; Komeda 1986; Kutsukake et al., 1990). This regulatory cascade possesses three classes of genes. Class I genes form the flhDC master operon at the top of the hierarchy, which encodes an $FlhD_2C_2$ transcriptional activator for the second class of gene expression. The majority of class II genes encode components of the flagellar export system and the basal body. The fliA gene at this second level encodes a sigma factor, σ^{28} (or RpoF), specific for flagellar genes. The operons of class III are positively regulated by σ^{28}, and negatively controlled by an anti-sigma factor, FlgM. The anti-sigma factor is retained inside the cell until the flagellar basal body and hook are completed (Hughes et al., 1993). At that time, the FlgM protein is exported to allow activation by σ^{28} of the transcription of class III genes, which encode the components of flagellar filament, (flagellin, FliC), hook-associated, motor, and chemotaxis proteins.

1.5.2 Regulatory cascade components of bacteria with one polar flagellum

In *V. cholerae* the class I gene at the top of the hierarchy encodes the FlrA protein, a σ^{54} associated transcription activator of the NtrC family (Prouty et al., 2001). This regulator, together with σ^{54}, activates the expression of the class II genes encoding structural components of the MS ring, the switch complex, and export apparatus, the σ^{28} flagellar-specific sigma factor, FliA, and FlrB and FlrC, the sensor kinase and transcriptional regulator, respectively, of a two-component signal-transducing system. The FlrC regulator along with σ^{54} activates the expression of the class III genes encoding the basal body, the hook and the flagellin essential for motility, FlaA. The sigma factor σ^{54} is required for the transcription of the class IV genes, including the motor genes and non-essential flagellin genes flaB, flaC, flaD, flaE (Figure 1.9). The absence of these last four genes does not affect flagellar function, but they may help to maintain the antigenic and environmental variations in flagellar filament composition (Klose and Mekalanos 1998a; McCarter 2001).

1.5.3 Spirochetes

There are many unknown factors with respect to motility and chemotaxis gene regulation in spirochetes. *T. pallidum, S. aurantia, T. denticola*, and *B. hyodysenteriae* have many motility promoter sequences homologous to σ^{28} recognition sequences (Charon et al., 1992; Li et al., 2000a). Thus, it is likely that flagellar gene regulation is similar in these spirochete species to that found in other bacteria. In contrast, all of the chemotaxis and motility gene promoters in *B. burgdorferi* are initiated by the housekeeping transcription factor σ^{70} (Ge et al., 1997; Gilmore et al., 2001; Li et al., 2000a). No motility-specific sigma factor recognition sequence is evident in genomic analysis. *B. burgdorferi* appears to be the only bacterial species that lacks transcriptional cascade control of motility gene expression by alternative sigma factors. One hypothesis is that motility and chemotaxis are so vital for the survival of *B. burgdorferi* in both the tick and mammal that many of the genes are constitutively expressed. In support of this hypothesis, *B. burgdorferi* expresses *flaB* message and produces periplasmic flagella in both these hosts (Das et al., 1997; Gilmore et al., 2001). Alternatively, perhaps *B. burgdorferi*, and possibly other spirochete species (Limberger et al., 1999), primarily rely on a translational control system to regulate motility and chemotaxis gene expression.

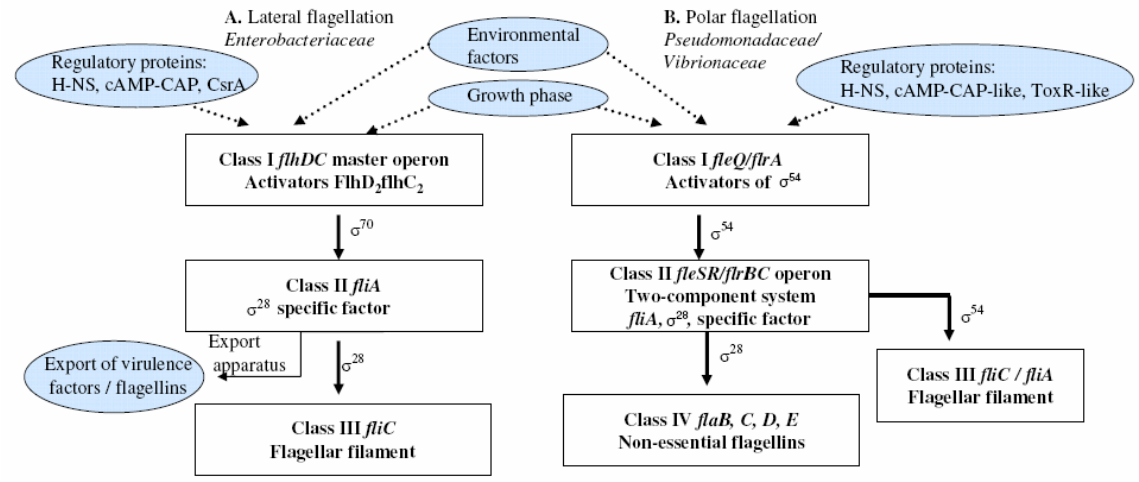


Figure 1.9 Flagellation cascade: Lateral (*Enterobacteriaceae* family) and polar (*Pseudomonadaceae, Vibrionaceae* families) flagellar cascades are compared and the factors controlling master regulator expression and the connections with other cellular processes are shown. **A**, The lateral flagellation cascade discovered in Enterobacteriaceae with the flhDC master operon at the top (class I) encoding the $FlhD_2C_2$ transcriptional activator of class II genes including flagellar-specific σ^{28} factor fliA gene. FliA (σ^{28}) is necessary for the transcription of class III genes including the fliC flagellin gene. **B**, The polar flagellation cascade identified in *Pseudomonadaceae* and *Vibrionaceae* with fleQ and flrA genes at the top. These genes encode σ^{54} associated NtrC-type transcriptional activators, FleQ and FlrA, respectively, which activate the transcription of class II genes, such as fleSR and flrBC. Both are two-component regulatory systems for class III genes that include flagellar filament genes fliC and flaA, in *Pseudomonadaceae* and *Vibrionaceae*. Another class II gene, fliA, encodes the σ^{28} flagellar-specific factor participating in the transcription of class IV genes, such as non-essential flagellin genes flaB, C, D, and E in *Vibrionaceae*. Other factors such as regulatory proteins, environmental signals and growth phase-controlling master regulatory gene expressions are shown in ovals at the top (Soutourina and Bertin 2003).

Preliminary results with insertion mutants in *flaA*, *flaB*, *flgE*, and *fliF* of *B. burgdorferi* support this possibility. For example, mutants in *flaB* of *B. burgdorferi* still synthesize similar amounts of *flaA* message as the wild type, but fail to produce even small amounts of FlaA protein. DNA microarray analysis for *B. burgdorferi* indicates that many of the transcripts encoding chemotaxis and motility genes are differentially regulated in the mammalian host compared to the tick (Revel et al., 2002). Genes besides those involved in motility and chemotaxis have been known for some time to be differentially regulated as a function of host (Pedersen et al., 1982). Both the environmental factors within the two hosts and the molecular details within the spirochete that promote this differential regulation of gene expression are just beginning to be elucidated (Hubner et al., 2001; Revel et al., 2002).

1.6 Chemotaxis-guided motility in bacteria

Motile bacteria use sophisticated chemotaxis signalling systems to direct their movements. In general, bacterial chemotactic signal transduction pathways have three basic elements: signal reception by bacterial chemoreceptors located on the membrane; signal transduction to relay the signals from membrane receptors to the motor; and signal adaptation to desensitize the initial signal input. The chemotaxis proteins involved in these signal transduction pathways have been identified and extensively studied, especially in the enterobacteria *Escherichia coli* and *Salmonella typhimurium.* Chemotaxis-guided bacterial movements enable bacteria to adapt better to their natural habitats *via* moving towards environments that contain higher concentrations of beneficial or lower concentrations of toxic, chemicals. The signalling pathway that is involved has long been viewed as a paradigm of histidine–aspartate-phosphorelay (HAP) signalling, and is one of the best-understood physiological processes in biology.

E. coli swim by rotating their five to eight helical flagella anticlockwise, which cause them to come together into a bundle that propels the cell forward. Switching the rotational direction of some flagellar motors, to clockwise disrupt this bundle and causes the cell to tumble. When the motors return to anticlockwise rotation, the cell is reoriented and swims off in a new direction (Armitage and Schmitt 1997; Turner et al., 2000). In homogeneous environments, swimming bacteria change direction about once a second, which produces random movement. In no homogeneous environments, the frequency of direction changing is controlled by positive or negative stimuli to bias the overall direction of movement. Although the number and location of flagella vary between bacterial species, the biasing of swimming always relies on chemosensory regulation of the rotary activity of the flagella motor (switching, stopping or slowing down) (Armitage and Schmitt 1997). Chemotactic signals are detected by dedicated transmembrane CHEMORECEPTORS the METHYL ACCEPTING CHEMOTAXIS PROTEINS (MCPs) and an MCP-like protein rather than by a fused sensory domain of an HPK (Histidine protein kinases) as is found for most HAP systems that regulate transcription. An adaptor protein, CheW, helps to link the MCPs to the cytoplasmic HPK, CheA, and two response regulators (RRs), compete for binding to CheA (Figure 1.10). One response regulator is a single-domain, flagellar- motor-binding protein, CheY, whereas the other, CheB, has two domains, one of which functions as a methylesterase and controls the adaptation of the MCP. Phosphorylated CheY (CheY–P) binds the switch protein FliM (Figure 1.10) on the flagellar motor and causes a reversal in the direction of motor rotation (Toker and Macnab 1997; Welch et al., 1993). In *E.*

coli, the phosphatase CheZ is required to increase the spontaneous dephosphorylation rate of CheY–P and allow rapid signal termination (McEvoy et al., 1999).

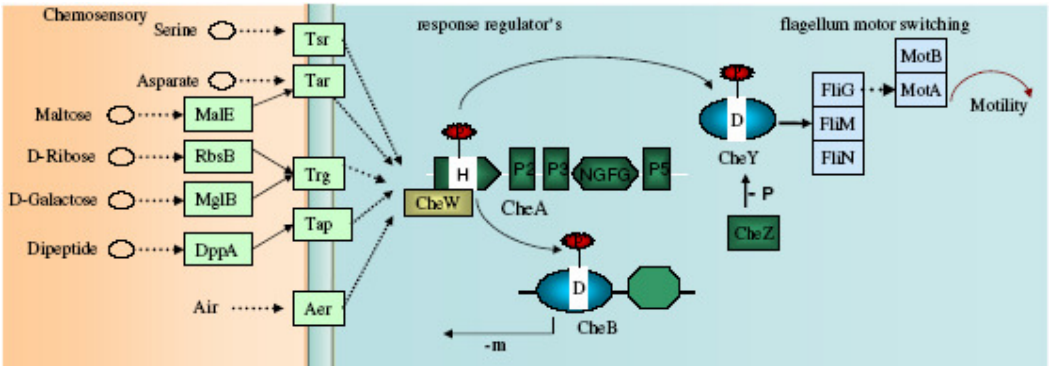


Figure 1.10 Bacterial chemosensory pathway: The chemosensory pathway of *E. coli*. The soluble HPK chemotaxis protein CheA has five domains per monomer that are designated P1–P5 from the N terminus to the C terminus. CheA senses changes through transmembrane chemoreceptor, which induce the *trans*-autophosphorylation of dimeric CheA on a His residue of the HPt domain. Two response regulators compete for this phosphoryl group: CheY, a single-domain, motor-binding protein (fliM), which controls flagellar motor switching through fliG, and CheB, acts as methylesterase, which controls the adaptation of the chemoreceptors.

Table 1.6 The Protein Data Bank (PDB) accession number of the known structures of chemosensory system and flagellum apparatus

X: No homologous proteins found in *T. pallidum*

Protein name	Structure	PDB accession	*Treponema* homologs
maltose-binding protein	maltose-binding protein	1ANF	X
MCP	periplasmic domain	1WAT	TP0040
MCP	cytoplasmic domain	1QU7	TP0040
CheA	P1 domain	1I5N	TP0363
CheA	P2 domain	1EAY	TP0363
CheA	P3–P5 domains	1B3Q	TP0363
CheW	CheW	1KOS	TP0364
CheY	CheY	4TMY	TP0366
CheR	CheR	1BC5	TP0630
CheB	CheB	1A2O	TP0631
CheZ with phosphorylated CheY bound	CheY-cheZ in complex	1KMI	X - TP0366
FliG	C-terminal	1QC7	TP0400, TP0026
FlgE	FlgE	2BGY, 1WLG	TP0727
FlhDC	FlhD and FlhC in complex	2AVU	X
FliS	FliS	1ORJ	TP0943
FliN	FliN	1O6A	TP0720

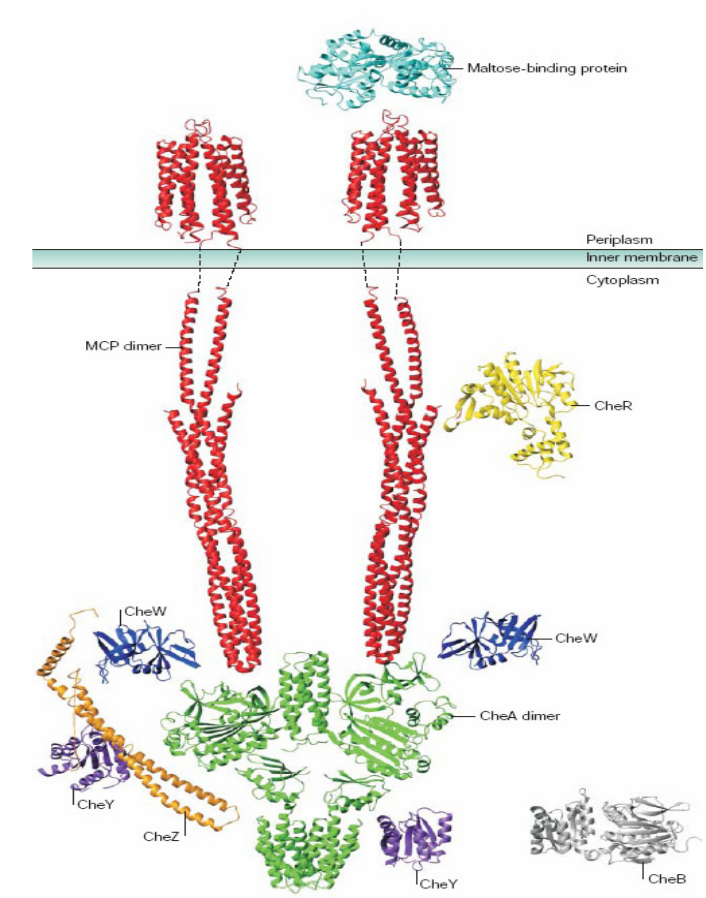


Figure 1.11 The structures of components of the chemosensory system of *Escherichia coli*: Two methyl-accepting chemotaxis protein (MCP) dimers are shown interacting with two CheW monomers and a CheA dimer. CheY, CheB and CheR are all depicted as monomers. The figure is a graphical representation and is not meant to imply that the domains or proteins interact in the manner depicted or that their orientations are as shown. The dashed lines represent the transmembrane domains and the HAMP domains ('Histidine kinases, Adenylyl cyclases, Methyl-binding proteins and Phosphatases' domains), for which there are, as yet, no structures (from Wadhams and Armitage 2004).

1.7 Protein-protein interactions

Protein-protein interactions affect all processes in a cell. It has been proposed that most proteins in a given cell are connected through an extensive network where non-covalent interactions are continuously forming and dissociating.

As elementary constituents of cellular protein complexes and pathways, protein-protein interactions are key determinants of protein function and protein interactions have been shown to be powerful predictors of protein function. The function of proteins is directly addressed by studies of protein-protein interactions (PPIs) or their co-occurrence in complexes. The function of unknown proteins may be discovered through their interaction with protein partners of known function.

Although a number of methods are available for high-throughput analysis of protein–protein interactions (PPIs), the most commonly used are the yeast-two-hybrid (Y2H) system and a combination of protein-complex purification and subsequent analysis by mass spectrometry (Gavin et al., 2002; Ho et al., 2002; Mann et al., 2001; Titz et al., 2004). The Y2H is a powerful tool for the identification of PPIs, which can be applied in a high-throughput manner to detect interactions across the entire proteome of an organism. The first genome-wide protein–protein interaction studies of a free-living organism, *S. cerevisiae*, have been carried out by Uetz et al., (Uetz et al., 2000). Rain et al. compiled a partial protein-protein interaction map of the human gastric pathogen *Helicobacter pylori*, using a high-throughput strategy of the yeast two-hybrid

assay to screen 261 *H. pylori* proteins against a highly complex library of genome-encoded polypeptides (Rain et al., 2001). Recently the interactomes of viruses (Kang et al., 2004; Uetz and Finley 2005), *C. elegans* (Li et al., 2004), *D. melanogaster* (Formstecher et al., 2005) and human (Rual et al., 2005; Stelzl et al., 2005), have been studied using the Y2H. Such interaction maps detect binary interactions among proteins and can therefore provide clues not only about the biological function of individual proteins but also about the functional associations with other processes in a cell.

1.7.1 Known protein-protein interactions of the chemotaxis and flagellum apparatus

Various efforts have been made to identity all the components of the flagellar complex and their interactions. Our comprehensive literature review using PubMed revealed 62 protein-protein interactions between flagellar components (Table 1.7). These protein-protein interactions have been identified in different bacterial species by various methods: affinity chromatography, immunoblot, Co-IP, genetic suppressor mutant screens and Y2H (see table 1.7). The structures of fliS, fliC, fliN, and flgE and C-terminal domains fliG have been determined by X-ray crystallography (Table 1.6).

Table 1.7 Known protein-protein interactions of chemotaxis and motility proteins

Proteins A and B interact. COGs A and B refer to clusters of orthologous groups of proteins A and B according to the COG database (Tatusov et al., 2003). NOGs: non-supervised orthologous groups, these are novel orthologous groups, which are fully automatically generated, and they do not have any manually curated functional annotation (Doerks et al., 2004)

PubMed ID	Protein A name	Protein B name	Protein A *T.pallidum* ortholog	Protein B *T.pallidum* ortholog	COG A	COG B	Method	Species
7578071	CheA	CheY	TP0363	TP0366	COG0643	COG0784	isothermal titration calorimetry	*E. coli*
8820640	CheAS	CheZ	TP0363	-	COG0644	COG3143	Co-IP (in vitro)	*E. coli*
377295	CheC	CheZ	TP0720	-	COG1886	COG3143	supressor mutations	*E. coli*
7623663	CheY	CheZ	TP0366	-	COG0784	COG3143	Mutation	*E. coli*
1400175	CheY	FliG	TP0366	TP0026	COG0784	COG1536	supressor mutations	*E. coli*
11135671	CheY	FliM	TP0366	TP0721	COG0784	COG1868	crystal structure	*E. coli*
11673434	FleN	FleQ	TP0712	TP0082	COG0455	COG3604	Y2H & Co-IP (in vitro)	*P.aeruginosa*
11673434	FleN	FleQ	TP0712	TP0519	COG0455	COG2204	Y2H & Co-IP (in vitro)	*P.aeruginosa*
10809679	FlgB	FliE	TP0396	TP0398	COG1815	COG3951	various methods	*S. typhimurium*
11554792	FlgJ	FlgB	TP0959	TP0396	COG3951	COG1815	various methods	*S. typhimurium*
11554792	FlgJ	FliE	TP0959	TP0398	COG3951	COG3951	various methods	*S. typhimurium*
9095196	FlgM	sigma 28	TP0917	TP0709	COG2239	COG1191	crystal structure	A. aeolicus
10320579	FlgN	FlgK	-	TP0660	COG3418	COG1256	affinity blotting	*E. coli*
10320579	FlgN	FlgL	-	TP0659	COG3418	COG1344	affinity blotting	*E. coli*
15516571	FlhA	FliH	TP0714	TP0401	COG1298	COG1317	Affinity blottig	*S. typhimurium*
15516571	FlhA	FliI	TP0714	TP0402	COG1298	COG1157	Affinity blottig	*S. typhimurium*
15516571	FlhA	FliJ	TP0714	TP0403	COG1298	COG2882	Affinity blottig	*S. typhimurium*
15516571	FlhA	FliO	TP0714	TP0719	COG1298	COG3190	Affinity blottig	*S. typhimurium*
15516571	FlhA	FliP	TP0714	TP0718	COG1298	COG1338	Affinity blottig	*S. typhimurium*
15516571	FlhA	FliQ	TP0714	TP0717	COG1298	COG1987	Affinity blottig	*S. typhimurium*
10940035	FlhB	FlgD	TP0715	TP0728	COG1377	COG1843	affinity blotting	*S. typhimurium*
10940035	FlhB	FliC	TP0715	TP0792	COG1378	COG1344	affinity blotting	*S. typhimurium*

10940035	FlhB	FliC	TP0715	TP0868	COG1379	COG1344	affinity blotting	*S. typhimurium*
10940035	FlhB	FliC	TP0715	TP0870	COG1380	COG1344	affinity blotting	*S. typhimurium*
11204784	FlhF	FlhF	TP0713	TP0713	COG1419	COG1419	Y2H	*X. oryzae*
11204784	FlhF	XA21	TP0713	-	COG1419	-	Y2H	*X. oryzae*
8986772	FliC	FliC	TP0868	TP0868	COG1344	COG1344	electron cryomicroscopy	*S. typhimurium*
11327763	FliC	FliS	TP0868	TP0943	COG1344	COG1516	immunoblott	*S.typhimurium*
1551848	FliE	FliE	TP0398	TP0398	COG3951	COG3951	autoradiography	*S. typhimurium*
11160096	FliF	FlhA	TP0399	TP0714	COG1766	COG1298	supressor mutations	*S. typhimurium*
10809678	FliG	FliF	TP0026	TP0399	COG1536	COG1766	Mutation	*S. typhimurium*
15126479	FliG	FliF	TP0400	TP0399	COG1536	COG1766	Mutation	*E. coli*
8757288	FliG	FliG	TP0026	TP0026	COG1536	COG1536	Co-IP	*E. coli*
8631704	FliG	FliM	TP0026	TP0721	COG1536	COG1868	Y2H	*E. coli*
8757288	FliG	FliN	TP0026	TP0720	COG1536	COG1886	Co-IP	*E. coli*
10440379	FliG	MotA	TP0026	TP0725	COG1536	COG1291	structural model	*T. maritima*
10998179	FliH	FliH	TP0401	TP0401	COG1317	COG1317	affinity chromatography	*S. typhimurium*
10998179	FliH	FliI	TP0401	TP0402	COG1317	COG1157	affinity chromatography	*S. typhimurium*
10350613	FliI	fliC	TP0402	TP0792	COG1157	COG1344	affinity blotting	*S. typhimurium*
10350613	FliI	fliC	TP0402	TP0868	COG1157	COG1344	affinity blotting	*S. typhimurium*
10350613	FliI	fliC	TP0402	TP0870	COG1157	COG1344	affinity blotting	*S. typhimurium*
10350613	FliI	hook protein	TP0402	TP0727	COG1157	COG1749	affinity blotting	*S. typhimurium*
12949107	FliJ	FlhA	-	TP0714	COG2882	COG1298	affinity blotting	*S. typhimurium*
12949107	FliJ	FliH	-	TP0401	COG2882	COG1317	affinity blotting	*S. typhimurium*
15757683	FliK	FlhB	-	TP0715	COG3144	COG1377		*S.typhimurium*
8206846	FliM	FliF	TP0721	TP0399	COG1868	COG1766	ultracentrifugation	*E. coli*
8757288	FliM	FliM	TP0721	TP0721	COG1868	COG1868	Co-IP	*E. coli*
9791106	FliM	FliN	TP0721	TP0720	COG1868	COG1886	Co-IP (in vitro)	*E. coli*
8757288	FliN	FliN	TP0720	TP0720	COG1886	COG1886	Co-IP	*E. coli*
12958592	FliS	FliC	TP0943	TP0792	COG1516	COG1344	Crystallographic	*Aquifex aeolicus*
12958592	FliS	FliC	TP0943	TP0870	COG1516	COG1344	Crystallographic	*Aquifex aeolicus*
11327763	FliS	FliS	TP0943	TP0943	COG1516	COG1516	immunoblott	*S. typhimurium*
11169117	FliT	FliD	-	TP0872	NOG08749/ NOG41787	COG1345	gel filtration	*S. typhimurium*
15491362	FliX	FlbD	-	TP0726	NOG42184	COG1582	yeast two-hybrid	*C. crescentus*
9545371	HAP2	HAP2	TP0872	TP0872	COG1345	COG1345	electron micrographs	*S. typhimurium*
8757288	MotA	FliM	TP0725	TP0721	COG1291	COG1868	Co-IP	*E. coli*
8627625	MotB	FliG	TP0724	TP0026	COG1360	COG1536	supressor mutations	*E. coli*
15101977	MotC	MotB	-	TP0724	NOG06999	COG1360	affinity blotting	*S. meliloti*
15101977	MotD	FliM	-	TP0721	COG1886	COG1868	affinity blotting	*S. Meliloti*
10783392	PomA	PomB	TP0725	TP0724	COG1291	COG1360	immunoblotting	*V. alginolyticus*
15968056	sigma(54)	HP0958.	TP0111	TP0494	COG1508	COG1579	yeast two-hybrid, mutant	*H.pylori*
10783392	PomA	PomA	TP0725	TP0725	COG1291	COG1291	immunoblotting	*V. alginolyticus*

1.8 Aims of my PhD work

- Although most components of the bacterial flagellar apparatus have been identified, I do not know of any systematic attempt to map the interactions of these components among each other and with the remainder of the proteome. Our lab therefore set out to map all protein-protein interactions of *Treponema pallidum* (see below why this species has been chosen). The flagellar interaction screen was part of a systematic genome-wide screen for all protein-protein interactions in a bacterium. Thus my specific goals were as follows:
- Characterize the interactome of the *Treponema pallidum* chemotaxis and flagellum apparatus i.e., identification of protein-protein interactions of the *Treponema pallidum* chemotaxis and flagellum apparatus using the yeast two-hybrid system
- Analysis of the protein-protein interaction network, in particular with respect to an understanding of the global organisation of protein interaction of flagellum apparatus with other pathways and processes in the cell.
- Prediction of functions for the proteins of yet unknown function based on their interactions.
- Transfer of these predictions to the genomes of model species e.g. *E. coli or B. subtilis* and experimental verification of the predictions either in these model systems or in vitro.
- Mutational analysis of motility genes in *E. coli* as an independent benchmark and dataset for comparison with *T. pallidum*.

1.9 *Treponema pallidum* as a model organism for interactome studies

For mapping the whole interactome of an organism (an organism's total set of protein–protein interactions), organisms with a small genome are a good choice because ideally all possible pair-wise protein combinations should be tested. However, this number increases exponentionaly with the genome size. For example, *E. coli* ~4000 predicted ORFs, and thus there are 4000*4000 ≈ 16 million combinations to be tested. For *Treponema pallidum* this number is 1041*1041 ≈ 1 million combinations, i.e. only 6% of the number of *E. coli*!

In addition to its small genome size, containing 1041 predicted ORFs (Fraser et al., 1998), all open-reading frames have been cloned into a versatile vector system (McKevitt et al., 2003). Of these 1041 predicted ORFs of *T. pallidum*, 885 ORFs have homologs with >= 40% similarity in *S. typhimurium*, 906 ORFs have such homologs in *B. subtilis* and 853 ORFs in *E. coli*, (http://www.tigr.org). So the protein-protein interaction data of *T. pallidum* flagellum can be used to predict protein interactions in these model organisms. Additionally we can learn about the common and specific flagellum components between spirochetes and other model organisms, because motility in spirochetes is quite different and more complex than the well-studied paradigms of *E.coli and S. typhimurium* (Charon and Goldstein 2002).

1.10 Experimental strategy

1.10.1 The yeast two-hybrid system

In order to map all protein-protein interactions of an organims, the experimental strategy needs to be applied in a high-throughput manner. Currently, there are only two methods which allow such high-throughput approaches, namely protein complex purification and subsequent analysis by mass spectrometry (MS) and the two-hybrid system. MS was not applicable to *T. pallidum* because this organism cannot be grown easily and its pathogenicity places severe restrictions on its maintenance. I thus decided to apply an automated two-hybrid system to map protein interactions in *T. pallidum*.

1.10.2 The principle

The Y2H method is a genetic screening system for protein-protein interaction detection in yeast. The Y2H system is based on the observation that protein domains can be separated, recombined, and still retain their properties. In particular, transcription factors can frequently be split into DNA-binding and activation domains. In the two-hybrid system, a DNA-binding domain (in this case, from the yeast Gal4 protein) is fused to a protein B (for bait) for which one wants to find interacting partners (Figure 1.12). A transcriptional activation domain is then fused to one or more or all the predicted ORFs P (for prey) of an organism and the bait and prey fusion proteins are then co-expressed in the same cell. Usually both protein fusions are expressed from plasmids that can be manipulated easily and then transformed into yeast cells. If the two proteins B and P interact, a transcription factor is reconstituted which in turn activates one or more appropriate reporter genes. The expression of the most critical reporter allows the cell to grow only under certain conditions. For example, the HIS3 reporter encodes imidazoleglycerolphosphate (IGP) dehydratase, a critical enzyme in histidine biosynthesis. In a screening strain lacking an endogenous copy of HIS3, expression of a HIS3 reporter gene is driven by a promoter that contains a Gal4p-binding site, so the bait protein fusion can bind to it. However, since the bait fusion does not contain a transcriptional activation domain it remains inactive. If a protein P with an attached activation domain binds to the bait, this activation domain can recruit the basal transcription machinery, and expression of the reporter gene ensues. These cells can now grow in the absence of histidine in the media because they can synthesize their own.

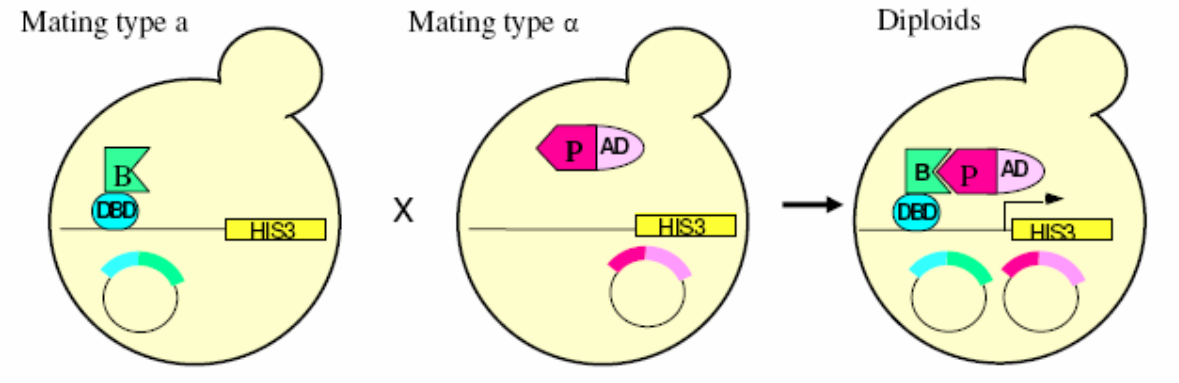


Figure 1.12 Yeast two-hybrid principle: A protein of interest 'B' is expressed in yeast as a fusion to a Gal4p DNA-binding domain (DBD, "bait"; circles denote expression plasmids). Another protein or library of proteins of interest 'P' is fused to Gal4p transcriptional activation domain (AD, "prey"). The two yeast strains are mated to combine the two fusion proteins in one cell. If proteins 'B' and 'P' interact in the resulting diploids cells, they reconstitute a transcription factor which activates a reporter gene (in this case HIS3) and therefore allows the cell to grow on selective media (media lacking histidine).

1.10.3 Array-based two-hybrid screens

In an array, a number of defined prey proteins are tested for interactions with a bait protein (Figure 1.13). Usually, the bait protein is expressed in one yeast strain and the prey is expressed in another yeast strain of different mating type. The two strains are mated so that the two proteins are expressed in the resulting diploid cell (Figure 1.12). The assays are done side-by-side, so they can be well controlled, i.e. compared. As the identity of the preys is usually known, no sequencing is required after positives have been identified. However, the prey clones need to be obtained or made up-front. This can be done for a few genes or for a whole genome, e.g. an ORFeome (i.e. all ORFs of a genome).

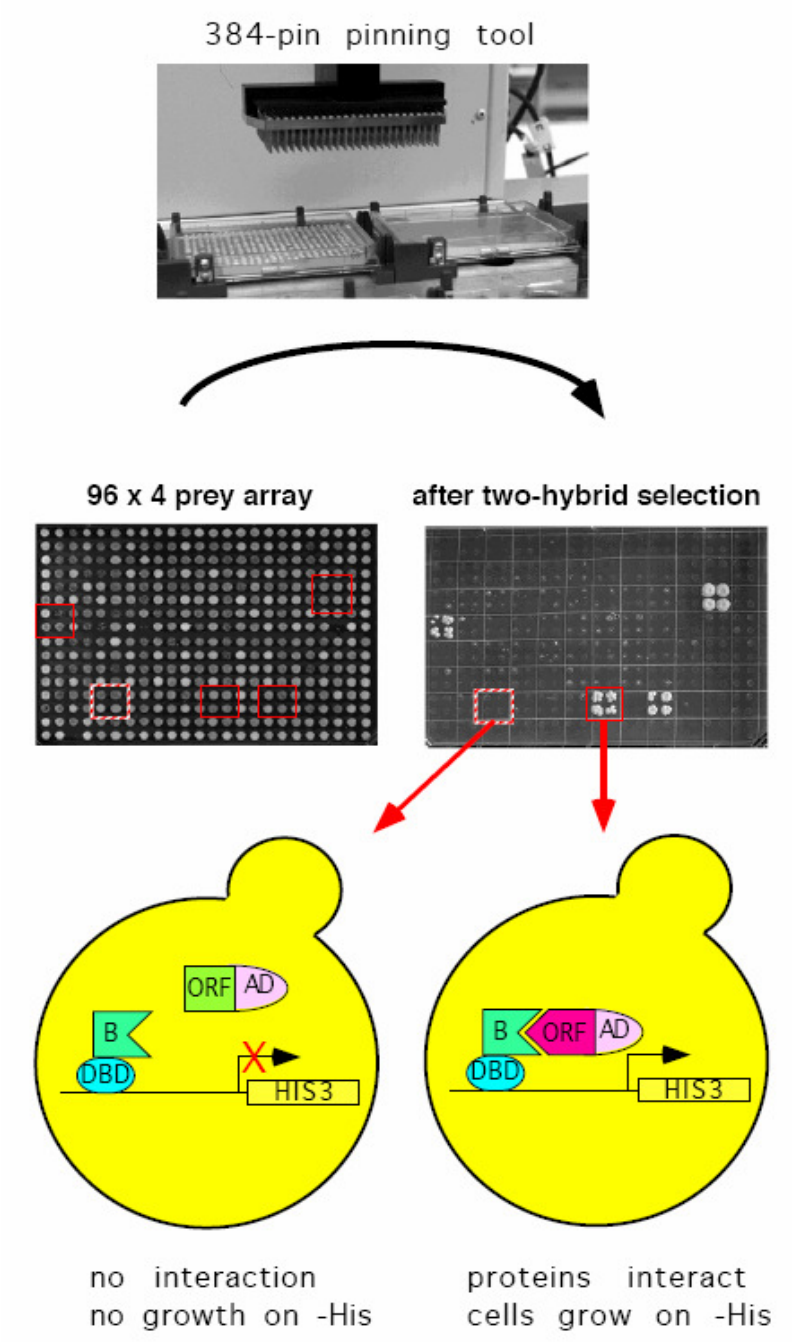


Figure 1.13 Scheme of an array-based two-hybrid screen

In the array-based Y2H system a bait protein is tested against the whole systematic prey library. This library consists of individual yeast colonies at specific positions of an array (e.g., in 384 format, middle left); each colony carries a specific prey ORF. Systematic testing is done by robotic transfer of yeast cells starting with the mating of bait and prey strains; diploid cells are selected on specific plates and, finally, transferred to plates selecting for the activation of a reporter gene; here the activation of the His3p gene is detected by scoring the growth on His deficient plates (middle right). The red rectangle on the selective plate marks a positive interaction between the bait and the prey at this specific position of the array (test is done in quadruplicates). Adapted from Rajagopala *et al* (Rajagopala et al., in press).

1.10.4 Advantages of array screens

Since many two-hybrid tests are done in parallel under the same condition, their results can be directly compared. In an array, each element has a known identity and thus it is immediately clear which two proteins are interacting when positives are selected. In addition, it is often immediately clear if an interaction is stronger than another one. Most importantly, since all of these assays are done in an ordered array, background signals can be easily distinguished from true signals (Figure 1.13). In fact, in some cases only an array screen may do the job. For example, if you have a bait protein that activates transcription on its own, a carefully controlled array may be the only way to distinguish between signal and background (see Figure 1.13). Similarly, weak interactors may only be detectable when compared with even weaker or no background.

1.10.5 Other applications relevant in this context

Originally, the two-hybrid system was invented to demonstrate the association of two proteins (Fields and Song 1989). Later, it was demonstrated that completely new protein interactions can be identified with this system, even when there are no candidates for an interaction with given bait. Over time, it has become clear that the ability to conveniently perform unbiased library screens is the most powerful application of the system. With whole-genome arrays, such unbiased screens can be expanded to defined, non-redundant sets of proteins. The system can also be exploited to map binding domains (Rain et al., 2001) or map interactions within a protein complex (Cagney et al., 2001). Finally, recent large-scale projects have been successful in systmatically mapping interactions within whole proteomes (yeast: Uetz et al. 2000, Ito et al. 2001, worm: Li et al., 2004 fly: Giot et al. 2003, and humam: Rual et al., 2005; Stelzl et al., 2005). These studies have shown for the first time that most proteins in a cell are actually connected to each other.

2. Materials and Methods

2.1 Materials

2.1.1 Instruments

Name	Source
Biomek 2000 (Robot)	Beckman Coulter, USA
OmniTray	Nalge Nunc International, USA
Agarose gel electrophoresis chamber	Peqlab, Erlangen
Bacterial/yeast incubator	Heraeus, Stuttgart
Bacterial/yeast shaker	Infors, Bottmingen, Switzerland
Developing machine	Kodak, New Haven, USA
Cold centrifuge 5810R	Eppendorf, Hamburg
PCR thermocycler PTC-200	MJ Research, Waltham, USA
pH meter	Eppendorf, Hamburg
Photometer	Eppendorf, Hamburg
PAG electrophoresis chamber	Amershan, Freiberg
E-PAGE, Pre-Cast Gel System	Invitrogen
Thermomixer	Eppendorf, Hamburg
Table centrifuge	Heraeus, Stuttgart
Vortex Genie2	Bender und Hobein, Karlsruhe
Odyssey	LI-COR Biosciences, Homburg

2.1.2 Consumable Materials

Name	Source
Eppendorf tube	Eppendorf, Hamburg
Film	Amershan, Freiberg
Glass beads	Sigma, Steinheim
Nitrocellulose membrane	Biorad, München
PCR tube	Roth, Karlsruhe
Whatman paper	Bender und Hobein, Karlsruhe
Sterile Filters (Filter pore size 0.22 µm)	Millipore, Schwalbach
PVDF membrane	Millipore, Schwalbach

2.1.3 General Chemicals

Name	Source
3-Amino-1,2,4-triazole (3-AT)	Sigma, Steinheim
Acrylamide	Roth, Karlsruhe
Agarose	Peqlab, Erlangen
Ampicillin	Roth, Karlsruhe
APS	Roth, Karlsruhe
Bromophenolblue	Sigma, Steinheim
BAS	Promega, Mannheim
Calcium acetate	Roth, Karlsruhe
Calcium chloride	Roth, Karlsruhe
Chloroform	Roth, Karlsruhe
DMSO	Fluka, Buchs, Swiss
DTT	Roth, Karlsruhe
EDTA	Roth, Karlsruhe
Erythromycin	Roth, Karlsruhe
Ethanol	Roth, Karlsruhe

Ethidium bromide	Roth, Karlsruhe
Galactose	Roth, Karlsruhe
Glucose	Roth, Karlsruhe
Glycin	Roth, Karlsruhe
Glycerol	Roth, Karlsruhe
HEPES	Roth, Karlsruhe
IPTG (Isopropyl-ß-D-thiogalactopyranosid)	Sigma, Steinheim
Isopropanol	Roth, Karlsruhe
Kanamycin	Roth, Karlsruhe
L-Arabinose	Sigma, Steinheim
Methanol	Roth, Karlsruhe
Sodium chloride	Roth, Karlsruhe
Sodium hydroxide	Roth, Karlsruhe
PEG	Sigma, Steinheim
Phleomycin	Sigma, Steinheim
Phenol	Roth, Karlsruhe
PIPES (piperazine-1, 2-bis[ethanesulfonic acid)	Sigma, Steinheim
Ponceau S	Sigma, Steinheim
SDS	Roth, Karlsruhe
TEMED	Sigma, Steinheim
Tris-Base	Roth, Karlsruhe
Triton X-100	Sigma, Steinheim
Brij 58 (polyoxyethylene 20 cetyl ether)	Sigma, Steinheim
Tween -20	Roth, Karlsruhe

2.1.4 Chemicals for media

Name	**Source**
Adenine	Merck, Darmstadt
Arginine	Merck, Darmstadt
Bacto-Agar	Nordwald, Hamburg
Bacto-Yeast-Extract	Roth, Karlsruhe
Bacto-Pepton	Roth, Karlsruhe
Histidine	Merck, Darmstadt
Isoleucine	Merck, Darmstadt
Leucine	Merck, Darmstadt
Lysine	Merck, Darmstadt
Methionine	Merck, Darmstadt
Phenylalanine	Merck, Darmstadt
SD (Difco Yeast Nitrogen Base)	Becton-Dickinson, Sparks, USA
Threonine	Sigma, Steinheim
Tryptophan	Merck, Darmstadt
Uracil	Sigma, Steinheim

2.1.5 DNA and Protein Ladders

Name	**Source**
1 kb DNA ladder	peqlab
Protein marker	peqlab

2.1.6 Other chemicals

Name	Source
dNTP	Sigma, Steinheim
ECL Western Blotting Analysis System	Amersham, Freiburg
IgG-Sepharose 6 Fast Flow	Amersham, Freiburg
Milk powder	Saliter, Obergünzburg
Ni-NTA Magnetic Agarose Beads	Qiagen, Hilden
PMSF	Serva, Heidelberg
Phenol-chloroform extract	Roth
Aprotinin (protease inhibitor)	Sigma
Antipain (protease inhibitor)	Sigma

2.1.7 Kits

Name	Source
Miniprep DNA purification	Qiagen, Hilden
Plasmid minipreparation 96-well	Millipore, Schwalbach
High Pure PCR-Product Purification Kit	Roche, Mannheim

2.1.8 Media for bacterial culture

2.1.8.1 LB liquid medium

0.5% (w/v) Bacto-Yeast-Extract
1% (w/v) Bacto-Peptone
1% (w/v) NaCl

The medium was autoclaved and an appropriate amount of antibiotic (Ampicilin 100 µg/ml, Kanamycine 50 µg/ml, Phleomycine 10 µg/ml) was added to the medium after it was cooled down to 50°C.

2.1.8.2 LB agar plates

The same recipe as for the liquid medium was used. In addition, 2% (w/v) Bacto-Agar was added before autoclaving.

2.1.8.3 Bacterial swarming plates

The same recipe as for the liquid medium was used. In addition, 0.25% (w/v) Bacto-Agar was added before autoclaving. If necessary, an appropriate amount of IPTG or L-Arabinose was added to the medium after it was cooled down to 50°C

2.1.9 Media for yeast culture

YEPD liquid medium

Yeast extract	10 g
Peptone	20 g
Glucose	20 g

Make up to 1 liter with sterile water and autoclave.

YEPD solid medium

Yeast Extract	10 g
Peptone	20 g
Glucose	20 g
Agar	16 g

Make up to 1 liter with sterile water and autoclave. After autoclaving add 4 ml of 1% Adenine solution (1% in 0.1 M NaOH) and pour 40 ml into each sterile Omnitray (Nalge Nunc International) plate under sterile hood and let them solidify before use.

Minimal media for yeast selective plates

Medium concentrate

Yeast Nitrogen Base	8.5 g
Ammonium sulfate	25 g
Glucose	100 g
Dropout mix (see below)	7 g

Make up to 1 liter with sterile water, and sterile filter (Filter pore size 0.22 µm, Millipore).

Procedure to make selective plates:

For one liter of medium autoclave 16 g agar in 800 ml water, cool the medium to 60º to 70º C then add 200 ml medium concentrate. Depending on the required selective plates you have to add the missing amino acids and / or 3AT (3-amino-1, 2, 4-triazol).

-Trp Plates: 8,3 ml Leucine and 8,3 ml Histidine from the stock solution (see below)

-**Leu Plates:** 8,3 ml Tryptophan + 8,3 ml Histidine solution from the stock solution

-Leu-Trp Plates: 8,3 ml Histidine from the stock solution

-Leu-Trp-His plates: nothing needs to be added

-Leu-Trp-His + 3mM 3AT plates: 6 ml of 3AT (3-amino-1,2,4-triazole, from 0.5 M stock solution) to a final concentration of 3mM.

Dropout mix (-His, -Leu, -Trp)

Amino acid	**Quantity**
Methionine	1 g
Arginine	1 g
Phenylalanine	2,5 g
Lysine	3 g
Tyrosine	3 g
Isoleucine	4 g
Glutamic acid	5 g
Aspartic acid	5 g
Valine	7,5 g
Threonine	10 g
Serine	20 g
Adenine	1 g
Uracil	1 g

Mix all components and store under dry and sterile conditions.

Amino acid stock solution:
Histidine (His): Dissolve 4 g of Histidine in 1 liter sterile water and sterile filter
Leucine (Leu): Dissolve 7.2 g of Leucine in 1 liter sterile water and sterile filter
Tryptophan (Trp): Dissolve 4.8 g of Tryptophan in 1 liter sterile water and sterile filter

2.1.10 General buffers and Solutions

6 x DNA loading buffer	**2 x SDS loading buffer**
30% (v/v) Glycerol 0.01 (w/v) Bromophenolblue 0.001 (w/v) RNAse	120 mM Tris-HCl pH 6.8 4% (w/v) SDS 20% (v/v) Glycerol 100 mM DTT (freshly added) 0.01 (w/v) Bromophenolblue
SDS-PAGE Laemmli buffer	**Transfer buffer**
25 mM Tris 250 mM Glycin 0.1% (w/v) SDS	24 mM Tris 190 mM Glycin 20% (v/v) Methanol 0.05% (w/v) SDS
TBS	**Plasmid Minipreparation solution I**
20 mM Tris-HCl 150 mM NaCl 0.2% (v/v) Tween20, adjust to pH 7.6	25 mM Tris-HCl pH 8.0 10 mM EDTA pH 8.0 50 mM Glucose
Minipreparation solution II	**Minipreparation solution III**
200 mM NaOH 0.1% (w/v) SDS	3 M Potassium acetate
TE buffer	**TAE buffer**
10 mM Tris-HCl pH 8.0 1 mM EDTA pH 8.0	40 mM Tris-Acetate 1 mM EDTA pH 8.0
Overlay assay blocking buffer	
2% non-fat powdered milk 0.1% Tween-20 in phosphate buffered saline, pH 7.4	

2.1.11 Plasmids

Name	Description	Source
Treponema pallidum ORFeome	All predicted *T. pallidum* ORFs (1041) cloned into pUniD/V5-His-TOPO	McKevitt et al. 2003
pLp-GBKT7-Amp	Yeast two-hybrid bait expression plasmids (DBD fusion)	This study
pLp-GADT7	Yeast two-hybrid prey expression plasmids (AD fusion)	Clontech
pAS1	Yeast two-hybrid bait expression (Cre-loxP) plasmids (DBD fusion)	This study
pGADT7g	Yeast two-hybrid Gateway prey expression plasmids (AD fusion)	Uetz et al. 2006
pOBD2	Yeast two-hybrid bait expression plasmids (DBD fusion)	Cagney et al. 2000
pETM-30	N-His, N-GST, C-His, Protein Expression vector	G. Stier et al
pGBKT7g	Yeast two-hybrid Gateway bait expression plasmids (AD fusion)	Uetz et al. 2006
pDG148-Stu	P_{spac} regulated protein expression in *B. subtilis*	Joseph et al. 2001
pDG-yviF	HA-tagged yviF cloned into pDG148-Stu	This study

pUniD/V5-His-TOPO	Entry vector for UPS system	Invitrogen
pUniD-yviF, -hag, -yvzB	*B. subtilis* ORFs cloned into pUniD/V5-His-TOPO	This study
pMM110	GST fusion protein expression in *E. coli*	Liu et al. 2000
pMM-TP0658, -yviF	pUniD constructs recombined with pMM110	This study
pHB-HA3	HA-tagged protein expression in *E. coli*	Liu et al. 2000
pHB-HA3-TP0561, -TP0711, -TP0792, -TP0868, -TP0870, -hag, -yvzB	pUniD constructs recombined with pHB-HA3	This study
pAC28, pEGST	co-expression of proteins in *E. coli*	Kholod et al. 2001

2.1.12 Yeast strains

Name	Genotype	Source
Y187	(MATα, ura3- 52, his3- 200, ade2- 101, trp1- 901, leu2- 3, 112, gal4Δ, met–, gal80Δ,URA3::GAL1UAS -GAL1TATA -lacZ)	Harper et al. 1993
AH109	(MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80ΔLYS2::GAL1UAS-GAL1TATA-HIS3, GAL2UAS-GAL2TATA-ADE2, URA3::MEL1UAS-MEL1 TATA-lacZ)	James et al. 1996

2.1.13 Bacterial strains

strain	description	source
E. coli DH5α	Chemically/heat Competent Cells	Invitrogen
E. coli BL21 (DE3)	T7 RNA polymerase gene under the control of IPTG	Stratagene
E. coli PIR	*E. coli* R6K gamma ori	NEB
E. coli BD3.1	Competent cells suitable for propagation of plasmids containing the *ccd*B gene	Invitrogen
E. coli knock out library	BW25113 *E. coli* strains	http://ecoli.aist-nara.ac.jp
RP437	*E. coli* strains	Courtesy: Victor Sourjik
MC4100	*E. coli* strains	Wexler et al. 2000
B. subtilis 168	trpC2	Bacillus Genetic Stock Center (BGSC)
B. subtilis Δhag	argF4 flaC51 hag-1 hisA1 ura	BGSC
B. subtilis Δmot	(SPbc2) motA::Tn917 trpC2	BGSC
B. subtilis ΔflgM	flgMΔ80 pheA1 trpC2	BGSC
B. subtilis Δupp	strain used for yviF deletion	Fabret et al. 2002
B. subtilis ΔyviF	yviF deletion by phleomycin-upp cassette integration (Fabret et al., 2002)	This study
pDG148-Stu	P_{spac} regulated protein expression in *B. subtilis*	Joseph et al. 2001
pDG-yviF	HA-tagged yviF cloned into pDG148-Stu	This study

2.1.14 PCR-Primers

Name	Sequence
hagHAforward	TAC GAC GTC CCA GAC TAC GCT GTT GAC ATG GCT AAA GAG ATG AG
hagHAreverse	AGC GTA GTC TGG GAC GTC GTA TGG GTA AAT TGT GTG CTC TAG ACG ATT TTG TAC
hagN2Aforward	AGC GCT TCT GGT GAA GCT TTG ACA GCT GCT GAG
hagN2Areverse	CTC AGC AGC TGT CAA AGC TTC ACC AGA AGC GCT

yviF_forward	AAT TCT CGA GGA ATG ATC ATT CAT ACG AAG TA
yviF_reverse	AAT TGA GCT CCT AGC ATG ATT CTC CTC CAA
hag_forward	AAT TCT CGA GGA ATG AGA ATT AAC CAC AAT AT
hag_reverse	AAT TGA GCT CTT AAC GTA ATA ATT GAA GTA
yvzB_forward	AAT TCT CGA GGA ATG GAT GCG CTT ATT GAG GA
yvzB_reverse	AAT TGA GCT CTT AAC GTA ACA ATT GAA GCA
pUniD_forward	CTA TCA ACA GGT TGA ACT G
pUniD_reverse	CAG TCG AGG CTG ATA GCG AGC T
L_{59}reverse	GGT TGA GCT CAG AGG CCG CGG ATT TGG CTG C
N_{60}forward	TTG GCT CGA GGA AAC CAG GCA TCC ACC AAT GC
S_{141}reverse	GGT TGA GCT CAG GAG AAG CGG CCC GTG AGC A
V_{118}forward	TTG GCT CGA GGA GTG GCA GAG GTA GAC CGC AT
S_{141}reverse	GGT TGA GCT CAG GAG AAG CGG CCC GTG AGC A
I_{200}forward	TTG GCT CGA GGA ATC GGC ACC ATC GAT GCT GC
L_{231}forward	TTG GCT CGA GGA CTT GAC ATC GCT GCG GAG AA
S_{199}reverse	GGT TGA GCT CAG CTC TTG TTG GCC GAG TCT G
yviF_p1	CAT GGG TGT TGG AGG AAG G
yviF_p2	AGT CGA CCT GCA GGC ATG CAA GCT GTT CTT CTT TTA TG T TCA TTT GGC C
yviF_p3	CGA GCT CGA ATT CAC TGG CCG TCG TAC CAT GGC CAA ATG A AC ATA AAA GAA GAA ACA AAG CAT CCG ATT GGA GG
yviF_p4	ATC GTT TAT ATC GAC TAA GTC G
pHB-HA3_pDGforward	AAG GAG GAA GCA GGT ATG GCA GGT TAC CCA TAC GAC
pHB-HA3_pDGreverse	GAC ACG CAC GAG GTC AGT CGA GGC TGA TAG CGA GCT
TP0658_pEGSTforward	AAT TGG ATC CAT GGA GAT TCA GAC GAA GAC GC
TP0658_pEGSTreverse	AAT TCT CGA GTC AAC ATT GTT CCT GCG CCC TTC
TP0868_pAC28forward	AAT TGG ATC CAT GAT TAT CAA TCA CAA CAT GAG
TP0868_pAC28reverse	AAT TGA GCT CCG GAG AAT TGA GAG AAT CGA C
yhbE-P1	GATTTGCTATGACTTCGAAG
yhbE-P1-cont	GGAAGAACTGATGCCGGGCG
yhbE-P2	AGTCGACCTGCAGGCATGCAAGCTCTTAGAACTTCCTGATCCATT
yhbF-P3	CGAGCTCGAATTCACTGGCCGTCGCTGGTGATAAATGGATCAGGAA GTTCTAAGCCGCATTCACAGATTAAAGA
yhbF-P4	TATCCTCAACGTCGACAGAG

2.1.15 Enzymes

Name	**Supplier**
Restriction Endonucleases	Promega Mannheim and New England Biolabs, Frankfurt
DNA ligase	Promega, Mannheim
Cre recombinase	New England Biolabs, Frankfurt
BP Clonase	Invitrogen, Karlsruhe
LR Clonase	Invitrogen, Karlsruhe
Taq polymerase	Promega, Mannheim
Pfu polymerase	Promega, Mannheim
Elongase (DNA polymerase)	Invitrogen, Karlsruhe

2.1.16 Antibodies

Name	Supplier
Anti-HA, mouse monoclonal (primary antibody)	Covance, Freiburg
Anti-GST	Sigma-Aldrich, München
Goat-anti-Mouse (secondary antibody)	Biorad, München

2.2 Methods

2.2.1 Plasmid constructions

2.2.1.1 Preparation of "Ultra Competent" *E. coli* Cells

This protocol reproducibly generates competent cultures of *E. coli* that yield $1x10^8$ to $3x10^8$ transformed colonies/mg of plasmid DNA. The protocol works optimally when the bacterial culture is grown at 18°C.

Required material:

DMSO
Plasmid DNA
Inoue transformation buffer (chilled to 0°C before use)

A. Prepare 0,5M PIPES (pH6.7) (piperazine-1, 2-bis[ethanesulfonic acid]) by dissolving 15,1 g of PIPES in 80 ml sterile H_2O. Adjust the pH to 6.7 with 5 M KOH, and then add H_2O to bring 100 ml and sterile-filter the solution. Divide into aliquots and store at -20°C

B. Preparation of Inoue transformation buffer by dissolving all of the solutions listed below in 800 ml of sterile H_2O and then add 20 ml of 0.5 M PIPS (pH 6.7). Adjust the volume of the Inoue transformation buffer to 1 liter with sterile H_2O and sterile filter.

Reagent	Amount per liter	Final concentration
$MnCl_2.4H_2O$	10,88 g	55 mM
$CaCl_2.2H_2O$	2,20 g	15 mM
KCl	18,65 g	250 mM
PIPES (0.5M, pH 6.7)	20 ml	10 mM

SOB

Bacto-Tryptone	20 g
Yeast Extract	5 g
NaCl	0,5 g

Shake until the solutes have dissolved in 950 ml deionised H_2O. Add 10 ml of a 250 mM solution of KCl. Adjust the pH of the medium to 7.0 with 5 N NaOH. Adjust the volume of the solution to 1 liter with deionised H_2O. Sterilize the solution by autoclaving. Just before use, add 5 ml of a sterile solution of 2 M $MgCl_2$.

Procedure: Preparation of Competent *E. coli* Cells

- Pick a single bacterial colony from a plate that has been incubated for 16-20 hrs. at 37°C and transfer it to 25 ml of LB media
- Incubate the culture for 6-8 hrs at 37°C with vigorous shaking (250-300 rpm)
- Inoculate the starter bacterial culture to three 1-liter flasks, each containing 250 ml of SOB. The first flask receives 10 ml of starter culture, the second 4 ml and the third 2 ml
- Incubate all 3 flasks overnight at 18-22°C with moderate shaking
- The following morning, read the OD_{600} of all 3 cultures. Continue to monitor every 45 minutes
- When the OD_{600} of one of the cultures reaches 0.55 transfer the culture vessel to an ice-water bath for 10 min. Discard the other 2 cultures
- Harvest the cells by centrifugation at 2500 g (3900 rpm in a Sorvall GSA rotor) for 10 min at 4°C
- Pour off the medium and store the open centrifuge tube on a stack of paper towels for 2 min. Use a vacuum aspirator to remove any drops of remaining medium adhering to the walls of the centrifuge tube
- Resespend the cells gently in 20ml of ice-cold Inoue buffer
- Add 1,5 ml of DMSO
- Mix the bacterial suspension by swirling and store it on ice for 10 minutes
- Working quickly, dispense the aliquots (50 µl) of the suspensions into chilled, sterile microfuge tubes or 96 deep well plates. Immediately snap-freeze the competent cells by immersing the tubes in a bath of liquid nitrogen
- Store tubes at -70 °C until needed

2.2.1.2 Transformation

- Remove a tube or 96 deep well plate of competent cells form the -70 °C freezer
- Thaw the cells by keeping the tube or 96 well plate on ice
- Add the transforming DNA (up to 25 ng per 50 µl competent cells), swirl to mix and incubate on ice for 30 minutes
- Transfer the tubes into a 42°C water both or heat block for 60-90 seconds
- Rapidly transfer the tubes to ice for 1-2 minutes
- Add 600 µl of SOB medium (with out antibiotic) and shake the cultures at 37°C for 1 hrs
- Centrifuge the cells at 13000 rpm for 1 min and resuspend in 200 µl sterile H_2O
- Transfer the suspension on to a LB plate with appropriate antibiotic
- Let the plates dry in sterile hood and incubate them at 37°C
- Colonies should appear in 12-16 hrs

2.2.1.3 Small scale plasmid preparation

- Pick isolated single bacterial colonies from the LB plate and inoculate with 2 ml of LB liquid medium with appropriate antibiotic

- Incubate overnight with shaking at 37°C
- Pellet 1 ml of bacterial sample by centrifugation at 8,000 rpm for 5 minutes
- Resespend the pellet with 100 µl of miniprep solution I
- Add 200 µl lysis solution (miniprep solution II), mix gently by inverting the Eppendorf tube and incubate on ice for 5 minutes
- Add 150 µl of miniprep solution III, mix gently and incubate on ice for 10 minutes. Clear off the protein, genomic DNA complex by centrifugation at 13,000 rpm for 15 minutes at 4°C. Transfer the supernatant into fresh Eppendorf tube
- Precipitate plasmid DNA by adding two volumes of isopropanol to the supernatant and incubate at room temperature for 5 minutes
- Centrifuge at 13,000 rpm for 15 minutes at 4°C
- Wash the pellet with 70% ethanol, air dry the pellet and dissolve in 50 µl dH_2O

2.2.1.4 Plasmid preparation (96-well plasmid preparation kit)

The Montage Plasmid Miniprep 96-well kit was used for high-purity plasmid miniprep preparation. The purification was carried out according to the protocol as recommended by the supplier (Millipore, Schwalbach).

2.2.1.5 Determination of nucleic acid concentration

The concentration of DNA was determined by the spectroscopic measurement of their optical density (OD) at 260 nm and 280 nm. The OD_{260} value of one is equivalent to 50 µg/ml of double stranded DNA. Pure DNA in aqueous solution should have an OD_{260}/OD_{280} ratio of 1.6-1.8.

2.2.1.6 Restriction endonuclease digestion of DNA

Usually 2-3 units of a restriction enzyme for each µg of DNA were used. 1 to 10 µg of DNA was digested in buffer recommended by the supplier. The reaction was carried out between 2 to 4 hours at the temperature recommended by the supplier. The quality of the digestion was checked by DNA agarose gel electrophoresis.

2.2.1.7 Nucleic acid analysis by agarose gel electrophoresis

The required amount of agarose (final concentration between 0.8 to 1.5 %) was dissolved in 1X TAE buffer and boiled in a microwave oven until the agarose dissolved. Ethidium bromide was added at a concentration of 0.3 µg/ml. The molten gel was poured into a horizontal gel chamber (peqlab). Combs with the appropriate number and size of the teeth were used to make the loading slots. After the gel had solidified, the gel was immersed in 1X TAE buffer, the samples were loaded onto the gel in loading buffer. Run the gel at 50-100 V in room temperature for 1 hour. After separation of the samples DNA was visualized by transillumination with 302 nm ultraviolet radiation.

2.2.1.8 Isolation and purification of DNA from agarose gels

The DNA band of interest was isolated electrophoretically by running the gel until DNA band of interest is isolated from adjacent contaminating fragments. A Roche kit for the purification of DNA fragments from agarose gel was used to isolate the DNA of interest from the agarose gel.

The purification was carried out according to the protocol recommended by the supplier (Roche, Mannheim).

2.2.1.9 Ligation

In all cases, the insert and vector were loaded on an agarose gel to check the DNA content before ligation. Ligation was performed in a total volume of 20 µl with an insert to vector ratio of 4:1 in 1x ligation buffer with one unit of T4 DNA ligase and incubated at 22°C overnight for sticky end ligation, and 16°C for overnight for blunt end ligation.

2.2.1.10 Polymerase chain reaction (PCR)

Usually all PCRs were performed in a total volume of 50 µl, 10 ng template DNA, 200 µM dNTPs, 0.2 µM of each each primer, 1x reaction buffer with one unit of GoTaq DNA polymerase or pfu DNA polymerase. The reaction was carried out in a MJ thermocycler, using specific cycling parameters depending on the application.

PCR conditions for amplification of ORFs from plasmid DNA

1	µl	Plasmid DNA (20-40 ng)
1	µl	dNTP mix (each 10 mM)
10	µl	5x Pfu Polymerase Buffer
1,5	µl	$MgCl_2$ (50mM)
1	µl	Forward Primer (10 pmol)
1	µl	Reverse Primer (10 pmol)
0,5	µl	Pfu Polymerase (1 unit)
34	µl	Water
50	µl	Total

PCR programme:

2 min 95°C

(60 sec 94°C
1 min 54°C…
2 min 72°C) x 30

10 min 72°C

Hold on 4°C

2.2.1.11 Yeast colony PCR

This protocol allows rapid detection of transformation success when primers are available to allow determination of correct gene products by size.

Prepare PCR master mix as described above, instead of genomic or plasmid DNA, select isolated yeast colonies to pick. Pick the colonies using sterile pipette tip and mix in a correspondingly labelled PCR reaction tube. Making sure to begin your PCR protocol with an extended time at 95° for 10 minutes.

2.2.2 Construction of *T. pallidum* Y2H prey and baits expression clones

2.2.2.1 Univector Plasmid fusion System:

We used the Univector Plasmid fusion System (UPS) for the construction of *T. pallidum*, yeast two-hybrid bait and prey expression plasmids. The UPS uses Cre–*lox* site-specific recombination to catalyze plasmid fusion between the univector a plasmid containing the gene of interest and host vectors containing regulatory information. Fusion events are genetically selected and place the gene under the control of new regulatory elements (Liu et al., 1998).

In order to generate the *T. pallidum* Y2H library, I used the *T. pallidum* ORFeome library (pUniD) as donor univector and pLp-GADT7 (Clontech) as prey, and pLp-GBKT7-Amp as Y2H bait expression vectors.

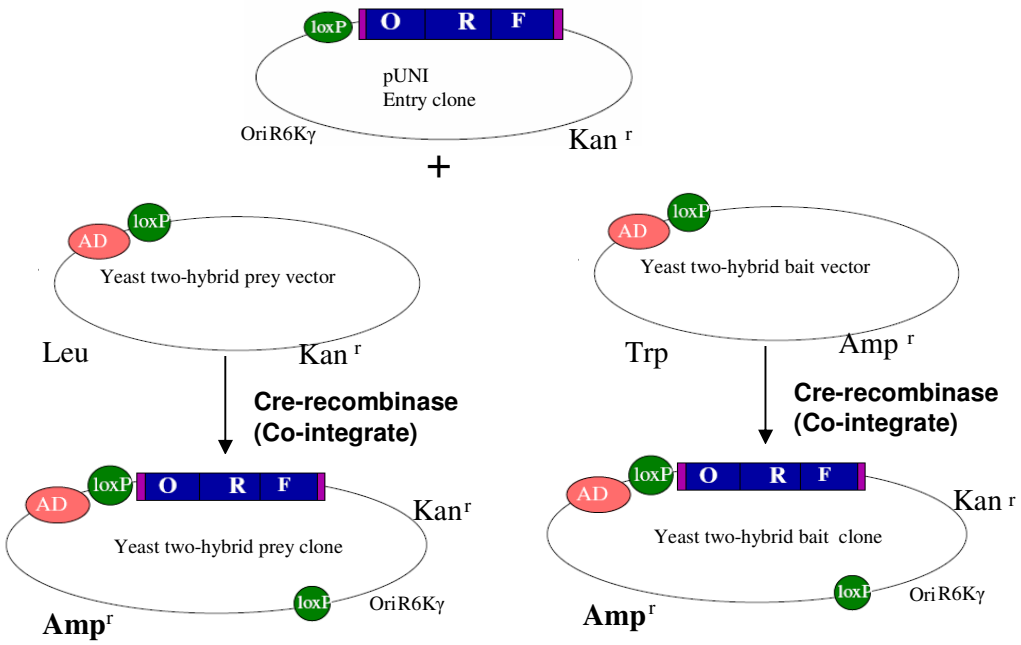


Figure 2.1 Univector plasmid fusion system. Schematic illustration of Cre-*loxP*–mediated site-specific recombination cloning of the *T. pallidum* ORFs form the pUNI vector. (McKevitt et al., 2003) into yeast two-hybrid expression vectors. Cre-*loxP*–mediated site-specific recombination fuses the pUNI and yeast two-hybrid expression plasmids, pLp-GADT7 (clontech) and pLp-GBKT7-Amp (Titz et al, unpublished) at the *loxP* site. As a result, the gene of interest is placed under the control of the two-hybrid expression plasmid promoter and fused to Gal4-AD and Gal4-BD domains in the two-hybrid prey and bait expression vectors, respectively.

2.2.2.2 Cre Recombination

Sample	Volume
10X Cre recombinase buffer	2 µl
Cre recombinase (NEB)	1 µl
pUniD (entry vector) 100 ng	--
Y2H expression vector (pLp-GADT7) 100 ng	--
Sterile water	--
Total volume	20 µl

Incubate the reaction at 37° C for 1 hour. Then to terminate the reaction incubate the reaction at 70° for 10 minutes. Use 10 µ of the reaction for transformation in to competent *E. coli* cells

2.2.2.3 Yeast Transformation for bait and prey construction

This protocol is suitable for 100 yeast transformations, and may be scaled up or down as needed. Selection of the transformed yeast cells requires leucine or tryptophan-free media ("-Leu" or "-Trp", depending on the selective marker on the plasmid). Moreover, at least one of the haploid strains must contain a two-hybrid reporter gene under GAL4 control.

Materials required:

- Salmon sperm DNA
- DMSO
- Competent host yeast strains, e.g. AH109 (for baits), and Y187 (for preys)
- Lithium Acetate (0.1M)
- Selective plates (depending on the selective markers)
- 96PEG solution

Carrier DNA (salmon sperm DNA): Dissolve 7.75 mg/ml salmon sperm DNA (e.g. Sigma D1626) in water and store at –20°C following a 15 min 121°C autoclave cycle.

Preparation of 96 PEG solution (100 ml): mix 45.6 g PEG (Sigma P3640), 6.1 ml of 2 M LiOAc (Lithium acetate), 1.14 ml of 1 M Tris pH 7.5 and 232 µl 0.5 M EDTA; make up to 100 ml with sterile water and autoclave.

Preparation of competent yeast cells:

- Inoculate 50 ml YEPD liquid medium with 200 µl liquid stock of yeast strains (e.g. AH109, Y187 or any other appropriate yeast strain; we use Y187 strains for preys and AH109 for baits) in a 250 ml flask and grow overnight with shaking at 30 ºC (minimum 15 h, max. 24 h)
- Spin out cells in 50 ml conical tube (3500 rpm, 5 min at room temperature), pour off supernatant and dissolve the pellet by adding 2 ml LiOAc (0.1M) and transfer resuspended yeast to two 1.5 ml microfuge tubes. Spin out yeast and resuspend in a total volume of 1.8 ml LiOAc (0.1M).

Preparation of "CT110" for yeast transformation

Materials required:

20.73 ml 96PEG
0.58 ml boiled salmon sperm DNA (boil frozen salmon sperm DNA at 95ºC for 5 min and cool on ice before use)
2.62 ml DMSO
20 ng prey or bait plasmid constructs

- Mix the above listed solutions in a 50 ml Falcon tube; add DMSO last and mix quickly after adding by shaking hard and vortex for 30 sec.
- Add all the competent yeast cells prepared above and mix hard by hand or by vortexing for 1 minute. Immediately pipette 245 µl into each of 96 wells of a 96-well dish (e.g. Costar 3596).
- Now add 50-100 ng of plasmid or 5 µl of PCR products (in case of co-transformation and homologous recombination in yeast) and positive control (empty vector) and negative

control (only CT110). Seal the 96 well plate with plastic or aluminium tape and vortex for 4 minutes.

- Incubate at 42°C for 30 minutes.
- Spin the 96 well plate for 10 min at 2000 rpm; discard the supernatant and aspirate with 8 channel wand or by tapping on cotton napkin a couple of times. Add 150 µl of sterile water to all 96 wells, resuspend and plate them on selective plates (35 mm) with -Leu (or -Trp) agar. Incubate the plates at 30ºC for 2-3 days. After one day the colonies start to appear; pick colonies after 2-3 days and make glycerol (20%) stocks (-80°C).

2.2.2.4 The *T. pallidum* Y2H prey array construction

The *Treponema pallidum* prey clones were kindly supplied by Tim Palzkill (Baylor College of Medicine, Houston, USA). After transformation of the prey clones into haploid yeast cells (Y187), a couple of yeast prey colonies were picked from the selective plate (-Leucin), and grown in 96 well YEPD liquid medium over night. The template prey array was made by pinning the overnight culture on to a single-well microtiter plate (e.g., OmniTray; Nalge Nunc) containing solid YEPD medium lacking Leucin, in 96 well format using the sterile replicator (HRD pinning head of Biomek 2000, Beckman Coulter). This way ~965 prey clones (out of 1041 ORFs) of *T. pallidum* were arrayed on 11 single well plates containing 96 clones per plate (Figure 2.2). In parallel glycerol (20% glycerol) stock were made for longer storage.

2.2.2.5 Preparing prey array for screening

Using the sterile replicator of a robotic workstation (HRD pinning head of Biomek 2000, Beckman Coulter) we transferred the yeast prey array from selective plates to single-well microtiter plates containing solid YEPD medium and grow the array overnight in a 30ºC incubator. We converted the 96 well prey arrays in to 384 well arrays by quadruplicating the 96 well prey arrays (96*4 =384). So each prey clone is represented 4 times in an array. The quadruplicated array was used for the Y2H screening (Figure 2.3).

Note: The template prey array should be kept on selective plates, and back up copies should be made every 4 to 6 weeks. Before Y2H screening the prey array should be grown on YEPD solid medium. The prey yeast clones tend to loose the plasmid when the array is stored longer on rich medium (YPED); this may lead to loss of diploids with both plasmids in subsequent matings. Ideally, a new prey copy on YEPD should be made from the template prey array every 3-4 weeks.

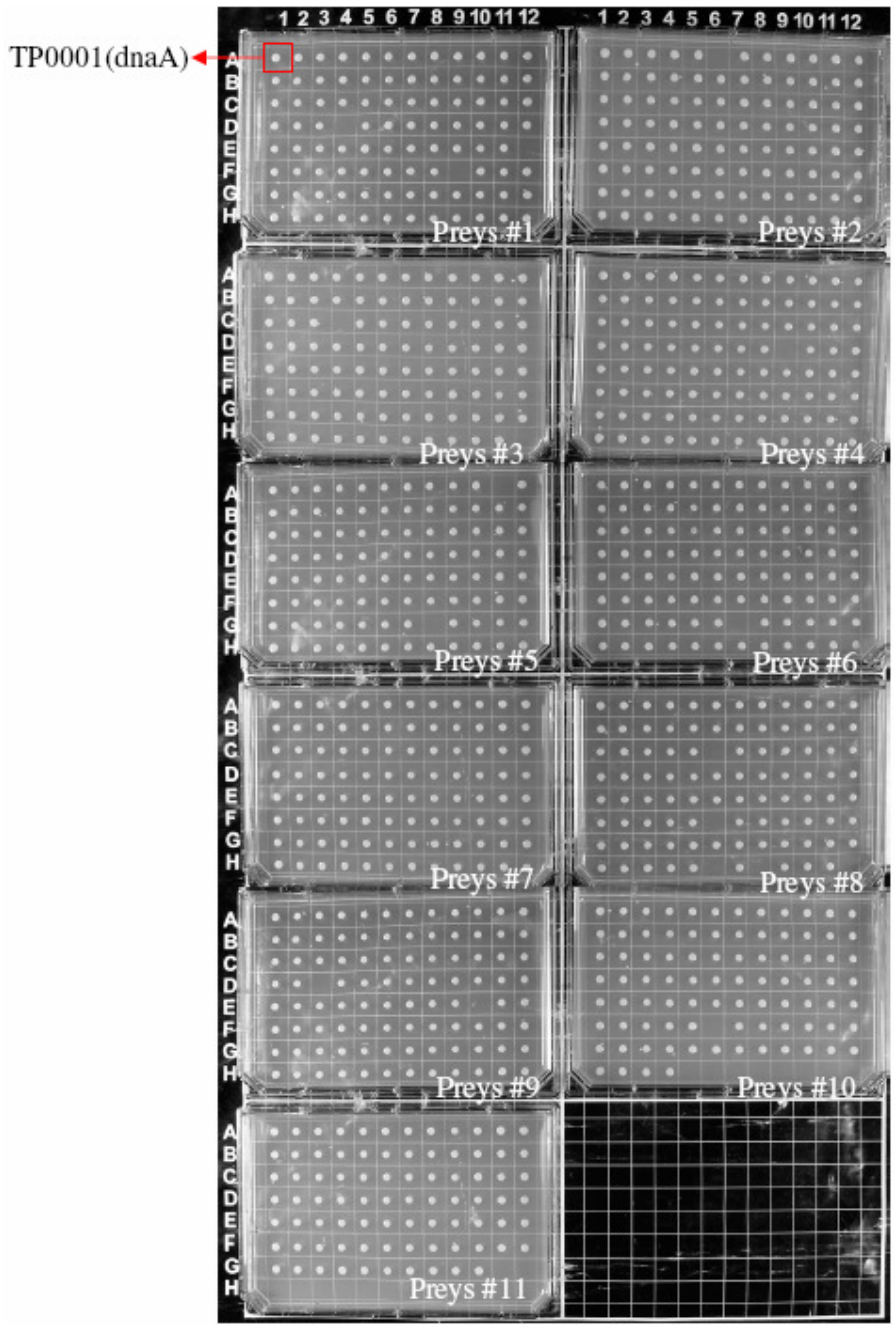


Figure 2.2 The *T. pallidum* prey array: All the Y2H prey clones of *T. pallidum* were arrayed on 11 single-well omnitray selective plate (-Leu). Each yeast colony contains a unique *T. pallidum* prey clone.

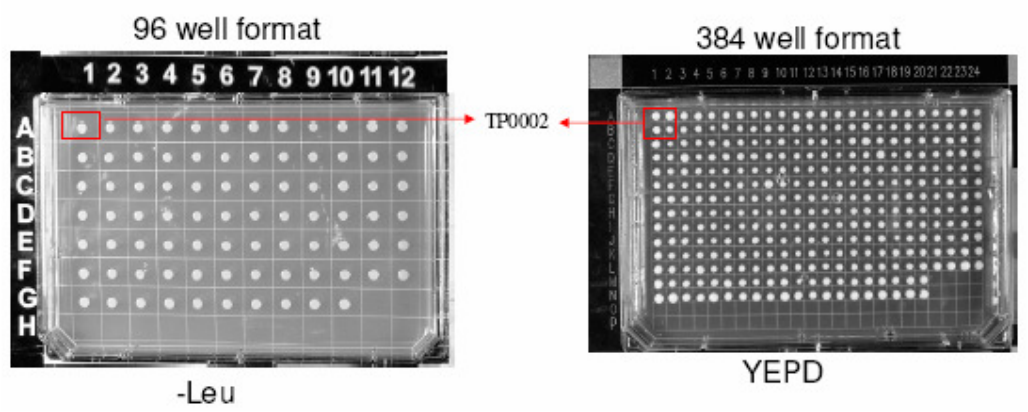


Figure 2.3 The conversion of 96 well format Y2H prey array to 384 well formats. A, *T. pallidum* preys plate 11 in 96 well format on selective plate (-Leu). B, The *T. pallidum* preys as 384 well format on YEPD medium, after conversion from 96 well formats to 384 well formats (as quadruplicates).

2.2.3 Yeast two-hybrid Screening

2.2.3.1 Semi-automated Yeast Two-hybrid screening

Materials required

- 20% (v/v) bleach (1 % Sodium hypochlorite)
- 95% (v/v) ethanol
- Single-well microtiter plate (e.g., OmniTray; Nalge Nunc) containing solid YEPD + Adenine medium, -Leu –Trp, –His –Leu –Trp and –His –Leu –Trp + different concentrations of 3AT.
- 384-Pin Replicator for manual screening (Nalge Nunc International) or Robot with HDR tool.
- Bait liquid culture (DBD fusion-expressing yeast strain)
- Yeast prey array on solid YEPD plates

Screening Procedure

Sterilization:

Sterilize a 384-Pin Replicator by dipping the pins into 20% bleach for 20 s, sterile water for 1 sec, 95 % ethanol for 20 s, and sterile water again for 1 s. Repeat this sterilization after each transfer.

Preparing prey array for screening:

Use the sterile replicator to transfer the yeast prey array from selective plates to single-well microtiter plates containing solid YEPD medium and grow the array overnight in 30ºC incubator.
If duplicate or quadruplicate colonies were not used to construct the array (see array construction), the entire experiment should be done using duplicate arrays
Ideally, the template prey array should be kept on selective plates.

Preparing bait liquid culture (DBD fusion-expressing yeast strain):
Inoculate 20 ml of liquid YEPD medium in a 250 ml conical flask with a bait strain and grow overnight in 30ºC shaker.
If the bait strains are frozen, it is streaked or pinned on a selective plate (-Trp) solid medium and grown 1-2 days at 30ºC. Baits from this plate are then used to inoculate the liquid YEPD medium.

Mating procedure:

- Pour the overnight liquid bait culture into sterile Omnitray plate. Dip the sterilized pins of the pin-replicator (thick pins should be used to pin baits) into the bait liquid culture and place directly onto a fresh single-well microtiter plate containing solid YEPD media. Repeat with the required number of plates and allow the yeast to dry onto the plates for 10-20 minutes
- Pick up the prey array (i.e., AD) yeast colonies with sterilized pins (thin pins should be used to pin the preys) and transfer them directly onto the baits pinned onto the YEPD plate, so that each of the 384 bait spots per plate receives different prey yeast cells (i.e., a different AD fusion protein). Incubate 1-2 days at 30ºC to allow mating.

Mating will take place in <15 hr, but a longer period is recommended, because some bait strains show poor mating efficiency. Adding adenine to the bait culture before mating increases mating efficiency of some baits.

Selection of Diploids:

For the selection of diploids, transfer the colonies from YEPD mating plates to single-well microtiter plates containing -Leu-Trp medium using the sterilized pinning tool (thin pins should be used in this step). Grow for 2-3 days at 30ºC until the colonies are >1 mm in diameter

This step is an essential control step because only diploid cells containing the Leu2 and Trp1 markers on the prey and bait vectors, respectively, will grow on this medium. This step also helps recovery of the colonies and increases the efficiency of the next selection step.

Interaction selection:

Transfer the colonies from –Leu-Trp plates to a single-well microtiter plate containing solid –His –Leu –Trp agar, using the sterilized pinning tool. If the baits are self-activating, they have to be transferred to –His –Leu –Trp + a specific concentration of 3AT (see bait self activation test). Incubate at 30ºC for 6-10 days.

Score the interactions by looking for growing colonies that are significantly above background by size and that are present as quadruplicate colonies.

The plates should be examined every day. Most two-hybrid positive colonies appear within 3 to 5 days, but occasionally positive interactions can be observed later. Very small colonies are usually designated as background; however, there is no absolute measure to distinguish between the background and real positives. When there are many (i.e., > 30) large colonies per array of

6000 positions, we consider these baits as "random" activators. In this case the screen should be repeated.
Scoring can be done manually or using automated image analysis procedures. When using image analysis, care must be taken not to score contaminated colonies as positives.

2.2.3.2 Bait self activation test

The aim of this test is to measure the background reporter activation (here: HIS3) by bait proteins in absence of an interacting prey protein. Prior to the two-hybrid screening, the bait yeast strains should be examined for self-activation. Self-activation is defined as a detectable bait-dependent reporter gene activation in the absence of any prey interaction partner. Weak to intermediate-strength self-activator baits can be used in two-hybrid array screens because the corresponding bait-prey interactions confer stronger signals than the self-activation background. In case of the HIS3 reporter gene the self-activation background can be titrated by adding 3-AT, a competitive inhibitor of HIS3. Self-activation of all the baits is examined on plates containing different concentrations of 3-AT. The lowest concentration of 3-AT that suppresses growth in this test is used for the interaction screen (see below) because it avoids background growth whereas true interactions are still detected.

Materials required:

- full medium and selective media agar in single-well microtiter plates (Omnitray plates, Nunc)
- YEPD plates
- -LT plates
- Selective plates without Trp, Leu, and His, but with different concentrations of 3-AT, e.g., 0 mM, 1 mM, 3 mM, 10 mM, 50 mM and 100 mM (-LTH/3-AT plates)
- Prey strain carrying the empty prey plasmid, e.g., Y187 strain with pLP-GADT7 plasmid (Clontech)

Procedure:

1. Bait strains are arrayed onto a single-well Omnitray agar plate; either the standard 96-spot format or the 384-spot format is used.

Baits are first inoculated at the different positions of a 96-well plate as liquid culture, and then cells are transferred (manually or with robot) to solid agar single-well plates (Omnitray plates). In this step the 96-well format was converted into the 384-well format, this will position each bait in quadruplicates on the 384-well formatted plate. Full media agar (YEPD agar) can be used, however, for long term storage of the array selective agar (-Trp) is suggested to prevent loss of plasmids.

2. The arrayed bait strains are mated with a prey strain carrying the empty prey plasmid, e.g., Y187 strain with pLP-GADT7 plasmid (Clontech). Mating is conducted according to the standard Y2H screening protocol (described above).

Note: Compared to the Y2H screening protocol bait and prey strains are exchanged during mating, i.e. preys are pinned first and then an array of baits is pinned on top of them.

3. After selecting for diploid yeast cells (on –LT agar) the cells are transferred to media selecting for the His3p reporter gene activity. The -LTH transfer is done to several selective plates with increasing concentrations of the competitive inhibitor of His3p, 3-Aminotriazole (3-AT).

Suggested are 3-AT concentrations of 0, 1, 3, 10, 25, 50 and 100 mM.

4. These –LTH/3-AT plates are incubated for 1 week at 30°C. The self-activation level of each bait is assessed: the lowest 3-AT concentration that completely prevents colony growth is noted (Figure 2.4). As this concentration of 3-AT suppresses reporter activation in absence of an interacting prey this 3-AT concentration is added to –LTH plates in the actual interaction screen (Y2H screening protocol).

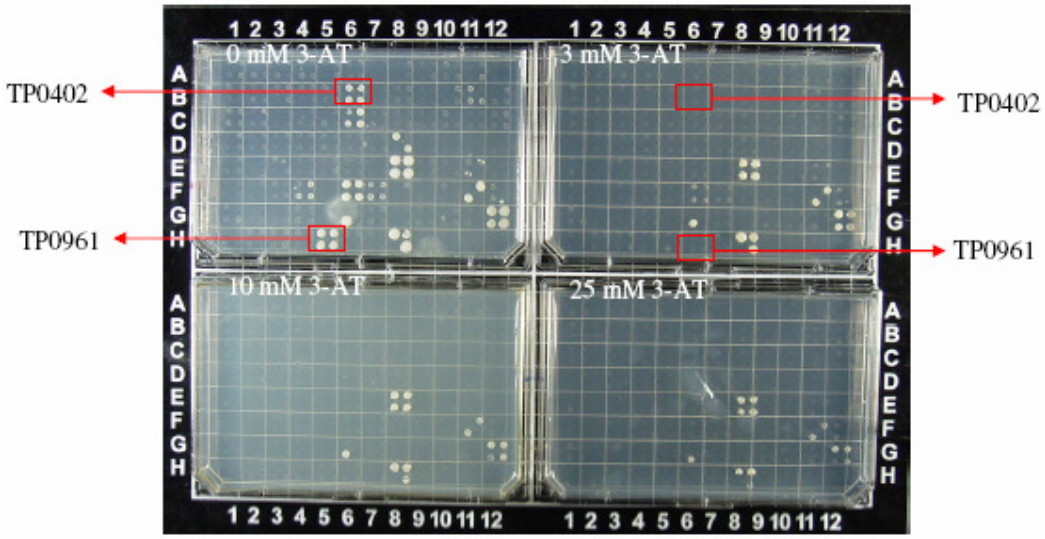


Figure 2.4 Self activation test of flagellum baits: Flagellum baits (in pLP-GBKT7-amp and pAS1) in AH109 yeast strains are screened with empty prey plasmid, here: Y187 strain with pLP-GADT7 plasmid for self activation test. After mating the diploids were pinned on different concentration of 3-AT. The growth of yeast on –LTH plates indicates self activation. The lowest concentration of 3-AT that suppresses growth in this test is noted and the same concentration of 3-AT is added in actual interaction screens. For example, TP0402 self activates on 0 mM 3-AT and the self activation can be suppressed by titrating with 3 mM 3-AT. So –LTH plates with 3 mM 3-AT plates should be used for the interaction screening of TP0402. Similarly, the activation strength of all the baits was noted.

2.2.3.3 Filtering raw Y2H data

As with many assay systems, the two-hybrid system has the potential to produce false positives (that is, reporter gene activity and subsequent colony growth where no specific protein-protein interaction is involved). Frequently, such false positives are associated with bait proteins that act as transcriptional activators. False positives may also be caused by proteins that have the propensity to take part in non-specific interactions (for largely unknown reasons). Filtering of raw Y2H results significantly improves the data quality of the protein-protein interactions. For filtering the raw Y2H data I have considered three parameters. First, protein interactions that can not be reproduced should be discarded. Second, for each prey the number of different interacting

baits is calculated. Preys interacting with a large number of baits –are assumed to be unspecific and thus can be discarded. However, the concrete cut-off number depends also on the nature of baits that are screened: if a large family of proteins is screened it is not surprising that many of them find the same bait. As a rough guideline the number of baits interacting with a certain prey should not be larger than 5% of the bait number. The third parameter is the background activation activity of the tested bait. The activation strength of interaction pairs must be significantly higher than with all other (background) pairs. In principle, at least with the HIS3 reporter, no activation (no colony growth) should be observed in non-interacting pairs (Figure 3.1).

Table 2.1 Raw Y2H interactions of fliS

List of raw Y2H interactions of fliS bait. **Score:** number of positive colonies out of 4 independent Y2H tests. **Prey count**: number of different interacting baits.

Prey	Description	Score	Prey count	Confidence score
TP0060	ribosomal protein L9 (rplI)	2 of 4	36	false positives
TP0961	flagellar basal-body rod protein (flgG)	4 of 4	78	false positives
TP0974	hypothetical protein	4 of 4	11	high confidence
TP0989	P26	2 of 4	100	false positives
TP0993	rare lipoprotein A, putative	3 of 4	108	false positives
TP0868	flagellar filament 34.5 kDa core protein (flaB1)	4 of 4	1	high confidence
TP0792	flagellar filament 33 kDa core protein (flaB2)	4 of 4	2	high confidence
TP0048	conserved hypothetical protein	4 of 4	1	high confidence
TP0870	flagellar filament 31 kDa core protein (flaB3)	4 of 4	1	high confidence
TP0943	flagellar protein (fliS)	4 of 4	15	high confidence
TP0197	ribosomal protein L29 (rpmC) _	3 of 4	10	high confidence
TP0209	ribosomal protein L36 (rpmJ)	3 of 4	24	high confidence
TP0258	conserved hypothetical protein	4 of 4	82	false positives
TP0277	carboxyl-terminal protease (ctp)	1 of 4	1	low confidence
TP0383	conserved hypothetical protein	3 of 4	20	high confidence
TP0398	flagellar hook-basal body complex protein (fliE)	3 of 4	41	false positives
TP0660	flagellar hook-associated protein 1 (flgK)	4 of 4	3	high confidence
TP0661	hypothetical protein	3 of 4	51	false positives
TP0788	hypothetical protein	4 of 4	48	false positives
TP0907	conserved hypothetical protein	4 of 4	64	false positives
TP0396	flagellar basal-body rod protein (flgB)	1 of 4	18	low confidence
TP0945	ribulose-phosphate 3-epimerase (cfxE)	3 of 4	6	high confidence

2.2.3.4 An example of filtering raw Y2H data

Raw Y2H interactions of flagellum protein fliS are listed in Table 2.1. The prey count listed in the table is the result of the Y2H screening of ~500 *T. pallidum* baits (including selected flagellum baits) with its prey library. We considered preys interacting with more than 25 baits (out of ~500 baits screened) as false positives (these are sticky preys unspecifically binding to many baits). We categorised the Y2H interactions into high confidence and low confidence

interactions based on the reproducibility of interactions. High confidence interactions are those that show a positive signal at least 3 times out of 4 independent Y2H tests. Low confidence interactions are those show positive signal 1 to 2 times out of 4 independent Y2H tests. For example, Y2H screening of flagella protein fliS yielded 22 interactions (Table 2.1). After filtering the raw data yielded 11 high confidence and 2 low confidence interactions (Table 2.1). 9 interactions are considered as false positives because of high prey count (>25).

2.2.4 Protein expression and purification in *E. coli*

Required material:

- BL21 (DE3) *E. coli* cells containing the respective clone (pMM110-GST plasmid)
- LB liquid medium with appropriate antibiotic
- IPTG (Stock solution : 1M in dH_2O)
- Reduced Glutathione (10 mM in 50 mM Tris-HCl pH 8.0)
- Glutathion Sepharose 4 Fast Flow (Amersham Biosciences)

20X KPE Solution

Chemical	Amount
KH_2PO_4	46 g
K_2HPO_4	243 g

Dissolve in 1 liter sterile water and sterile filter

1 X SB medium

Bactotryptone	12 g
Yeast Extract	24 g
Glycerol 100%	4 ml

Make up to 950 ml with sterile water, autoclave, and cool to 50° C; add 50 ml of 20X KPE

Lysis buffer:

50 mM Tris-HCL, pH 8.0

0.3 M NaCl

1 mM EDTA

Lysis buffer and Brij-Lysozyme (50 ml)

45 ml Lysis buffer
5 ml 1% brij58 (Sigma)
100 µg/ml Lysozyme (Sigma)

Proteinase inhibitors

PMSF (200 mM stock in isopropanol)→ Dilute 1:200
Aprotinin + Antipain (5µg/ml stock in DMSO) → Dilute 1:5000

Elution buffer (for GST purification)

120 mM	NaCl
100 mM	Tris/HCl pH 8.0
1 mM	PMSF
20 mM	Glutathion (reduced)
5 mM	DTT

Procedure:

- Inoculate 1 liter of SB medium with appropriate antibiotic with pre-culture of *E. coli* BL21 (DE3) cells transformed with expression plasmid with protein of interest as GST fusion. For example, pMM110 vector with TP0658-GST.
- Shake at 37°C until the cells density reach OD_{600} of ~0,6
- Induce fusion protein expression by adding 0,5 mM IPTG (final concentration)
- After induction continue growth by shaking 4 to 6 hrs at 30° C
- Centrifuge bacteria at 3500 rpm for 10 minutes (pellet can be frozen at -20)
- Resuspend pellets in ice cold 50 ml lysis buffer with protease inhibitor, and incubate on ice for 30 minutes
- Lyse the cells on ice using a probe sonicator: (2 x 10 pulses hold 4, 40 volts)
- Centrifuge at 12 000 rpm at 4 °C for 10 minutes
- To the 50 ml probe add 1-2 ml of washed sepharose beads
- Rotate 1 h at 4°C
- Wash beads 3x with PBS
- Elute the GST fusion protein with 10 mM reduced glutathione.

2.2.5 SDS PAGE

SDS Polyacrylamide gel (for 2 gels 10 cm x 8 cm with 1mm spacers)

Required material:

- 1 M Tris-HCl pH 8.8 (Roth 3029.1)
- 1 M Tris-HCl pH 6.8
- 0,5 M EDTA
- 20 % SDS
- 30 % Acrylamide solution (Roth)
- 10 % Ammonium persulfate (APS)
- TEMED (Sigma)
- Gel Chambers (2 gels 10 cm x 8 cm with 1mm spacers)
- Protein molecular weight marker

Procedure:

Resolution Gel (15% Acrylamide)

H_2O	3,33 ml
Tris-HCl pH 8.8	2,53 ml
Acrylamide	4 ml
SDS 20%	50 µl
Temed	13,33 µl
APS (10%)	50 µl

Pour into Gel chamber (so that it fills up 2/3 of the chamber) and cover with water using a pipette (for the removal of air bubbles) and let solidify.

Stacking Gel

H_2O	4,35 ml
Tris-HCl pH 6.8	750 µl
Acrylamide	900 µl
SDS 20%	30 µl
Temed	11, 25 µl

Remove water and pour stacking gel. Immediately insert the combs and let solidify

Coomassie Blue Staining of Acrylamide Gels

Required Material:

Coomassie Blue stain:

Methanol	50 %
Acetic acid	10 %
Coomassie Brilliant Blue R250	0.2 %

Destain solution:

Methanol	30 %
Acetic acid	10 %

Procedure:

- After separation of protein samples, incubate the gel in Coomassie Blue stain for 30 minutes
- Recover Coomassie stain (can be used several times)
- Incubate the Coomassie-stained gel in destain solution (change several times the solution) until the background is sufficiently reduced
- Rinse in the gel in water and dry the gel for longer storage

2.2.6 Western Blot Transfer (semi-dry)

Required material:

- Transfer buffer (see buffers of protocol)
- Ponceau S (commercial)
- Blot Chamber
- Whatman paper
- Nitrocellulose membrane (Immobilon-P, Millipore)

Procedure:

- Cut blot paper (8 sheets of Whatman paper per gel) to the size of your gel and equilibrate in transfer buffer
- Cut blotting membrane to the size of your gel, and activate in methanol for 2 minutes (in case of PVDF membrane), then equilibrate in transfer buffer
- Disassemble electrophoresis chamber, lift short plate form saucer plate and carefully transfer gel (cut off stacking gel) into a tray with transfer buffer, shortly equilibrate
- Build up blot: To bottom platinum anode of Trans-Blot-Semi-Dry-place:
  Pre-wet filter paper (4 sheets of Whatman paper)
  Pre-wet membrane
  Gel
  Pre-wet filter paper (4 sheets of Whatman paper)
  Roll out air bubbles
- Secure safety cover and connect to power supply
- Run 1 gel at 15V for 60 min, 2 gels at 20 V for 60 min
- Stop transfer, discard filter paper and briefly wash blot in dH_2O
- Check transfer by staining blot in Ponceau S solution and destain in dH_2O

2.2.7 Overlay assay

Blot overlays are a standard and very useful method for studying interactions between proteins in vitro. In principle, a blot overlay is similar to a Western blot. For both procedures, samples are run on SDS-PAGE gels, transferred to nitrocellulose or PVDF, and then overlaid with a soluble protein that may bind to one or more immobilized proteins on the blot. In the case of a Western blot, the overlaid protein is antibody. In the case of a blot overlay, the overlaid protein is a probe of interest, often a fusion protein that is easy to detect (in this study GST, HA and Myc fusion protein were used).

2.2.7.1 Experimental Procedure:

SDS-PAGE and blotting of protein sample: The purpose of this step is to immobilize the samples of interest on nitrocellulose or PVDF membrane. It is very important to keep the blot clean during the handling steps involved in the transfer procedure, since contaminants can contribute to increased background problems later on during detection of the overlaid probe.

- Place gel in SDS-PAGE apparatus and fill chamber with 1X running buffer.

- Mix purified or total cell lysate of HA-tagged fusion proteins with SDS-PAGE sample buffer to a final concentration of approximately 0.1 µg/µl of fusion protein.
- Load 20 µl of fusion protein (2 µg total) in each lane of the gel and molecular weight markers.
- Run gel for approximately 80 min at 150 V using the power supply.
- Stop gel, turn off the power supply, remove the gel from its protective casing, and place in transfer buffer.
- Activate the PVDF membrane i.e. cut the required size of membrane and place in the tray containing MeOH for 2 minutes, then transfer the membrane to sterile water 1 minute and to transfer buffer to equilibrate.
- Put PVDF membrane and gel together in transfer apparatus, and transfer proteins from gel to PVDF membrane using a power supply for 60 min at 25 V.

Overlay: During the overlay step, the probe is incubated with the blot and unbound probe is then washed away.

Note: GST and hexahistidine-tagged fusion proteins should be purified as extensively as possible. If the probe has many contaminants, this may contribute to increasing the background during the detection step, making visualization of the specifically bound probe more difficult.

- Block blot in blocking buffer (see buffers) for at least 30 minutes
- Add GST or His fusion protein of interest to a concentration of 25 nM in 10 ml blocking buffer.
- Incubate purified GST fusion proteins (for example GST-TP0658) with blot for 1 hr at room temperature while rocking slowly.
- Discard GST fusion protein solution and wash blot three times for 5 minutes each with 10 ml of blocking buffer while rocking slowly.
- Add anti-GST antibody at 1:1000 dilution (approximately 200 ng/ml final) to 10 ml blocking buffer.
- Incubate anti-GST antibody with blot for 1 hr while rocking slowly.
- Discard anti-GST antibody solution and wash blot three times for 5 minutes each with 10 ml of blocking buffer while rocking slowly.
- Add goat anti-mouse HRP-coupled secondary antibody at 1:2000 dilution to 10 ml blocking buffer.
- Incubate secondary antibody with blot for 1 hr while rocking slowly.
- Discard secondary antibody solution and wash blot three times for 5 minutes each with 10 ml of blocking buffer while rocking slowly.
- Wash blot one time for 5 minutes with phosphate buffered saline, pH 7.4.

2.2.7.2 Detection of overlaid proteins:

The final step of the overlay is to detect the probe that is bound specifically to proteins immobilized on the blot. In viewing different exposures of the visualized probe, an effort should be made to obtain the best possible signal-to-noise ratio.

- Incubate blot with enhanced chemiluminescence's solution for 60 seconds.
- Remove excess ECL solution from blot and place blot in clear plastic sheet.
- Tape sheet into autoradiography cassette.

- Move to darkroom and place one sheet of film into autoradiography cassette with blot.
- Expose film for 5-2000 seconds, depending on intensity of signal and develop film in standard film developer.

Notes: *In my experience, increased concentration of milk in blocking buffer from 1% to 4% helps to reduce the background signal.Blocking the blot in blocking buffer over night at 4° C reduce unspecific binding.Non-specific background signal will increase linearly with time of exposure of film. Thus, shorter exposures may have more favourable signal-to-noise ratios.*

2.2.8 Peptide competitive inhibition assay

- The HA-tagged flagellin proteins were transferred onto the PVDF membrane.
- Peptides for competition experiments were synthesized using standard peptide chemistry protocols (courtesy of Dr. Olaf Zwernemann).
- The purified GST-TP0658 was pre-incubated with different concentrations of the peptide VGLDIAAENLQAAESRIRD (0 mM, 1 mM, 10 mM and 100 mM) or a control peptide (DRRLADHFLGKI) in overlay buffer for 2h at 4°C.
- The overlay assay was carried out (see overlay assay).

2.2.9 *Bacillus subtilis* genomic DNA isolation

- Inoculate isolated bacterial colony, in to 3 ml LB (with antibiotic) grow them at 37° C overnight.
- Spin the cells; Re-suspend in 500 µl TE; add lysozyme 50 µl (1mg/ml stock), and incubate at room temperature for 30 minutes.
- Add 25 µl SDS (20%); vortex.
- Add equal volume to phenol-chloroform extract (Phenol:chloroform:iso-amyl alcohol = 22:24:1).
- Centrifuge the sample at 13000 rpm for 5 minutes.
- Transfer the upper phase in to new fresh Eppendorf tube (repeat the phenol extraction step one more time).
- Precipitate the genomic DNA by adding equal volume of iso-propanal with Na OAC (0.3M).
- Centrifuge at 4° C for 10 minutes.
- Wash the pellet with 80% ethanol, air dry and dissolve in 200 µl sterile water.

2.2.10 Gene deletion in *B. subtilis* and *E. coli*

In *Bacillus subtilis* gene deletion is achieved by the integration of a phleomycin-cassette as described by Fabret *et al.*(Fabret et al., 2002). This novel approach relies on the use of *upp*, which encodes uracil phosphoribosyl-transferase, as a counter-selectable marker. For example, the yviF gene deletion was constructed using yviF_p1/yviF_p2 and yviF_p3/yviF_p4 primers to PCR-amplify 1.5 kb the flanking regions of the yviF gene (Figure 2.5).

Phleomycin-cassette DNA was mixed with both PCR products and subjected to a joining PCR reaction using the primer pair yviF_p1/yviF_p4. *B. subtilis* 168 Δupp competent cells were transformed and selected for phleomycin resistance. The presence of the phleomycin-cassette at the correct locus in the chromosome was checked by PCR.

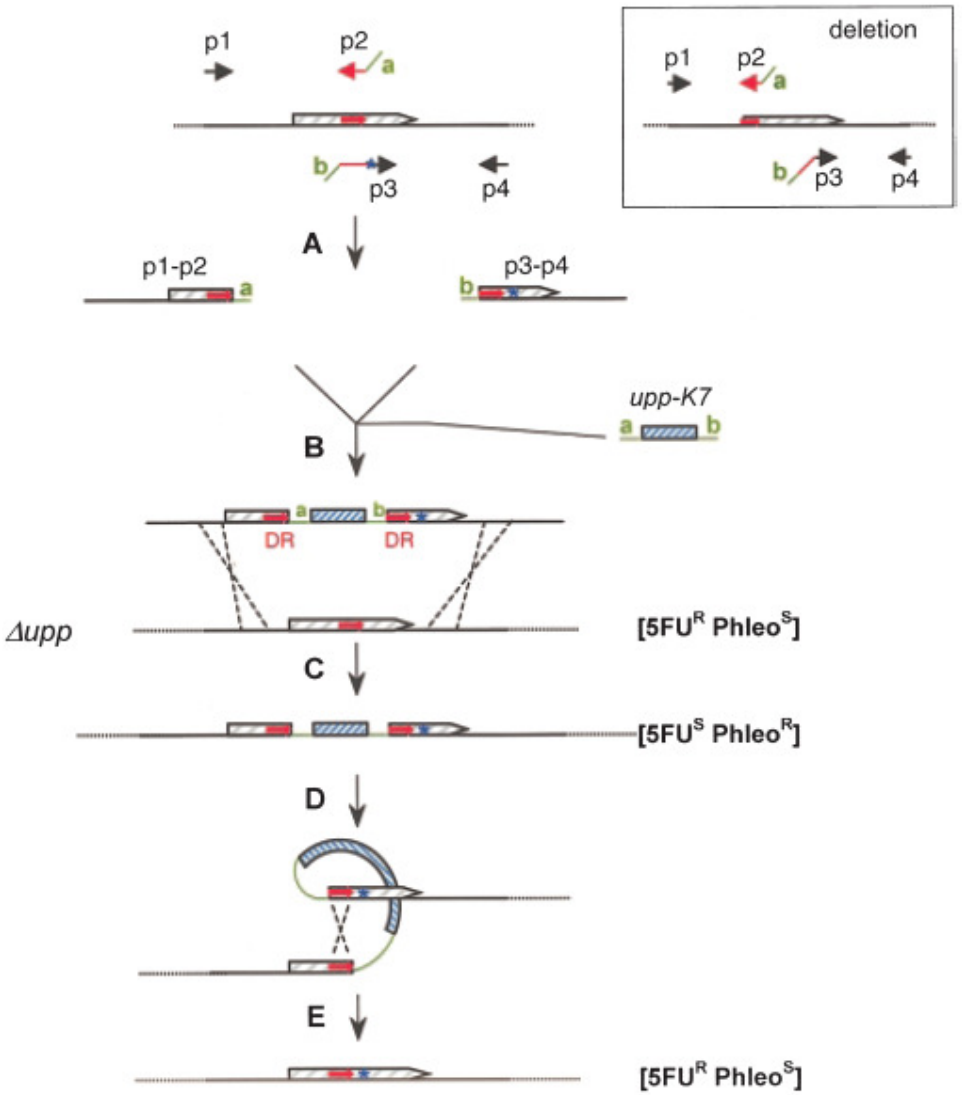


Figure 2.5 Mutagenesis using the *upp*-cassette. A DNA fragment located upstream of the site to mutate, and a fragment encompassing that site are polymerase chain reaction (PCR)-generated using the p1-p2 and p3-p4 primer combinations respectively (A). Primer 2 comprises a 5'-region homologous to the a-end of the *upp*-cassette (green line), and a 3'-region homologous to the chromosome (red line). Primer 3 is composed of: (i) a 5'-region homologous to the b-end of the *upp*-cassette (green line); (ii) a 30 nt region corresponding to the chromosome sequence designated by →, and homologous to the p2-end of fragment p1-p2 (red line); (iii) a 3'-region homologous to the chromosome (black line) that carries the mutation to be introduced in the chromosome (*). Importantly, the 5'-regions of p2 and p3 are designed so that the *upp*-cassette genes are codirectionally transcribed with the targeted chromosomal gene. DNA fragments p1-p2 and p3-p4 are joined with the *upp*-cassette in a joining PCR reaction (B). Homologous integration of this fragment in a Δ*upp* strain, which is $5FU^R$ $Phleo^S$, provokes the replacement of the chromosomal gene by the *upp*-tagged mutated copy and yields a $5FU^S$ $Phleo^R$ strain (C). The excision of the *upp*-cassette (D) through recombination between the 30 bp direct repeats → generates a $5FU^R$ $Phleo^S$ strain that carries only the desired mutation with no other modification in its chromosome (E). In the inset, the primer positions to generate a deletion are indicated. The 30 nt central region of p3, which creates the DR, is homologous to the remote p2 priming site (from Fabret et al., 2002)

I used the *E. coli* knock-out library (BW25113 *E. coli* strains) from http://ecoli.aist-nara.ac.jp (Courtesy of Prof. H. Mori) for the motility screens. The yncE and TatD mutants were constructed in RP437 *E. coli* strains. The gene deletions are constructed as described by Datsenko and Wanner (2000).

2.2.11 Bacterial swarming assay (motility assay)

Swarming plates were prepared as LB plates with 0.25% agar. Bacteria were incubated overnight at 37°C in liquid culture and spotted onto the agar plates using tooth picks. Swarming plates were evaluated after incubating for 7h at 37°C.

2.2.12 Rescue of the ΔyviF mutant phenotype

The vector pDG-yviF was created by ligation-independent cloning as described in Joseph *et al.*(Joseph et al., 2001). The pUni entry clone of yviF (pUniDyviF) was recombined with host vector pHB-HA3. The resulting HA-tagged yviF construct served as template for a PCR reaction with the primer pair pHB-HA3_pDGforward / pHB-HA3_pDGreverse. The PCR product was cloned into the vector pDG148-Stu. This plasmid was transformed to ΔyviF strain and the swarming assay was caried out on swarming plates with 1mM IPTG for the induction of the yviF expression.

2.2.13 Co-expression of TP0658 and TP0868

TP0658 was PCR-amplified from its pUni clone (McKevitt et al., 2003), pUniD-TP0658, using the primer pair TP0658_pEGSTforward/TP0658_pEGSTreverse. The PCR product was cloned into pEGST (Kholod and Mustelin 2001) using BamHI and XhoI generating a GST fusion of TP0658. TP0868 was PCR-amplified from its pUni clone using the primer pair TP0868_pAC28forward/TP0868_pAC28reverse. The PCR product was cloned into pAC28 (Kholod and Mustelin 2001) using BamHI and SacI generating a His-tag fusion of TP0868. *E. coli* BL21/DE3 cells (Novagen) were co-transformed with both plasmids. After overnight incubation at 37°C–relying on the basal level expression of the promoters-the cells were lysed, proteins separated by SDS-PAGE, blotted, and the blot probed for GST (antibody G1160, Sigma-Aldrich, München, Germany) or His (His probeH-15, Santa Cruz Biotechnology, CA, USA).

2.2.14 Test of yviF mutant for flagellin expression

B. subtilis Δupp (wild type for yviF cassette integration) cells and ΔyviF cells were transformed with empty expression vector, pDG148-Stu, or yviF rescue plasmid, pDG-yviF, and compared. For induction of yviF expression the bacterial cultures were incubated for 3h at 37°C with 1mM IPTG. Cells were harvested and lysed using standard procedures. Equal amounts of protein as judged by Coomassie staining of a SDS-PAGE gel were loaded. In a Western Blot the cell lysates were probed for the HA-tag (anti-HA antibody HA.11, Covance Research Products, CA, USA) and flagellin (using a GST-yviF overlay, described above).

2.2.15 Peptide Synthesis

SPOT synthesis which was originally introduced by Frank (1992) is an easy and flexible technique for simultaneous and parallel chemical synthesis of peptides at distinct positions on a membrane support (Frank 1992).The SPOT method has opened up unlimited opportunities to synthesize and subsequently screen large arrays of synthetic peptides (Landgraf et al., 2004; Otte et al., 2003). Peptide arrays prepared by SPOT synthesis can be used to study molecular recognition events, such as epitope mapping, the analysis of protein-protein and the identification of biologically active peptides. Peptide arrays can be applied to precisely depict molecular recognition events on the single amino acid level.

Peptide Synthesis Procedure

Required Materials:

Whatman paper (hardened low ash Nr. 1450 917)
DMF (Dimethylfromamide)
NMP (N-methyl pyrrolidone)
DIC (Diisopropyl carbodiimide)
Piperidine (20% in DMF)
HOBt (Hydroxy Benzotriazole)
Fmoc-AA-OH
Fmoc- β-Ala-Opfp
Acetic Anhydride
TFA (Trifluoroacetic acid)
Ethanol
TIBS (Triisobutylsilane)
dH_2O
NMI (1-Methylimidazole)
Triisoporpylsilan

Preparation of β-**Alanine membranes**

- Cut 10x15 cm Whatman paper
- Incubate in a metal dish following solution:

β-alanine OH	2,56 g
DMF	40 ml
DIC	1496 µl
NMI	1270 µl

- Incubate 20 ml pro membrane overnight shaking gently
- Wash 3 x with DMF
- Incubate membranes 20 min with 20% piperidine (removal of F-moc group)
- Wash 5 x in DMF
- Wash 2 x in EtOH

- Let dry overnight

Amino Acids Aliquots

- Weigh 0.5 mmol of each amino acid
- Add for each amino acid 1ml of HOBt (2.3 g in 20 ml NMP)
- Dissolve amino acids and fill up to 1.5 ml with NMP
- Store at -20°C

Peptide Synthesis:

Thaw amino acids overnight at RT

Activate amino acidswith 240µl **Activation Stock Solution:**

0.4 ml DIC
2.0 ml NMP
2.8 ml DMF

Mix amino acids and activate for 30 min at RT

Centrifuge 3 min at 4000 rpm

Pipet solution (avoiding precipitated amino acids) to the synthesizer vials in correct order

Fill up:

DMF

Ethanol

20% Piperidine solution

Capping solution: 0.3ml Acetic Anhydride in 15ml DMF

- Activate membrane in DMF
- Start Program
- For β-alanine membranes spacer spotting dissolve 0.0573 g Fmoc.β.Ala-Opfp in 400 µl DMSO
- Thaw amino acids aliquots overnight
- Activate amino acids
- Discard old vials and pipet amino acids into new ones
- Fill up Capping and Piperidine solutions
- Thaw amino acids overnight at RT
- Activate amino acids
- Discard old vials and pipet amino acids into new ones
- Fill up Capping and Piperidine solutions

- Thaw amino acids overnight at RT
- Let membrane dry overnight

TFA removal of all side chain protecting groups (for β-Alanine membranes)

- Incubate the membranes without shaking in 90% TFA solution

0.5 g Phenol
1.5 ml TIBS
2.5 ml DCM
45 ml TFA

- Wash 4x 3 min with DCM
- Wash 3x 3 min with DMF
- Wash 3x 3 min with EtOH
- Let Dry
- Incubate the membranes without shaking in 50% TFA solution

0.5 g Phenol
1 ml dH_2O
1.5 ml TIBS
22.5 ml DCM
25 ml TFA

- Wash 4x 3 min with DCM
- Wash 3x 3 min with DMF
- Wash 3x 3 min with EtOH
- Let Dry

TFA removal of all side chain protecting groups (for PEG membranes = Intavis membranes)

Pro membrane: 9 ml TFA
0.5 ml Triisoporpylsilan
0.5 ml dH_2O

- Incubate for 1 h in the mixed solution, shake gently in between
- Wash 4x 3 min with DCM
- Wash 3x 3 min with DMF
- Wash 3x 3 min with EtOH
- Let Dry

Positive control peptides

Three control peptides that bind to the anti-GST antibody were used:

1. QRALAKDLIVPRRP
2. LAKDLIVPRRPEWN
3. DLVIRPPRPPKVLGL

Protein incubation

Activate the membrane 10 min with methanol

Wash 3 x 5 min with 1x TBS

10x TBS (1 liter)

Tris base 24.2 g
NaCl 292.2 g
pH to 7.5 with HCl

Block 3 hrs with **Blocking Solution**

1 g skim powder milk
2.5 g Saccharose
In 50ml 1 x TBS

Incubate protein 10 µg/ml in blocking solution (- Tween) overnight rotating at 4°C

Wash 3 x 5 min with 1x TBS

Incubate 3 hrs at RT with primary antibody in blocking solution + 0.05% Tween

Wash 3 x 5 min with 1x TBS

Incubate 1.5 hrs at RT with secondary antibody in blocking solution + 0.05% Tween

Wash 3 x 5 min with 1x TBS

ECL: - Mix 1 vol. of solution A and 1 vol. of solution B

- Pipete the mixture on to the blot and let activate for 1 min

- Put between blot and film a cling film and expose

2.16 Flagellum phylogenetic supertree construction

2.16.1 Selection of conserved flagellum proteins and sequence processing

- Proteins involved in 'Flagellar assembly' were downloaded from the KEGG database (Kanehisa et al., 2006) and selected. In total sequences of 48 protein families which cover up to 32 species were selected.
- FASTA formatted protein sequences were aligned using CLUSTALW (default parameters).
- Manual inspection of the alignments. If a family contained recent paralogs (paralogs which are more similar to each other than to proteins of other species), one protein was randomly chosen and removed. If there were early paralogs (paralogs which are more similar to proteins from other species than to its own), only the most similar compared to the majority of proteins was retained.
- The PIR formatted outputs were submitted to GBLOCKS (Castresana 2000) (default parameters). This program searches for highly informative phylogenetic blocks. A block contains sites which are conserved in at least 50% of the taxa and is flanked by highly conserved sites (conserved in at least 85% of the taxa). Protein families with less than 4

taxa or no conserved gblocks sites were excluded from further analysis (9 families): flbD flgA flgB flgC flgD flgE flgF flgG flgH flgI flgK flgL flgM flgN flhA flhB flhC flhF flhG fliA fliD fliD fliE fliF fliG fliI fliK fliN fliY fliP fliQ fliR fliS fliT motA motB.

- Subsequently, files of the remaining 35 protein families were used to construct maximum parsimony (MP) and maximum likelihood (ML) trees.

2.16.2 Protein family trees

2.16.2.1 Construction of maximum parsimony bootstrap consensus trees

- The PIR formatted blocks were converted into NEXUS format by the READSEQ program.
- The NEXUS files were executed by PAUP*, version win-4b10(Swofford 2003). For each family a bootstrap analysis with 100 bootstrap replicates was performed using a heuristic search based on the MP method. In total 35 bootstrap consensuses (50% majority-rule) trees were constructed. These trees were compiled into a single tree file and gene names were translated into species names.

2.16.2.2 Construction of maximum likelihood (ML) bootstrap consensus trees

- The PIR formatted blocks were converted into PHYLIP format by the READSEQ program.
- The PHYLIP files were bootstrapped by SEQBOOT (Felsenstein 2005) with 100 bootstrap replicates.
- Maximum likelihood distance matrices were computed by TREE-PUZZLE 5.2 (Schmidt et al., 2002) under a gamma law-based model of substitution (alpha parameter estimated by TREE-PUZZLE, eight gamma rate categories) in combination with PUZZLEBOOT.
- Consensus trees were generated by NEIGHBOR and CONSENSE (Felsenstein 2005) (50% majority-rule).
- Consensus trees were compiled into a single tree file and gene names were translated into species names.

2.16.3 Supertree construction

- To look for underepresentated taxa, co-occurrence matrices of the MP and ML trees were computed by the CLANN supertree software (Creevey and McInerney 2005). The co-occurrence matrices indicated that *Streptomyces coelicolor* and *Chlamydia trachomatis serovar* have significantly lower co-occurrence values than the majority of taxa. Therefore, those taxa have been removed from further analysis and the supertree files were constructed for the remaining 30 taxa.
- A matrix representation using a parsimony (MRP) approach was used to represent the bootstrapped consensus trees as a binary matrix. The MRP matrices of the MP and ML bootstrapped consensus tree were constructed with CLANN.
- For each matrix a bootstrapped (100 bootstrap replicates) consensus tree (50% majority-rule) was generated by PAUP* (Swofford 2003) using a heuristic search based on the MP method.
- The resulting two trees were merged by CLANN (50% majority-rule).

3. Results

3.1 Yeast two-hybrid (Y2H) screening of *T. pallidum* flagellum and chemotaxis proteins

3.1.1 Construction of a *Treponema pallidum* yeast two-hybrid (Y2H) proteome array

A prerequisite for the Y2H screening is the existence of a Y2H proteome library. We used the ORFeome (i.e. all ORFs of a genome) of *Treponema pallidum* (McKevitt et al., 2003) to construct Y2H bait and prey expression plasmids. The *Treponema pallidum* prey clones (ORFs with Gal4p activation domain fusions) were constructed by transferring all predicted ORFs of *Treponema pallidum* from a Univector Plasmid fusion System (UPS) entry vector (pUniD), into the Y2H prey vectors pLP-GADT7 vector (Clontech), using Cre recombination (Figure 2.1, methods). The prey constructs were transformed into haploid yeast cells. Finally, individual yeast colonies, each carrying one specific prey construct, were arrayed on YEPD agar plates in a 96-format. For Y2H screening individual prey constructs were subsequently arrayed in quadruplicates in a 384-well format (see methods).

3.1.2 Selection of proteins associated with bacterial motility and chemotaxis for Y2H screening

To have a comprehensive set of bacterial motility and chemotaxis proteins, I chose 45 proteins annotated with ‘‘chemotaxis and motility’’ from the TIGR *Treponema pallidum* proteome database (http://www.tigr.org). 17 proteins were selected from the dataset of Rain et al. (2001). These proteins were hypothesized to have a function in motility because of their Y2H protein-protein interactions with known flagellum proteins (Rain et al., 2001). 15 proteins were selected based on predictions by the String and/or KEGG database(s) (Kanehisa et al., 2002; von Mering et al., 2003). These “predictions” were usually based on physical interactions or indirect evidence such as participation in the same metabolic pathway or cellular process. Such associations were usually derived from high-throughput experimental data and from predictions based on genomic context analysis. In total 75 proteins were selected (Table 3.1), and these proteins were used as baits (ORFs with Gal4p DNA binding domain fusions) for Y2H screening. Baits were also constructed by recombination-based transfer of the ORFs from an entry vector (pUniD) into the Y2H bait vector pLP-GBKT7 (Clontech), using Cre recombination (see methods). The bait constructs were also transformed into haploid yeast cells. After a self activation test the baits were used for array-based genome-wide Y2H screening (see self activation test in methods).

3.1.3 Array-based two-hybrid screening of motility and chemotaxis proteins

In the course of this project I screened all *T. pallidum* flagellum and chemotaxis baits (Table 3.1) against the whole genome prey array (see methods for details). An example of a screen with the protein FlaB3 flagellin of *T. pallidum* (a homolog of FliC in *E. coli*) is shown in Figure 3.1. Only those cells are growing which are indicative of a two-hybrid interaction, e.g. between FlaB3 and FliS, TP0658, and FlaB1 (see below).

Table 3.1 List of *T. pallidum* flagellum and chemotaxis baits selected for Y2H screening and number of Y2H interactions identified

M, Classified as chemotaxis and motility in TIGR (http://www.tigr.org/). **R,** associated with motility based on interactions found by Rain et al. (Rain et al., 2001). **S**, association with motility proteins in STRING, KEGG (K) or both (**KS**). Baits screening status, **Yes**: baits have been successfuly screened, **No**: baits not available. "Y2H interactions as baits" is the number of interactions found with the protein as bait; "Y2H interactions total" includes interactions with the protein as prey.

Common name	ORF	Type	Description	Baits Screening status	Y2H Interactions as a bait	Y2H Interactions total
cheA	TP0363	M	chemotaxis histidine kinase	no	0	0
cheB	TP0631	KS	protein-glutamate methylesterase	yes	0	0
cheR	TP0630	M	chemotaxis protein methyltransferase	yes	15	16
cheW	TP0364	M	purine-binding chemotaxis protein	yes	0	0
cheW	TP0439	M	purine-binding chemotaxis protein	yes	0	0
cheX	TP0365	M	chemotaxis protein	yes	0	0
cheY	TP0366	M	chemotaxis response regulator	yes	0	0
clpA	TP0801	R	ATP-dependent Clp protease subunit A	yes	0	0
csrA	TP0657	S	carbon storage regulator	yes	2	2
filG	TP0026	M	flagellar motor switch protein	yes	0	8
flaA	TP0249	M	flagellar filament outer layer protein	no	0	0
flaA	TP0664	M	flagellar filament outer layer protein	yes	2	3
flaB1	TP0868	M	flagellar filament 34.5 kDa core protein	yes	3	5
flaB2	TP0792	M	flagellar filament 33 kDa core protein	yes	4	5
flaB3	TP0870	M	flagellar filament 31 kDa core protein	yes	19	20
flbD	TP0726	M	flagellar protein	yes	0	4
flgB	TP0396	M	flagellar basal-body rod protein	no	0	5
flgC	TP0397	M	flagellar basal-body rod protein	no	0	3
flgD	TP0728	M	flagellar hook assembly scaffolding protein	yes	12	12
flgE	TP0727	M	flagellar hook protein	yes	2	3
flgG	TP0960	M	flagellar basal-body rod protein	yes	0	0
flgG	TP0961	M	flagellar basal-body rod protein	yes	19	19
flgK	TP0660	M	flagellar hook-associated protein 1	yes	4	6
flgL	TP0659	M	flagellar hook-associated protein 3	yes	3	3
flhA	TP0714	M	flagellar biosynthesis protein	yes	2	2
flhB	TP0715	M	flagellar biosynthetic protein	yes	1	1
flhF	TP0713	M	flagellar-associated GTP-binding protein	yes	1	1
fliD	TP0872	M	flagellar filament cap protein	yes	0	0
fliE	TP0398	M	flagellar hook-basal body complex protein	yes	15	16
fliF	TP0399	M	flagellar basal-body M ring protein	yes	3	4
fliG	TP0400	M	flagellar motor switch protein	yes	12	12
fliH	TP0401	M	flagellar assembly protein	yes	1	1
fliI	TP0402	M	flagellum-specific ATP synthase	yes	3	3
fliL	TP0722	M	flagellar protein	yes	2	2
fliM	TP0721	M	flagellar motor switch protein	yes	3	3
fliQ	TP0717	M	flagellar biosynthetic protein	yes	1	1
fliP	TP0718	M	flagellar biosynthetic protein	yes	0	0
fliR	TP0716	M	flagellar biosynthetic protein	yes	3	3
fliS	TP0943	M	flagellar protein	yes	11	13

fliY	TP0720	M	flagellar motor switch protein	yes	19	19
groEL	TP0030	R	heat shock protein	yes	5	5
lepA	TP0510	R	GTP-binding membrane protein (lepA)	no	0	0
lig	TP0634	R	DNA ligase (lig)	no	0	0
Mcp1	TP0040	M	methyl-accepting chemotaxis protein	yes	0	0
Mcp2	TP0488	M	methyl-accepting chemotaxis protein	no	0	0
Mcp2	TP0639	M	methyl-accepting chemotaxis protein	yes	0	0
Mcp2	TP0640	M	methyl-accepting chemotaxis protein	yes	3	5
metK	TP0794	R	S-adenosylmethionine synthetase	yes	3	3
motA	TP0725	M	flagellar motor rotation protein	yes	1	2
motB	TP0724	M	flagellar motor rotation protein	yes	0	0
nrdA	TP1008	R	ribonucleoside-diphosphate reductase, subunit alpha	yes	0	0
nrdB	TP0053	R	ribonucleoside-diphosphate reductase, subunit beta	yes	0	1
polA	TP0105	R	DNA polymerase I	yes	0	0
recN	TP0442	R	DNA repair protein	no	0	0
rplB	TP0192	R	ribosomal protein L2	yes	0	0
SigG	TP0218	S	sigma factor SigG regulation protein, putative	yes	2	2
tap1	TP0729	K	treponemal aqueous protein	yes	0	0
thdF	TP0550	R	thiophene and furan oxidation protein	yes	0	0
tig	TP0506	R	trigger factor	yes	2	2
topA	TP0394	R	DNA topoisomerase I	yes	2	2
ylxH	TP0712	K	ATP-binding protein	yes	0	1
ylxH	TP0853	S	ATP-binding protein	yes	0	0
	TP0378	M	flagellar protein, putative	yes	1	1
	TP0403	M	flagellar protein, putative	yes	2	3
	TP0408	R	hypothetical protein	no	0	0
	TP0464	R	conserved hypothetical protein	yes	2	2
	TP0567	KS	conserved hypothetical protein	yes	3	3
	TP0658	S	transmembrane protein, putative	yes	5	8
	TP0709	KS	RNA polymerase sigma factor	yes	1	1
	TP0710	S	conserved hypothetical protein	yes	0	0
	TP0719	S	hypothetical protein	yes	0	0
	TP0764	S	conserved hypothetical protein	yes	0	0
	TP0939	R	pyruvate oxidoreductase	yes	1	3
	TP0959	KS	hypothetical protein (flgJ)	yes	6	6
	TP0965	R	membrane fusion protein, putative	yes	1	1

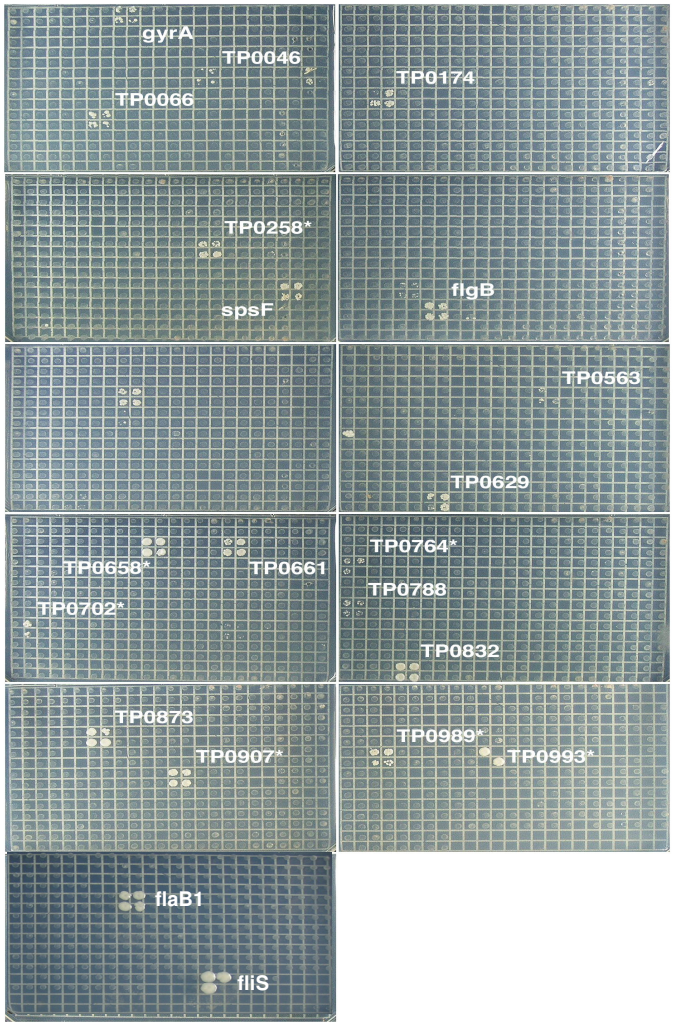


Figure 3.1 An example of an array-based Y2H screen: The bait TP0870 (FlaB3, homologous to FliC in *E.coli*) was screened against the *Treponema pallidum* prey library (preys as quadruplicates on 11 Omnitray plates as 384-well format). Yeast colonies indicate protein-protein interactions.

3.1.4 Y2H protein-protein interactions of flagellum and chemotaxis proteins

After filtering the raw Y2H protein-protein interaction data of the 75 motility baits revealed 204 unique high confidence and 64 low confidence protein-protein interactions (Table 3.2). In total 268 protein-protein interactions, involving 174 unique proteins was identified for flagellar and chemotaxis proteins in the *T. pallidum* Y2H array screens (Table 3.2).

Table 3.2 List of Y2H interactions of flagellum and chemotaxis baits

Baits are marked in bold letter and the interacting preys listed below. **Prey count**: is the number of different interacting baits. **: high confidence interactions. *: low confidence interactions (the parameters used to classfy, high confidence, and low confidence interactions are described in methods).

Bait/Prey	**Defination**	**Prey count**	**Confidence Score**
TP0030	**heat shock protein (groEL)**		
TP0050	conserved hypothetical protein	10	**
TP0987	hypothetical protein	2	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0726	flagellar protein (flbD)	8	**

TP0946	glucose-inhibited division protein B (gidB)	10	**
TP0046	**hypothetical protein**		
TP0398	flagellar hook-basal body complex protein (fliE)	24	**
TP0218	**sigma factor SigG regulation protein, putative**		
TP0269	conserved hypothetical protein	3	**
TP0561	conserved hypothetical protein	24	**
TP0377	conserved hypothetical protein	3	*
TP0365	**chemotaxis protein (cheX)**		
TP0877	conserved hypothetical protein	5	*
TP0377	**conserved hypothetical protein**		
TP0403	flagellar protein, putative	1	**
TP0378	flagellar protein, putative	1	**
TP0378	**flagellar protein, putative**		
TP0002	DNA polymerase III, subunit beta (dnaN)	5	**
TP0281	hypothetical protein	22	*
TP0394	**DNA topoisomerase I (topA)**		
TP0208	preprotein translocase subunit (secY)	3	**
TP0763	hypothetical protein	3	**
TP0140	K+ transport protein (ntpJ)	2	*
TP0209	ribosomal protein L36 (rpmJ)	24	*
TP0540	anti-sigma F factor antagonist (spoIIAA)	1	*
TP0398	**flagellar hook-basal body complex protein (fliE)**		
TP0005	DNA gyrase, subunit A (gyrA)	10	**
TP0046	hypothetical protein	22	**
TP0050	conserved hypothetical protein	10	**
TP0064	hypothetical, protein	8	**
TP0066	hypothetical protein	11	**
TP0092	RNA polymerase sigma factor E (rpoE)	12	**
TP0979	conserved hypothetical protein	1	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0162	Holliday junction DNA helicase (ruvB)	4	**
TP0396	flagellar basal-body rod protein (flgB)	18	**
TP0530	V-type ATPase, subunit E, putative	12	**
TP0579	hypothetical protein	5	**
TP0629	hypothetical protein	7	**
TP0383	conserved hypothetical protein	20	**
TP0197	ribosomal protein L29 (rpmC)	10	*
TP0399	**flagellar basal-body M ring protein (fliF)**		
TP0561	conserved hypothetical protein	24	**
TP0708	hypothetical protein	4	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0400	**flagellar motor switch protein (fliG)**		
TP0014	hypothetical protein	4	**
TP0066	hypothetical protein	11	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0209	ribosomal protein L36 (rpmJ)	24	**
TP0399	flagellar basal-body M ring protein (fliF)	2	**

TP0648	conserved hypothetical protein	3	**
TP0091	cysteinyl-tRNA synthetase (cysS)	2	**
TP0092	RNA polymerase sigma factor E (rpoE)	12	**
TP0233	anti-sigma F factor antagonist, putative	14	**
TP0443	conserved hypothetical protein	3	**
TP0665	hypothetical protein	2	**
TP0743	ribosomal protein L27 (rpl27)	1	**
TP0401	**flagellar assembly protein (fliH)**		
TP0026	flagellar motor switch protein (fliG)	8	**
TP0402	**flagellum-specific ATP synthase (fliI)**		
TP0413	phosphoglucomutase	2	**
TP0640	methyl-accepting chemotaxis protein (mcp2)	5	**
TP0833	hypothetical protein	22	**
TP0403	**flagellar protein, putative**		
TP0868	flagellar filament 34.5 kDa core protein (flaB1)	1	**
TP0586	leucyl-tRNA synthetase (leuS)	21	**
TP0088	conserved hypothetical protein	10	*
TP0421	**conserved hypothetical protein**		
TP0939	pyruvate oxidoreductase	6	**
TP0095	hypothetical protein	8	**
TP0711	conserved hypothetical protein	12	**
TP0439	**purine-binding chemotaxis protein (cheW)**		
TP0046	hypothetical protein	22	*
TP0160	prolyl-tRNA synthetase (proS)	17	*
TP0712	ATP-binding protein (ylxH)	3	*
TP0800	serine-type D-Ala-D-Ala carboxypeptidase (dacC)	2	*
TP0464	conserved hypothetical protein		*
TP0426	V-type ATPase, subunit A (atpA)	1	*
TP0888	riboflavin kinase/FMN adenylyltransferase (ribF)	1	*
TP0471	hypothetical protein	2	*
TP0477	glucose-6-phosphate dehydrogenase isozyme (devB)	2	*
TP0592	hypothetical protein	1	*
TP0464	**conserved hypothetical protein**		
TP0712	ATP-binding protein (ylxH)	3	**
TP0739	conserved hypothetical protein	1	**
TP0506	**trigger factor (tig)**		
TP0086	conserved hypothetical protein	2	**
TP0605	translation elongation factor TS (tsf)	2	**
TP1008	ribonucleoside-diphosphate reductase, subunit alpha (nrdA)	2	*
TP0622	hypothetical protein	2	*
TP0767	translation elongation factor G (fusA)	2	*
TP0712	ATP-binding protein (ylxH)	3	*
TP0281	hypothetical protein	22	*
TP0567	**conserved hypothetical protein**		
TP0421	conserved hypothetical protein	4	**
TP0675	hypothetical protein	3	**

TP0708	hypothetical protein	4	**
TP0630	**chemotaxis protein methyltransferase (cheR)**		
TP0046	hypothetical protein	22	**
TP0064	hypothetical, protein	8	**
TP0088	conserved hypothetical protein	10	**
TP0092	RNA polymerase sigma factor E (rpoE)	12	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0121	conserved hypothetical protein	5	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0162	Holliday junction DNA helicase (ruvB)	4	**
TP0305	CTP synthase (pyrG)	3	**
TP0330	cell division protein (ftsH)	3	**
TP0396	flagellar basal-body rod protein (flgB)	18	**
TP0586	leucyl-tRNA synthetase (leuS)	21	**
TP0851	hypothetical protein	1	**
TP0877	conserved hypothetical protein	5	**
TP0209	ribosomal protein L36 (rpmJ)	24	**
TP0843	heat shock protein, putative	4	*
TP0640	**methyl-accepting chemotaxis protein (mcp2)**		
TP0064	hypothetical, protein	8	**
TP0088	conserved hypothetical protein	10	**
TP0397	flagellar basal-body rod protein (flgC)	19	**
TP0046	hypothetical protein	22	*
TP0657	**carbon storage regulator (csrA)**		
TP0064	hypothetical protein	1	**
TP0088	conserved hypothetical protein	20	**
TP0397	conserved hypothetical protein	13	*
TP0046	**transmembrane protein, putative**		
TP0620	tpr protein I (tprI)	1	**
TP0736	hydrogenase, gamma chain (hydG)	1	**
TP0831	arginyl-tRNA synthetase (argS) }	1	**
TP0939	pyruvate oxidoreductase	9	**
TP0941	hypothetical protein	1	**
TP0731	conserved hypothetical protein	1	*
TP0659	**flagellar hook-associated protein 3 (flgL)**		
TP0002	DNA polymerase III, subunit beta (dnaN)	5	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0726	flagellar protein (flbD)	8	**
TP0260	hypothetical protein	9	*
TP0660	**flagellar hook-associated protein 1 (flgK)**		
TP0465	hypothetical protein	3	**
TP0626	exonuclease, putative	9	**
TP0383	conserved hypothetical protein	20	**
TP0518	conserved hypothetical protein	19	**
TP0024	conserved hypothetical protein	13	*
TP0067	conserved hypothetical protein	5	*
TP0080	quinoline 2-oxidoreductase	1	*
TP0997	protease IV (sppA)	4	*
TP0751	hypothetical protein	4	*

TP0664	**flagellar filament outer layer protein (flaA)**		
TP0233	anti-sigma F factor antagonist, putative	14	**
TP0247	N-acetylmuramoyl-L-alanine amidase (amiA)	9	**
TP0709	**RNA polymerase sigma factor**		
TP0974	hypothetical protein	11	**
TP0713	**flagellar-associated GTP-binding protein (flhF)**		
TP0711	conserved hypothetical protein	16	**
TP0714	**flagellar biosynthesis protein (flhA)**		
TP0377	conserved hypothetical protein	3	**
TP0708	hypothetical protein	4	**
TP0561	conserved hypothetical protein	24	*
TP0715	**flagellar biosynthetic protein (flhB)**		
TP0561	conserved hypothetical protein	24	**
TP1016	basic membrane protein (tpn39b)	2	*
TP0716	**flagellar biosynthetic protein (fliR)**		
TP0561	conserved hypothetical protein	24	**
TP0725	flagellar motor rotation protein (motA)	2	**
TP0774	Mg2+ transport protein (mgtC)	2	**
TP0833	hypothetical protein	22	*
TP0717	**flagellar biosynthetic protein (fliQ)**		
TP0561	conserved hypothetical protein	24	**
TP0720	**flagellar motor switch protein (fliY)**		
TP0193	ribosomal protein S19 (rpsS)	1	**
TP0209	ribosomal protein L36 (rpmJ)	24	**
TP0233	anti-sigma F factor antagonist, putative	14	**
TP0833	hypothetical protein	22	**
TP0048	conserved hypothetical protein	1	**
TP0066	hypothetical protein	11	**
TP0088	conserved hypothetical protein	10	**
TP0100	thioredoxin, putative	3	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0397	flagellar basal-body rod protein (flgC)	19	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0287	conserved hypothetical protein	13	**
TP0579	hypothetical protein	5	**
TP0629	hypothetical protein	7	**
TP0660	flagellar hook-associated protein 1 (flgK)	3	**
TP0700	hypothetical protein	1	**
TP0772	hypothetical protein	3	**
TP0782	hypothetical protein	2	**
TP0843	heat shock protein, putative	4	**
TP0721	**flagellar motor switch protein (fliM)**		
TP0066	hypothetical protein	11	**
TP0092	RNA polymerase sigma factor E (rpoE)	12	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0722	**flagellar protein (fliL)**		
TP0260	hypothetical protein	9	**
TP0561	conserved hypothetical protein	24	**
TP0725	**flagellar motor rotation protein (motA)**		

TP0725	flagellar motor rotation protein (motA)	2	**
TP0727	**flagellar hook protein (flgE)**		
TP0209	ribosomal protein L36 (rpmJ)	24	**
TP0292	outer membrane protein (tpn50)	1	**
TP0728	**flagellar hook assembly scaffolding protein (flgD)**		
TP0554	phosphoglycolate phosphatase (gph)	15	**
TP0579	hypothetical protein	5	**
TP0586	leucyl-tRNA synthetase (leuS)	21	**
TP0629	hypothetical protein	7	**
TP0726	flagellar protein (flbD)	8	**
TP0727	flagellar hook protein (flgE)	2	**
TP0005	DNA gyrase, subunit A (gyrA)	10	**
TP0046	hypothetical protein	22	**
TP0050	conserved hypothetical protein	10	**
TP0100	thioredoxin, putative	3	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0396	flagellar basal-body rod protein (flgB)	18	**
TP0792	**flagellar filament 33 kDa core protein (flaB2)**		
TP0064	hypothetical, protein	8	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0658	transmembrane protein, putative	5	**
TP0943	flagellar protein (fliS)		**
TP0794	**S-adenosylmethionine synthetase (metK)**		
TP0002	DNA polymerase III, subunit beta (dnaN)	5	**
TP0046	hypothetical protein	22	**
TP0955	hypothetical protein	2	**
TP0951	ribosomal protein L34 (rpmH)	1	*
TP0868	**flagellar filament 34.5 kDa core protein (flaB1)**		
TP0050	conserved hypothetical protein	10	**
TP0658	transmembrane protein, putative	5	**
TP0943	flagellar protein (fliS)	18	**
TP0870	**flagellar filament 31 kDa core protein (flaB3)**		
TP0658	transmembrane protein, putative	5	**
TP0873	hypothetical protein	1	**
TP0005	DNA gyrase, subunit A (gyrA)	10	**
TP0046	hypothetical protein	22	**
TP0050	conserved hypothetical protein	10	**
TP0066	hypothetical protein	11	**
TP0174	hypothetical protein	1	**
TP0396	flagellar basal-body rod protein (flgB)	18	**
TP0629	hypothetical protein	7	**
TP0630	chemotaxis protein methyltransferase (cheR)	2	**
TP0702	conserved hypothetical protein	2	**
TP0711	conserved hypothetical protein	16	**
TP0053	ribonucleoside-diphosphate reductase, subunit beta (nrdB)	1	**
TP0981	sensory transduction histidine kinase, putative	4	**
TP0640	methyl-accepting chemotaxis protein (mcp2)	5	**
TP0760	penicillin-binding protein (pbp)	2	**

TP0257	glycerophosphodiester phosphodiesterase (glpQ)	2	**
TP0943	flagellar protein (fliS)	18	**
TP0383	conserved hypothetical protein	20	**
TP0518	conserved hypothetical protein	19	**
TP0162	Holliday junction DNA helicase (ruvB)	4	*
TP0382	hypothetical protein	2	*
TP0186	oxygen-independent coproporphyrinogen III oxidase, putative	3	*
TP0554	phosphoglycolate phosphatase (gph)	15	*
TP0445	4-methyl-5(b-hydroxyethyl)-thiazole monophosphate biosynthesis enzyme (thiJ)	8	*
TP0591	HPr kinase (ptsK)	3	*
TP0698	hypothetical protein	3	*
TP0031	hypothetical protein	1	*
TP0853	ATP-binding protein (ylxH)	2	*
TP0887	ribosomal protein S15 (rpsO)	1	*
TP0891	translation initiation factor 2 (infB)	2	*
TP0470	conserved hypothetical protein	2	*
TP0510	GTP-binding membrane protein (lepA)	2	*
TP0992	conserved hypothetical protein	3	*
TP0836	hypothetical protein	1	*
TP0586	leucyl-tRNA synthetase (leuS)	21	*
TP0939	**pyruvate oxidoreductase**		
TP0396	flagellar basal-body rod protein (flgB)	18	**
TP0943	**flagellar protein (fliS)**		
TP0974	hypothetical protein	11	**
TP0943	flagellar protein (fliS)	15	**
TP0197	ribosomal protein L29 (rpmC)	10	**
TP0209	ribosomal protein L36 (rpmJ)	24	**
TP0660	flagellar hook-associated protein 1 (flgK)	3	**
TP0945	ribulose-phosphate 3-epimerase (cfxE)	10	**
TP0870	flagellar filament 31 kDa core protein (flaB3)	20	**
TP0868	flagellar filament 34.5 kDa core protein (flaB1)	1	**
TP0792	flagellar filament 33 kDa core protein (flaB2)	2	**
TP0048	conserved hypothetical protein	1	**
TP0383	conserved hypothetical protein	20	**
TP0277	carboxyl-terminal protease (ctp)	1	*
TP0396	flagellar basal-body rod protein (flgB)	18	*
TP0959	**hypothetical protein**		
TP0005	DNA gyrase, subunit A (gyrA)	10	**
TP0046	hypothetical protein	22	**
TP0088	conserved hypothetical protein	10	**
TP0162	Holliday junction DNA helicase (ruvB)	4	**
TP0260	hypothetical protein	9	**
TP0467	hypothetical protein	2	**
TP0961	flagellar basal-body rod protein (flgG)	30	*
TP0961	**flagellar basal-body rod protein (flgG)**		
TP0005	DNA gyrase, subunit A (gyrA)	10	**
TP0024	conserved hypothetical protein	13	**

TP0050	conserved hypothetical protein	10	**
TP0026	flagellar motor switch protein (fliG)	8	**
TP0397	flagellar basal-body rod protein (flgC)	19	**
TP0236	transcription antitermination protein (nusG)	7	**
TP0314	hypothetical protein	2	**
TP0160	prolyl-tRNA synthetase (proS)	17	**
TP0287	conserved hypothetical protein	13	**
TP0341	UDP-N-acetylmuramate--alanine ligase (murC)	8	**
TP0664	flagellar filament outer layer protein (flaA)	14	**
TP0726	flagellar protein (flbD)	8	**
TP0751	hypothetical protein	4	**
TP0757	polypeptide deformylase (def)	16	**
TP0772	hypothetical protein	3	**
TP0773	periplasmic serine protease DO (htrA)	17	**
TP0782	hypothetical protein	2	**
TP0877	conserved hypothetical protein	5	**
TP0911	conserved hypothetical protein	1	**
TP0965	membrane fusion protein, putative	5	**
TP0514	excinuclease ABC, subunit A (uvrA)	4	**
TP0763	hypothetical protein	3	*
TP0767	translation elongation factor G (fusA)	2	*
TP0965	**membrane fusion protein, putative**		
TP0431	conserved hypothetical protein	1	*
TP0432	hypothetical protein	1	*
TP0536	hypothetical protein	1	*
TP0965	membrane fusion protein, putative	10	*
TP1008	**ribonucleoside-diphosphate reductase, subunit alpha (nrdA)**		
TP0068	conserved hypothetical protein	2	*
TP0141	methylated-DNA-protein-cysteine S-methyltransferase (dat)	1	*
TP0192	ribosomal protein L2 (rplB)	1	*

Table 3.3 List of flagellum interactions which are found in both directions (i.e., bait → prey, and prey → bait)

Bait	Description	Prey	Description
TP0398	flagellar hook-basal body complex protein (fliE)	TP0046	hypothetical protein
TP0943	flagellar protein (fliS)	TP0870	flagellar filament 31 kDa core protein (flaB3)
TP0943	flagellar protein (fliS)	TP0868	flagellar filament 34.5 kDa core protein (flaB1)
TP0943	flagellar protein (fliS)	TP0792	flagellar filament 34.5 kDa core protein (flaB2)

3.1.5 Comparison of *T. pallidum* Y2H interactions with literature data

A comprehensive literature review for known flagellum protein-protein interactions in PubMed yielded 62 interactions (Table 1.7). In order to compare the overlap between know flagellum interactions with our Y2H interactions, interologs were predicted for *T. pallidum* (i.e., pairs of interacting proteins that have interacting homologs in another species). 54 interologs were predicted (including paralogues) for *T. pallidum* flagellum proteins, and 15 of these interologs

are reproduced in the Y2H screens (Table 3.5). However, many interactions reported in the literature were not found in this Y2H screens such as interactions that depend on post translational modifications. For example, phosporylated CheY binds to FliM and its phosphatase cheZ. The electrostatic interactions between MotA with the C-terminal region of FliG were detected by mutational studies and a detailed structural model has been published (Lloyd et al., 1999). These interactions have been mainly identified by diverse methods ranging from affinity purification to genetic suppressor screens.

Table 3.4 Protein-protein interactions overlapping with known interactions

PubMed ID	Species	Litrature Interaction		*T. pallidum* Y2H		*C. jejuni* Y2H		*H. pylori* Y2H		*E. coli* MS	
10783392	*V. alginolyticus*	PomA	PomA	TP0725	TP0725	-	-	-	-	-	-
8757288	*E. coli*	MotA	FliM			Cj0337c	Cj0060c	-	-	-	-
10998179	*S. typhimurium*	FliH	FliI	-	-	-	-	HP0353	HP1420	-	-
10998179	*S. typhimurium*	FliH	FliH	-	-	-	-	HP0353	HP0353	-	-
8986772	*S. typhimurium*	fliC	fliC	-	-	-	-	-	-	-	-
11327763	*S.typhimurium*	FliC	FliS	TP0792	TP0943	Cj1339c	Cj0549	HP0601	HP0753	-	-
11327763	*S.typhimurium*	FliC	FliS	TP0659	TP0943	Cj0720c	Cj0549	HP0115	HP0753	-	-
11327763	*S.typhimurium*	FliC	FliS	TP0868	TP0943	Cj0887c	Cj0549	HP0295	HP0753	-	-
11327763	*S.typhimurium*	FliC	FliS	TP0870	TP0943	Cj1338c	Cj0549	-	-	-	-
11204784	*X. oryzae*	FlhF	FlhF	-	-	Cj0064c	Cj0064c	-	-	-	-
15968056	*H. pylori*	σ^{54}	HP0958	-	-	-	-	HP0714	HP0958	-	-
11327763	*S. typhimurium*	FliS	FliS	TP0943	TP0943	-	-	-	-	-	-
10440379	*T. maritima*	FliG	MotA	-	-	-	-	-	-		
8757288	*E. coli*	FliG	FliG	TP0400	TP0400	-	-	-	-	-	-
8757288	*E. coli*	FliG	FliG	TP0400	TP0026	-	-	-	-	-	-
8757288	*E. coli*	FliG	FliG	TP0026	TP0026	-	-	-	-	-	-
10809678	*S. typhimurium*	FliG	FliF	TP0400	TP0399	-	-	-	-	-	-
10809678	*S. typhimurium*	FliG	FliF	TP0026	TP0399	-	-	-	-	-	-
8631704	*E. coli*	FliG	FliM	TP0400	TP0721	-	-	-	-	-	-
8631704	*E. coli*	FliG	FliM	TP0026	TP0721	-	-	-	-	-	-
8757288	*E. coli*	FliG	FliN	TP0400	TP0720	-	-	-	-	-	-
8757288	*E. coli*	FliG	FliN	TP0026	TP0720	-	-	-	-	-	-
10320579	*E. coli*	FlgN	FlgK	-	-	-	-	-	-	b1070	b1082
10320579	*E. coli*	FlgN	FlgL	-	-	-	-	-	-	b1070	b1923
10320579	*E. coli*	FlgN	FlgL	-	-	-	-	-	-	b1070	b1083

MS: Mass spectrometry complex purifications data.

3.1.6 Comparison of Y2H interactions from *T. pallidum, C. jejuni, and H. pylori*

We compared the known flagellum interactions with two other independent high-throughput systematic protein-protein interaction studies (Table 1.7). Recently, Russell L. Finley and co-workers completed a LexA-based yeast two-hybrid interactome of *Campylobacter jejuni*, (Russell L. Finley, pers. comm) which identified 761 protein-protein interactions for motility proteins. Our curated literature dataset predicted 29 interologs for *Campylobacter jejuni.* Of these, 6 interactions were found in the *Campylobacter jejuni* Y2H screens (Table 3.4).

Rain et al, published a partial protein interaction map for *H. pylori,* testing 261 baits against a genomic fragment library (Rain et al., 2001). Only 8 orthologus baits were in common to both *T. pallidum* and *H. pylori* Y2H studies (Table 3.6), and 14 interactions for these baits were

identified for *H. pylori* and 36 for *T. pallidum*, of which 6 interactions were overlapping in both datasets (i.e. 6/36 = 17% of the *Treponema* interactions and 6/14 = 43% of the *Helicobacter* interactions). Surprisingly, even though the methods used for *Campylobacter jejuni* and our *T. pallidum* protein interaction studies are similar, the overlap between the two dataset is very low (4 out of 131 = 3 %) (Table 3.8). Possible explanations for these observations are discussed in the discussion.

Table 3.5 Comparison of Y2H interactions from *T. pallidum, C. jejuni, and H. pylori* with literature interactions

Literature Interactions	***T. pallidum***	***H. pylori***	***C. jejuni***	***E. coli***
Predicted literature interactions	***64***	***56***	***53***	***65***
Predicted literature PPIs reproduced	***15***	***6***	***6***	***3***
% of literature PPIs have been confirmed by different high throughput studies	*23,4*	*11*	*11*	*4,6*
Normalisation (taking only cmmon bait proteins into account)				
Predicted literature PPIs of which one of the protein is part of the interacting bait list	***54***	***19***	***29***	***57***
% of predicted literature PPIs of which a protein is part of the interacting bait list	***27,8***	***32***	***21***	***5,3***

PPIs: Protein-protein interactions

Table 3.6 Baits common in *T. pallidum* and *H. pylori* Y2H studies

H. pylori **baits**	***T. pallidum*** **baits**	**Description**	**COG**
HP1067	TP0366	chemotaxis response regulator (cheY)	COG0784
HP0391	TP0439	purine-binding chemotaxis protein (cheW-2)	COG0835
HP0353	TP0401	flagellar assembly protein (fliH)	COG1317
HP0601	TP0792	flagellar filament 33 kDa core protein (flaB2)	COG1344
HP0115	TP0659	flagellar hook-associated protein 3 (flgL)	COG1344
HP0753	TP0943	flagellar protein (fliS)	COG1516
HP0870	TP0727	flagellar hook protein (flgE)	COG1749
HP1559	TP0396	flagellar basal-body rod protein (flgB)	COG1815

COG: Cluster of orthologous group.

Table 3.7 Comparison of Y2H interactions from *T. pallidum, C. jejuni, and H. Pylori*

	Overlapping interactions					Predicted flagellum interactions					Percent overlap				
	T. pallidum	*H. pylori*	*C. jejuni*	*E. coli*	*Tp+Hp+Cj*	*T. pallidum*	*H. pylori*	*C. jejuni*	*E. coli*	*Tp+Hp+Cj*	*T. pallidum*	*H. pylori*	*C. jejuni*	*E. coli*	*Tp+Hp+Cj*
T. pallidum		6	4	5	3		124	131	131	118		4,84	3.05	3,82	2,54
H. pylori	4		3	2	3	58		105	79	54	6,90		2,86	2,53	5,56
C. jejuni	4	4		6	2	162	328		342	152	2,47	1,22		1,75	1,32
E. coli	4	4	6			336	504	611			1,19	0,79	0,98		

Predicted interactions based on other datasets. Examples are in red. For example, 124 interactions are predicted for *T. pallidum* based on the *H. pylori* Y2H interaction dataset, assuming that orthologs of interacting proteins ("interologs") interact as well. 5 interactions from the Arifuzzaman et al. dataset (5 out of 131 = 3.82%) are overlapping with our *T. pallidum* Y2H dataset, and a comparison with the *Campylobacter jejuni* flagellum interactions revealed that 4 interactions (4 out of 131 = 3%) are overlapping in both datasets (Table 3.7).

Table 3.8 Protein-protein interactions overlapping with different datasets

T. pallidum	***H. pylori***	***C. jejuni***	***E. coli***
flaB1 - TP0658	flaA - HP1154	-	-
flaB2 - TP0658	flaA - HP1377	-	-
flaB3 - TP0658	-	-	-
flaB1 - fliS	flaA - fliS	flaC - fliS	-
flaB2 - fliS	flaB - fliS	flaA - fliS	-
flaB3 - fliS	-	flaB - fliS	-
TP0100-flgD	-	Cj0864 - flgD	-
-	motB - HP1464	motB - Cj1648	-
proS - groEL	-	-	proS - mopA
proS - flgL	-	-	proS - fliC
proS - flaB2	-	-	
cheW-2 - dacC	-	-	cheW - dacA
rplB - nrdA	-	-	rplB - nrdE
-	rpl2 - deaD	-	rplB - rhlE
-		-	rplB - srmB
-	cheW - tlpA	-	cheW - tar
-	-	-	cheW - tsr
-	-	groEL - Cj1313	mopA - rimJ
-	-	xerD - flgG2	intC - flgG
-	-	groEL - fliH	mopA - fliH
-	-	fliL - Cj0579c	fliL - ybeC
-	-	nuoC - fliM	nuoC - fliM
-	-	rplB - fliM	rplB - fliM

Comparisons of the protein-protein interactions overlap between different high throughput protein-protein interactions datasets. Genes which do not have assigned common name are indicated by systematic ORF name.

3.1.7 Comparison of Y2H interactions with complex purification data from *E. coli*

Arifuzzaman et al. carried out a large-scale comprehensive pull-down assay using a His-tagged *E. coli* ORF clone library, and 2,667 out of 4,339 proteins were successfully analyzed for their putative interaction with target proteins on Ni^{2+}-NTA columns. Co-purified putative interacting proteins were identified by MALDI-TOF MS. In total they identified 11,201 binary "protein-protein interactions" according to spoke model (Prof. H. Mori, pers. comm.), the 'spoke' model counts only associations between the bait and interactors, 564 of these "interactions" involved at least one motility protein. Published interactions predicted 57 interologs for *E. coli* of which 3 were reproduced in the *E. coli* pull down approach (Table 3.5).

Butland et al. (2005) purified complexes from *E. coli* using His-tagged proteins and identified their components by MS. Unfortunately only one flagellum protein was affinity-tagged, namely FliY (TP0720) which copurified together with mopA, yaiL, ybhK, ygiF, and yqjD while FliM (TP0721) has been co-purified with ycbL.

In summary, our Y2H screens reproduced 27.8% (15 out of 54 interactions) of the known previously published interactions, whereas the co-purification study by (Arifuzzaman et al.) reproduced 5.3 % (3 out of 57 interactions). Finley et al. found 21% (6 out of 29) of the known protein-protein interactions (Table.3.5).

3.1.8 Comparison of Y2H interactions with computationally predicted interactions

I used protein interactions predicted for the *T.pallidum*, *C. jejuni* and *H. pylori* know flagellum proteins by the String database (von Mering et al., 2003) and compared them to my Y2H screen results (Table 3.9). I considered functional association predictions based on three types of genomic context associations: conserved genomic neighbourhood, gene fusion events and co-occurrence of genes across genomes. The searches aim to identify pairs of genes which appear to be under common selective pressures during evolution (more than expected by chance), and which are therefore thought to be functionally associated. STRING assigns a confidence score to each predicted association. The benchmarked confidence scores in STRING generally correspond to the probability of finding the linked proteins within the same KEGG pathway.

Of these predicted protein-protein interaction associations (confidence scores > 0.7), 3 % (19/620) were confirmed by our *T. pallidum* Y2H screens, 1.8 % (11/620) and 1.15% were confirmed by *C. jejuni* and *H. pylori* Y2H screens, respectively (Figure 3.2). As we expected, the percentage of confirmed predicted functional associations increased with increased confidence scores of the prediction. Interestingly, when I combine the PPIs of three datasets, the confirmed string predictions increased significantly (form 3 to 6%). This indicates that the interactions identified in *T. pallidum, C. jejuni* and *H. pylori* Y2H screens are complementary to each other, at least for the functional associations predicted by STRING (i.e., no significant overlapping interactions between different Y2H datasets) (Figure 3.2).

Table 3.9 Comparison of *T. pallidum, C. jejuni, and H. pylori* Y2H interactions with computationally predicted interactions by the STRING database

		***T. pallidum* Y2H interactions**		***C. Jejuni* Y2H interactions**		***H. pylori* Y2H interactions**		***T. pallidum, C. jejuni, and H. pylori***
Stringency	i-COG predicted	i-COGs found	Found/ predicted	i-COG predicted	Found/ predicted	i-COGs found	Found/ predicted	Found/ predicted
STRING 0.1	1236	26	2,10	14	1,13	7	0,57	3,8
STRING 0.4	883	23	2,60	12	1,36	7	0,79	4,75
STRING 0.7	620	19	3,06	11	1,77	7	1,13	5,96
STRING 0.9	350	14	4,00	6	1,71	7	2,00	7,71

i-COG predicted: Interacting cluster of orthologous groups (orthologous protein families) predicted for *T. pallidum* flagellum baits by the String database.

i-COGs found: Interacting COGs predictions confirmed by Y2H (Table 5.1 in Appendix).

Found/ predicted: Percentage of predicted interactions confirmed by Y2H.

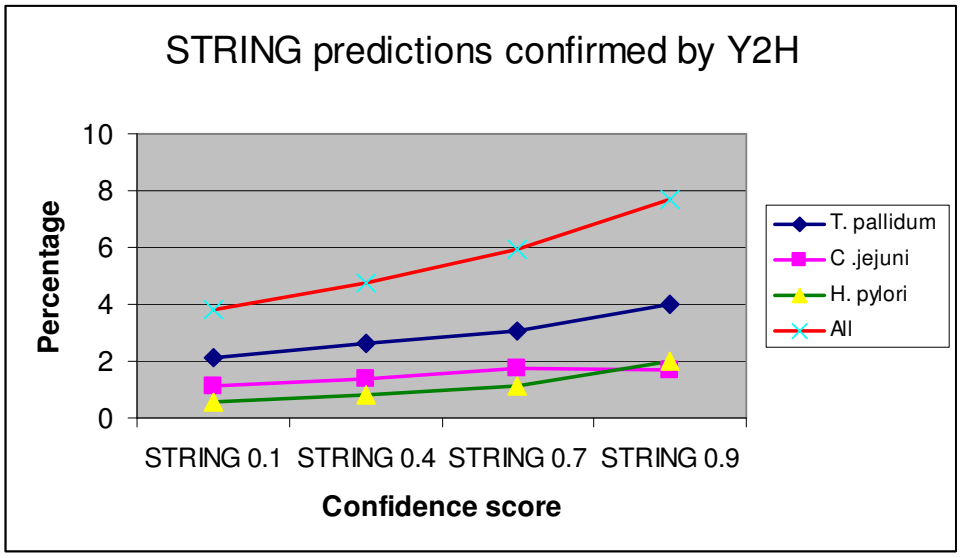


Figure 3.2 Graph showing percentage of computationally predicted protein interaction associations confirmed by Y2H screens. For example, my *T. pallidum* Y2H screens confirmed 4% of the functional association predictions based on genomic context associations by string database

3.2 The *Treponema pallidum* flagellum interaction map

3.2.1 Interconnection of flagellum with other cellular processes

For obtaining a broader overview of the associations of the flagellum apparatus with other functional classes, I looked for the enrichment of different functional classes in the interaction dataset. As expected, flagellum proteins tend to interact with known motility proteins and are enriched in the interaction dataset i.e., 20 % of the identified interactions are with known flagellum proteins (Figure 3.3), indicating Y2H can detect the intricate protein–protein interactions between different components of flagellum.

Strikingly, conserved proteins with unknown function (or no annotation) are highly enriched in the flagellum interaction dataset. This finding indicates the possible role of these proteins in bacterial motility. This finding motivated further analysis of these unknown proteins as potential new components of the bacterial flagellum, or at least new motility-related genes (see below). Additionally, functional associations with proteins involved in tRNA aminoacylation, DNA replication, and electron transport were found.

One interesting metabolic connection is implied by the interaction between NifJ (TP0939), a pyruvate oxidoreductase, and flgB, a flagellar rod protein, as an orthologues interaction (interolog) was also detected in *Helicobacter pylori* by another Y2H study (Rain et al., 2001) This connection is, furthermore, supported by the interaction of pyruvate oxidoreductase with TP0421, a protein newly classified as a motility related gene in this study (see below).

Finally, connections of flagella proteins with ribosomal and DNA metabolism (external flagella proteins) proteins were found (Table 3.2). However, these were not followed up and remain to be investigated.

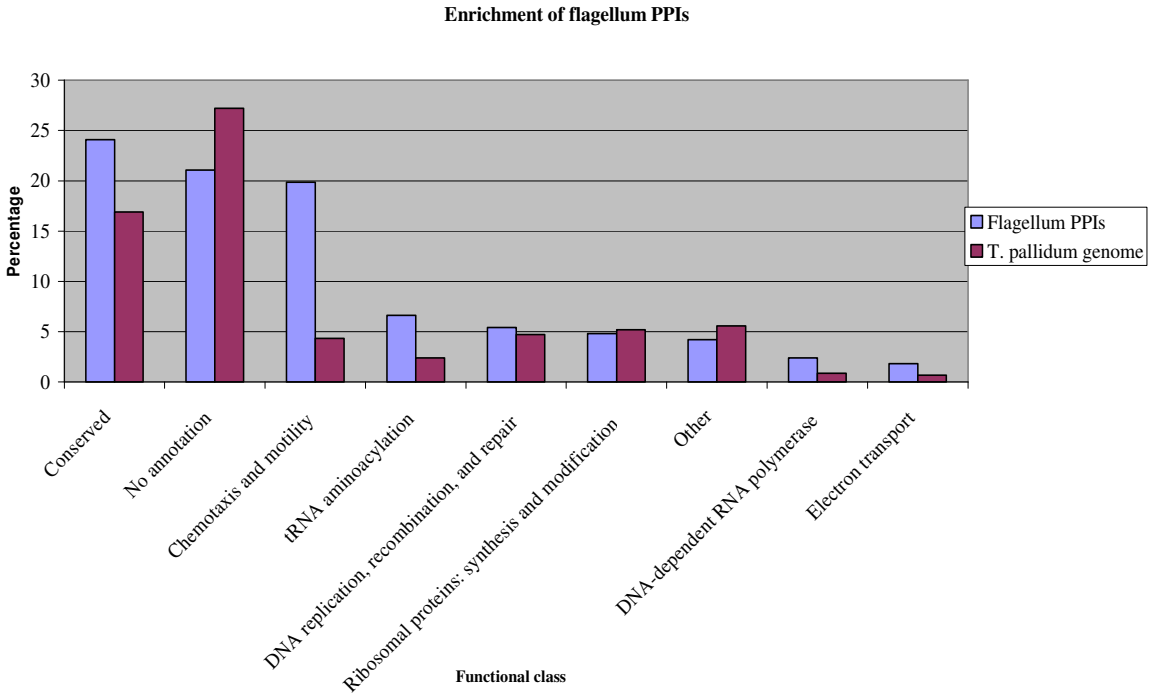
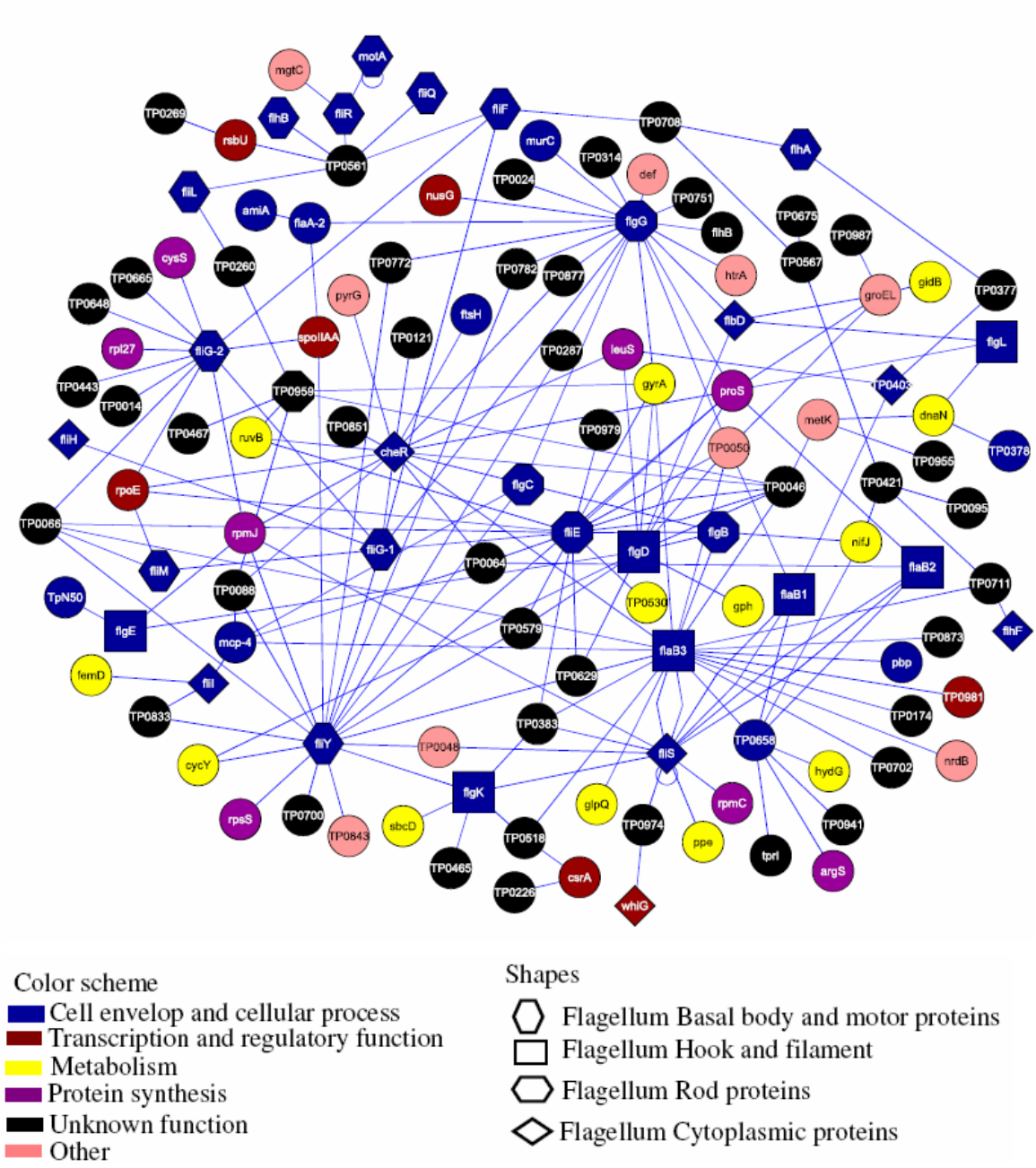


Figure 3.3 Histogram showing enrichment of the different functional classes of proteins interacting , with known flagellum proteins, identified in *T. pallidum* Y2H screens. For example, known chemotaxis and motility proteins are highly enriched in the interaction dataset (20% of the total interactions compared to known chemotaxis and motility genes of the genome < 5%). The percentage of flagellum PPIs is the percentage of preys belonging to different functional classes. *T. pallidum* genome is the percentage of all genes that have been assigned to these functional classes (classification based on TIGR annotations).



Color scheme
- Cell envelop and cellular process
- Transcription and regulatory function
- Metabolism
- Protein synthesis
- Unknown function
- Other

Shapes
- Flagellum Basal body and motor proteins
- Flagellum Hook and filament
- Flagellum Rod proteins
- Flagellum Cytoplasmic proteins

Figure 3.4 Y2H protein-protein interaction map of the *T. pallidum* flagellum apparatus: Proteins are indicated as nodes, protein interactions are represented as solid edges. *T. pallidum* proteins were assigned into six functional classes depicted in different colours based on TIGR annotations for the corresponding ORF or its ortholog. The known flagellum proteins are depicted by different shapes (nodes) based on the localization in the flagellum complex; proteins of unknown function are coloured black. The interaction map is generated using Cytoscape (Shannon et al., 2003).

3.2.2 Intra-flagellum interactions

Self-assembly is used in biological systems to construct large molecular complexes and cellular organelles. The assembly mechanism of the bacterial flagellum system involves dynamic conformational changes and partial folding. The bacterial flagellum is a classical example of such a mechanism (Namba and Vonderviszt 1997; Vonderviszt et al., 1989). Intricate protein–protein interactions between different components of the flagellum complex take place and result in the assembly of a complex multi-protein structure tightly orchestrated in time and space. In my Y2H screening of *T. pallidum* flagellum baits I identified 19 novel intra-flagellum protein interactions which are depicted in the Figure 3.5.

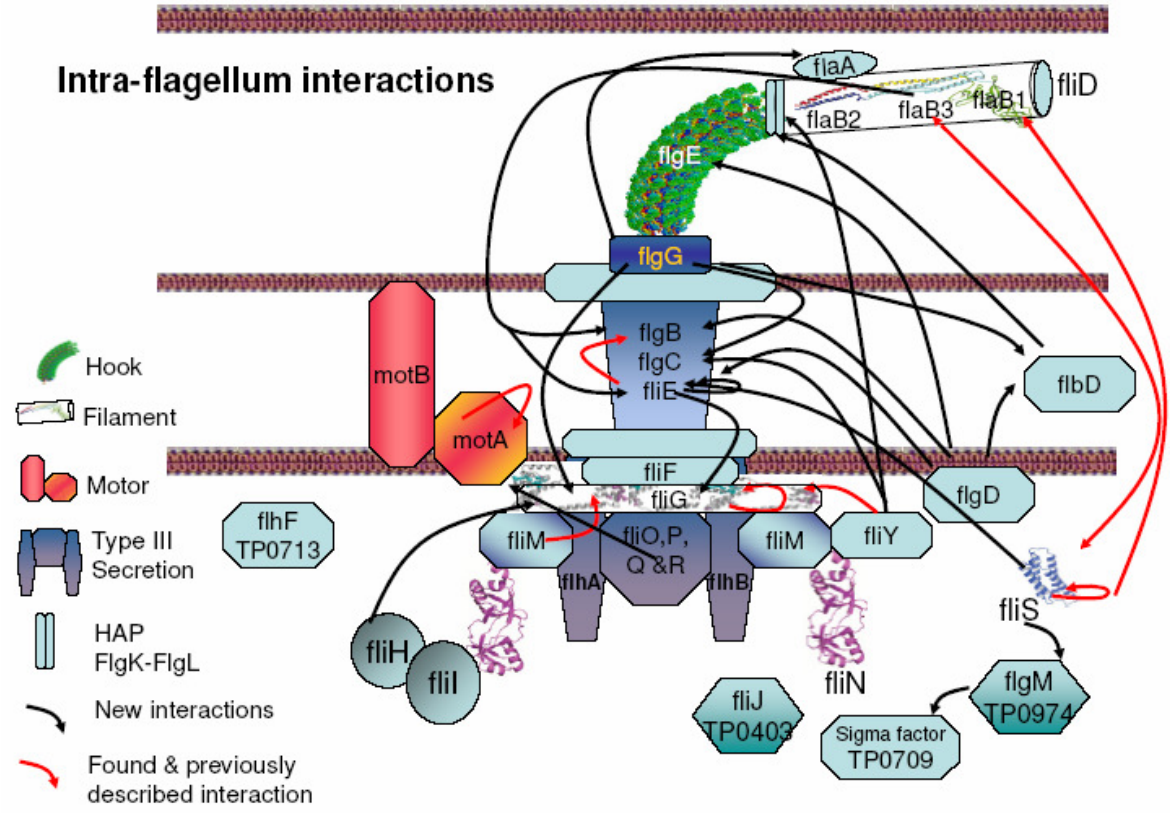


Figure 3.5 Y2H intra-flagellum interactions of *T. pallidum*. The interactions indicated by red arrows are previously known interactions reproduced in our Y2H screens. Black arrows are novel interactions identified. Proteins of known structure are represented by ribbon diagrams (except for FlgE which is shown as a higher-order structure).

3.2.3 New components of the bacterial flagellum complex

Interestingly, my screens with flagellar baits identified many interactions with conserved hypothetical proteins, i.e. proteins of as yet unknown function. These interactions thus predict a function of these proteins in motility. The *T. pallidum* does not allow us to verify such predictions experimentally, but we can use our interaction data to predict interactions in more tractable organisms such as *E. coli* or *B. subtilis*. We identified 72 uncharacterized proteins interacting with proteins known to be involved in flagellum function (Table 5.2 in appendix). To

characterise the functional relevance of the identified *T. pallidum* Y2H protein-protein interactions during bacterial motility, orthologs for all the *T. pallidum* interacting proteins were predicted for *E. coli* and *B. subtilis*. I have identified othologous target genes for gene deletion experiments based on COGs (cluster of orthologous groups of proteins) and in total 23 COGs were found in *E. coli* and 30 COGs in *B. subtilis* (Table 3.10). Bacterial motility assays were done for all the *E. coli* deletion mutants and 7 *B. subtilis* mutants. Of these 5 *B. subtilis* mutants and 10 *E. coli* mutants showed a motility phenotype (Table 3.12 and Table 3.13). For TP0712, conserved hypothetical proteins associated with flagella by our Y2H interactions, an altered motility phenotype has been reported in the literature (Table 3.13), confirming the functional relevance of identified protein-protein interaction in bacterial motility.

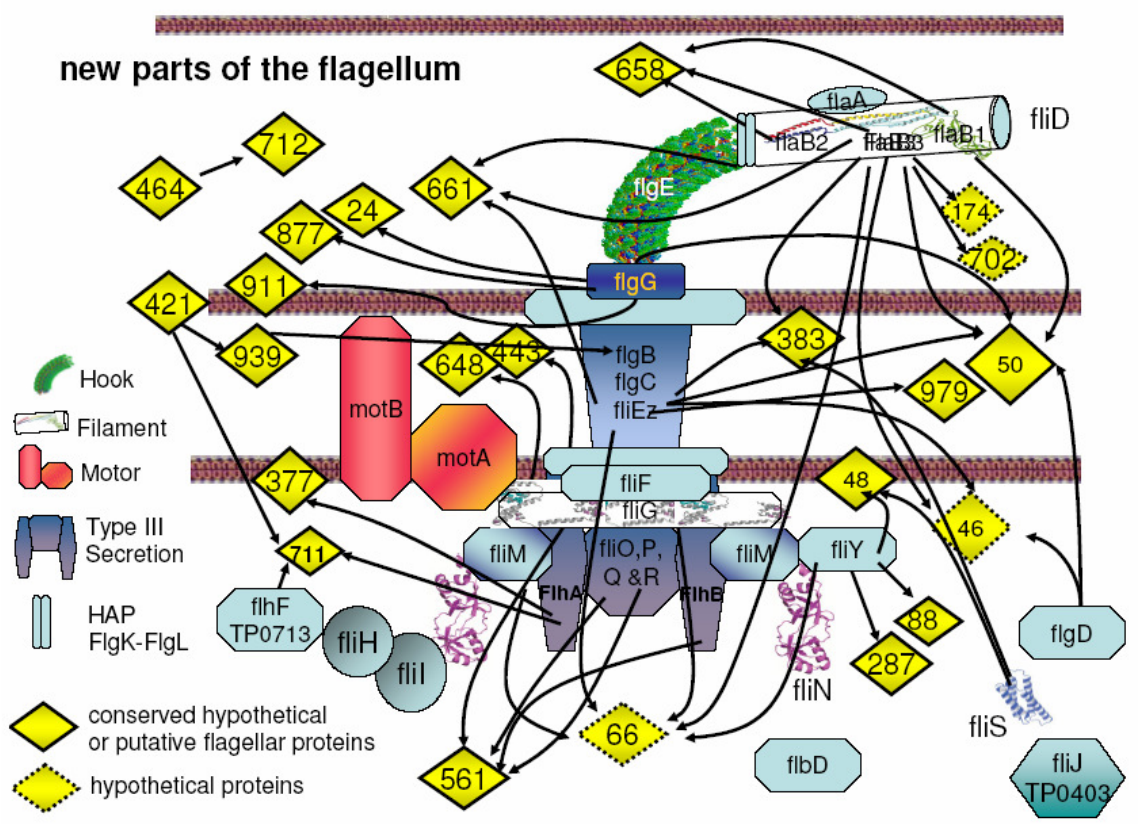


Figure 3.6 New components of the bacterial flagellum complex: Graphical representation of new components of the bacterial flagellum, i.e. proteins of yet unknown function interacting with known flagellum proteins in our Y2H screening of the *T. pallidum* proteome. The interactions are indicated by black arrows (direction: bait to prey) the numbered yellow boxes correspond to the systematic name of the protein (*T. pallidum).* For example, 561 corresponds to TP0561. Known protein structures of orthologs are shown as ribbon diagrams (except for FlgE which is shown as a higher-order structure).

Table 3.10 List of hypothetical and conserved hypothetical proteins interacting with known motility proteins and their *E. coli* and *B. subtilis* orthologs

Interacting Bait(s)	*T. pallidum* Prey	Prey Description	COG	*E. coli* Ortholog	*B. subtilis* Ortholog
TP0567	TP0675	hypothetical protein	COG3735	ybaP	
TP0567	TP0421	hypothetical protein	COG3391	b1452	ywhL ywhK
TP0377	TP0403	flagellar protein, putative	COG2882	fliJ	fliJ
flgG-2	TP0911	conserved hypothetical protein	COG2257		ylqH
groEL fliE flgD flaB1 flaB3 flgG-2	TP0050	hypothetical protein	COG2236		
fliY cheR	TP0843	heat shock protein, putative	COG2214	cbpA	
cheR flgG-2 cheX	TP0877	hypothetical protein	COG2206		
flaB3	TP0981	sensory transduction histidine kinase, putative	COG2199	b1489 ycdT yaiC yeaI	ytrP ydaK yhcK
cheR flgD metK flaB3 TP0959 cheW-2 mcp2-3	TP0046	hypothetical protein	COG1774		yaaT
flaB2 flaB1 flaB3	TP0658	transmembrane protein, putative	COG1699		yviF
fliY fliS	TP0048	hypothetical protein	COG1664		yhbF yhbE
csrA flgK flaB3	TP0518	hypothetical protein	COG1564		yloS
TP0965	TP0431	transcriptional activator, putative	COG1521		yacB
TP0218 fliF flhB fliR fliQ fliL flhA	TP0561	hypothetical protein	COG1512		ydjH
cheR	TP0121	hypothetical protein	COG1509	yjeK	kamA
fliG-2 flaA-2 fliY	TP0233	anti-sigma F factor antagonist, putative	COG1366		yezB rsbR spoIIAA yetI
TP0464	TP0739	hypothetical protein	COG1316		ywtF lytR yvhJ
TP0421 TP0658	TP0939	pyruvate ferredoxin oxidoreductase	COG1013		
TP0965	TP0965	membrane fusion protein, putative	COG0845	acrE yhiU yhiI acrA ylcD b0795 b2074	yknX yvrP
nrdA	TP0068	hypothetical protein	COG0820	yfgB	yloN
flaB3	TP0702	hypothetical protein	COG0739	yebA nlpD	yunA spoIIQ spoIVFA
fliY flgG-2	TP0782	predicted coding region	COG0739	yebA nlpD	yunA spoIIQ spoIVFA
TP0218	TP0269	hypothetical protein	COG0621	yleA yliG	ymcB yqeV
flgG-2 flgK	TP0024	hypothetical protein	COG0569	trkA	yjbQ yuaA ykqB
fliY flgD	TP0100	thioredoxin, putative	COG0526	dsbA trxA trxC dsbE	yneN trxA ydfQ P ykvV bdbA ybdE
TP0658	TP0731	hypothetical protein	COG0494	wcaH mutT yqiE yrfE b1134 b2299 yffH	yvcI nudF mutT ytkD
cheR	TP0330	cell division protein, putative	COG0465	hflB	yjoB ftsH

flgK	TP0067	hypothetical protein	COG0457	nfrA ygeG	rapA rapC rapG
TP0421	TP0095	hypothetical protein	COG0457	nfrA ygeG	rapA rapG
flaB3	TP0470	hypothetical protein	COG0457	nfrA ygeG	rapA rapG
TP0464	TP0471	hypothetical protein	COG0457	nfrA ygeG	rapA rapG
tig	TP0622	hypothetical protein	COG0457	nfrA ygeG	rapA rapG
fliG-2	TP0648	hypothetical protein	COG0457	nfrA ygeG	rapA rapG
flgK	TP0626	exonuclease, putative	COG0420	sbcD	yhaO
fliE	TP0979	putative deoxyribonuclease	COG0084	ycfH	yabD

3.3 Bacterial swarming assay to detect motility phenotype

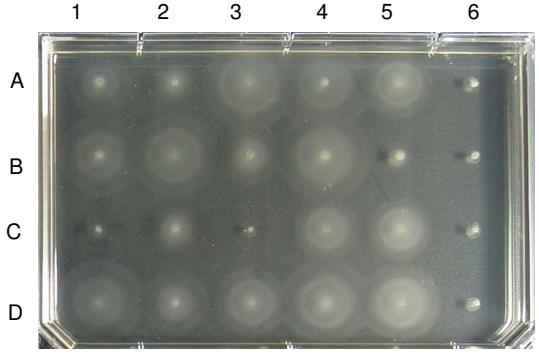
Motility of most bacterial species depends on the proper function of the flagellar apparatus. Hauser first described swarming in *Proteus* species in 1885 (Williams and Schwarzhoff 1978). Swarmer cells need to express their flagellar genes in order to move. In *E. coli* and *S. typhimurium* the flagellar genes are expressed in a hierarchical manner (Soutourina and Bertin 2003). The chemotaxis system is important for regulating swarming movement in all bacteria. Bacterial swarming assays have been used to examine the proper function of the flagellum and chemotaxis signalling. Bacteria will swarm optimally on 0.5% to 2% agar and generally do not swarm at agar concentrations above 2%.

In order to check the functional relevance of the identified Y2H protein-protein interactions during bacterial motility, I carried out a motility phenotype assay. I used a collection of *E. coli* gene deletions (courtesy of H. Mori) to identify genes involved in motility. Mutants that showed a motility defect in swarming assays were then compared to the set of proteins known to be involved in chemotaxis, flagellum function, or associated with flagellar proteins in our Y2H screens (Table 3.11 and Table 3.12). Out of 49 proteins known to be involved in flagellum /chemotaxis in *E. coli*, 37 gene mutants show non motile phenotype and 6 show a reduced motility phenotype, indicating these proteins are crucial for bacterial motility (Table 3.11 and Table 3.12). Unexpectedly, fliE, fliH, flgH, fliZ, fliL, and fliT *E. coli* mutants do not show a motility defect under the conditions tested.

FliE is a flagellar basal body protein which is essential for flagellin secretion and flagellar assembly (Reed et al., 2002). FliE subunits constitute a junction zone between the MS ring and the rod and the proximal rod structure also consists of FlgB subunits. A mutant allele of fliE caused extremely poor flagellation and swarming in *Salmonella* (Minamino et al., 2000). Unexpectedly in *E. coli* the FliE mutant does not show any motility defect under tested conditions.

The fliZ gene encodes a positive regulatory factor for the class 2 flagellar operons in *Salmonella typhimurium*. The fliZ mutation reduced the amounts of excreted flagellar proteins and reduced the invasion ability of *S. enterica* serovar *typhimurium* to HEp-2 cells. fliZ is necessary for full virulence (Iyoda et al., 2001). Interestingly, the fliZ mutant of *E.coli* does not show any motility defect. Thus, while fliZ is necessary for virulence, it is not essential for motility.

A



B

	1	2	3	4	5	6
A	ΔentF	ΔmoeB	ΔccmC	ΔsspA	ΔphoE	**ΔflgD**
B	ΔentE	ΔgrxA	ΔmenE	ΔhemX	**ΔppiD**	**ΔflgF**
C	**ΔlipA**	ΔabgB	**ΔubiX**	Δfre	ΔsfmC	**ΔflgI**
D	ΔnadA	ΔpaaG	ΔhemF	ΔmenG	ΔybgP	**ΔflgK**

Figure 3.7 Bacterial Swarming assay. **A**, *E. coli* deletion mutants were screened for a motility phenotype on bacterial swarming plates (LB medium with 0.25% agar in Omnitray plates). The mutants were grown overnight in 1 ml LB liquid medium, and inoculated on swarming agar plates using the HDR tool of a Biomek 2000 robot, and incubated at 37°C for 7 hours. The motility phenotype was analysed by measuring the diameter of swarming using Irfanview software. **B**, The names of the *E. coli* gene deletions corresponding to the swarming plate in A. the mutants considered to be non-motile are indicated in bold letters.

Table 3.11 Known flagellum/chemotaxis genes and their deletion phenotype

ECO_ID: *Escherichia coli* K-12 MG1655 gene, **JW_ID**: *E.coli* (BW25113 strains) mutant ID (source http://ecoli.aist-nara.ac.jp), **TPA_ORF**: *T. pallidum* ORF name, **Mean percent motility:** mean percentage of the swarming of 2 independent mutants compared to wild type. For example, the cheA mutant swarming diameter is 6.6 % compared to wild type (100%)**.** Mutants that do not show a motility phenotype are indicated in bold letters. Mutants with a swarming diameter < 20% are scored as non-motile, and < 50 % is scored as reduced motility.

Gene name	TPA_ORF	ECO_ID	JW_ID	Phenotype	Mean percent motility
cheA	TP0363	B1888	JW1877	non motile	6.67
cheB	TP0631	B1883	JW1872	non motile	13.56
cheR	TP0630	B1884	JW1873	non motile	37.72
cheW	TP0364	B1887	JW1876	non motile	6.67
cheY	TP0366	B1882	JW1871	non motile	6.67
cheZ		B1881	JW1870	non motile	10.08
flgA		B1072	JW1059	non motile	45.01
flgB	TP0396	B1073	JW1060	non motile	6.67
flgC	TP0397	B1074	JW1061	non motile	6.67
flgD	TP0728	B1075	JW1062	non motile	6.67
flgE	TP0727	B1076	JW1063	non motile	6.67
flgF		B1077	JW1064	non motile	10.08

flgG	TP0960	B1078	JW1065	non motile	6.67
flgJ	TP0959	B1081	JW1068	non motile	22.08
flgK	TP0660	B1082	JW1069	non motile	6.76
flgL		B1083	JW1070	non motile	34.71
flgM		B1071	JW1058	reduced	46.79
flgN		B1070	JW1057	non motile	27.92
flhA	TP0714	B1879	JW1868	non motile	28.9
flhB	TP0715	B1880	JW1869	non motile	6.67
flhC		B1891	JW1880	non motile	26.02
flhD		B1892	JW1881	reduced	46.47
flhE		B1878	JW1867	reduced	27.27
fliA	TP0709	B1922	JW1907	non motile	6.67
fliC	TP0792	B1923	JW1908	non motile	7.09
fliD	TP0872	B1924	JW1909	non motile	7.18
fliF	TP0399	B1938	JW1922	non motile	6.67
fliG	TP0400	B1939	JW1923	non motile	6.67
fliH		B1940	JW1924	reduced	29.22
fliI	TP0402	B1941	JW1925	non motile	6.67
fliK		B1943	JW1927	non motile	45.41
fliM	TP0721	B1945	JW1929	non motile	6.67
fliN	TP0720	B1946	JW1930	non motile	15.63
fliO		B1947	JW5316	non motile	10.08
fliP	TP0718	B1948	JW1932	non motile	6.67
fliQ	TP0717	B1949	JW1933	non motile	45.81
fliR	TP0716	B1950	JW1934	non motile	6.67
fliS	TP0943	B1925	JW1910	non motile	6.96
motA	TP0725	B1890	JW1879	non motile	6.67
MotB	TP0724	B1889	JW1878	non motile	33.91
tap	TP0040	B1885	JW1874	non motile	7.22
tar	TP0040	B1886	JW1875	reduced	35.15
tsr	TP0040	B4355	JW4318	reduced	32.49
fliZ		**B1921**	**JW1906**	**normal**	**>90**
fliE	**TP0398**	**B1937**	**JW1921**	**normal**	**>90**
fliH	**TP0401**	**B1940**	**JW1924**	**normal**	**>90**
flgH		**B1079**	**JW5153**	**normal**	**>90**
fliL	**TP0722**	**B1944**	**JW1928**	**normal**	**>90**
fliT		**B1926**	**JW1911**	**normal**	**>90**

Note: The Sequencing project of E. coli revealed an insertion sequence (IS1) in MG1655 and an insertion sequence (IS5) in the W3110 strain upstream of flhD. In BW25113 no insertion sequence was found upstream of flhD. However, in some cases of our deletion derivatives we found an IS sequence at this position. It looks like a frequently inserted hot spot, and motility might be dependent on this insertion (Prof. H. Mori, pers. comm).

Table 3.12 *E. coli* orthologus of *T. pallidum* Y2H interacting preys showing motility phenotype

ECO_ID: *Escherichia coli* K-12 MG1655 gene, **JW_ID**: *E.coli* (BW25113 starins) mutant ID (source http://ecoli.aist-nara.ac.jp), **TPA_ORF**: *T. pallidum* ORF name. **Mean percent motility:** mean percentage of the swarming of 2 independent mutants tested, compared to wild type. For example, Δ yciM swarm 21.1 % compare to wild type (100%) (Figure 3.8 and 3.9). # These mutants were constructed in *E. coli* strain RP437.

Interacting Bait	Prey	Gene Name	ECO_ID	Annotation	Source	Phenotype	Mean percent motility
flagellar filament outer layer protein (flaA)	TP0247	amiA	B2435	N-acetylmuramoyl-L-alanine amidase (amiA)	Y2H	reduced	45,8
flgK	TP0067	bcsC	B3530	oxidase involved in cellulose synthesis	Y2H	reduced	58.97
flgG	TP0236	rfaH	B3842	transcriptional regulator	Y2H	non motile	
fliS	TP0945	rpe	B3386	ribulose phosphate 3-epimerase	Y2H	non motile	21.35
flgE, fliY, fliG, topA, fliS, and flgE	TP0209	rpmJ	B3299	50S ribosomal subunit protein L36	Y2H	reduced	13.98
fliE,	TP0979	ycfH	B1100	hypothetical protein	Y2H	reduced	37.95
cheW	TP0471	yciM	B1280	putative heat shock protein	Y2H	reduced	21.87
cheW and ATP-binding protein (ylxH)	TP0464	yggH	B2960	tRNA (m7G46) methyltransferase	Y2H	reduced	56.37
flaB3	TP0186	yggW	B2955	putative oxidase	Y2H	reduced	33.33
TP0567	TP0421	#yncE	B1452	Putative receptor	Y2H	reduced	
fliE	TP0979	#TatD	B4483	Magnesium-dependent DNase	Y2H	Increased	

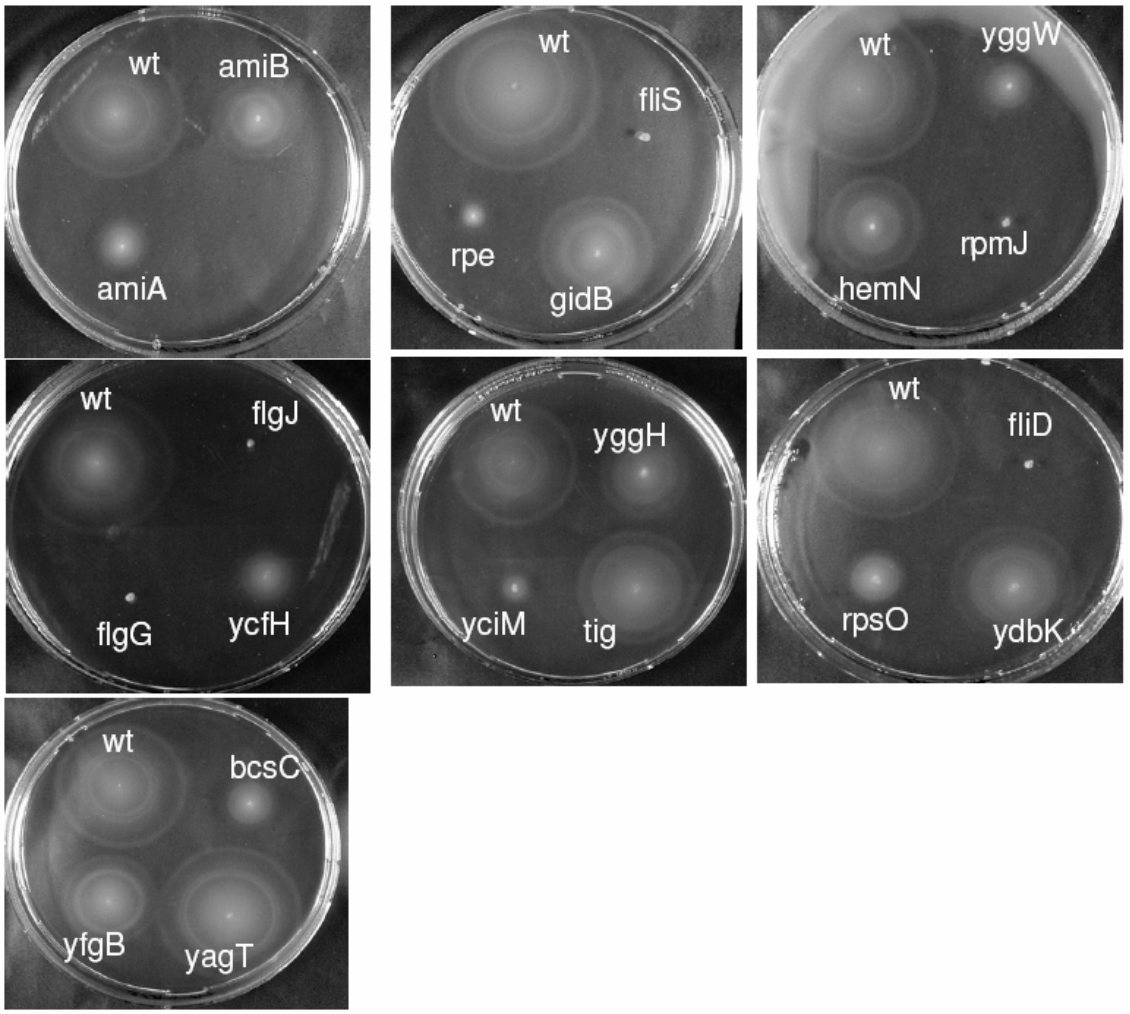


Figure 3.8 Bacterial Swarming assay: The *E. coli* deletion mutants were screened for a motility phenotype on bacterial swarming plates (LB medium with 0.25% agar). The mutants were grown overnight in 1 ml LB liquid medium and inoculated on swarming agar plates using tooth picks and incubated at 37°C for 7 hours. The motility phenotype was analysed by measuring the diameter of swarming using Irfanview software (Table 3.12).

3.3.2. *B.subtilis* and *E. coli* mutants:

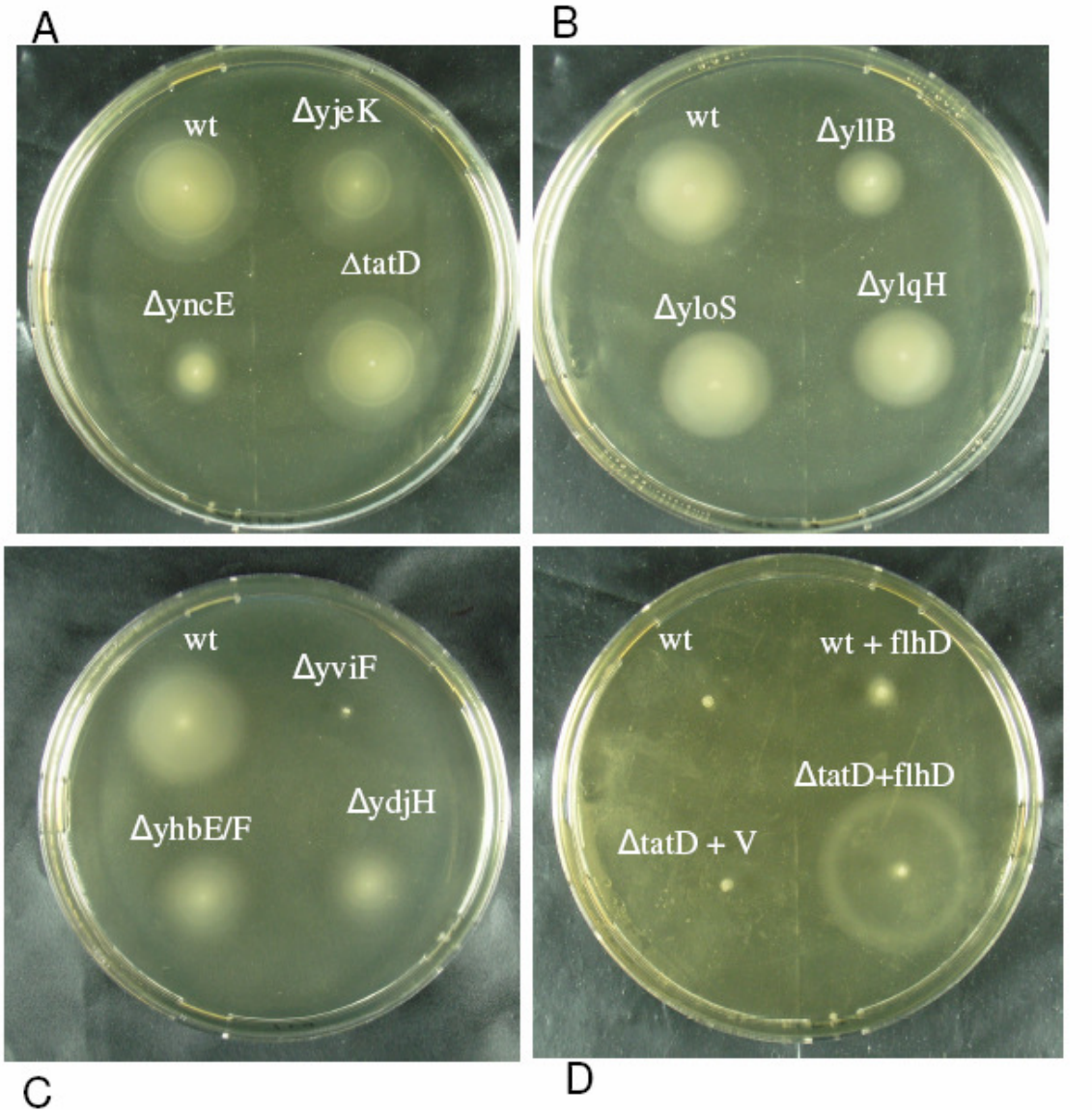


Figure 3.9 Swarming assays of *E. coli/B. subtilis* mutants whose *Treponema* homologs are flagellar interactors and thus implicated in motility (see experimental procedures for details). **A**, the *E. coli* wild type strain RP437 and the deletion mutants Δyjek, ΔyncE, and ΔtatD. **B**, the *B. subtilis* wild type 168 strain (wt), ΔyllB, ΔyloS, and ΔylqH mutants were tested with 1mM IPTG for the induction of downstream genes. **C**, *B. subtilis* wild type 168 strain (wt), ΔyviF, ΔyhbE/F (yhbE and yhbF double mutant), and ΔydjH mutants. **D**, the non-motile *E.coli* strains MC4100 (wt), MC4100 transformed with a plasmid expressing flhD (wt+flhD), a tatD triple mutant transformed with an empty plasmid (ΔtatD+V) and a tatD triple mutant transformed with a plasmid expressing flhD.

Table 3.13 Motility phenotype of the conserved hypothetical proteins interacting with flagellum baits

Prey ORF: Interacting Y2H prey (*T. pallidum*) with known flagellum baits. **Species** that show the motility phenotype on the right: **STM**: *S. typhimurium,* **BSU**: *B. subtilis*, **HPY**: *H. pylori*; **String**: association with motility genes by String database predictions derived by genome context associations (von Mering et al., 2003). Motility phenotype: "**+**" no motility defect, "**-/+**" reduced motility, "**-**" non motile.

Prey ORF	Bait ORF and common name	Speices	STRING associations	Mutant name	Source	Motility phenotype
TP0024	TP0961 (flgG)	STM		TH9226	This study	+
TP0046	TP0398 (fliE), TP0728 (flgD), TP0870 (flaB3), TP0794 (metK), TP0959	BSU		yaaT	Dr. Ogasawara (Bacteria)	-/+
TP0048	TP0720 (fliY),TP0943 (fliS)	BSU		yhbE and YhbF	This study	-/+
TP0377	TP0714 (flhA)		yes			
TP0383	TP0943 (fliS),TP0660, (flgK)	BSU		BFS2817 (yllB)	Dr. Errington (Bacteria)	-/+
TP0443	TP0400 (fliG)		yes			
TP0518	TP0657 (csrA),TP0660 (flgK)	BSU		BFS2824 (yloS)	Dr. Errington (Bacteria)	+
TP0561	TP0218 (SigG), TP0399 (fliF), TP0715 (flhB) TP0716 (fliR), TP0717 (fliQ), TP0722 (fliL)	BSU		ydjH	This study	-/+
TP0648	TP0400 (fliG)	STM		TH9229	This study	+
TP0658	TP0792 (flaB2), TP0868 (flaB1),TP0870 (flaB3)	BSU		yviF	This study	-
TP0702	TP0870 (flaB3)	STM		TH9231 & TH9212	This study	+
TP0712	TP0464	HPY	yes	HP1034	Observation (van Amsterdam and van der Ende 2004)	-
TP0877	TP0630 (cheR), TP0961(flgG)		yes			
TP0911	TP0961 (flgG)	BSU		ylqH	Dr. Errington (Bacteria)	+

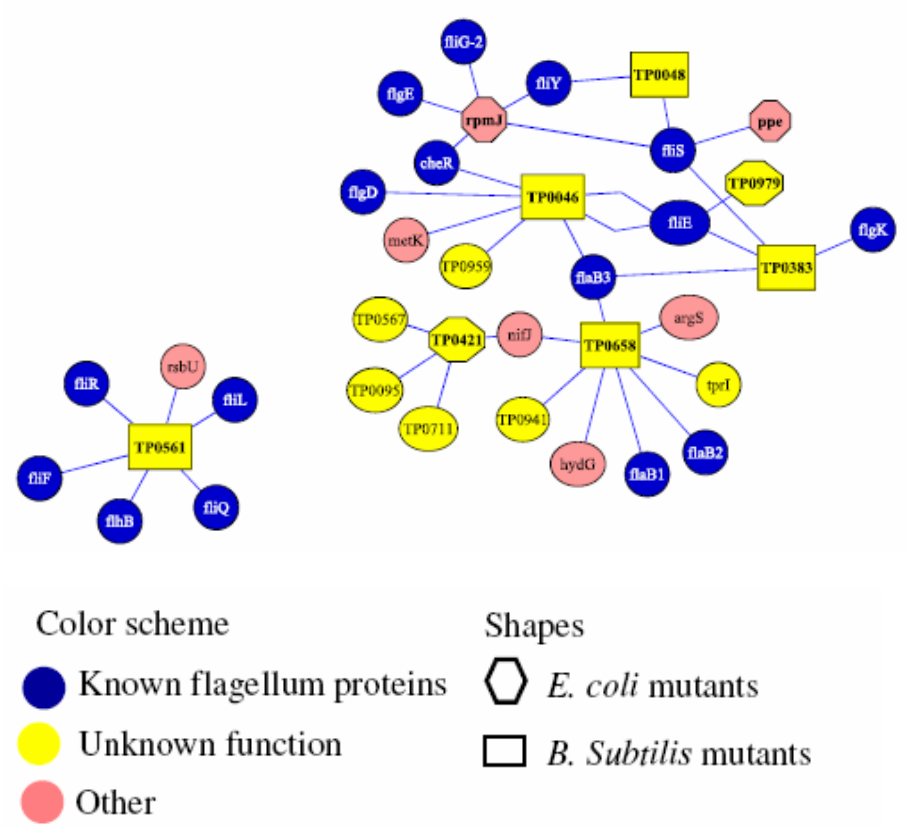


Figure 3.10 Protein-protein interaction map of conserved hypothetical proteins interacting with known flagellum baits and the mutants show motility phenotype. Proteins are indicated as nodes. The orthologous mutants showing a motility phenotype in *E. coli* are depicted as hexagonal shapes, and *B. subtilis* mutants with a phenotype as square shapes.

3.4 Bacterial flagellum phylogenetic supertree

To estimate the phylogenetic relationships of the bacterial flagella of different bacteria species, we constructed a flagellar protein supertree. Here we used combined alignments of 35 orthologous flagellum proteins conserved up to 32 species (see methods for supertree construction). Although individual protein trees are variable in their support of domain integrity, trees based on combined protein data sets strongly support separate monophyletic domains (Brown et al., 2001). Within the bacteria, based on the flagellum supertree spirochetes as monophyletic group (Figure 3.11). The spirochetes have the earliest derived flagellum, if we compare the flagellum supertree with the universal trees based on large combined protein sequence data sets, where spirochetes were placement as the earliest derived bacterial group (Brown et al., 2001).

The β/γ proteobacteria and α proteobacteria species are cluster together in the flagellum supertree, indicating more similarity in the flagellum apparatus with each other then other bacterial species. Similarly, the ε proteobacteria *H. pylori* and *C. jejuni* are clustered together indicating that these two bacteria have a very similar flagellar machinery.

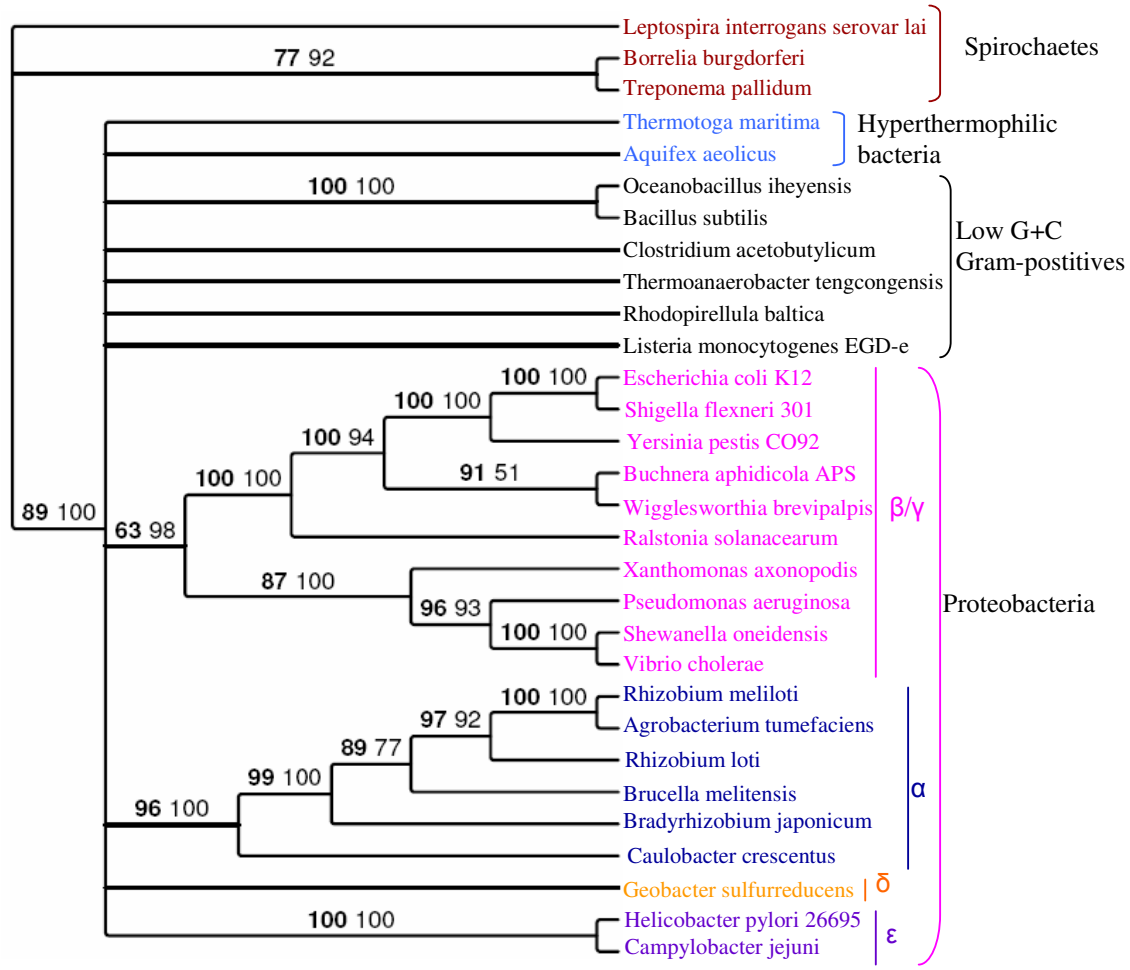
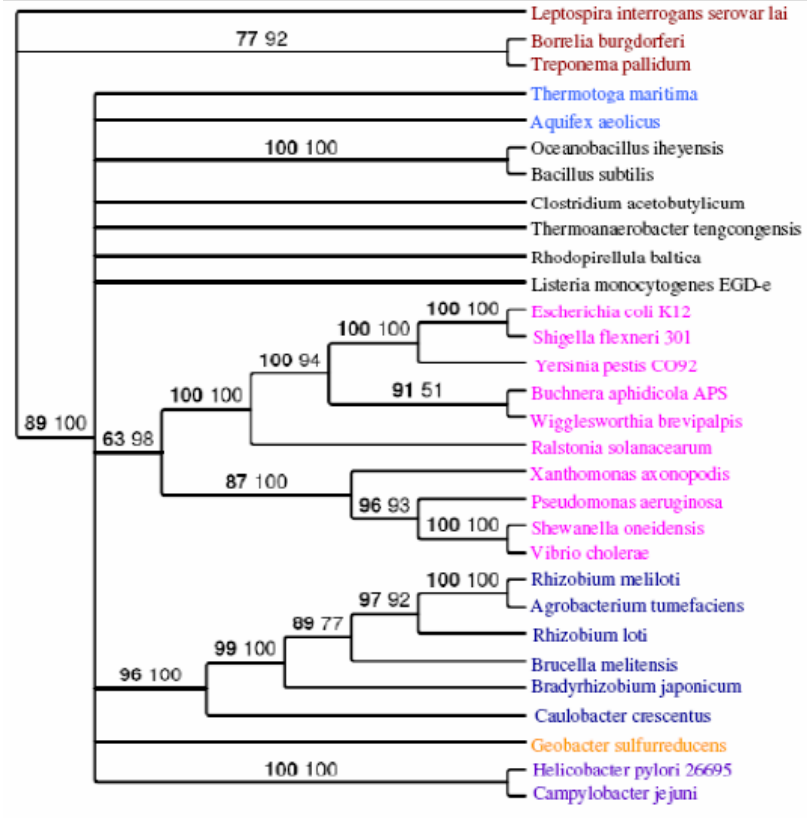


Figure 3.11 Bacterial flagellum supertree constructed from 35 flagellum proteins conserved across 30 species. Two methods have been used to generate single flagellum protein trees resulting in two supertrees generated by Clan software (Creevey and McInerney 2005). One supertree is based on trees made by Paup* and other made by TREE PUZZLE (ML with gamma distribution estimation of evolutionary rate heterogeneity). The consensus supertree was constructed by merging these two supertrees into one (using the Clan software). Numbers along the branches show percentage occurrence of nodes in 100 bootstrap replicates of MP (maximum parsimony) analyses of the supertrees based on parsimony trees (bold text), and based on maximum-likelihood trees (plain text).



GENOME COGs	MOT COGs*	*T.pallidum* interologs#
1254	50	84
579	39	73
634	39	104
1159	54	79
1081	41	55
1606	50	88
1767	50	91
1551	51	94
1314	48	91
1478	42	70
1493	40	80
1516	58	90
2140	55	86
1994	48	74
1994	56	80
559	26	32
580	31	51
2021	64	87
1864	61	87
2244	66	88
1917	65	79
1914	64	80
2085	43	70
1979	43	67
2178	49	69
1655	36	55
2099	52	74
1737	45	70
919	41	64
1088	48	71
	1455	

Figure 3.12 Bacterial flagellum supertree: The interologs were predicted for other bacterial species based on *T. pallidum* flagellum apparatus Y2H interactions and overlaid on flagellum supertree. Genome COGs, are the total number of functionally classified cluster of orthologous groups (orthologous protein families) of the respective species. For example, the *T.pallidum* genome has 634 orthologous protein families. MOT COGs, orthologous protein families involved in bacterial motility function (i.e., flagellum and chemotaxis). The *T. pallidum* flagellum apparatus Y2H screens yielded 104 interacting COGs. Based on these COGs interologs were predicted for other bacterial species. For example, out of 104 *T. pallidum* COGs, 90 interologs have been predicted for *E. coli*. In total we predicted 1455 interologs for 30 bacterial species.

3.5 Functional characterization of the interaction of TP0658 with flagellin (flaB1, flaB2 and flaB3)

3.5.1 TP0658 interacts with flagellin proteins in vivo and in vitro

TP0658 is a 17 kDa, protein of unknown function, associated with carbon storage regulator (CsrA) in an operon (Figure 3.13). CsrA is an RNA-binding protein that coordinates central carbon metabolism, activates flagellum biosynthesis and motility by activating the master operon for flagellum biosynthesis, flhDC (Romeo et al., 1993; Sabnis et al., 1995; Uchiyama 2003; Wei et al., 2001). The TP0658 ortholog Cj1075 in *Campylobacter jejuni* shows a motility defect when mutated (Golden and Acheson 2002), implicating a role of TP0658 in motility.

We found that all three flagellin proteins of *T. pallidum* (TP0868=flaB1, TP0792=flaB2, and TP0870=flaB3) interact with TP0658 in Y2H screens (Figure 3.14A). All three flagellin interactions of TP0658 could be verified in a far-western blot (overlay assay) (Figure 3.14B).

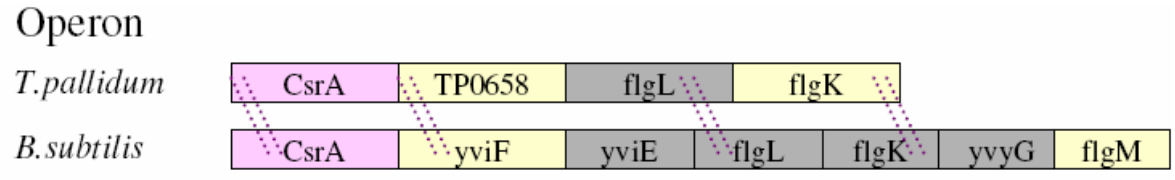
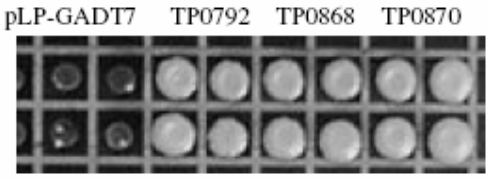
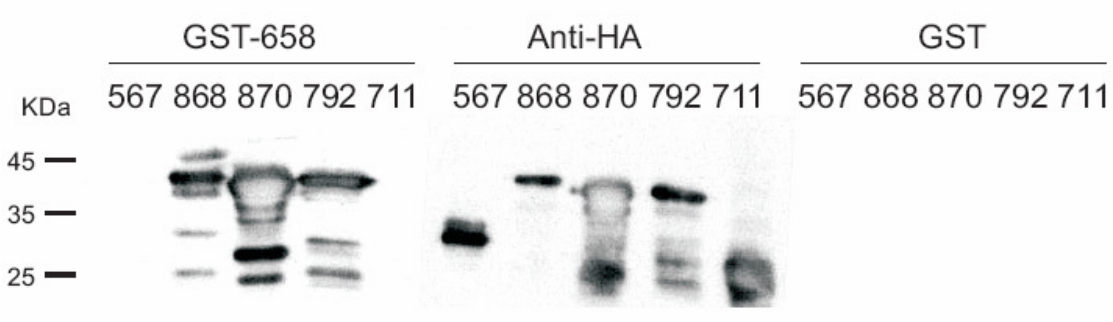


Figure 3.13 Carbon storage regulator A operon (CsrA) of *T. pallidum*, and *B. subtilis* showing the conservation of the TP0658-containing operon and its homolog in *B. subtilis*, yviF, in the CsrA operon.

A





B

Figure 3.14 *T. pallidum* protein TP0658 interacts with flagellin proteins (FlaBs): A, Retest of the Y2H interactions of TP0658 with *T. pallidum* flagellins, TP0792 (flaB2), TP0868 (flaB1), TP0870 (flaB3). TP0792, TP0868 and TP0870 were re-tested as a bait (DNA-binding protein fusion), and TP0658 as prey (activation domain fusions), and pLP-GADT7 vector as a negative control in quadruplicates. The rapidly growing cells on histidine-deficient medium indicate the Y2H protein-protein interaction. **B,** HA-tagged flagellin proteins of *T. pallidum* - TP0792 (FlaB2), TP0868 (FlaB1), and TP0870 (FlaB3) - were expressed in *E. coli* and tested for a protein interaction with GST-tagged, purified TP0658 in an overlay assay (GST-658) (see methods). The proteins TP0567 and TP0711 were included as negative controls. Probing with anti-HA and GST served as negative controls, the blot shows TP0658 interacting with all the flagellins (flaB1-flaB3) of *T. pallidum.*

3.5.2 The protein interactions of TP0658 with flagellins are conserved in other bacteria

Sequence orthologs of TP0658 were identified by the MBGD database (Uchiyama 2003). TP0658 is conserved in several bacterial species including *Spirochetes*, *Bacillales*, and *delta/epsilon Proteobacteria* (Table 3.14). To test whether the protein interactions of TP0658 with flagellins are conserved in other species, we used the *B. subtilis* ortholog of TP0658, namely yviF. We tested whether yviF interacts with the flagellin proteins of *B. subtilis*, hag and yvzB. Hag is the full length flagellin protein of *B. subtilis* (and a homolog of fliC in *E. coli*). yvzB represents an N-terminally truncated form with unknown function. The overlay assay showed that yviF interacts with both *B. subtilis* flagellin proteins. Additionally, yviF interacts with all three flagellin proteins of *T. pallidum* pointing to a conservation of the interaction epitope (Figure 3.15A). In addition, the interaction between yviF and hag could be confirmed *in vitro*, in an overlay assay using different *B. subtilis* deletion strains (Figure 3.15B). The truncated flagellin protein of *B. subtilis*, yvzB, could not be detected in this Western Blot (Figure 3.15B) probably due to insufficient yvzB expression under the conditions used or it may not be expressed at all.

Table 3.14 Ortholog proteins of TP0658 and Flagellins

***T. pallidum* genes**	**Species**	**Ortholog**	**Description**
TP0658	*B. subtilis*	yviF (BSU35380)	new assembly factor
	H. pylori	HP1154	
		HP1377	
	C. jejuni	CJ1075	
TP0792 (FlaB3) TP0868 (FlaB1) TP0870 (FlaB3)	*B. subtilis*	hag (BSU35360)	flagellin proteins
		yvzB (BSU35150)	
	E. coli	fliC (B1923)	
	H. pylori	flaA (HP0601)	
		flaB (HP0115)	
	C. jejuni	flaA (CJ1339C)	
		flaB (CJ1339C)	
		flaC (CJ1339C)	

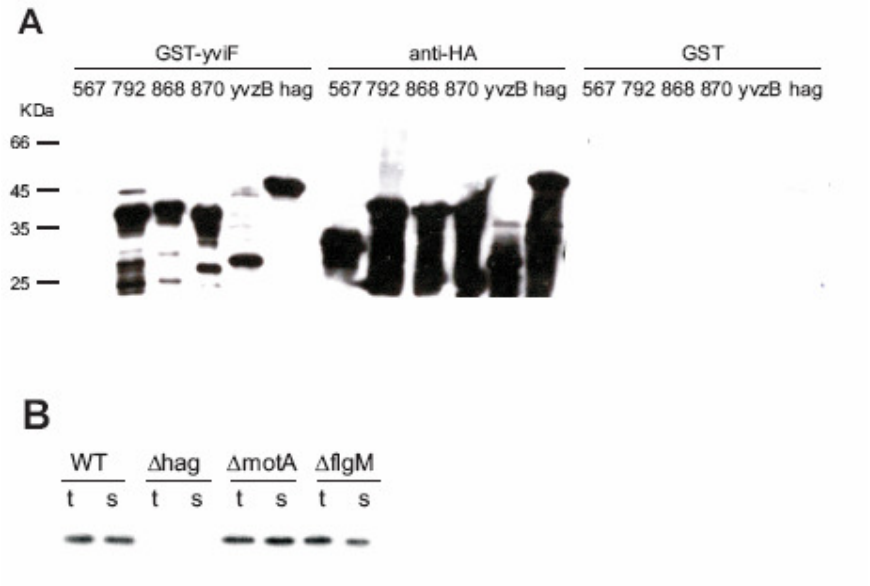


Figure 3.15 The *B.subtilis* ortholog of TP0658, yviF, interacts with flagellin proteins of *B.subtilis* and *T. pallidum*. **A**, HA-tagged flagellin proteins of *T. pallidum*, TP0792 (FlaB2), TP0868 (FlaB1), TP0870 (FlaB3), and *B. subtilis* flagellins, hag and yvzB, were expressed in *E. coli,* and tested for a protein interaction with GST-tagged yviF in an overlay assay (GST-yviF). The protein TP0567 was included as negative control. Probing with anti-HA and GST alone served as loading and negative controls respectively. **B**, Lysates of *B. subtilis* wild type strain (168), and a hag, motA and flgM mutant were tested for proteins interacting with purified GST-yviF in an overlay assay. The molecular mass of bands shown corresponds to hag (~ 31 kDa); total (t) and soluble (s) extracts are shown.

3.5.3 TP0658 interacts with C-terminal region of flagellin (flaB1)

As the structure of the flagellin protein is known (Samatey et al., 2001), we mapped the interaction epitope of the flagellin protein to get a structural insight into the function of this interaction, The multiple sequence alignment of the flagellins of *S. typhimurium, E. coli, T. pallidum, H. pylori, C. jejuni* and *B. subtilis* show high conservation in the N and C-terminus of flagellin in different species (Figure 3.16). In order to characterize the interaction epitope, I constructed 8 truncations of flaB1 of *T. pallidum* (Figure 3.17B) Combining systematic truncations of TP0868 (flaB1) with an overlay assay showed that TP0658 interacts with an epitope within the last 55 C-terminal amino-acids of TP0868 (L_{231} – C-terminus) (Figure 3.17C). In addition, the interaction of TP0658/yviF with the C-terminal half of flagellin is also supported by the interaction of yviF with the N-terminally truncated flagellin, yvzB, which is naturally lacking the region homologous of the first 110 amino-acids of flaB1 (TP0868) (Figure 3.15A).

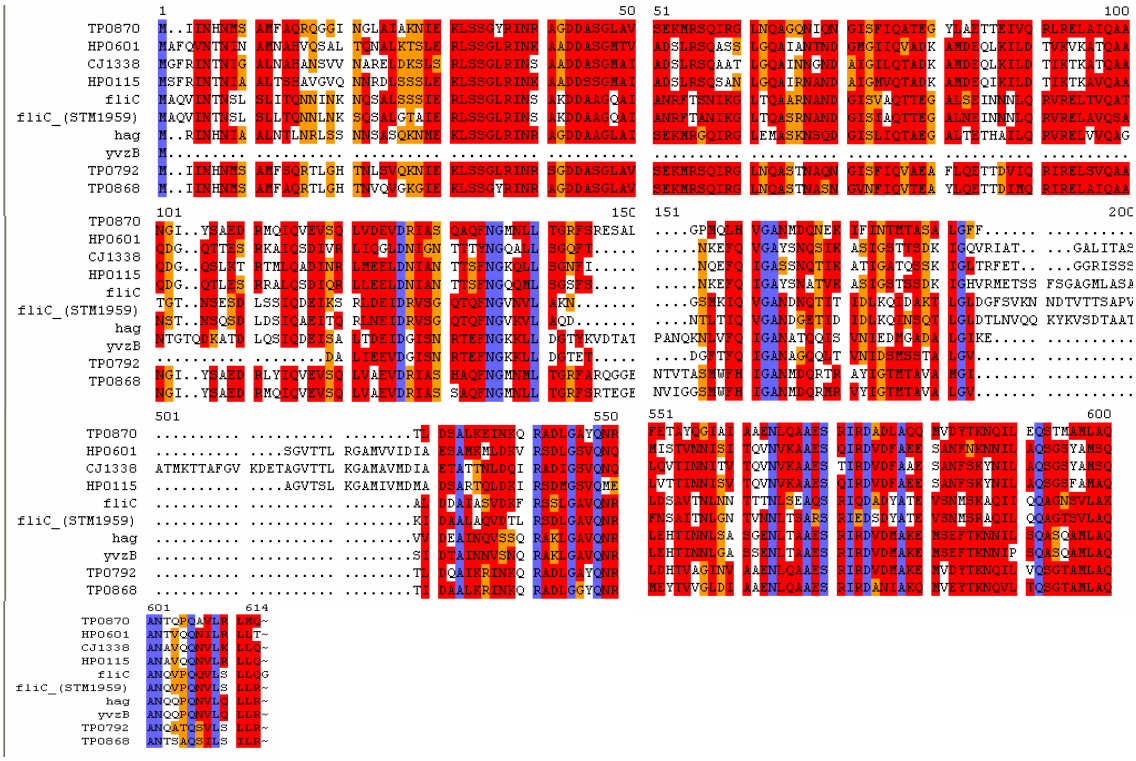


Figure 3.16 Multiple sequence alignment of flagellins: The multiple sequence alignment was performed using Clustal (Husar). The flagellins of *S. typhimurium* (fliC_STM1959), *E. coli* (fliC), *T. pallidum* (TP0792, TP0870, and TP0868), *B. subtilis* (hag and yvzB), *H. pylori* (HP0601), *and C. jejuni* (CJ1336) were used for the multiple sequence alignment. The identical amino acids are coloured in blue.

3.5.4 The interaction epitopes of TP0658 and FliS are similar

FliS is a flagellin-specific chaperone. The structure of fliS in complex with a fragment of fliC (flagellin) (Figure 3.18A) reveals that, like the type III secretion chaperones, flagellar export chaperones bind their target proteins in an extended conformation (Evdokimov et al., 2003). However, fliS adopts a novel fold that is clearly distinct from those of the type III secretion chaperones, indicating that they do not share a common evolutionary origin. It has been previously known that fliS binds to the C-terminal region of flagellins and prevents polymerization of flagellin (Ozin et al., 2003). The interaction of fliS in the C-terminal region of flaB1 (flagellin of *T. pallidum*) is confirmed in the Y2H assay (Figure 3.18B), confirming the conservation of the fliS-flagellin interaction in *T. pallidum*.

Strikingly, the identified interaction epitope of TP0658 is similar to the described interaction epitope of the fliS-flagellin interaction: fliS interacts with a region in the last 40 amino-acids of the C-terminus of *S. typhimurium* flagellin (Ozin et al., 2003). Even though TP0658/yviF and fliS do not show any protein sequence similarity with each other, they apper to compete for the same interaction epitope.

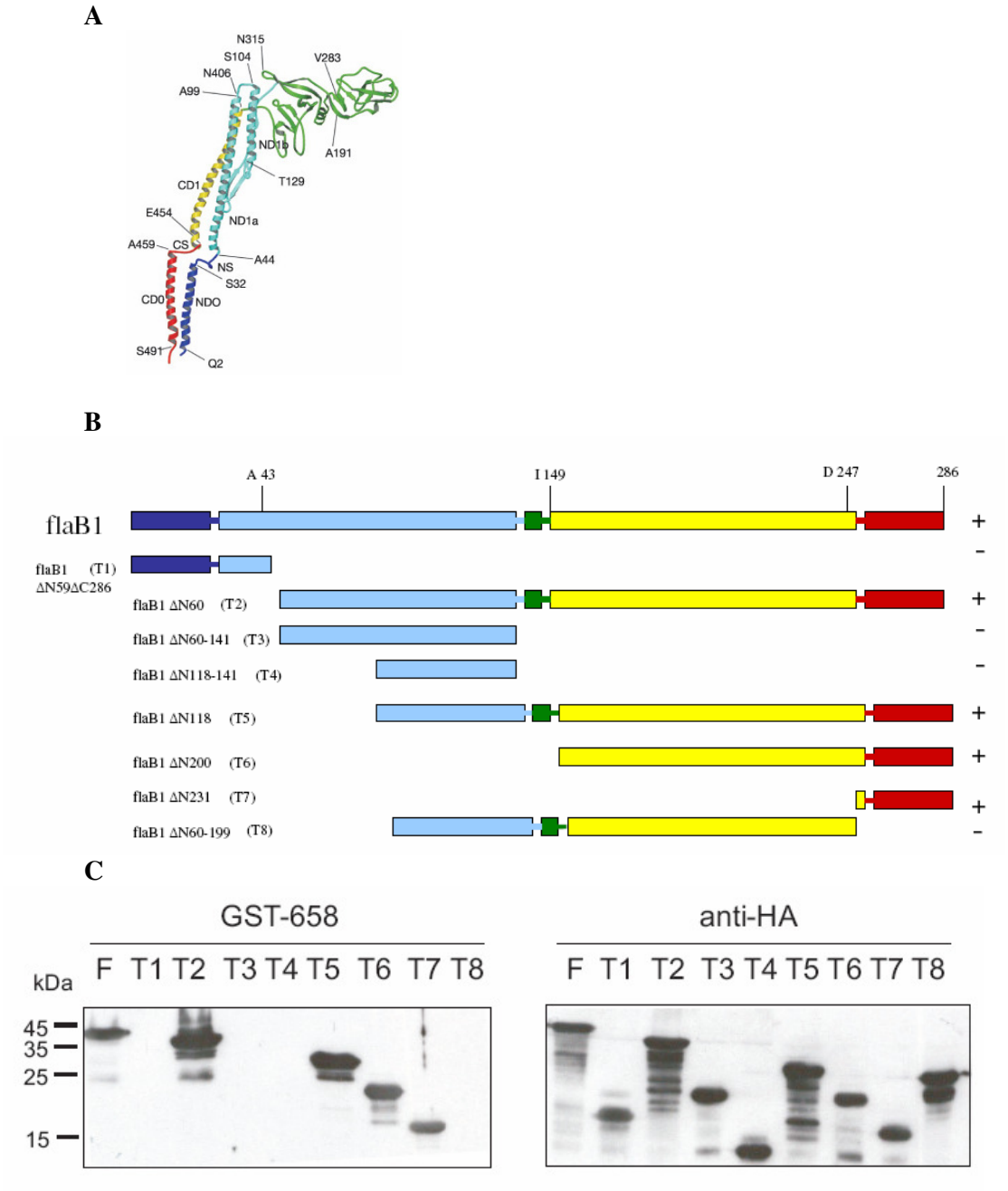


Figure 3.17 TP0658 interacts with the C-terminal region of flagellin (flaB1). A, the *S. typhimurium* flagellin (fliC), diagram of the Cα backbone. The chain is colored as follows: residues 1–44, blue; 44–179, cyan; 179–406, green; 406–454, yellow; 454–494, red (Yonekura et al., 2003). B, Schematic representation of TP0868 gene-truncation constructs, and results of interaction epitope mapping. The primary sequence of TP0868 is color-coded according to the diagram. Positive and negative interaction results for the fragments of flaB1 are indicated by " + " for positive interaction, and "- "for negative interaction (Picture courtesy of Bjoern Titz). C, Eight fragments of flaB1 (namely T1-T8) were tested for an interaction with purified GST-TP0658 protein in an overlay assay.

A B

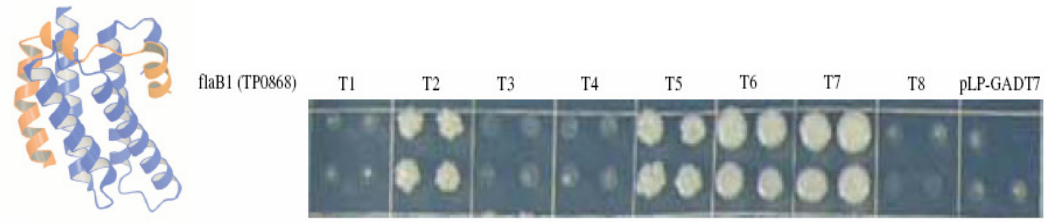


Figure 3.18 A, The structure of flagellar export chaperone, *Aquifex aeolicus* FliS (blue) in complex with residues 464-518 of fliC (orange) (from Evdokimov et al., 2003). B, Identification of the interaction between truncated fragments of TP0868 (T1-T8) in a Y2H assay: TP0868 truncation fragments (T1-T8) (see figure 3.17B) were tested as baits and fliS (TP0943) as prey. The pLP-GADT7 vector was used as a negative control. Yeast colonies in quadruplicates indicate the positive Y2H interactions.

3.5.5 TP0658 interacts with the sequence between L_{231} and D_{247} of TP0868 (flaB1)

For a detailed characterization of the interaction epitope of flagellin, we employed the SPOT peptide synthesis technology. Peptide arrays were used to identify the interacting epitope of flaB1 protein. The amino-acid sequence of the C-terminus of TP0868 was divided into peptides of 15 amino-acids with one amino-acid shifts. These peptides are synthesized on a cellulose membrane and the membrane is probed with purified GST-TP0658 fusion protein. The interacting peptide was detected by an anti-GST antibody followed by chemiluminescence. TP0658 interacted with peptides which form the sequence between L_{231} and D_{247} of flaB1 (TP0868) (Figure 3.19). The relevance of the interacting peptide was further demonstrated by a competitive peptide inhibition experiment (see below).

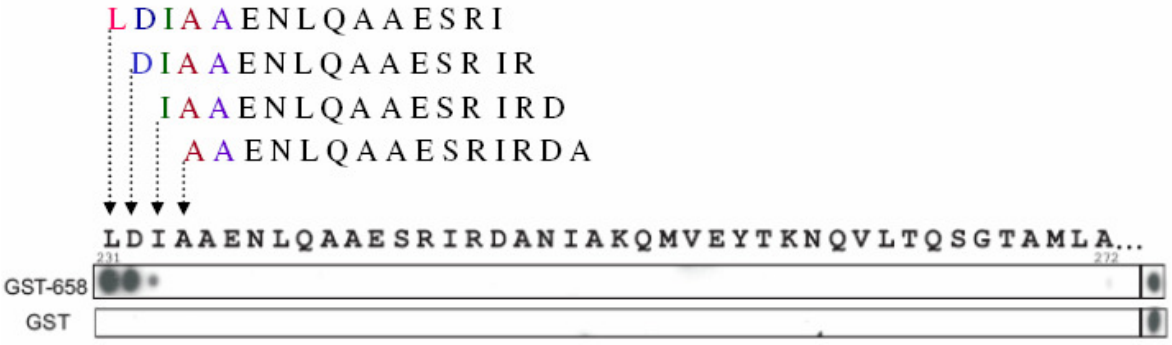
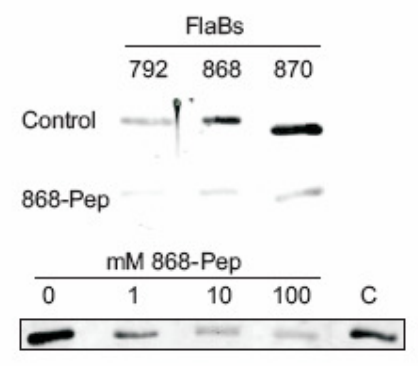


Figure 3.19 The C-terminal sequence of TP0868 (flaB1) was divided into peptides of 15 amino-acids with one amino-acid shifts (top panel showing an example of the amino acid sequence of each peptide spot). These peptides were directly synthesized onto a cellulose membrane and subsequently tested for an interaction with GST-658 or GST as a control. The first amino acid of each 15-mer peptide is shown above its spot. The last spot in each row is a positive control recognized by the anti-GST antibody (QRALAKDLIVPRRP). Only the first three peptides starting from L231 interacted, indicating an interaction epitope between L_{231} to D_{247}.

3.5.6 A C-terminal peptide of TP0868/flaB1 is sufficient to compete with TP0658 binding

We carried out a peptide competition assay to confirm the relevance of interacting peptides identified by the peptide spot approach. A synthetic peptide comprising the interaction epitope sequence VGL_{231}DIAAENLQAAESRIRD$_{247}$ was used for the competitive inhibition assay. The purified GST-TP0658 GST was pre-incubated with the respective concentration of inhibitory peptide VGLDIAAENLQAAESRIRD or an unrelated control peptide (DRRLADHFLGKI) in overlay buffer for 2h at 4°C and the overlay assay was carried out (see methods). A peptide comprising the interaction epitope sequence (VGL_{231}DIAAENLQAAESRIRD$_{247}$) was able to inhibit the binding of TP0658 to all three *T. pallidum* flagellin proteins and showed a dependency of the inhibition strength on the peptide concentration (Figure. 3.20).

Figure 3.20 Peptide competitive inhibition assay:
A synthetic peptide comprising the interaction epitope sequence of TP0868 (VGLDIAAENLQAAESRIRD), 868-Pep, was used in different concentrations, (1 mM, 10 mM, and 100 mM). The 868-Pep, was able to inhibit the binding of GST-658 to all three flagellin proteins, FlaB1 (TP0868), FlaB2 (TP0792), and FlaB3 (TP0870). Lower panel the inhibitory effect of 868-Pep was concentration-dependent (shown for TP0868, FlaB1) (Picture courtesy of Bjoern Titz).



3.5.7 A conserved Asparagine of hag (*B. subtilis*) and flaB1 (*T. pallidum*) is crucial for yviF/TP0658 binding

To identify single amino-acids crucial for binding of TP0658/yviF an alanine-scan was conducted using the peptide array approach. Each position of the identified peptide (VGLDIAAENLQAAESRIRD) was systematically replaced by alanine and the peptide membrane was probed with purified GST-TP0658 protein. The blot shows that I_{233} and N_{237} of the VGLDIAAENLQAAESRIRD interaction epitope are crucial for binding of TP0658 (Figure 3.21A). For additional verification and testing of evolutionary conservation of this interaction epitope in *B. subtilis* two targeted mutations of hag (flagellin protein of *B. subtilis*) were created. First, in the construct hag-N2A only the Asn residue at position 255 of hag was replaced by Alanine. Second, in the construct hag-HA the whole homologous interaction epitope N_{247}NLSASGENLTAAESRIRD$_{265}$ in hag was replaced by a HA-tag sequence (YPYDVPDYA) (Figure 3.21B). This residue is orthologus to N_{237} found to be crucial for TP0658 binding to TP0868 in the *T. pallidum* Peptide Spot analysis. Strikingly, an overlay with GST-yviF showed complete loss of binding of yviF to these mutated forms of hag (even when probing a large surplus of the mutants) (Figure 3.21C) – verifying the essentiality of Asn_{255} (Asn_{237} in TP0868) and demonstrating the evolutionary conservation of the interaction epitope identified for *T. pallidum* proteins.

The findings for the interaction epitope are summarized in Figure 3.21D. TP0658/yviF binds to an evolutionary conserved interaction region of flagellin proteins which is localized to the so

called Cα loop in the flagellin structure. A conserved Asn (position 237 of TP0868) is crucial for binding.

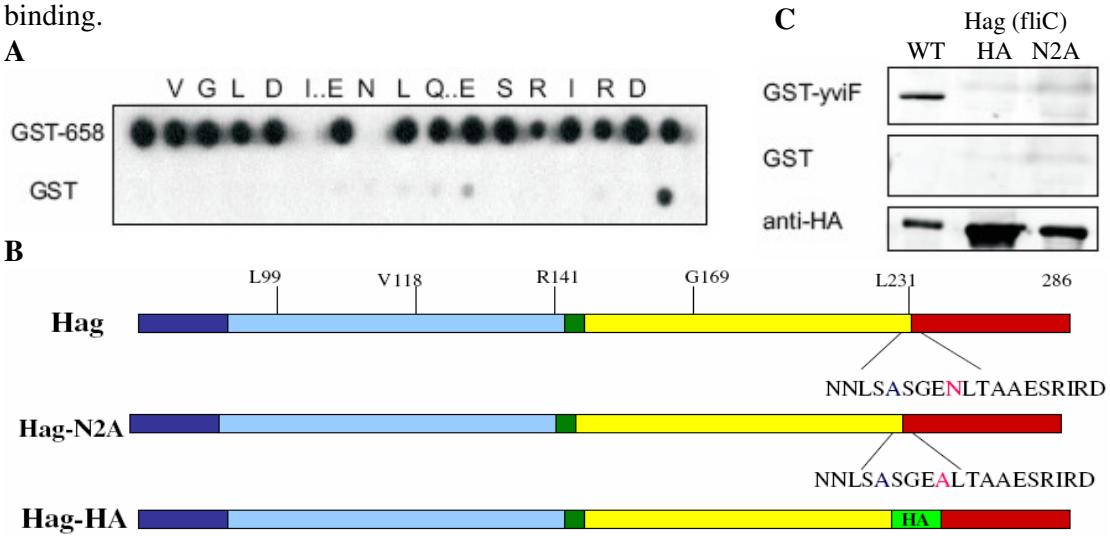


Figure 3.21 An extended interaction peptide (V_{229}-D_{247}) was tested in an alanine-scan showing that I_{233} and N_{237} are crucial for binding. A, the first and the last spot of the top row contain the wild type (VGLDIAAENLQAAESRIRD) and an antibody control sequence (QRALAKDLIVPRRP, recognized by anti-GST antibodies), respectively. The remaining peptide spots have the indicated amino-acid replaced by alanine, and ".." indicate positions not synthesized, because of naturally occurring alanine residues at these positions. **B**, Two targeted mutations of hag (flagellin protein of *B.subtilis*). First, in the construct hag-N2A only the Asparigin residue at position 255 of hag was replaced by Alanin. Second, in the construct hag-HA the whole interaction epitope N_{247}NLSASGENLTAAESRIRD$_{265}$ in hag, was replaced by a HA-tag sequence (YPYDVPDYA). C, Binding epitope mutants of hag (*B. subtilis* flagellin) were tested for binding to yviF in an overlay assay. The construct hag-HA has the whole interaction epitope replaced by a HA-tag. Hag.N2A has the crucial Asn_{255} residue replaced by alanine. Probing with GST protein and anti-HA antibodies served as controls (Picture **3.21** C, courtesy of Bjoern Titz).

3.5.8 yviF is involved in motility

A functional involvement of yviF in bacterial motility was tested using a *B. subtilis* mutant of yviF. The ΔyviF strain was created by specific integration of a phleomycin-upp cassette into the yviF locus as described by Fabret *et al.* (Fabret et al., 2002). However, despite several attempts the antibiotic resistance cassette could not be removed from the genome by the suggested counter-selection procedure. As an additional control the rescue of the phenotype by over-expression of yviF was conducted. The yviF mutant was non-motile in the swarming plate assay (Figure 3.22). Strikingly, only in presence of IPTG a plasmid for IPTG regulated expression of yviF (P_{spac} promoter) could rescue the motility phenotype, clearly proving that the reduced motility is due to lack of the yviF gene in the ΔyviF mutant strains.

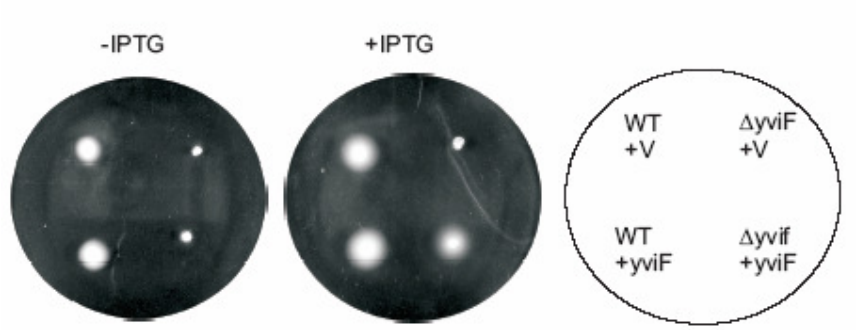


Figure 3.22 A **yviF mutant is non-motile.** The *B. subtilis* ΔyviF mutant was tested in a swarming assay without (-IPTG) and with (+IPTG) induction of HA-tagged yviF expression from a plasmid (+yviF). *B. subtilis* cells transformed with the empty vector, pDG148-Stu, served as controls (+V). The ΔyviF mutant has a clear swarming defect that can be rescued by yviF expression (Picture courtesy of Bjoern Titz).

3.3.9 TP0658/yviF is a new assembly factor of the flagellum

Concluding from the protein interaction epitope we thought TP0658/yviF might be a previously unidentified assembly factor of the flagellum, specifically for flagellin proteins. Consequently, we tested the direct effect of TP0658/yviF on flagellin proteins. Deletion of yviF leads to a strong reduction of the amount of flagellin protein detected in *B. subtilis* cells (Figure 3.23A), probably due to degradation of flagellin which cannot be incorporated into the nascent filament. Strikingly, a similar stabilization was found when TP0658 and TP0868 were co-expressed in *E. coli* supporting a direct stabilization effect of TP0658 mediated by the identified protein interaction (Figure 3.23B).

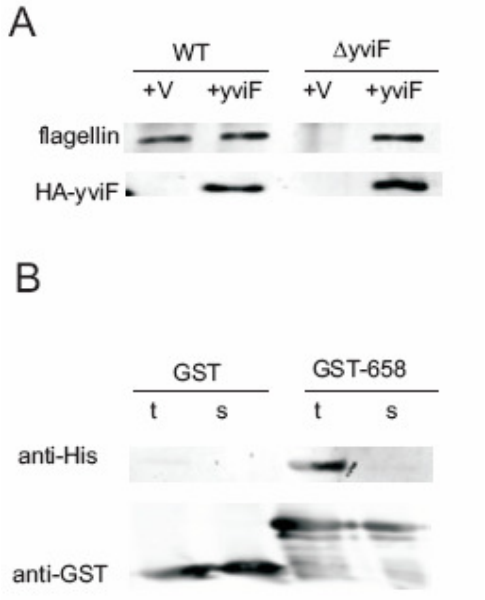


Figure 3.23 TP0658 and yviF stabilize flagellin. A, *B. subtilis* strains were tested for yviF expression from a plasmid (anti-HA) and flagellin (hag) expression (GST-yviF overlay). Wild type (WT) and ΔyviF cells carrying an empty plasmid, pDG148-Stu (+V), or the yviF expression plasmid, pDG-yviF (+yviF), are compared. B, His-tagged TP0868 (FlaB1) and GST-tagged TP0658 or GST alone were co-expressed in *E. coli*. Total (t) and soluble (s) lysates were tested. Co-expression of TP0658 leads to TP0868 stabilization (Picture courtesy of Bjoern Titz).

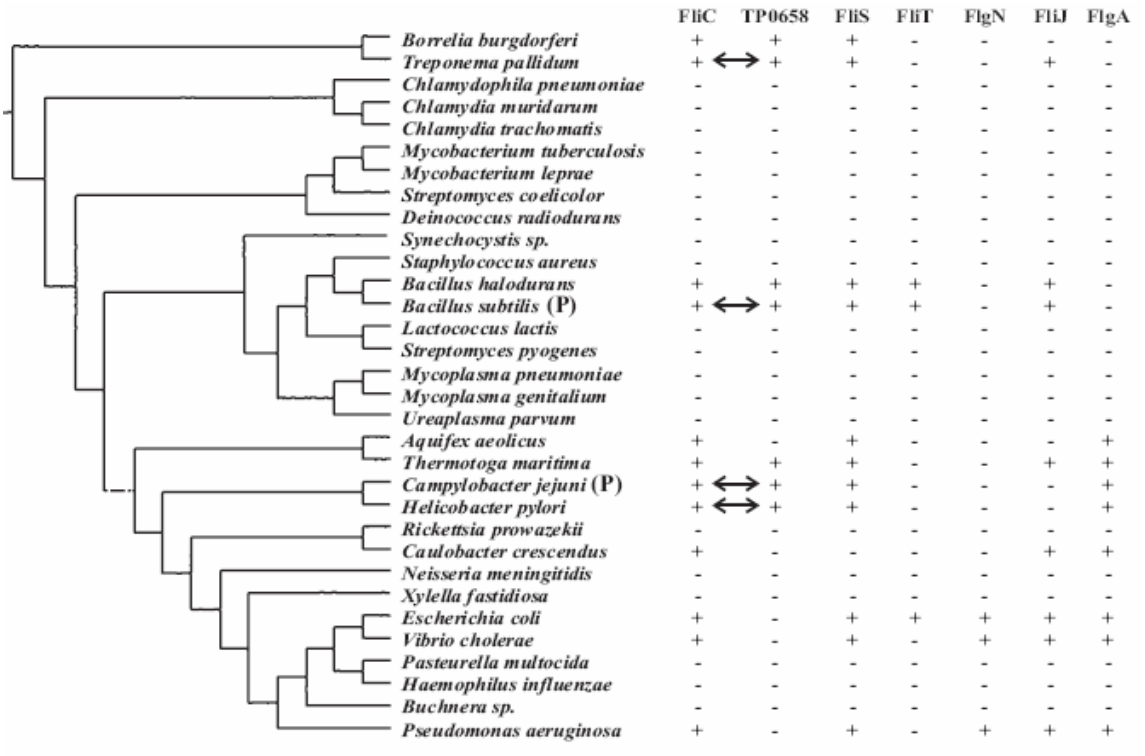


Figure 3.24 Conservation of TP0658 and its interactions in bacteria. A phylogenetic tree (Daubin et al., 2002) shows the presence of TP0658 and other flagella-related chaperones: FliS is the main chaperone for flagellin (FliC), FliT a chaperone for the filament capping protein FliD, FlgN for the hook-filament junction proteins FlgK and FlgL, and FliJ for rod and hook proteins (Auvray et al., 2001; Bennett et al., 2001; Fraser et al., 1999; Yokoseki et al., 1995). FlgA is an assembly factor for the P-ring. Interactions between FliC and TP0658 are indicated by double-arrows. Mutations of TP0658 homologs in *Bacillus subtilis* and *Campylobacter jejuni* are known to exhibit a motility phenotype (P).

3.5.10 Purification of TP0658 protein for Crystallization

TP0658 is a 17 kDa protein of previously unknown function. Even though fliS and TP0658 protein do not show any primary sequence similarity they are interacting in the same region of flaBs (flagellins). The structure of fliS has been determined by X-ray crystallography. It will be interesting to see if the two proteins are structurally similar despite a lack of primary sequence similarity. This motivated us to attempt expression and crystallization of TP0658. PCR-amplified the TP0658 from a pUniD vector with an endogenous stop codon and cloned the ORF into the pETM30 expression vector (N-His N-GST C-His). After cloning, the TP0658 gene sequence was verified by sequencing. The protein was expressed in *E. coli* BL21 (DE3), and the HIS-GST-TP0658 protein was purified by GST-beads (Amersham Biosciences) (Figure 3.25), followed by Ni-NTA Magnetic Agarose Beads (Qiagen), and finally the N-terminal HIS-GST tag was removed by HIS-TEV protease treatment (Figure 3.25).

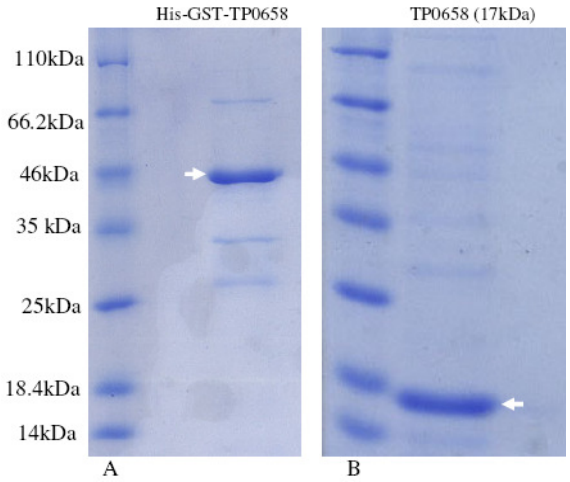


Figure 3.25 Coomassie Blue-stained SDS-polyacrylamide gel. A, Lane 1, molecular weight standards. Lane 2, GST purified TP0658 protein. From *E. coli* BL21 (DE3) expressing the N-HIS, N-GST-TP0658. B, A; Lane 1, molecular weight standards. Lane 2 purified TP0658 (17 kDa), after removing the N-HIS, N-GST tag by TEV protease.

4. Discussion

4.1 The *Treponema pallidum* flagellum interactome

The study presented here is the first comprehensive attempt to identify bacterial flagellum apparatus protein-protein interactions (PPIs) on a genome-wide scale. As elementary constituents of cellular protein complexes and pathways, protein interactions have been shown to be powerful predictors of protein function. This study mainly aimed to identify all the proteins associated with chemotaxis signalling and flagellum function in a qualitative way using the yeast two-hybrid system.

We selected *Treponema pallidum,* the syphilis spirochete, as a model organism for flagellum interactome analysis. The flagellar interaction screen was a pilot project for a systematic genome-wide screen for all protein-protein interactions in a bacterium. The *T. pallidum* was chosen because of its small genome size, encoding 1041 predicted ORFs (Fraser et al., 1998), and the availability of an ORFeome i.e., all open-reading frames cloned into a versatile vector system (McKevitt et al., 2003). The PPI data of *T. pallidum* flagellum can be transferred to other model organisms such as *E. coli* and *B. subtilis* because a significant number of proteins has homologs in other bacterial species (Figure 4.1). Additionally we can learn about the common and specific flagellum components in spirochetes and other model organisms, because motility in spirochetes has several unique features and is more complex than the well-studied paradigms of *E. coli and S. typhimurium* (Charon and Goldstein 2002), and there are many unknown factors with respect to motility and chemotaxis in spirochetes.

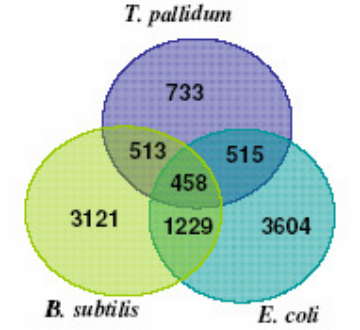


Figure 4.1 Venn diagram showing conserved orthologous groups (COGs) of proteins in *T. pallidum, E. coli* and *B. subtil*is. COG orthologues are identified using an all-against-all sequence comparison of the proteins encoded in completely sequenced genomes (source: String database v6.3)

4.1.2 How comprehensive is the interaction map of *T. pallidum* flagellum apparatus?

I collected all the proteins annotated with ‘‘chemotaxis and motility” as well as also proteins predicted to be functionally associated with the flagellum (Rain et al., 2001; Kanehisa et al., 2002; von Mering et al., 2003). In total 75 proteins were selected as flagellum baits (Table 3.1). A comprehensive genome-wide Y2H array-based two-hybrid screening was done with a *T. pallidum* Y2H proteome library, for all the selected flagellum and chemotaxis proteins, and the interacting proteins were identified.

Two-hybrid screening is challenging due to potentially high false-positive and false-negative error rates. Filtering of raw Y2H results significantly improves the data quality of the PPIs (see methods). Filtering the raw Y2H interaction data revealed 268 unique PPIs involving 174 unique proteins (Table 3.2). Of these identified 268 PPIs only 15 interologous interactions were previously described by other studies (Table 3.4).

4.1.3 False-negatives

In high-throughput PPI studies false positives (that is, reporter gene activity where no specific protein-protein interaction is involved), and false negatives (failure to detect physiologically relevant interactions) are nearly inevitable. Numerous two-hybrid studies showed that the two-hybrid system never detects all physiological interactions. However, this limitation is also true for any other method.

False negatives have been more difficult to address. Estimates of false-negative rates based on comparisons to protein interaction lists culled from the literature are biased by the examples selected for analysis and by systematic differences between screens.

Aloy and Russell showed that the two-hybrid system tends to detect transient interactions, whereas interactions within protein complexes are more efficiently detected using purified complexes in combination with mass spectrometric analysis (Aloy and Russell 2002). Edwards et al, estimated the number of false negatives in two-hybrid screens to be between 43% and 71%, based on a comparison of two-hybrid data and the crystal structures of the proteasome and the Arp2/3 complexes, whose protein pairs were systematically tested by two-hybrid analysis (Edwards et al., 2002).

In order to estimate the false negative rates of the *T. pallidum* flagellum apparatus Y2H screens, I compared the flagellum Y2H and the other known flagellum PPIs in several bacterial species. A comprehensive literature review for known flagellum protein-protein interactions in PubMed yielded 62 interactions (Table 1.7). Based on literature interactions 54 interologs were predicted for *T. pallidum* (i.e., pairs of interacting proteins that interact identically in two species). Strikingly, our Y2H screens reproduced 27.9% (15 out of 54 interactions) of the known interactions, with a false negative rate of 72% (Table 3.5).

I estimated the false-negative rate for other two independent high-throughput systematic PPIs studies. Recently, Russell L. Finley and coworkers completed a LexA-based yeast two-hybrid interactome of *Campylobacter jejuni*, (Russell L. Finley, pers. comm) which identified 761 protein-protein interactions for motility proteins. Our curated literature dataset predicted 29 interologs for *Campylobacter jejuni.* Of these, 21% (i.e., 6/29) interactions were found in the *Campylobacter jejuni* Y2H screens, with false negative rate of 79% (Table 3.5).

I compared the Arifuzzaman et al., complex purification data from *E. coli* (H. Mori, pers. comm) with known flagellum protein-protein interactions. In total they identified 11526 PPIs, among which 564 interactions belong to the flagellum subset, 57 interologs were predicted for *E. coli* based on known interactions. Of these only 5.3% (3/57) interactions were found in the *E. coli* pull-down approach (Table 3.5). Surprisingly, 94 % and 98 % of our (*T. pallidum)* and Russ Finley's (C. *jejuni*) two-hybrid interactions are not represented in the protein complexes, i.e. these interacting proteins are not found in the same complex. However, many interactions reported in the literature were not found in this Y2H screen. For example, chemotaxis proteins did not reproduce many known interactions, possibly because post-translational modification-dependent bacterial PPIs can not be detected in the yeast system. For example, phosporylated cheY binds to fliM and its phosphatase CheZ but we could not detect these interactions. Other possible reasons for the high false negative rate may be mislocalisation and sterical constraints, which may not be fulfilled by all protein interaction pairs. "True positives" can be defined operationally based on their reproducibility in related species, for example in *E. coli, S. typhimurium*, or others, as well as by independent methods that confirm these interactions (see

below). The high rate of false negatives in complex purification data from *E. coli* may be due to the large number of membrane-associated or transmembrane proteins of the flagellum apparatus which are difficult to purify.

4.1.4 Reliability of Y2H interactions

Several studies suggest that interaction networks assessed by two-hybrid screens can be fairly reliable. In the study on the initial version of human interactome by Stelzl and colleagues evaluated the quality of their Y2H data by verifying a random sample of interactions using co-immunoprecipitation assay, in which 116 protein pairs were tested with a success rate of 72/116 (62%) (Stelzl et al., 2005). Similarly, Rual and colleagues on the human interactome, found a verification rate of approximately 78% of their yeast two-hybrid interactions by an independent co-affinity purification assay, and significant correlations with other biological attributes (Rual et al., 2005). Recently, we carried out a systematic array based Y2H screening of Kaposi sarcoma–associated herpes virus (KSHV), and ~50% of the array-based Y2H, PPIs could be confirmed by co-immunoprecipitation, suggesting the accuracy of most of the two-hybrid results (Uetz et al., 2005). The failure of co-immunoprecipitation does not necessarily indicate that the two-hybrid results are false positive, as structural constraints might play a different role in co-immunoprecipitation.

4.1.5 Comparative analysis of the overlapping *T. pallidum* Y2H interactions with different high-throughput studies

In order to identify the overlapping interactions of different high throughput flagellum PPIs datasets, I compared *T. pallidum* flagellum interactions with other large-scale protein-protein interaction data-sets, as well as with 'functional association predictions'.

Rain et al. published a partial protein interaction map of *H. pylori,* testing 261 baits against a genomic fragment library (Rain et al., 2001). Only 8 orthologous baits were in common to both *T. pallidum* and *H. pylori* Y2H studies (Table 3.6). 14 interactions for these baits were identified for *H. pylori* and 36 for *T. pallidum*, and 6 interactions were overlapping in both datasets (Table 3.7) (i.e. 6/36 = 17% of the *Treponema* interactions and 6/14 = 43% of the *Helicobacter* interactions). Surprisingly, even though the methods used for *Campylobacter jejuni* and our *T. pallidum* protein interaction studies are similar, the overlap between the two dataset is very low (4 out of 131 = 3 %) (Table 3.7). A possible explanation for the small overlap may be the somewhat different screening protocols. Even though the principle of the method is the same in both Y2H studies, Finley et al. used a LexA-based Y2H system, in which the DNA-binding domain (DBD) is provided by the entire prokaryotic LexA protein. In addition, the reporter was usually lacZ and LEU2, and the AD (activation domain) was typically a bacterial sequence called B42 (Kolonin et al., 2000). By contrast, our Y2H study is based on the DNA-binding and activation domain of the yeast Gal4 protein.

The overlap between the *E. coli* PPIs based on pull-down assays (Arifuzzaman et al.) with our and Russ Finley's two-hybrid studies is 3.82% (using a spoke model for the complex data). The reason for very small overlap between different methods is largely unknown, although it may have to do with the significant number of flagellum proteins that are membrane-associated or transmembrane proteins. They are likely to give no or different results when compared to two-hybrid assays, simply based on their different localization and membrane affinity.

We systematically compared how many predicted functional associations for the known flagellum baits, based on three types of genomic context associations (conserved genomic neighborhood, gene fusion events, and co-occurrence of genes across species), can be validated by Y2H screens. The *T. pallidum* Y2H screens confirmed 4 %, the *C. jejuni,* and *H. pylori* Y2H screens ~ 2 %, of the protein-protein associations predicted by the String database (von Mering et al., 2003). Strikingly, when I combine the PPIs of the three datasets, the confirmed string predictions increased significantly (form 4% to 7%), which indicates that the interactions identified in *T. pallidum, C. jejuni* and *H. pylori* Y2H screens are complementary to each other, at least for the String-predicted functional associations (Figure 3.2).

4.2. Novel intra-flagellum interactions

The *T. pallidum* flagellum Y2H screening revealed 19 novel PPIs between known components of the flagellum (Figure 3.5) which revealed new structural and functional insights into the flagellum complex. The MS ring (FliF) and the type-III flagellar export system components assemble first, probably in a coordinated fashion. The rotor/switch assembles next. FliG – as also supported by the found FliF-FliG interaction – firmly attaches to the MS ring (Figure 3.5); FliG undergoes a self-interaction as it builds a homopolymeric complex of 26 subunits (Jones and Macnab 1990). The Y2H screening shows the direct interaction between FliG and FliM (a component of C-ring) and FliG and FliY (a motor switch protein containing a C-terminal SpoA domain like FliM) (Pallen et al., 2005a), indicating that C-ring proteins directly mount on the FliG structure. The flagellum motor proteins (motA, motB) – stator components of the flagellum motor – are thought to assemble after C-ring formation (Macnab 2003). The MotA dimer was identified in our screen supporting a similar interaction of *Vibrio* PomA, a functional ortholog of motA driven by sodium ions instead of protons (Sato and Homma 2000). Not much is known about the requirement of assembly factors for motor proteins, but the interaction between MotA and FliR points to an involvement of at least one flagellar export protein. FliR and FlhB are membrane proteins which physically interact with each other and are necessary for flagellar export (Van Arnam et al., 2004). FliE is probably the first protein using the type III export pathway, forms a complex of ~9 subunits (FliE-FliE interaction) (Muller et al., 1992), and is thought to mediate the connection between the MS ring and the rod (Macnab 2003). The indirect interaction with the MS-ring mediated by the FliE-FliG interaction needs to be evaluated in more detail; a direct interaction between FliE and FlgB supports previous findings that FlgB forms the proximal part of the rod (Minamino and MacNab 2000). FlgG, thought to be the most distal part of the rod (Okino et al., 1989), interacts with FlgC supporting the following order of rod proteins: FliE, FlgB, FlgC, and FlgG in contrast to common wisdom which has an inverted order of E-C-B-G. Alternatively, these proteins may not be organized serially but in a more complex structure.

Rod proteins are known to interact with enzymes having muraminidase activity like FlgJ (Hirano et al., 2001). Although only a low confidence Y2H interaction between FlgJ and FlgG was identified in the screen, the interaction between murC (UDP-N-acetylmuramate-alanine-ligase) and FlgG demonstrates that rod proteins interact with more than one peptidoglycan-modifying enzyme. As supported by the FlgE-FlgD interaction, FlgE monomers polymerize under the capping protein FlgD (Ohnishi et al., 1994). After addition of hook-associated proteins, the flagellin proteins assemble to form the filament, assisted by export chaperones like FliS (FliS-

flagellin interaction) and FliJ (FliJ-flagellin interaction). The export of other components relies on the presence of additional soluble factors, e.g. on the ATPase FliI and its regulator FliH. These proteins are thought to dock to the cytoplasmic parts of FlhAB (Gonzalez-Pedrajo et al., 2002). The interaction between FliH and FliG indicates an additional way of docking to the membrane.

FlgM is an anti-sigma factor regulating flagellin synthesis. Recently Pallen et al. (2005) identified a remote FlgM homolog in *T. pallidum*, TP0974 (Pallen et al., 2005a). The interaction between TP0974 and a sigma-factor (TP0709) is the first experimental evidence, that TP0974 is also functional in *T. pallidum* as an anti-sigma factor.

4.3 New components of the bacterial flagellum complex

Interestingly, my Y2H screens with flagellar baits identified many interactions with conserved hypothetical proteins, i.e. proteins of as yet unknown function. These interactions thus predict a function of these proteins in motility. *T. pallidum* does not allow us to verify such predictions experimentally, but we can use our interaction data to predict interactions in more tractable organisms such as *E. coli* or *B. subtilis*. We identified 72 uncharacterized proteins interacting with proteins known to be involved in flagellum function (Table 5.2 in appendix). To characterise the functional relevance of the identified *T. pallidum* Y2H protein-protein interactions during bacterial motility, orthologs of all *T. pallidum* interacting proteins were predicted in *E. coli* and *B. subtilis* and I have identified target genes for gene deletion experiments base on COGs (cluster of orthologous groups of proteins) (Table 3.10). Bacterial motility assays were done for all these deletion mutants in *E. coli* and for 7 *B. subtilis* mutants. Of these 5 *B. subtilis* mutants and 10 *E. coli* mutants showed an altered motility phenotype (Table 3.13 and Table 3.12). For TP0712 a conserved hypothetical proteins associated with flagella by our Y2H interactions, an altered motility phenotype has been reported in the literature (Table 3.13), supporting the functional relevance of identified protein-protein interaction in bacterial motility.

These included 8 uncharacterized proteins, for which I assigned a function in motility based on their motility defect in the orthologous gene mutant and their physical two-hybrid association with known flagellum proteins (Table 4.2).

Table 4.2 New components of the bacterial flagellum

T. pallidum ORF	Orthologous gene mutant	Interacting proteins	Motility phenotype
TP0046	yaaT (BSU)	cheR, flgD metK, flaB3, TP0959, cheW-2, mcp2-3	**-/+**
TP0048	yhbE and YhbF (BSU)	fliY, fliS	**-/+**
TP0383	yllB (BSU)	fliS, flgK	**-/+**
TP0421	yncE (ECO)	TP0567	**-/+**
TP0561	ydjH (BSU)	SigG, fliF, flhB, fliR, fliQ, fliL	**-/+**
TP0658	yviF (BSU)	flaB2, flaB1, flaB3	**-**
TP0712	HP1034 (HPY)	TP0464	**-**
TP0979	ycfH (ECO)	FliE	**-/+**

BSU: Orthologues protein in *B. subtilis*, ECO: *E. coli*, HPY: *H. pylori*. Phenotype: "-/+" reduced motility, "-" non-motile.

4.3.1 *B.subtilis* Mutants

Table 4.3 *B.subtilis* mutants

T.pallidum ORF	*B.subtilis* Gene name	Protein sequence identity	Motility Phenotype	Interacting proteins
TP0658	yviF	36%	-	flaB2, flaB1 flaB3
TP0383	yllB	33%	-/+	fliS, flgK
TP0911	ylqH	35%	-/+	flgG
TP0048	yhbE and yhbF	34%	-/+	fliY, fliS
TP0046	yaaT	39%	-/+	cheR, flgD metK, flaB3, TP0959, cheW-2, mcp2-3
TP0561	ydjH	<20%	-/+	SigG, fliF, flhB, fliR, fliQ, fliL

Phenotype: "-/+" reduced motility, "-" non-motile.

TP0658 (yviF in *B. subtilis*), for example, was found to interact with all three flagellin proteins, FlaB1-B3, of *T. pallidum*. Further evidence indicates that TP0658 might be an assembly factor for the flagellar filament (described below).

TP0383 / yllB (*B. subtilis*) is an uncharacterised protein, predicted to be functionally associated with cell division genes (the yllB gene is located next to the genes of cell division proteins ftsW, ftsZ and ftsA). TP0383 is highly conserved in *E. coli* and *B. subtilis*. The disruption of the gene had no detectable cell division phenotype (Daniel et al., 1996). The predicted functional association of TP0383 thus was not in agreement with the observed phenotype. The structure of its ortholog in *M. pneumoniae*, mraZ (= yabB), has been solved and shows a hexameric ring-like structure (Figure 4.2), supporting the self-interaction of TP0383 found in our Y2H screens (Chen et al., 2004). Interestingly, TP0383 specifically interacted with flagellin-specific chaperone fliS, hook associated protein flgK, rod protein fliE, and flagellin flaB3, as well as with an enzyme involved in cell wall metabolism, murF. A functional relevance of TP0383 for motility is demonstrated by the observed motility phenotype of the *B. subtilis* ortholog of TP0383, ylllB (Figure 3.9B).

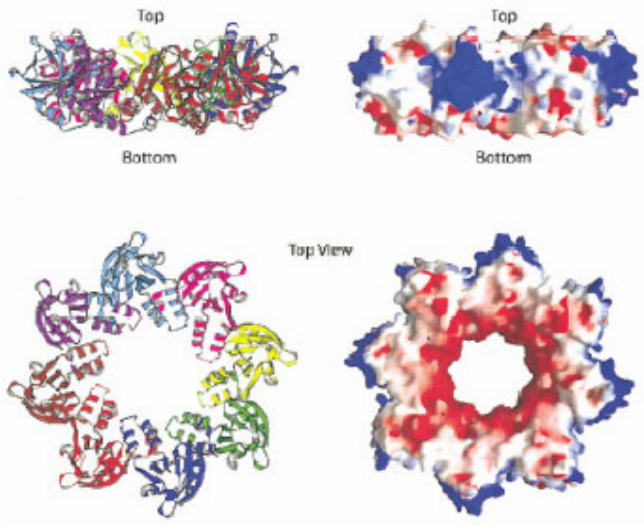


Figure 4.2 Overall structure of marZ (UPF0040) from Mycoplasma pneumoniae: Ribbon representations are shown on the left panels. The octamer is colored by chain. The electrostatic potential surface is shown on the right panels. Upper panel showing side view; lower panels showing top view (from Chen et al., 2004).

TP0911/ylqH shows homology to the cytoplasmic part of FlhB (Figure 4.3), one of the flagellum-specific export proteins that regulates substrate specificity. In our Y2H screens TP0911 specifically interacted with flagellar basal-body rod protein (flgG), suggesting a motility function. Using the *B. subtilis* ortholog, ylqH, find reduction in motility (Figure 3.9B) indicates ylqH might be an export protein.

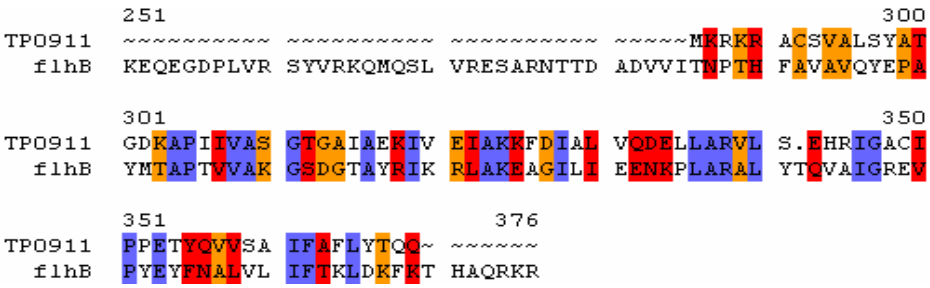


Figure 4.3 Multiple sequence alignment of TP0911 and flhB of *T. pallidum*. The sequence alignment was performed using the Clustal tool of Husar. The identical amino acids in both TP0911 and flhB are coloured in blue. TP0911 shows significant sequence similarity with the C-terminal cytosolic domain of flhB.

4.3.2 *E. coli* mutants

I used a collection of *E. coli* gene deletions (courtesy of H. Mori) to identify genes involved in motility. Mutants that showed a motility defect in swarming assays were then compared to the set of proteins known to be involved in chemotaxis, flagellum function, or associated with flagellar proteins in our Y2H screens. Interestingly out of 49 proteins known to be involved in flagellum / chemotaxis, 37 gene mutants show a non-motile phenotype, and 6 mutants show a phenotype with reduced motility, indicating these proteins are crucial for bacterial motility (Table 3.11). Additionally, 10 *E. coli* mutants, orthologs of *T. pallidum* which are associated with flagellar proteins in our Y2H screens, show a motility defect, implicating their function in motility (Table 3.12).

amiA and amiD: The N-acetylmuramoyl-L-alanine amidases (AmiA, AmiB and AmiC) are periplasmic enzymes capable of destroying cell wall cross-links by cleaving the peptide moiety from N-acetylmuramic acid. It was shown recently that all the Δami cells form chains, indicating that the amidases contribute to daughter cell separation by helping to split the septal murein (Heidrich et al., 2001). TP0247 (an ortholog of *E. coli* amiC and amiA) interacts with flaA, which forms the flagellar filament outer layer protein in spirochetes. The Y2H interaction of flaA and amiC, and the non-motile phenotype of the amiC mutant and the reduced motility of amiA in *E. coli* indicate a functional link between peptidoglycane remodelling and flagellar filament formation. However, the nature of this relation remains unclear.

4.3.3 Motility – gene regulation

In the past few years, a lot of data has been published about the regulation of motility in polarly and laterally flagellated bacteria. However, the mechanism of motility control by environmental factors and by some regulatory proteins remains largely unknown. The organisation of the flagellar system has been extensively studied in enterobacteria and multiple levels of *flhDC* regulation were observed, such as transcriptional and posttranscriptional control in *E. coli*. Our

Y2H screens revealed two putative proteins that may be involved in the regulation of flagellar gene expression.

RfaH is a virulence regulator of enterobacteria that functions as a transcriptional anti terminator (Artsimovitch and Landick 2002; Bailey et al., 2000). Originally, RfaH was discovered as a regulator involved in the synthesis of LPS-core in *Salmonella enterica* and *E. coli* (Lindberg and Hellerqvist 1980). RfaH-affected operons include those encoding the F-factor, hemin uptake receptor, as well as the toxins alpha-hemolysin (Bailey et al., 1992; Leeds and Welch 1996). Virulence of the prototype uropathogenic *E. coli* strain 536 was recently shown to be abolished through inactivation of rfaH (Nagy et al., 2002). RfaH plays a role in the infectious process of *E. coli* as early as the colonization of the intestinal tract. The role of RfaH in bacterial motility is unknown. The physical interaction of RfaH with flagellar class III gene (flgG) and the non-motile phenotype of RfaH mutant in *E. coli* indicate the potential regulatory role of RfaH in motility gene expression. This hypothesis may be tested by using DNA microarray measurements of RfaH and/or flgG mutants.

TP0979 Interacted with fliE. Its ortholog in *E. coli*, TatD, has two additional paralogs, ycfH and yjjV; both are functionally unassigned and unlinked. *ycfH* shows 29% amino acid sequence identity and *yjjV* shows 24% amino acid sequence identity to TatD. As TatD is localized in an operon with genes of the twin-arginine transport (Tat) system, a transport function of TatD was anticipated. But even a strain with all three TatD paralogs deleted did not show a Tat-related transport deficiency – leaving the question for TatD's function unanswered (Wexler et al., 2000). As we found no significant decrease in motility for the TatD single mutant, we thought to test the previously described triple mutant (all TatD paralogs deleted) for motility. Unfortunately, the triple mutant was constructed in a MC4100 strain background, a strain known to be non-motile, due to an FlhD (master regulator for motility) point mutation (Young et al., 1999). Unexpectedly, the triple mutant showed a slight rescue of motility. I investigated the relation between the FlhD point mutation and the TatD paralogs by expressing a functional FlhD construct both in the MC4100 strain and the triple mutant. Strong motility is observed in TatD triple mutant strains with FlhD expression pointing to a regulatory antagonism of FlhD and TatD paralogs (Figure 3.9D). Further experimental verifications are necessary to confirm the negative regulatory role of TatD paralogs in motility.

4.3.4 Motility – metabolism

One interesting metabolic connection is implied by the interaction between pyruvate oxidoreductase (TP0939) and flgB, a flagellar rod protein, as an orthologous interaction (interologs) was also detected in *Helicobacter pylori* by another Y2H study (Rain et al., 2001). TP0939 is pyruvate oxidoreductase involved in pyruvate synthesis (oxidized ferredoxin + coenzyme A + pyruvate = reduced ferredoxin + CO2 + acetyl-CoA). This connection is, furthermore, supported by the interaction of pyruvate oxidoreductase with TP0421 and TP0658, proteins newly classified as motility related gene in this study (Figure 4.4).

TP0939 is very large protein (1184 aa) and has 4 clear domains/motifs. In Arifuzzaman et al (unpublished) *E. coli* complex purification studies TP0939 homolog in *E. coli*, ybdK, was copurified together with marA (Multiple antibiotic resistance protein) and fhuF (Ferric iron reductase protein) proteins, these interactions were not detected in our Y2H studies. The deletion mutant in *E. coli* show no reduction in motility under the conditions studied. The Y2H

interaction of pyruvate oxidoreductase (TP0939) shows the physical connection between the metabolic enzymes with flagellum proteins. However, the molecular function of these interactions were remain to be investigated

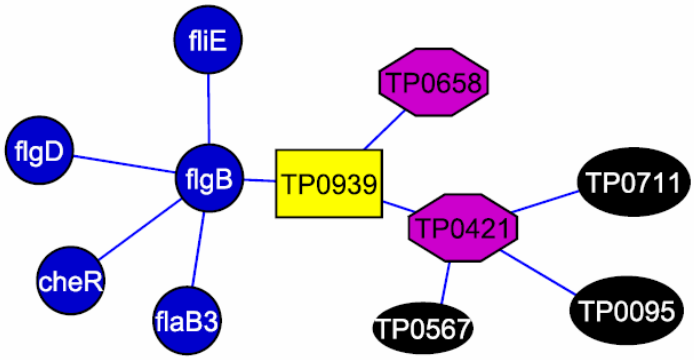


Figure 4.4 Protein-protein interactions of TP0939 (pyruvate oxidoreductase): Proteins are indicated as nodes. The orthologous mutants showing a motility phenotype in *E. coli / B. subtilis* are depicted as hexagonal shapes. Known flagellum proteins are coloured blue and proteins of unknown function are coloured block.

4.4 Variations of the theme – features of Spirochete's motility

The spirochete flagellum possesses unique features not found in other bacterial species, most prominently, the periplasmic localisation of two flagellum bundles fixed to the cell poles (Charon et al., 1992; Li et al., 2000a). During translational motility these two bundles perform an asymmetrical rotation; the molecular basis of this asymmetry, however, is unknown. Two hypotheses try to explain how spirochetes achieve asymmetric rotation. The first hypothesis states that chemotaxis plays an essential role in this asymmetry (Armitage 1999; Berg 1975; Bren and Eisenbach 2000). The second hypothesis claims that the motors at both cell poles are somehow differently organized in spirochetes than in other bacteria. One finding in opposition to the chemotaxis hypothesis is the absence of CheA in *B. burgdorferi*. These cells continually rotate their flagellar bundles in opposite direction and constantly run (Li et al., 2002) despite the assumed absence of phosphorylated CheY.

FliG: Several candidate effectors could be involved in flagellar asymmetric rotation. One such possible effector is fliG. The *T. pallidum, B. burgdorferi* and *T. denticola* have only one copy of the genes that encode the basal body and switch complex, the only exception being fliG. This protein is a central component of the bacterial flagellar motor switch. All sequenced spirochetes have two homologs encoding fliG. *T. pallidum* fliG-1 (TP0026) and fliG-2 (TP0400) show significant similarity in their protein sequences (Figure 4.5). However, they show different interacting proteins in our Y2H screen (Figure 4.6). Although the found differential interaction patterns do not explain the molecular basis of asymmetric behaviour, they may indicate the basis for this asymmetry, which might be, for example, mediated by FliG-2 recruiting additional proteins (of up to now unknown function) to the motor complex.

Apart from this asymmetry *Spirochetes* possess additional flagellar proteins, probably necessary for the unique localisation of the flagella: FlaA envelops the filament (Li et al., 2000b), and interacts with a peptidoglycane enzyme and amiA, probablycreates space for the filament. Its interaction with FlgG, a rod protein, indicates that also the rod might be surrounded by FlaA
Judged by the differential genomic distribution of different export chaperones (Pallen et al., 2005a), different functions and substrate specificities need to be assumed in different bacterial species. FliS is thought to be a flagellin-specific chaperone (Ozin et al., 2003), whereas FlgN and FliT are substrate-specific flagellar chaperones that prevent oligomerization of the HAPs (hook associated proteins) in *S. typhimurium*. Interestingly, flgN and fliT orthologs are not present in spirochetes (Pallen et al., 2005a). The interactions between fliS-flgK and fliS-flgB are the first experimental evidence that fliS may also function as a chaperone for flgB and flgK in spirochetes, partly substituting for flgN and fliT (Table 3.2).

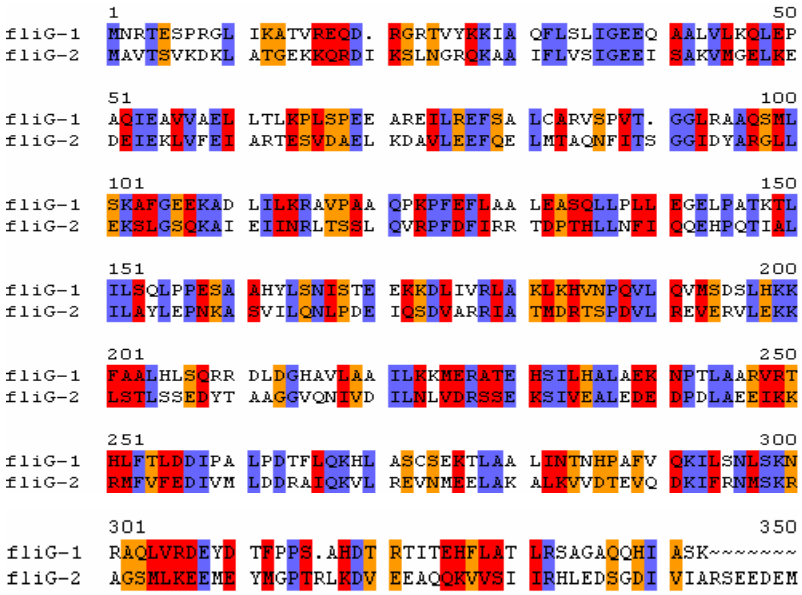


Figure 4.5 Multiple sequence alignment of *T. pallidum* fliG-1 and fliG-2: The sequence alignment was performed using the Clustal tool of Husar. The identical amino acids in both sequences are coloured in blue. The alignment of *T. pallidum* fliG-1 (TP0026) and fliG-2 (TP0400) show significant similarity over their entire protein sequences.

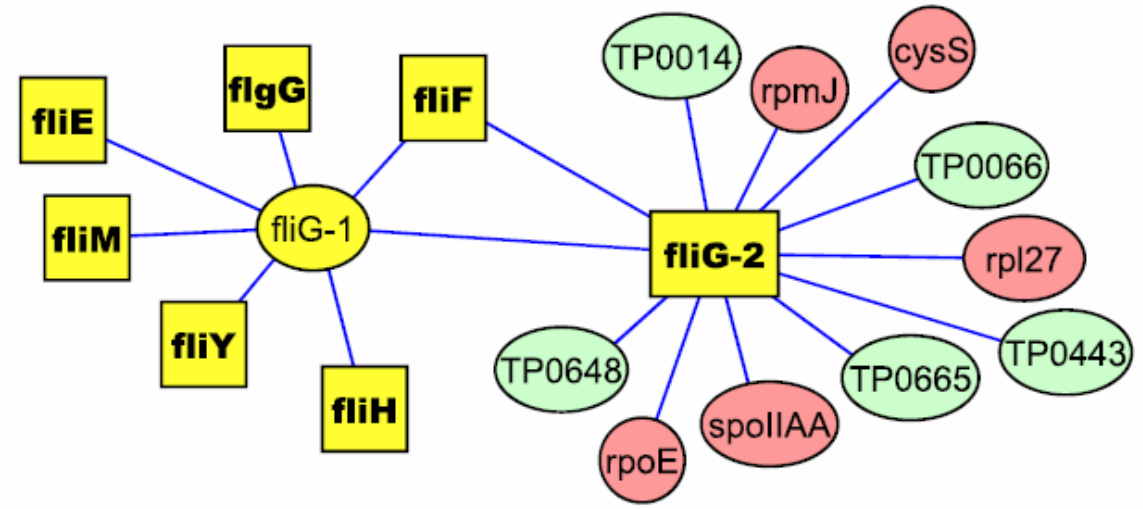


Figure 4.6 Y2H protein-protein interaction map of *T. pallidum* flagellar motor switch proteins (fliG-1 and fliG-2): Proteins are indicated as nodes, protein interactions are represented as solid edges. Known flagellum proteins are coloured yellow; proteins of yet unknown function are coloured light green. Baits are represented by bold letters and square boxes. The interaction partners of fliG-1 had other interactions also (not shown here).

4.5 Characterization of the interaction of TP0658/yviF with flagellins

TP0658 is a 17 kDa protein of unknown function, encoded together with a carbon storage regulator (CsrA) in an operon (Figure 3.13). We found that all three flagellin proteins of *T. pallidum* (FlaB1 = TP0868, FlaB2 = TP0792, and FlaB3 = TP0870) interact with TP0658 in Y2H screens (Figure 3.14A). I selected this interaction to characterize the interaction epitope biochemically and to investigate the molecular function of TP0658. All three flagellin interactions with TP0658 have been verified *in vitro* by far-western blots (overlay assay) (Figure 3.14B). TP0658 is conserved in several bacterial species including *Spirochetes*, *Bacillales*, and *delta/epsilon Proteobacteria*. To test whether the protein interactions of TP0658 with flagellins are conserved in other species, the interactions between its *B. subtilis* ortholog yviF and the *B. subtilis* flagellins hag and yvzB were verified biochemically. Additionally, yviF still undergoes an interaction with all three flagellin proteins of *T. pallidum* pointing to a conservation of the interaction epitope (Figure 3.15A).

4.6 TP0658/yviF is putative new assembly factor for bacterial flagellum

I identified a putative new assembly factor for the bacterial flagellum using systematic yeast two-hybrid screening. What is the molecular mechanism of TP0658 function? Several observations are reminiscent of the export chaperone FliS in *S. typhimurium* (Auvray et al., 2001; Ozin et al., 2003). First, the direct interaction between flagellin and TP0658. Second, lack of TP0658 leads to loss of flagellar filaments supporting a role of TP0658 in filament assembly and/or stabilization of flagellin in the cytosol. In fact, Golden & Acheson (Golden and Acheson 2002) showed that a mutant of *Campylobacter* CJ1075 was not only completely immotile but also had no detectable flagellin (FlaA). A direct effect on flagellin stability is also supported by artificially co-expression of TP0858 and flagellin (Figure 3.23). Third, FUGUE, a computer program for recognizing distant homologues by sequence structure comparison (Shi et al., 2001)

resulted in the structure of CesT as top hit when databases were searched for yviF homologs. CesT is also a type III secretion chaperone (Luo et al., 2001). However, the detailed molecular mechanism of TP0658 and FliS function remain unclear. They may act in parallel or subsequently and it can not be excluded that TP0658 might, for example, add a further layer of regulation to the process. Finally, type III secretion chaperones (TTSCs) have been shown to prevent degradation and aggregation of their cognate substrates prior to secretion (Bennett and Hughes 2000), just as TP0658 appears to do (Figure 3.23). Once secretion of the bound substrate is initiated, the TTSCs are thought to donate the substrates to the secretion apparatus. Unlike chaperones associated with virulent TTSSs, all flagella-associated TTSCs bind the carboxyterminal region of their partners (Bennett and Hughes 2000). There is also no obvious similarity between TP0658 and FliS, so it is unclear if their function is similar. To unravel these distinct functions of FliS and TP0658 it would be interesting to replace FliS by TP0658 in species such as *E. coli* which has only FliS but no TP0658 homolog. Interestingly, their seems to be no species which has only TP0658, thus phylogenetic evidence suggests rather non-overlapping functions (Figure 3.24).

Outlook

The study presented here mainly attempted to identify all the possible intra-flagellum interactions and flagellum associated proteins in a quantitative way using genome-wide protein-protein interaction screening, and the results obtained by genome-wide Y2H screening of *T. pallidum* flagellum baits opened new perspectives on the flagellum complex assembly and regulation.

The *T. pallidum* flagellum apparatus interactome revealed 253 novel PPIs, even thought *T. pallidum* does not allow us to verify the molecular function of the identified interactions experimentally, we can use the interaction data to predict interactions in more tractable organisms such as *E. coli* or *B. subtilis*, because most of the interacting proteins are conserved in other bacterial species and the interaction map can be used as a benchmark for studying the bacterial flagellum complex relationships in particular with respect to an understanding of the global relationships of the of flagellum apparatus with other pathways and processes in the cell.

There are many unknown factors with respect to motility and chemotaxis gene regulation in spirochetes. For example, the master regulators of flagellum genes are not identified. The *T. pallidum* flagellum interactome revealed several uncharacterized spirochete-specific proteins interacting with know flagellum proteins. Characterizing the function of the identified interactions would be a major step towards understanding the common and specific flagellum components between spirochetes and other model organisms, because motility in spirochetes is quite different and more complex than the well-studied paradigms of *E. coli and S. typhimurium*.

The flagellar apparatus has been intensely studied, not the least because motility mutants were so easy to obtain. However, it turned out to be much more difficult to gain detailed insights into the structure, function, dynamics, and evolution of the flagellar machinery. Structural genomics will reveal how the individual proteins look like at atomic detail and our interaction data will hopefully help to find crystallizable pairs of proteins or whole subcomplexes. This should also lead to a better understanding of the assembly of the flagellar motor. Our knowledge about individual steps in this process is rudimentary at best.

This study uncovered many new protein interactions but it did not shed light on the dynamics of these interactions. It remains to be studied in which order these interactions take place, which ones are physiologically relevant and how they are regulated. Post-translational modifications will certainly play a role here as much as assembly factors and chaperones do. Our analysis provided only a first step towards an understanding in this direction.

Finally, it will be exciting to find out how “supposedly irreducibly complex” systems such as the flagellar machinery can evolve. Although it seems to be clear that the flagellum is derived from a type three secretion system it remains mysterious which selective pressures lead to the addition of a multitude of proteins that made some ancestral cell motile. If we can figure out what else these acquired motility proteins did before their recruitment and how they fit into the developing motor, we would have solved a fundamental question in the evolution of life.

References:

Akeda, Y. and J.E. Galan. Chaperone release and unfolding of substrates in type III secretion *Nature* **437:** 911-915, 2005.

Akerley, B.J. and J.F. Miller. Flagellin gene transcription in Bordetella bronchiseptica is regulated by the BvgAS virulence control system *J Bacteriol* **175:** 3468-3479, 1993.

Akerley, B.J., D.M. Monack, S. Falkow, and J.F. Miller. The bvgAS locus negatively controls motility and synthesis of flagella in Bordetella bronchiseptica *J Bacteriol* **174:** 980-990, 1992.

Aldridge, P., J. Karlinsey, and K.T. Hughes. The type III secretion chaperone FlgN regulates flagellar assembly via a negative feedback loop containing its chaperone substrates FlgK and FlgL *Mol Microbiol* **49:** 1333-1345, 2003.

Allen-Vercoe, E. and M.J. Woodward. The role of flagella, but not fimbriae, in the adherence of Salmonella enterica serotype Enteritidis to chick gut explant *J Med Microbiol* **48:** 771-780, 1999.

Aloy, P. and R.B. Russell. The third dimension for protein interactions and complexes *Trends Biochem Sci* **27:** 633-638, 2002.

Armitage, J.P. Bacterial tactic responses *Adv Microb Physiol* **41:** 229-289, 1999.

Armitage, J.P. and R. Schmitt. Bacterial chemotaxis: Rhodobacter sphaeroides and Sinorhizobium meliloti--variations on a theme? *Microbiology* **143 (Pt 12):** 3671-3682, 1997.

Artsimovitch, I. and R. Landick. The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand *Cell* **109:** 193-203, 2002.

Auvray, F., J. Thomas, G.M. Fraser, and C. Hughes. Flagellin polymerisation control by a cytosolic export chaperone *J Mol Biol* **308:** 221-229, 2001.

Bailey, M.J., C. Hughes, and V. Koronakis. In vitro recruitment of the RfaH regulatory protein into a specialised transcription complex, directed by the nucleic acid ops element *Mol Gen Genet* **262:** 1052-1059, 2000.

Bailey, M.J., V. Koronakis, T. Schmoll, and C. Hughes. Escherichia coli HlyT protein, a transcriptional activator of haemolysin synthesis and secretion, is encoded by the rfaH (sfrB) locus required for expression of sex factor and lipopolysaccharide genes *Mol Microbiol* **6:** 1003-1012, 1992.

Bakeeva, L.E., K.M. Chumakov, A.L. Drachev, A.L. Metlina, and V.P. Skulachev. The sodium cycle. III. Vibrio alginolyticus resembles Vibrio cholerae and some other vibriones by flagellar motor and ribosomal 5S-RNA structures *Biochim Biophys Acta* **850:** 466-472, 1986.

Bennett, J.C. and C. Hughes. From flagellum assembly to virulence: the extended family of type III export chaperones *Trends Microbiol* **8:** 202-204, 2000.

Bennett, J.C., J. Thomas, G.M. Fraser, and C. Hughes. Substrate complexes and domain organization of the Salmonella flagellar export chaperones FlgN and FliT *Mol Microbiol* **39:** 781-791, 2001.

Berg, H.C. Chemotaxis in bacteria *Annu Rev Biophys Bioeng* **4:** 119-136, 1975.

Berg, H.C. and L. Turner. Torque generated by the flagellar motor of Escherichia coli *Biophys J* **65:** 2201-2216, 1993.

Bertin, P., E. Terao, E.H. Lee, P. Lejeune, C. Colson, A. Danchin, and E. Collatz. The H-NS protein is involved in the biogenesis of flagella in Escherichia coli *J Bacteriol* **176:** 5537-5540, 1994.

Bren, A. and M. Eisenbach. The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY *J Mol Biol* **278:** 507-514, 1998.

Bren, A. and M. Eisenbach. How signals are heard during bacterial chemotaxis: protein-protein interactions in sensory signal propagation *J Bacteriol* **182:** 6865-6873, 2000.

Brown, J.R., C.J. Douady, M.J. Italia, W.E. Marshall, and M.J. Stanhope. Universal trees based on large combined protein sequence data sets *Nat Genet* **28:** 281-285, 2001.

Brown, P.N., C.P. Hill, and D.F. Blair. Crystal structure of the middle and C-terminal domains of the flagellar rotor protein FliG *Embo J* **21:** 3225-3234, 2002.

Brown, P.N., M.A. Mathews, L.A. Joss, C.P. Hill, and D.F. Blair. Crystal structure of the flagellar rotor protein FliN from Thermotoga maritima *J Bacteriol* **187:** 2890-2902, 2005.

Butland, G., J.M. Peregrin-Alvarez, J. Li, W. Yang, X. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan et al. Interaction network containing conserved and essential protein complexes in Escherichia coli *Nature* **433:** 531-537, 2005.

Cagney, G., P. Uetz, and S. Fields. Two-hybrid analysis of the Saccharomyces cerevisiae 26S proteasome *Physiol Genomics* **7:** 27-34, 2001.

Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis *Mol Biol Evol* **17:** 540-552, 2000.

Charon, N.W. and S.F. Goldstein. Genetics of motility and chemotaxis of a fascinating group of bacteria: the spirochetes *Annu Rev Genet* **36:** 47-73, 2002.

Charon, N.W., E.P. Greenberg, M.B. Koopman, and R.J. Limberger. Spirochete chemotaxis, motility, and the structure of the spirochetal periplasmic flagella *Res Microbiol* **143:** 597-603, 1992.

Chen, S., J. Jancrick, H. Yokota, R. Kim, and S.H. Kim. Crystal structure of a protein associated with cell division from Mycoplasma pneumoniae (GI: 13508053): a novel fold with a conserved sequence motif *Proteins* **55:** 785-791, 2004.

Chi, B., R.J. Limberger, and H.K. Kuramitsu. Complementation of a Treponema denticola flgE mutant with a novel coumermycin A1-resistant T. denticola shuttle vector system *Infect Immun* **70:** 2233-2237, 2002.

Claret, L. and C. Hughes. Rapid turnover of FlhD and FlhC, the flagellar regulon transcriptional activator proteins, during Proteus swarming *J Bacteriol* **182:** 833-836, 2000.

Creevey, C.J. and J.O. McInerney. Clann: investigating phylogenetic information through supertree analyses *Bioinformatics* **21:** 390-392, 2005.

Dangl, J.L. and J.D. Jones. Plant pathogens and integrated defence responses to infection *Nature* **411:** 826-833, 2001.

Daniel, R.A., A.M. Williams, and J. Errington. A complex four-gene operon containing essential cell division gene pbpB in Bacillus subtilis *J Bacteriol* **178:** 2343-2350, 1996.

Das, S., S.W. Barthold, S.S. Giles, R.R. Montgomery, S.R. Telford, 3rd, and E. Fikrig. Temporal pattern of Borrelia burgdorferi p21 expression in ticks and the mammalian host *J Clin Invest* **99:** 987-995, 1997.

Datsenko, K.A. and B.L. Wanner. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products *Proc Natl Acad Sci U S A* **97:** 6640-6645, 2000.

Daubin, V., M. Gouy, and G. Perriere. A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history *Genome Res* **12:** 1080-1090, 2002.

DePamphilis, M.L. and J. Adler. Purification of intact flagella from Escherichia coli and Bacillus subtilis *J Bacteriol* **105:** 376-383, 1971.

Doerks, T. Mering, CV. and Bork, P. Functional clues for hypothetical proteins based on genomic context analysis in prokaryotes *Nucleic Acids Research* 32(21):6321-6326. 2004

Eaton, K.A., D.R. Morgan, and S. Krakowka. Motility as a factor in the colonisation of gnotobiotic piglets by Helicobacter pylori *J Med Microbiol* **37:** 123-127, 1992.

Edwards, A.M., B. Kus, R. Jansen, D. Greenbaum, J. Greenblatt, and M. Gerstein. Bridging structural biology and genomics: assessing protein interaction data with known complexes *Trends Genet* **18:** 529-536, 2002.

Ely, B., T.W. Ely, W.B. Crymes, Jr., and S.A. Minnich. A family of six flagellin genes contributes to the Caulobacter crescentus flagellar filament *J Bacteriol* **182:** 5001-5004, 2000.

Emerson, S.U., K. Tokuyasu, and M.I. Simon. Bacterial flagella: polarity of elongation *Science* **169:** 190-192, 1970.

Evdokimov, A.G., J. Phan, J.E. Tropea, K.M. Routzahn, H.K. Peters, M. Pokross, and D.S. Waugh. Similar modes of polypeptide recognition by export chaperones in flagellar biosynthesis and type III secretion *Nat Struct Biol* **10:** 789-793, 2003.

Fabret, C., S.D. Ehrlich, and P. Noirot. A new mutation delivery system for genome-scale approaches in Bacillus subtilis *Mol Microbiol* **46:** 25-36, 2002.

Fan, F., K. Ohnishi, N.R. Francis, and R.M. Macnab. The FliP and FliR proteins of Salmonella typhimurium, putative components of the type III flagellar export apparatus, are located in the flagellar basal body *Mol Microbiol* **26:** 1035-1046, 1997.

Felsenstein, J. PHYLIP (Phylogeny Inference Package) version 3.6., U.o.W. Distributed by the author *Department of Genome Sciences, Seattle., Editor*, 2005.

Fenchel, T. Microbial behavior in a heterogeneous world *Science* **296:** 1068-1071, 2002.

Fields, S. and O. Song. A novel genetic system to detect protein-protein interactions *Nature* **340:** 245-246, 1989.

Formstecher, E., S. Aresta, V. Collura, A. Hamburger, A. Meil, A. Trehin, C. Reverdy, V. Betin, S. Maire, C. Brun et al. Protein interaction mapping: a Drosophila case study *Genome Res* **15:** 376-384, 2005.

Francis, N.R., G.E. Sosinsky, D. Thomas, and D.J. DeRosier. Isolation, characterization and structure of bacterial flagellar motors containing the switch complex *J Mol Biol* **235:** 1261-1270, 1994.

Frank, R. SPOT synthesis: an easy technique for the positionally addressble, parallel chemical synthesis on a membrane support. *Tetrahedron* **48 (42): 9217-9232.**, 1992.

Fraser, C.M., S.J. Norris, G.M. Weinstock, O. White, G.G. Sutton, R. Dodson, M. Gwinn, E.K. Hickey, R. Clayton, K.A. Ketchum et al. Complete genome sequence of Treponema pallidum, the syphilis spirochete *Science* **281:** 375-388, 1998.

Fraser, G.M., J.C. Bennett, and C. Hughes. Substrate-specific binding of hook-associated proteins by FlgN and FliT, putative chaperones for flagellum assembly *Mol Microbiol* **32:** 569-580, 1999.

Fraser, G.M., B. Gonzalez-Pedrajo, J.R. Tame, and R.M. Macnab. Interactions of FliJ with the Salmonella type III flagellar export apparatus *J Bacteriol* **185:** 5546-5554, 2003.

Galan, J.E. and A. Collmer. Type III secretion machines: bacterial devices for protein delivery into host cells *Science* **284:** 1322-1328, 1999.

Gavin, A.C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat et al. Functional organization of the yeast proteome by systematic analysis of protein complexes *Nature* **415:** 141-147, 2002.

Ge, Y., I.G. Old, I. Saint Girons, and N.W. Charon. Molecular characterization of a large Borrelia burgdorferi motility operon which is initiated by a consensus sigma70 promoter *J Bacteriol* **179:** 2289-2299, 1997.

Ge, Y., C. Li, L. Corum, C.A. Slaughter, and N.W. Charon. Structure and expression of the FlaA periplasmic flagellar protein of Borrelia burgdorferi *J Bacteriol* **180:** 2418-2425, 1998.

Gilmore, R.D., Jr., M.L. Mbow, and B. Stevenson. Analysis of Borrelia burgdorferi gene expression during life cycle phases of the tick vector Ixodes scapularis *Microbes Infect* **3:** 799-808, 2001.

Golden, N.J. and D.W. Acheson. Identification of motility and autoagglutination Campylobacter jejuni mutants by random transposon mutagenesis *Infect Immun* **70:** 1761-1771, 2002.

Gonzalez-Pedrajo, B., G.M. Fraser, T. Minamino, and R.M. Macnab. Molecular dissection of Salmonella FliH, a regulator of the ATPase FliI and the type III flagellar protein export pathway *Mol Microbiol* **45:** 967-982, 2002.

Hayashi, F., K.D. Smith, A. Ozinsky, T.R. Hawn, E.C. Yi, D.R. Goodlett, J.K. Eng, S. Akira, D.M. Underhill, and A. Aderem. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5 *Nature* **410:** 1099-1103, 2001.

Heidrich, C., M.F. Templin, A. Ursinus, M. Merdanovic, J. Berger, H. Schwarz, M.A. de Pedro, and J.V. Holtje. Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of Escherichia coli *Mol Microbiol* **41:** 167-178, 2001.

Hirano, T., T. Minamino, and R.M. Macnab. The role in flagellar rod assembly of the N-terminal domain of Salmonella FlgJ, a flagellum-specific muramidase *J Mol Biol* **312:** 359-369, 2001.

Hirano, T., S. Yamaguchi, K. Oosawa, and S. Aizawa. Roles of FliK and FlhB in determination of flagellar hook length in Salmonella typhimurium *J Bacteriol* **176:** 5439-5449, 1994.

Hirota, N. and Y. Imae. Na+-driven flagellar motors of an alkalophilic Bacillus strain YN-1 *J Biol Chem* **258:** 10577-10581, 1983.

Ho, Y., A. Gruhler, A. Heilbut, G.D. Bader, L. Moore, S.L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier et al. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry *Nature* **415:** 180-183, 2002.

Homma, M., D.J. DeRosier, and R.M. Macnab. Flagellar hook and hook-associated proteins of Salmonella typhimurium and their relationship to other axial components of the flagellum *J Mol Biol* **213:** 819-832, 1990a.

Homma, M., K. Kutsukake, M. Hasebe, T. Iino, and R.M. Macnab. FlgB, FlgC, FlgF and FlgG. A family of structurally related proteins in the flagellar basal body of Salmonella typhimurium *J Mol Biol* **211:** 465-477, 1990b.

Hubner, A., X. Yang, D.M. Nolen, T.G. Popova, F.C. Cabello, and M.V. Norgard. Expression of Borrelia burgdorferi OspC and DbpA is controlled by a RpoN-RpoS regulatory pathway *Proc Natl Acad Sci U S A* **98:** 12724-12729, 2001.

Hughes, K.T., K.L. Gillen, M.J. Semon, and J.E. Karlinsey. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator *Science* **262:** 1277-1280, 1993.

Irikura, V.M., M. Kihara, S. Yamaguchi, H. Sockett, and R.M. Macnab. Salmonella typhimurium fliG and fliN mutations causing defects in assembly, rotation, and switching of the flagellar motor *J Bacteriol* **175:** 802-810, 1993.

Iyoda, S., T. Kamidoi, K. Hirose, K. Kutsukake, and H. Watanabe. A flagellar gene fliZ regulates the expression of invasion genes and virulence phenotype in Salmonella enterica serovar Typhimurium *Microb Pathog* **30:** 81-90, 2001.

Jones, C.J. and R.M. Macnab. Flagellar assembly in Salmonella typhimurium: analysis with temperature-sensitive mutants *J Bacteriol* **172:** 1327-1339, 1990.

Jones, C.J. and S. Aizawa. The bacterial flagellum and flagellar motor: structure, assembly and function *Adv Microb Physiol* **32:** 109-172, 1991.

Josenhans, C. and S. Suerbaum. The role of motility as a virulence factor in bacteria *Int J Med Microbiol* **291:** 605-614, 2002.

Joseph, P., J.R. Fantino, M.L. Herbaud, and F. Denizot. Rapid orientated cloning in a shuttle vector allowing modulated gene expression in Bacillus subtilis *FEMS Microbiol Lett* **205:** 91-97, 2001.

Joys, T.M. The flagellar filament protein *Can J Microbiol* **34:** 452-458, 1988.

Kanehisa, M., S. Goto, S. Kawashima, and A. Nakaya. The KEGG databases at GenomeNet *Nucleic Acids Res* **30:** 42-46, 2002.

Kanehisa, M., S. Goto, M. Hattori, K.F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa. From genomics to chemical genomics: new developments in KEGG *Nucleic Acids Res* **34:** D354-357, 2006.

Kang, S.H., W.S. Lim, and K.H. Kim. A protein interaction map of soybean mosaic virus strain G7H based on the yeast two-hybrid system *Mol Cells* **18:** 122-126, 2004.

Kapatral, V., J.W. Olson, J.C. Pepe, V.L. Miller, and S.A. Minnich. Temperature-dependent regulation of Yersinia enterocolitica Class III flagellar genes *Mol Microbiol* **19:** 1061-1071, 1996.

Kennedy, M.J., E.L. Rosey, and R.J. Yancey, Jr. Characterization of flaA- and flaB- mutants of Serpulina hyodysenteriae: both flagellin subunits, FlaA and FlaB, are necessary for full motility and intestinal colonization *FEMS Microbiol Lett* **153:** 119-128, 1997.

Kholod, N. and T. Mustelin. Novel vectors for co-expression of two proteins in E. coli *Biotechniques* **31:** 322-323, 326-328, 2001.

Klose, K.E. and J.J. Mekalanos. Differential regulation of multiple flagellins in Vibrio cholerae *J Bacteriol* **180:** 303-316, 1998a.

Klose, K.E. and J.J. Mekalanos. Distinct roles of an alternative sigma factor during both free-swimming and colonizing phases of the Vibrio cholerae pathogenic cycle *Mol Microbiol* **28:** 501-520, 1998b.

Kolonin, M.G., J. Zhong, and R.L. Finley. Interaction mating methods in two-hybrid systems *Methods Enzymol* **328:** 26-46, 2000.

Komeda, Y. Fusions of flagellar operons to lactose genes on a mu lac bacteriophage *J Bacteriol* **150:** 16-26, 1982.

Komeda, Y. Transcriptional control of flagellar genes in Escherichia coli K-12 *J Bacteriol* **168:** 1315-1318, 1986.

Kubori, T., N. Shimamoto, S. Yamaguchi, K. Namba, and S. Aizawa. Morphological pathway of flagellar assembly in Salmonella typhimurium *J Mol Biol* **226:** 433-446, 1992.

Kubori, T., Y. Matsushima, D. Nakamura, J. Uralil, M. Lara-Tejero, A. Sukhan, J.E. Galan, and S.I. Aizawa. Supramolecular structure of the Salmonella typhimurium type III protein secretion system *Science* **280:** 602-605, 1998.

Kutsukake, K., Y. Ohya, and T. Iino. Transcriptional analysis of the flagellar regulon of Salmonella typhimurium *J Bacteriol* **172:** 741-747, 1990.

Landgraf, C., S. Panni, L. Montecchi-Palazzi, L. Castagnoli, J. Schneider-Mergener, R. Volkmer-Engert, and G. Cesareni. Protein interaction networks by proteome peptide scanning *PLoS Biol* **2:** E14, 2004.

Leeds, J.A. and R.A. Welch. RfaH enhances elongation of Escherichia coli hlyCABD mRNA *J Bacteriol* **178:** 1850-1857, 1996.

Li, C., C.J. Louise, W. Shi, and J. Adler. Adverse conditions which cause lack of flagella in Escherichia coli *J Bacteriol* **175:** 2229-2235, 1993.

Li, C., A. Motaleb, M. Sal, S.F. Goldstein, and N.W. Charon. Spirochete periplasmic flagella and motility *J Mol Microbiol Biotechnol* **2:** 345-354, 2000a.

Li, C., L. Corum, D. Morgan, E.L. Rosey, T.B. Stanton, and N.W. Charon. The spirochete FlaA periplasmic flagellar sheath protein impacts flagellar helicity *J Bacteriol* **182:** 6698-6706, 2000b.

Li, C., R.G. Bakker, M.A. Motaleb, M.L. Sartakova, F.C. Cabello, and N.W. Charon. Asymmetrical flagellar rotation in Borrelia burgdorferi nonchemotactic mutants *Proc Natl Acad Sci U S A* **99:** 6169-6174, 2002.

Li, S., C.M. Armstrong, N. Bertin, H. Ge, S. Milstein, M. Boxem, P.O. Vidalain, J.D. Han, A. Chesneau, T. Hao et al. A map of the interactome network of the metazoan C. elegans *Science* **303:** 540-543, 2004.

Limberger, R.J., L.L. Slivienski, J. Izard, and W.A. Samsonoff. Insertional inactivation of Treponema denticola tap1 results in a nonmotile mutant with elongated flagellar hooks *J Bacteriol* **181:** 3743-3750, 1999.

Lindberg, A.A. and C.G. Hellerqvist. Rough mutants of Salmonella typhimurium: immunochemical and structural analysis of lipopolysaccharides from rfaH mutants *J Gen Microbiol* **116:** 25-32, 1980.

Liu, Q., M.Z. Li, D. Leibham, D. Cortez, and S.J. Elledge. The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes *Curr Biol* **8:** 1300-1309, 1998.

Lloyd, S.A., F.G. Whitby, D.F. Blair, and C.P. Hill. Structure of the C-terminal domain of FliG, a component of the rotor in the bacterial flagellar motor *Nature* **400:** 472-475, 1999.

Luo, Y., M.G. Bertero, E.A. Frey, R.A. Pfuetzner, M.R. Wenk, L. Creagh, S.L. Marcus, D. Lim, F. Sicheri, C. Kay et al. Structural and biochemical characterization of the type III secretion chaperones CesT and SigE *Nat Struct Biol* **8:** 1031-1036, 2001.

Lux, R., A. Moter, and W. Shi. Chemotaxis in pathogenic spirochetes: directed movement toward targeting tissues? *J Mol Microbiol Biotechnol* **2:** 355-364, 2000.

Lux, R., J.N. Miller, N.H. Park, and W. Shi. Motility and chemotaxis in tissue penetration of oral epithelial cell layers by Treponema denticola *Infect Immun* **69:** 6276-6283, 2001.

Macnab, R.M. Genetics and biogenesis of bacterial flagella *Annu Rev Genet* **26:** 131-158, 1992.

Macnab, R.M. The bacterial flagellum: reversible rotary propellor and type III export apparatus *J Bacteriol* **181:** 7149-7153, 1999.

Macnab, R.M. How bacteria assemble flagella *Annu Rev Microbiol* **57:** 77-100, 2003.

Macnab, R.M. Type III flagellar protein export and flagellar assembly *Biochim Biophys Acta* **1694:** 207-217, 2004.

Macnab, R.M. and M.K. Ornston. Normal-to-curly flagellar transitions and their role in bacterial tumbling. Stabilization of an alternative quaternary structure by mechanical force *J Mol Biol* **112:** 1-30, 1977.

Mann, M., R.C. Hendrickson, and A. Pandey. Analysis of proteins and proteomes by mass spectrometry *Annu Rev Biochem* **70:** 437-473, 2001.

McCarter, L.L. Polar flagellar motility of the Vibrionaceae *Microbiol Mol Biol Rev* **65:** 445-462, table of contents, 2001.

McEvoy, M.M., A. Bren, M. Eisenbach, and F.W. Dahlquist. Identification of the binding interfaces on CheY for two of its targets, the phosphatase CheZ and the flagellar switch protein fliM *J Mol Biol* **289:** 1423-1433, 1999.

McKevitt, M., K. Patel, D. Smajs, M. Marsh, M. McLoughlin, S.J. Norris, G.M. Weinstock, and T. Palzkill. Systematic cloning of Treponema pallidum open reading frames for protein expression and antigen discovery *Genome Res* **13:** 1665-1674, 2003.

Meister, M., G. Lowe, and H.C. Berg. The proton flux through the bacterial flagellar motor *Cell* **49:** 643-650, 1987.

Minamino, T. and R.M. Macnab. Components of the Salmonella flagellar export apparatus and classification of export substrates *J Bacteriol* **181:** 1388-1394, 1999.

Minamino, T. and R.M. MacNab. Interactions among components of the Salmonella flagellar export apparatus and its substrates *Mol Microbiol* **35:** 1052-1064, 2000.

Minamino, T., S. Yamaguchi, and R.M. Macnab. Interaction between FliE and FlgB, a proximal rod component of the flagellar basal body of Salmonella *J Bacteriol* **182:** 3029-3036, 2000.

Moens, S. and J. Vanderleyden. Functions of bacterial flagella *Crit Rev Microbiol* **22:** 67-100, 1996.

Muller, V., C.J. Jones, I. Kawagishi, S. Aizawa, and R.M. Macnab. Characterization of the fliE genes of Escherichia coli and Salmonella typhimurium and identification of the FliE

protein as a component of the flagellar hook-basal body complex *J Bacteriol* **174:** 2298-2304, 1992.

Nagy, G., U. Dobrindt, G. Schneider, A.S. Khan, J. Hacker, and L. Emody. Loss of regulatory protein RfaH attenuates virulence of uropathogenic Escherichia coli *Infect Immun* **70:** 4406-4413, 2002.

Namba, K. and F. Vonderviszt. Molecular architecture of bacterial flagellum *Q Rev Biophys* **30:** 1-65, 1997.

Nambu, T. and K. Kutsukake. The Salmonella FlgA protein, a putativeve periplasmic chaperone essential for flagellar P ring formation *Microbiology* **146 (Pt 5):** 1171-1178, 2000.

O'Brien, E.J. and P.M. Bennett. Structure of straight flagella from a mutant Salmonella *J Mol Biol* **70:** 133-152, 1972.

O'Toole, G.A. and R. Kolter. Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development *Mol Microbiol* **30:** 295-304, 1998.

Ohnishi, K., Y. Ohto, S. Aizawa, R.M. Macnab, and T. Iino. FlgD is a scaffolding protein needed for flagellar hook assembly in Salmonella typhimurium *J Bacteriol* **176:** 2272-2281, 1994.

Okino, H., M. Isomura, S. Yamaguchi, Y. Magariyama, S. Kudo, and S.I. Aizawa. Release of flagellar filament-hook-rod complex by a Salmonella typhimurium mutant defective in the M ring of the basal body *J Bacteriol* **171:** 2075-2082, 1989.

Otte, L., U. Wiedemann, B. Schlegel, J.R. Pires, M. Beyermann, P. Schmieder, G. Krause, R. Volkmer-Engert, J. Schneider-Mergener, and H. Oschkinat. WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains *Protein Sci* **12:** 491-500, 2003.

Ottemann, K.M. and J.F. Miller. Roles for motility in bacterial-host interactions *Mol Microbiol* **24:** 1109-1117, 1997.

Ozin, A.J., L. Claret, F. Auvray, and C. Hughes. The FliS chaperone selectively binds the disordered flagellin C-terminal D0 domain central to polymerisation *FEMS Microbiol Lett* **219:** 219-224, 2003.

Pallen, M.J., C.W. Penn, and R.R. Chaudhuri. Bacterial flagellar diversity in the post-genomic era *Trends Microbiol* **13:** 143-149, 2005a.

Pallen, M.J., S.A. Beatson, and C.M. Bailey. Bioinformatics, genomics and evolution of non-flagellar type-III secretion systems: a Darwinian perspective *FEMS Microbiol Rev* **29:** 201-229, 2005b.

Patterson-Delafield, J., R.J. Martinez, B.A. Stocker, and S. Yamaguchi. A new fla gene in Salmonella typhimurium--flaR--and its mutant phenotype-superhooks *Arch Mikrobiol* **90:** 107-120, 1973.

Pedersen, N.S., C.S. Petersen, M. Vejtorp, and N.H. Axelsen. Serodiagnosis of syphilis by an enzyme-linked immunosorbent assay for IgG antibodies against the Reiter treponeme flagellum *Scand J Immunol* **15:** 341-348, 1982.

Pratt, L.A. and R. Kolter. Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili *Mol Microbiol* **30:** 285-293, 1998.

Prouty, M.G., N.E. Correa, and K.E. Klose. The novel sigma54- and sigma28-dependent flagellar gene transcription hierarchy of Vibrio cholerae *Mol Microbiol* **39:** 1595-1609, 2001.

Rain, J.C., L. Selig, H. De Reuse, V. Battaglia, C. Reverdy, S. Simon, G. Lenzen, F. Petel, J. Wojcik, V. Schachter et al. The protein-protein interaction map of Helicobacter pylori *Nature* **409:** 211-215, 2001.

Rajagopala, S.V., B. Titz, and P. Uetz. Array-based yeast two-hybrid screening for protein-protein interactions *Methods in Enzymology*, in press.

Reed, K.A., M.E. Hobert, C.E. Kolenda, K.A. Sands, M. Rathman, M. O'Connor, S. Lyons, A.T. Gewirtz, P.J. Sansonetti, and J.L. Madara. The Salmonella typhimurium flagellar basal

body protein FliE is required for flagellin production and to induce a proinflammatory response in epithelial cells *J Biol Chem* **277:** 13346-13353, 2002.

Revel, A.T., A.M. Talaat, and M.V. Norgard. DNA microarray analysis of differential gene expression in Borrelia burgdorferi, the Lyme disease spirochete *Proc Natl Acad Sci U S A* **99:** 1562-1567, 2002.

Romeo, T. Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB *Mol Microbiol* **29:** 1321-1330, 1998.

Romeo, T., M. Gong, M.Y. Liu, and A.M. Brun-Zinkernagel. Identification and molecular characterization of csrA, a pleiotropic gene from Escherichia coli that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties *J Bacteriol* **175:** 4744-4755, 1993.

Rosey, E.L., M.J. Kennedy, and R.J. Yancey, Jr. Dual flaA1 flaB1 mutant of Serpulina hyodysenteriae expressing periplasmic flagella is severely attenuated in a murine model of swine dysentery *Infect Immun* **64:** 4154-4162, 1996.

Rosey, E.L., M.J. Kennedy, D.K. Petrella, R.G. Ulrich, and R.J. Yancey, Jr. Inactivation of Serpulina hyodysenteriae flaA1 and flaB1 periplasmic flagellar genes by electroporation-mediated allelic exchange *J Bacteriol* **177:** 5959-5970, 1995.

Rual, J.F., K. Venkatesan, T. Hao, T. Hirozane-Kishikawa, A. Dricot, N. Li, G.F. Berriz, F.D. Gibbons, M. Dreze, N. Ayivi-Guedehoussou et al. Towards a proteome-scale map of the human protein-protein interaction network *Nature*, 2005.

Sabnis, N.A., H. Yang, and T. Romeo. Pleiotropic regulation of central carbohydrate metabolism in Escherichia coli via the gene csrA *J Biol Chem* **270:** 29096-29104, 1995.

Sadziene, A., D.D. Thomas, V.G. Bundoc, S.C. Holt, and A.G. Barbour. A flagella-less mutant of Borrelia burgdorferi. Structural, molecular, and in vitro functional characterization *J Clin Invest* **88:** 82-92, 1991.

Samatey, F.A., H. Matsunami, K. Imada, S. Nagashima, and K. Namba. Crystallization of a core fragment of the flagellar hook protein FlgE *Acta Crystallogr D Biol Crystallogr* **60:** 2078-2080, 2004a.

Samatey, F.A., K. Imada, S. Nagashima, F. Vonderviszt, T. Kumasaka, M. Yamamoto, and K. Namba. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling *Nature* **410:** 331-337, 2001.

Samatey, F.A., H. Matsunami, K. Imada, S. Nagashima, T.R. Shaikh, D.R. Thomas, J.Z. Chen, D.J. Derosier, A. Kitao, and K. Namba. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism *Nature* **431:** 1062-1068, 2004b.

Sato, K. and M. Homma. Multimeric structure of PomA, a component of the Na+-driven polar flagellar motor of vibrio alginolyticus *J Biol Chem* **275:** 20223-20228, 2000.

Schmidt, H.A., K. Strimmer, M. Vingron, and A. von Haeseler. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing *Bioinformatics* **18:** 502-504, 2002.

Shaikh, T.R., D.R. Thomas, J.Z. Chen, F.A. Samatey, H. Matsunami, K. Imada, K. Namba, and D.J. Derosier. A partial atomic structure for the flagellar hook of Salmonella typhimurium *Proc Natl Acad Sci U S A* **102:** 1023-1028, 2005.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks *Genome Res* **13:** 2498-2504, 2003.

Shi, J., T.L. Blundell, and K. Mizuguchi. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties *J Mol Biol* **310:** 243-257, 2001.

Shi, W., C. Li, C.J. Louise, and J. Adler. Mechanism of adverse conditions causing lack of flagella in Escherichia coli *J Bacteriol* **175:** 2236-2240, 1993.

Smith, K.D. and A. Ozinsky. Toll-like receptor-5 and the innate immune response to bacterial flagellin *Curr Top Microbiol Immunol* **270:** 93-108, 2002.

Sockett, H., S. Yamaguchi, M. Kihara, V.M. Irikura, and R.M. Macnab. Molecular analysis of the flagellar switch protein FliM of Salmonella typhimurium *J Bacteriol* **174:** 793-806, 1992.

Soutourina, O., A. Kolb, E. Krin, C. Laurent-Winter, S. Rimsky, A. Danchin, and P. Bertin. Multiple control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the flhDC master operon *J Bacteriol* **181:** 7500-7508, 1999.

Soutourina, O.A. and P.N. Bertin. Regulation cascade of flagellar expression in Gram-negative bacteria *FEMS Microbiol Rev* **27:** 505-523, 2003.

Soutourina, O.A., E. Krin, C. Laurent-Winter, F. Hommais, A. Danchin, and P.N. Bertin. Regulation of bacterial motility in response to low pH in Escherichia coli: the role of H-NS protein *Microbiology* **148:** 1543-1551, 2002.

Stanton, T.B., E.G. Matson, and S.B. Humphrey. Brachyspira (Serpulina) hyodysenteriae gyrB mutants and interstrain transfer of coumermycin A(1) resistance *Appl Environ Microbiol* **67:** 2037-2043, 2001.

Stelzl, U., U. Worm, M. Lalowski, C. Haenig, F.H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koeppen et al. A human protein-protein interaction network: a resource for annotating the proteome *Cell* **122:** 957-968, 2005.

Suzuki, H., K. Yonekura, and K. Namba. Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis *J Mol Biol* **337:** 105-113, 2004.

Suzuki, H., K. Yonekura, K. Murata, T. Hirai, K. Oosawa, and K. Namba. A structural feature in the central channel of the bacterial flagellar FliF ring complex is implicated in type III protein export *J Struct Biol* **124:** 104-114, 1998.

Suzuki, T., T. Iino, T. Horiguchi, and S. Yamaguchi. Incomplete flagellar structures in nonflagellate mutants of Salmonella typhimurium *J Bacteriol* **133:** 904-915, 1978.

Swofford, D.L. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) *Sinauer Associates: Sunderland, Massachusetts.*, 2003.

Tatusov, R.L., N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Mekhedov, A.N. Nikolskaya et al. The COG database: an updated version includes eukaryotes *BMC Bioinformatics* **4:** 41, 2003.

Thomas, N.A., S.L. Bardy, and K.F. Jarrell. The archaeal flagellum: a different kind of prokaryotic motility structure *FEMS Microbiol Rev* **25:** 147-174, 2001.

Titz, B., M. Schlesner, and P. Uetz. What do we learn from high-throughput protein interaction data? *Expert Rev Proteomics* **1:** 111-121, 2004.

Toker, A.S. and R.M. Macnab. Distinct regions of bacterial flagellar switch protein FliM interact with FliG, FliN and CheY *J Mol Biol* **273:** 623-634, 1997.

Tominaga, A., M.A. Mahmoud, T. Mukaihara, and M. Enomoto. Molecular characterization of intact, but cryptic, flagellin genes in the genus Shigella *Mol Microbiol* **12:** 277-285, 1994.

Turner, L., W.S. Ryu, and H.C. Berg. Real-time imaging of fluorescent flagellar filaments *J Bacteriol* **182:** 2793-2801, 2000.

Uchiyama, I. MBGD: microbial genome database for comparative analysis *Nucleic Acids Res* **31:** 58-62, 2003.

Uetz, P. Two-hybrid arrays *Curr Opin Chem Biol* **6:** 57-62, 2002.

Uetz, P. and R.L. Finley, Jr. From protein networks to biological systems *FEBS Lett* **579:** 1821-1827, 2005.

Uetz, P., Y.A. Dong, C. Zeretzke, C. Atzler, A. Baiker, B. Berger, S.V. Rajagopala, M. Roupelieva, D. Rose, E. Fossum et al. Herpesviral protein networks and their interaction with the human proteome *Science* **311:** 239-242, 2006.

Uetz, P., L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart et al. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae *Nature* **403:** 623-627, 2000.

van Amsterdam, K. and A. van der Ende. Helicobacter pylori HP1034 (ylxH) is required for motility *Helicobacter* **9:** 387-395, 2004.

Van Arnam, J.S., J.L. McMurry, M. Kihara, and R.M. Macnab. Analysis of an engineered Salmonella flagellar fusion protein, FliR-FlhB *J Bacteriol* **186:** 2495-2498, 2004.

von Mering, C., M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, and B. Snel. STRING: a database of predicted functional associations between proteins *Nucleic Acids Res* **31:** 258-261, 2003.

Vonderviszt, F., S. Kanto, S. Aizawa, and K. Namba. Terminal regions of flagellin are disordered in solution *J Mol Biol* **209:** 127-133, 1989.

Wadhams, G.H. and J.P. Armitage. Making sense of it all: bacterial chemotaxis *Nat Rev Mol Cell Biol* **5:** 1024-1037, 2004.

Wei, B.L., A.M. Brun-Zinkernagel, J.W. Simecka, B.M. Pruss, P. Babitzke, and T. Romeo. Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli *Mol Microbiol* **40:** 245-256, 2001.

Welch, M., K. Oosawa, S. Aizawa, and M. Eisenbach. Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria *Proc Natl Acad Sci U S A* **90:** 8787-8791, 1993.

Wexler, M., F. Sargent, R.L. Jack, N.R. Stanley, E.G. Bogsch, C. Robinson, B.C. Berks, and T. Palmer. TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export *J Biol Chem* **275:** 16717-16722, 2000.

Williams, A.W., S. Yamaguchi, F. Togashi, S.I. Aizawa, I. Kawagishi, and R.M. Macnab. Mutations in fliK and flhB affecting flagellar hook and filament assembly in Salmonella typhimurium *J Bacteriol* **178:** 2960-2970, 1996.

Williams, F.D. and R.H. Schwarzhoff. Nature of the swarming phenomenon in Proteus *Annu Rev Microbiol* **32:** 101-122, 1978.

Wilson, D.R. and T.J. Beveridge. Bacterial flagellar filaments and their component flagellins *Can J Microbiol* **39:** 451-472, 1993.

Yokoseki, T., K. Kutsukake, K. Ohnishi, and T. Iino. Functional analysis of the flagellar genes in the fliD operon of Salmonella typhimurium *Microbiology* **141 (Pt 7):** 1715-1722, 1995.

Yonekura, K., S. Maki-Yonekura, and K. Namba. Structure analysis of the flagellar cap-filament complex by electron cryomicroscopy and single-particle image analysis *J Struct Biol* **133:** 246-253, 2001.

Yonekura, K., S. Maki-Yonekura, and K. Namba. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy *Nature* **424:** 643-650, 2003.

Yonekura, K., S. Maki, D.G. Morgan, D.J. DeRosier, F. Vonderviszt, K. Imada, and K. Namba. The bacterial flagellar cap as the rotary promoter of flagellin self-assembly *Science* **290:** 2148-2152, 2000.

Young, G.M., M.J. Smith, S.A. Minnich, and V.L. Miller. The Yersinia enterocolitica motility master regulatory operon, flhDC, is required for flagellin production, swimming motility, and swarming motility *J Bacteriol* **181:** 2823-2833, 1999.

Zhao, R., N. Pathak, H. Jaffe, T.S. Reese, and S. Khan. FliN is a major structural protein of the C-ring in the Salmonella typhimurium flagellar basal body *J Mol Biol* **261:** 195-208, 1996.

Appendix:

Table 5.1 List of computationally predicted associations by STRING database confirmed by *T. palliudm, C. jejuni, and H. pylori* Y2H screens

Predicted STRING associations confirmed experimentally by *Treponema pallidum* Y2H							
STRING 0.1		STRING 0.4		STRING 0.7		STRING 0.9	
Protein A	Protein B	Protein A	Protein B	Protein A	Protein B	Protein A	Protein B
TP0026	TP0398	TP0026	TP0398	TP0026	TP0398	TP0026	TP0398
TP0396	TP0398	TP0396	TP0398	TP0396	TP0398	TP0396	TP0398
TP0399	TP0400	TP0399	TP0400	TP0399	TP0400	TP0399	TP0400
TP0026	TP0399	TP0026	TP0399	TP0026	TP0399	TP0026	TP0399
TP0026	TP0401	TP0026	TP0401	TP0026	TP0401	TP0026	TP0401
TP0402	TP0640	TP0402	TP0640	TP0403	TP0868	TP0716	TP0725
TP0403	TP0868	TP0403	TP0868	TP0396	TP0630	TP0026	TP0720
TP0026	TP0630	TP0026	TP0630	TP0716	TP0725	TP0397	TP0720
TP0396	TP0630	TP0396	TP0630	TP0026	TP0720	TP0026	TP0721
TP0397	TP0640	TP0716	TP0725	TP0397	TP0720	TP0727	TP0728
TP0659	TP0726	TP0026	TP0720	TP0660	TP0720	TP0396	TP0728
TP0716	TP0725	TP0397	TP0720	TP0026	TP0721	TP0792	TP0943
TP0026	TP0720	TP0660	TP0720	TP0726	TP0728	TP0868	TP0943
TP0397	TP0720	TP0026	TP0721	TP0727	TP0728	TP0870	TP0943
TP0660	TP0720	TP0726	TP0728	TP0396	TP0728	TP0396	TP0870
TP0026	TP0721	TP0727	TP0728	TP0792	TP0943	TP0630	TP0870
TP0726	TP0728	TP0396	TP0728	TP0868	TP0943	TP0660	TP0943
TP0727	TP0728	TP0792	TP0943	TP0870	TP0943		
TP0396	TP0728	TP0868	TP0943	TP0396	TP0870		
TP0792	TP0943	TP0870	TP0943	TP0630	TP0870		
TP0868	TP0943	TP0396	TP0870	TP0660	TP0943		
TP0870	TP0943	TP0630	TP0870	TP0396	TP0943		
TP0396	TP0870	TP0640	TP0870				
TP0630	TP0870	TP0660	TP0943				
TP0640	TP0870	TP0439	TP0712				
TP0660	TP0943	TP0396	TP0943				
TP0439	TP0712						
TP0853	TP0870						
TP0396	TP0943						
Predicted STRING associations confirmed experimentally by *Campylobacter jejuni*Y2H							
STRING 0.1		STRING 0.4		STRING 0.7		STRING 0.9	
Protein A	Protein B	Protein A	Protein B	Protein A	Protein B	Protein A	Protein B
Cj0063c	Cj0064c	Cj0063c	Cj0064c	Cj0063c	Cj0064c	Cj0063c	Cj0064c
Cj0549	Cj1339c	Cj0549	Cj1339c	Cj0549	Cj1339c	Cj0549	Cj1339c
Cj0549	Cj0720c	Cj0549	Cj0720c	Cj0549	Cj0720c	Cj0549	Cj0720c
Cj0549	Cj1338c	Cj0549	Cj1338c	Cj0549	Cj1338c	Cj0549	Cj1338c
Cj0059c	Cj0060c	Cj0059c	Cj0060c	Cj0059c	Cj0060c	Cj0059c	Cj0351
Cj0059c	Cj0351	Cj0059c	Cj0351	Cj0059c	Cj0351	Cj0336c	Cj1408
Cj0547	Cj0720c	Cj0547	Cj0720c	Cj0547	Cj0720c	Cj0060c	Cj0337c
Cj0336c	Cj1408	Cj0336c	Cj1408	Cj0336c	Cj1408	Cj0060c	Cj0320
Cj0285c	Cj0528c	Cj0285c	Cj0528c	Cj0285c	Cj0528c		
Cj0526c	Cj0793	Cj0526c	Cj0793	Cj0064c	Cj0285c		
Cj0064c	Cj0285c	Cj0064c	Cj0285c	Cj0060c	Cj0337c		
Cj0060c	Cj0337c	Cj0060c	Cj0337c	Cj0060c	Cj0923c		

Cj0060c	Cj0923c	Cj0060c	Cj0923c	Cj0060c	Cj0320		
Cj0284c	Cj0547	Cj0060c	Cj0320				
Cj0060c	Cj0320						
Cj0351	Cj1222c						

Table 5.2 List of hypothetical and conserved hypothetical proteins interacting with known motility proteins and their *E. coli* and *B. subtilis* orthologs (Cluster of orthologous groups).

T. pallidum **Prey**	**Prey Description**	**Interacting Bait(s)**	**COG**	***E. coli*** **Orthologs**	***B. subtilis*** **Orthologs**
TP0675	hypothetical protein	TP0567	COG3735	ybaP	
TP0421	hypothetical protein	TP0567	COG3391	b1452	ywhL ywhK
TP0403	flagellar protein, putative	TP0377	COG2882	fliJ	fliJ
TP0911	conserved hypothetical protein	flgG-2	COG2257		ylqH
TP0050	hypothetical protein	groEL fliE flgD flaB1 flaB3 flgG-2	COG2236		
TP0843	heat shock protein, putative	fliY cheR	COG2214	cbpA	
TP0877	hypothetical protein	cheR flgG-2 cheX	COG2206		
TP0981	sensory transduction histidine kinase, putative	flaB3	COG2199	b1489 ycdT yaiC yeaI	ytrP ydaK yhcK
TP0046	hypothetical protein	cheR flgD metK flaB3 TP0959 cheW-2 mcp2-3	COG1774		yaaT
TP0658	transmembrane protein, putative	flaB2 flaB1 flaB3	COG1699		yviF
TP0048	hypothetical protein	fliY fliS	COG1664		yhbF yhbE
TP0518	hypothetical protein	csrA flgK flaB3	COG1564		yloS
TP0431	transcriptional activator, putative	TP0965	COG1521		yacB
TP0561	hypothetical protein	TP0218 fliF flhB fliR fliQ fliL flhA	COG1512		ydjH
TP0121	hypothetical protein	cheR	COG1509	yjeK	kamA
TP0233	anti-sigma F factor antagonist, putative	fliG-2 flaA-2 fliY	COG1366		yqhA yetI
TP0739	hypothetical protein	TP0464	COG1316		ywtF lytR yvhJ
TP0939	pyruvate ferredoxin oxidoreductase	TP0421 TP0658	COG1013		
TP0965	membrane fusion protein, putative	TP0965	COG0845	acrE yhiU yhiI acrA ylcD b0795 b2074	yknX yvrP
TP0068	hypothetical protein	nrdA	COG0820	yfgB	yloN
TP0702	hypothetical protein	flaB3	COG0739	yebA nlpD	yunA spoIIQ spoIVFA
TP0782	predicted coding region	fliY flgG-2	COG0739	yebA nlpD	yunA spoIIQ spoIVFA
TP0269	hypothetical protein	TP0218	COG0621	yleA yliG	ymcB yqeV
TP0024	hypothetical protein	flgG-2 flgK	COG0569	trkA	yjbQ yuaA ykqB
TP0100	thioredoxin, putative	fliY flgD	COG0526	dsbA trxA trxC dsbE	yneN ykvV bdbA ybdE
TP0731	hypothetical protein	TP0658	COG0494	wcaH yffH	yvcI nudF mutT ytkD
TP0330	cell division protein, putative	cheR	COG0465	hflB	yjoB ftsH
TP0067	hypothetical protein	flgK	COG0457	nfrA ygeG	rapA rapG
TP0095	hypothetical protein	TP0421	COG0457	nfrA ygeG	rapA rapG
TP0470	hypothetical protein	flaB3	COG0457	nfrA ygeG	rapA rapG
TP0471	hypothetical protein	TP0464	COG0457	nfrA ygeG	rapA rapG
TP0622	hypothetical protein	tig	COG0457	nfrA ygeG	rapA rapG
TP0648	hypothetical protein	fliG-2	COG0457	nfrA ygeG	rapA rapG
TP0626	exonuclease, putative	flgK	COG0420	sbcD	yhaO
TP0979	putative deoxyribonuclease	fliE	COG0084	ycfH	yabD

TP0031	hypothetical protein	flaB3	-		
TP0064	hypothetical protein	fliE cheR mcp2-3 flaB2	-		
TP0066	hypothetical protein	fliE fliG-2 fliY fliM flaB3	-		
TP0086	hypothetical protein	tig	-		
TP0088	hypothetical protein	cheR mcp2-3 fliY TP0959 TP0403	-		
TP0174	predicted coding region	flaB3	-		
TP0226	predicted coding region	csrA	-		
TP0260	predicted coding region	fliL TP0959 flgL	-		
TP0281	hypothetical protein	TP0378 tig	-		
TP0287	hypothetical protein	fliY flgG-2 csrA	-		
TP0314	hypothetical protein	flgG-2	-		
TP0377	hypothetical protein	flhA TP0218	-		
TP0382	hypothetical protein	flaB3	-		
TP0432	predicted coding region	TP0965	-		
TP0443	hypothetical protein	fliG-2	-		
TP0465	predicted coding region	flgK	-		
TP0467	hypothetical protein	TP0959	-		
TP0579	predicted coding region	fliE fliY flgD	-		
TP0592	hypothetical protein	TP0464	-		
TP0629	predicted coding region	fliE fliY flgD flaB3	-		
TP0665	hypothetical protein	fliG-2	-		
TP0698	predicted coding region	flaB3	-		
TP0700	predicted coding region	fliY	-		
TP0708	predicted coding region	fliF TP0567 flhA	-		
TP0711	hypothetical protein	TP0421 flhF flaB3	-		
TP0751	predicted coding region	flgG-2 flgK	-		
TP0763	predicted coding region	topA flgG-2	-		
TP0772	predicted coding region	fliY flgG-2	-		
TP0833	predicted coding region	fliI fliY fliR	-		
TP0836	predicted coding region	flaB3	-		
TP0851	predicted coding region	cheR	-		
TP0873	predicted coding region	flaB3	-		
TP0941	predicted coding region	TP0658	-		
TP0955	predicted coding region	metK	-		
TP0974	predicted coding region	TP0709 fliS	-		
TP0987	predicted coding region	groEL	-		
TP0992	hypothetical protein	flaB3	-		

Table 5.3 List of *E. coli* mutants screened for motility phenotype

(Mean percent motility: is the mean percentage of the swarming of 2 independent mutants tested, compared to wild type).

TPA_ORF	**Source**	**ECO_ID**	**JW_ID**	**Gene name**	**Motility Phenotype**	Mean percent motility
TP0030	Y2H	B4143	JW4103	groL	-	83.95
TP0040	KFP	B3072	JW3043	aer	-	79.49
TP0040	KFP	B1421	JW1417	trg	-	89.29
TP0040	KFP	B1885	JW1874	tap	non motile	7.22

TP0040	KFP	B1886	JW1875	tar	reduced	35.15
TP0040	KFP	B4355	JW4318	tsr	reduced	32.49
TP0067	Y2H	B3530	JW5862	bcsC	reduced	58.97
TP0068	Y2H	B2517	JW2501	yfgB	-	72.36
TP0080	Y2H	B0286	JW0280	yagT	-	84.29
TP0080	Y2H	B2868	JW2836	xdhC	-	72.58
TP0100	Y2H	B2195	JW2183	dsbE	-	70.40
TP0105	KFP	B3863	JW3835	polA	-	79.00
TP0121	Y2H	B4146	JW4106	yjeK	reduced	64.07
TP0140	Y2H	B1363	JW1358	trkG		85.93
TP0140	Y2H	B3849	JW5576	trkH	-	66.30
TP0141	Y2H	B1335	JW1329	ogt	-	83.04
TP0141	Y2H	B2213	JW2201	ada	-	71.54
TP0162	Y2H	B1860	JW1849	ruvB	-	81.52
TP0186	Y2H	B2955	JW2922	yggW	reduced	54.34
TP0186	Y2H	B3867	JW3838	hemN	-	70.63
TP0209	Y2H	B3299	JW3261	rpmJ	reduced	13.98
TP0247	Y2H	B4169	JW4127	amiB	-	63.27
TP0247	Y2H	B2435	JW2428	amiA	reduced	45.80
				amiC	-	65.77
TP0257	Y2H	B2239	JW2233	glpQ	-	75.90
TP0277	Y2H	B1830	JW1819	prc	-	29.29
TP0363	KFP	B1888	JW1877	cheA	non motile	6.67
TP0364	KFP	B1887	JW1876	cheW	non motile	6.67
TP0366	KFP	B1882	JW1871	cheY	non motile	6.67
TP0383	Y2H	B0081	JW0079	mraZ	-	65.10
TP0396	KFP	B1073	JW1060	flgB	non motile	6.67
TP0397	KFP	B1074	JW1061	flgC	non motile	6.67
TP0399	KFP	B1938	JW1922	fliF	non motile	6.67
TP0400	KFP	B1939	JW1923	fliG	non motile	6.67
TP0402	KFP	B1941	JW1925	fliI	non motile	6.67
TP0445	Y2H	B0424	JW5057	thiJ	-	82.67
TP0464	Y2H	B2960	JW2927	yggH	reduced	67.37
TP0471	Y2H	B1280	JW1272	yciM	reduced	34.15
TP0506	Y2H	B0436	JW0426	tig	-	95.36
TP0510	Y2H	B2569	JW2553	lepA	-	55.01
TP0514	Y2H	B4058	JW4019	uvrA	-	65.71
TP0550	KFP	B3706	JW3684	trmE	-	65.60
TP0554	Y2H	B3385	JW3348	gph	-	68.76
TP0561	Y2H	B2778	JW5445	ygcG	reduced	75.49
TP0626	Y2H	B0398	JW0388	sbcD	-	79.71
TP0630	KFP	B1884	JW1873	cheR	non motile	37.72
TP0631	KFP	B1883	JW1872	cheB	non motile	13.56
TP0660	KFP	B1082	JW1069	flgK	non motile	6.76
TP0709	KFP	B1922	JW1907	fliA	non motile	6.67
TP0714	KFP	B1879	JW1868	flhA	non motile	28.90
TP0715	KFP	B1880	JW1869	flhB	non motile	6.67

TP0716	KFP	B1950	JW1934	fliR	non motile	6.67
TP0717	KFP	B1949	JW1933	fliQ	non motile	45.81
TP0718	KFP	B1948	JW1932	fliP	non motile	6.67
TP0720	KFP	B1946	JW1930	fliN	non motile	15.63
TP0721	KFP	B1945	JW1929	fliM	non motile	6.67
TP0725	KFP	B1890	JW1879	motA	non motile	6.67
TP0727	KFP	B1076	JW1063	flgE	non motile	6.67
TP0728	KFP	B1075	JW1062	flgD	non motile	6.67
TP0731	Y2H	B3397	JW3360	nudE	reduced	40.20
TP0774	Y2H	B3508	JW5670	yhiD	-	77.94
TP0792	KFP	B1923	JW1908	fliC	non motile	7.09
TP0800	Y2H	B2010	JW5329	dacD	-	83.95
TP0800	Y2H	B0632	JW0627	dacA	-	79.44
TP0800	Y2H	B0839	JW0823	dacC	-	86.07
TP0801	KFP	B0882	JW0866	clpA	-	88.23
TP0872	KFP	B1924	JW1909	fliD	non motile	7.18
TP0887	Y2H	B3165	JW3134	rpsO	reduced	59.74
TP0939	Y2H	B1378	JW1372	ydbK	-	60.67
TP0943	KFP	B1925	JW1910	fliS	non motile	6.96
TP0945	Y2H	B3386	JW3349	rpe	reduced	21.35
TP0946	Y2H	B3740	JW3718	gidB	-	67.80
TP0959	KFP	B1081	JW1068	flgJ	non motile	22.08
TP0960	KFP	B1078	JW1065	flgG	non motile	6.67
TP0979	Y2H	B1100	JW1086	ycfH	reduced	68.98
TP0981	Y2H	B1535	JW1528	ydeH	-	70.41
TP0997	Y2H	B1766	JW1755	sppA	-	86.03
	KFP	B1881	JW1870	cheZ	non motile	10.08
	KFP	B1072	JW1059	flgA	non motile	45.01
	KFP	B1077	JW1064	flgF	non motile	10.08
	KFP	B1083	JW1070	flgL	non motile	34.71
	KFP	B1071	JW1058	flgM	reduced	46.79
	KFP	B1070	JW1057	flgN	non motile	27.92
	KFP	B1891	JW1880	flhC	non motile	26.02
	KFP	B1943	JW1927	fliK	non motile	45.41
	KFP	B1947	JW5316	fliO	non motile	10.08
TP0724	KFP	B1889	JW1878	MotB	non motile	33.91
	KFP	B1878	JW1867	flhE	reduced	27.27
	KFP	B1892	JW1881	flhD	reduced	46.47
	KFP	B1940	JW1924	fliH	reduced	29.22
	KFP	B1921	JW1906	fliZ	-	
	KFP	B1937	JW1921	fliE	-	
	KFP	B1079	JW5153	flgH	-	
TP0722	KFP	B1944	JW1928	fliL	-	
	KFP	B1926	JW1911	fliT	-	
TP0247	Y2H	B2435		amiA	reduced	41.93
TP0247	Y2H	B2817		amiC	-	65.77
TP0236	Y2H	B3842		rfaH	non motile	21

TP0945	Y2H	B3386		rpe	non motile	21.34
TP0209	Y2H	B3299		rpmJ	reduced	20.54
TP0979	Y2H	B1100		ycfH	reduced	37.95
TP0471	Y2H	B1280		yciM	reduced	21.87
TP0464	Y2H	B2960		yggH	reduced	
TP0186	Y2H	B2955		yggW	reduced	33.33

6.5 Additional projects not described here in detail

During my PhD work I carried out yeast two-hybrid (Y2H) screens of *Sulfolobus spindle-shaped virus* 1 (SSV1) and Varicella Zoster Virus (VZV) as side projects. These projects were not directly related to my PhD work which focused on the "Interactome" of *T. pallidum* using the Y2H system. However, the study of viral interactomes were interesting because of the small genome sizes and certain technical and biological differences of these screens (see below).

***Sulfolobus spindle-shaped virus* 1 (SSV1)** is the type virus of the *Fuselloviridae* family and the first high-temperature virus (archaebacterial virus) genome to be completely sequenced (Wexler et al., 2000). Sulfolobus *spindle-shaped virus* (SSV1, SSV2, SSV RH and SSV K1), and their *Sulfolobus* hosts are emerging as a model system for examining archaea and life at high temperatures. However, the viruses of these organisms are poorly understood. Even though SSV1 is considerd to be the most well-charaterised *Sulfolobus* virus, the function of 28 of the 34 SSV1 virus proteins remains mysterious (Stedman et al., 2003). In order to study the functions of SSV1 proteins more systematically, we decided to screen its proteome for protein-protein interactions. Although this does not involve host-virus interactions, we expected a number of interactions that should provide some evidence for possible functions. Additionally, this study was technically challenging as we did not know before whether the Y2H system is able to detect PPIs of thermophilic organisms.

Varicella-zoster virus (VZV) is a human herpesvirus that commonly causes chicken pox (varicella), usually in young children. Following primary infection a lifelong latent infection is established, and the virus often reactivates in adulthood or senescence to cause shingles (zoster). Although the VZV genome contains ORFs that are known or predicted to encode more than 70 distinct proteins, functions have been assigned to only about half of these VZV gene products (Cohen 2001). In many cases, the contributions of VZV genes to viral replication are presumed because of their partial sequence homologies with ORFs in Herpes Simplex Virus type 1 (HSV-1), which is the prototype of the alpha-herpesviruses (Roizman 2001). Studies have revealed a significant number of interactions between herpesviral and host proteins. Surprisingly little is known about interactions among herpesviral proteins. In order to study the functions of VZV proteins more systematically, we decided to gerate a genomewide protein interaction map for human VZV.

The results from these studies and another one on transcriptional activators in yeast are provided here as reprints (VZV, yeast) and as a manuscript currently under review (SSV1)

Supplementary Table

Table S1: Predicted interologs based on *T. pallidum* Y2H interactions

T. pallidum	*B. subtils*	*B. burgdorferi*	*E. coli*	*H pylori*	*S. typhimurium*
groL-orf4	-	BB0649-BB0282	-	-	-
groL-gidB	BSU06030-BSU41000	BB0649-BB0177	b4143-b3740	HP0010-HP1063	-
groL-TP0987	-	-	-	-	-
groL-TP0050	-	BB0649-BB0103	-	HP0010-HP0735	-
topA-secY	BSU16120-BSU01360	-	b1763-b3300	HP0116-HP1300	STM1298-STM3420
	BSU04260-BSU01360	-	-	HP0440-HP1300	-
topA-TP0763	-	-	-	-	-
fliE-TP0788	-	-	-	-	-
fliE-TP0064	-	-	-	-	-
fliE-ruvB	-	BB0292-BB0022	b1937-b1860	HP1557-HP1059	STM1968-STM1894
fliE-TP0979	BSU16200-BSU00390	BB0292-BB0194	b1937-b1100	HP1557-HP1573	STM1968-STM1202
	-	-	-	-	STM1968-STM4564
fliE-TP0066	-	-	-	-	-
fliE-rpoE	BSU16200-BSU27120	-	b1937-b2573	-	STM1968-STM2640
	BSU16200-BSU12560	-	b1937-b4293	-	-
	BSU16200-BSU38700	-	-	-	-
	BSU16200-BSU14730	-	-	-	-
	BSU16200-BSU00980	-	-	-	-
	BSU16200-BSU26840	-	-	-	-
	BSU16200-BSU09520	-	-	-	-
	BSU16200-BSU23100	-	-	-	-
	BSU16200-BSU01730	-	-	-	-
fliE-TP0530	-	-	-	-	-
fliE-mraZ	BSU16200-BSU15130	-	b1937-b0081	-	STM1968-STM0119
fliE-TP0661	-	-	-	-	-
fliE-TP0629	-	-	-	-	-
fliE-flgB	BSU16200-BSU16180	BB0292-BB0294	b1937-b1073	HP1557-HP1559	STM1968-STM1174
fliE-fliE	BSU16200-BSU16200	BB0292-BB0292	b1937-b1937	HP1557-HP1557	STM1968-STM1968
fliE-gyrA	BSU16200-BSU00070	BB0292-BB0435	b1937-b3019	HP1557-HP0701	STM1968-STM3174
	BSU16200-BSU18100	BB0292-BB0035	b1937-b2231	-	STM1968-STM2272
fliE-TP0579	-	-	-	-	-
fliE-flaB3	BSU16200-BSU35360	BB0292-BB0147	b1937-b1923	HP1557-HP0601	STM1968-STM1959
	BSU16200-BSU35400	BB0292-BB0182	b1937-b1083	HP1557-HP0115	STM1968-STM1184
	BSU16200-BSU35150	-	-	HP1557-HP0295	STM1968-STM2771
fliE-fliG-1	BSU16200-BSU16220	BB0292-BB0290	b1937-b1939	HP1557-HP0352	STM1968-STM1970
	-	BB0292-BB0221	-	-	-
fliE-TP0050	-	BB0292-BB0103	-	HP1557-HP0735	-
fliF-TP0561	BSU16210-BSU06200	-	-	-	-
fliF-TP0708	-	-	-	-	-
fliF-mgtE	BSU16210-BSU13300	BB0291-BB0380	-	-	-
fliF-fliG-1	BSU16210-BSU16220	BB0291-BB0290	b1938-b1939	HP0351-HP0352	STM1969-STM1970
	-	BB0291-BB0221	-	-	-
fliG-2-TP0233	BSU16220-BSU07180	-	-	-	-

	BSU16220-BSU04670	-	-	-	-
	BSU16220-BSU23470	-	-	-	-
	BSU16220-BSU04680	-	-	-	-
	BSU16220-BSU04710	-	-	-	-
	BSU16220-BSU13200	-	-	-	-
	BSU16220-BSU19450	-	-	-	-
	BSU16220-BSU24760	-	-	-	-
	BSU16220-BSU07170	-	-	-	-
fliG-2-rpmJ	BSU16220-BSU01400	BB0290-BB0499	b1939-b3299	HP0352-HP1297	STM1970-STM3419
	-	BB0221-BB0499	-	-	STM1970-STM0470
fliG-2-TP0066	-	-	-	-	-
fliG-2-cysS	BSU16220-BSU00940	BB0290-BB0599	b1939-b0526	HP0352-HP0886	STM1970-STM0537
	-	BB0221-BB0599	-	-	-
fliG-2-rpoE	BSU16220-BSU27120	-	b1939-b2573	-	STM1970-STM2640
	BSU16220-BSU12560	-	b1939-b4293	-	-
	BSU16220-BSU38700	-	-	-	-
	BSU16220-BSU14730	-	-	-	-
	BSU16220-BSU00980	-	-	-	-
	BSU16220-BSU26840	-	-	-	-
	BSU16220-BSU09520	-	-	-	-
	BSU16220-BSU23100	-	-	-	-
	BSU16220-BSU01730	-	-	-	-
fliG-2-TP0443	-	-	-	-	-
fliG-2-rpmA	BSU16220-BSU27940	BB0290-BB0780	b1939-b3185	HP0352-HP0297	STM1970-STM3303
	-	BB0221-BB0780	-	-	-
fliG-2-TP0648	BSU16220-BSU12430	BB0290-BB0542	b1939-b0568	HP0352-HP0275	STM1970-STM3616
	BSU16220-BSU03770	BB0290-BB0106	b1939-b2851	HP0352-HP1479	STM1970-STM2886
	BSU16220-BSU27490	BB0290-BB0133	-	-	STM1970-STM1403
	BSU16220-BSU28240	BB0290-BB0459	-	-	STM1970-STM1228
	BSU16220-BSU25830	BB0290-BB0261	-	-	STM1970-STM1399
	BSU16220-BSU05010	BB0290-BB0171	-	-	-
	BSU16220-BSU02820	BB0290-BB0537	-	-	-
	BSU16220-BSU36380	BB0290-BB0058	-	-	-
	BSU16220-BSU18910	BB0290-BB0210	-	-	-
	BSU16220-BSU22590	BB0290-BB0170	-	-	-
	BSU16220-BSU37460	BB0290-BB0326	-	-	-
	BSU16220-BSU36690	BB0290-BB0195	-	-	-
	BSU16220-BSU34810	BB0290-BB0714	-	-	-
	BSU16220-BSU06140	BB0290-BB0209	-	-	-
	BSU16220-BSU06830	BB0290-BB0298	-	-	-
	BSU16220-BSU40300	BB0221-BB0542	-	-	-
	-	BB0221-BB0106	-	-	-
	-	BB0221-BB0133	-	-	-
	-	BB0221-BB0459	-	-	-
	-	BB0221-BB0261	-	-	-
	-	BB0221-BB0171	-	-	-
	-	BB0221-BB0537	-	-	-
	-	BB0221-BB0058	-	-	-
	-	BB0221-BB0210	-	-	-

	-	BB0221-BB0170	-	-	-
	-	BB0221-BB0326	-	-	-
	-	BB0221-BB0195	-	-	-
	-	BB0221-BB0714	-	-	-
	-	BB0221-BB0209	-	-	-
	-	BB0221-BB0298	-	-	-
fliG-2-TP0665	-	-	-	-	-
fliG-2-fliF	BSU16220-BSU16210	BB0290-BB0291	b1939-b1938	HP0352-HP0351	STM1970-STM1969
	-	BB0221-BB0291	-	-	-
fliG-2-TP0014	-	-	-	-	-
fliG-2-fliG-1	BSU16220-BSU16220	BB0290-BB0290	b1939-b1939	HP0352-HP0352	STM1970-STM1970
	-	BB0290-BB0221	-	-	-
	-	BB0221-BB0290	-	-	-
	-	BB0221-BB0221	-	-	-
fliG-2-rplI	BSU16220-BSU40500	BB0290-BB0112	b1939-b4203	HP0352-HP0514	STM1970-STM4394
	-	BB0221-BB0112	-	-	-
fliH-fliG-1	BSU16230-BSU16220	BB0289-BB0290	b1940-b1939	HP0353-HP0352	STM1971-STM1970
	-	BB0289-BB0221	-	-	-
fliI-TP0413	BSU16240-BSU01770	BB0288-BB0004	b1941-b3176	HP1420-HP1275	STM1415-STM2104
	BSU16240-BSU09310	BB0288-BB0835	b1941-b2048	HP1420-HP0075	STM1415-STM3294
	-	-	-	-	STM1415-STM2083
	-	-	-	-	STM2894-STM2104
	-	-	-	-	STM2894-STM3294
	-	-	-	-	STM2894-STM2083
	-	-	-	-	STM1972-STM2104
	-	-	-	-	STM1972-STM3294
	-	-	-	-	STM1972-STM2083
fliI-mcp4	BSU16240-BSU31240	BB0288-BB0597	b1941-b4355	HP1420-HP0599	STM1415-STM1919
	BSU16240-BSU13950	BB0288-BB0680	b1941-b1421	HP1420-HP0082	STM1415-STM3138
	BSU16240-BSU03440	BB0288-BB0596	b1941-b1886	HP1420-HP0103	STM1415-STM3152
	BSU16240-BSU10380	BB0288-BB0681	b1941-b3072	HP1420-HP0099	STM1415-STM1626
	BSU16240-BSU33690	BB0288-BB0578	b1941-b1885	-	STM1415-STM3577
	BSU16240-BSU07360	-	-	-	STM1415-STM1657
	BSU16240-BSU31260	-	-	-	STM1415-STM4533
	BSU16240-BSU31230	-	-	-	STM1415-STM3216
	BSU16240-BSU18610	-	-	-	STM2894-STM1919
	BSU16240-BSU31250	-	-	-	STM2894-STM3138
	-	-	-	-	STM2894-STM3152
	-	-	-	-	STM2894-STM1626
	-	-	-	-	STM2894-STM3577
	-	-	-	-	STM2894-STM1657
	-	-	-	-	STM2894-STM4533
	-	-	-	-	STM2894-STM3216
	-	-	-	-	STM1972-STM1919
	-	-	-	-	STM1972-STM3138
	-	-	-	-	STM1972-STM3152
	-	-	-	-	STM1972-STM1626
	-	-	-	-	STM1972-STM3577
	-	-	-	-	STM1972-STM1657

	-	-	-	-	STM1972-STM4533
	-	-	-	-	STM1972-STM3216
fliI-TP0833	-	-	-	-	-
TP0403-recR	BSU16250-BSU00210	-	b1942-b0472	-	STM1973-STM0486
TP0403-flaB1	BSU16250-BSU35360	-	b1942-b1923	-	STM1973-STM1959
	BSU16250-BSU35400	-	b1942-b1083	-	STM1973-STM1184
	BSU16250-BSU35150	-	-	-	STM1973-STM2771
trmB-TP0739	BSU29900-BSU35840	-	-	-	-
	BSU29900-BSU35650	-	-	-	-
	BSU29900-BSU35520	-	-	-	-
TP0567-TP0421	-	-	-	-	-
TP0567-TP0708	-	-	-	-	-
TP0567-mgtE	-	-	-	-	-
TP0567-TP0675	-	-	-	-	-
TP0567-flaB3	-	-	-	-	-
cheR-TP0877	-	BB0040-BB0374	-	-	-
	-	BB0670-BB0374	-	-	-
	-	BB0414-BB0374	-	-	-
cheR-pyrG	-	BB0040-BB0575	b1884-b2780	-	STM1918-STM2953
	-	BB0670-BB0575	-	-	-
	-	BB0414-BB0575	-	-	-
cheR-TP0330	-	BB0040-BB0789	b1884-b3178	-	STM1918-STM3296
	-	BB0670-BB0789	-	-	-
	-	BB0414-BB0789	-	-	-
cheR-TP0064	-	-	-	-	-
cheR-ruvB	-	BB0040-BB0022	b1884-b1860	-	STM1918-STM1894
	-	BB0670-BB0022	-	-	-
	-	BB0414-BB0022	-	-	-
cheR-rpoE	-	-	b1884-b2573	-	STM1918-STM2640
	-	-	b1884-b4293	-	-
cheR-TP0088	-	-	-	-	-
cheR-flgB	-	BB0040-BB0294	b1884-b1073	-	STM1918-STM1174
	-	BB0670-BB0294	-	-	-
	-	BB0414-BB0294	-	-	-
cheR-TP0851	-	-	-	-	-
cheR-fliE	-	BB0040-BB0292	b1884-b1937	-	STM1918-STM1968
	-	BB0670-BB0292	-	-	-
	-	BB0414-BB0292	-	-	-
cheR-leuS	-	BB0040-BB0251	b1884-b0642	-	STM1918-STM0648
	-	BB0670-BB0251	-	-	-
	-	BB0414-BB0251	-	-	-
cheR-flaB3	-	BB0040-BB0147	b1884-b1923	-	STM1918-STM1959
	-	BB0040-BB0182	b1884-b1083	-	STM1918-STM1184
	-	BB0670-BB0147	-	-	STM1918-STM2771
	-	BB0670-BB0182	-	-	-
	-	BB0414-BB0147	-	-	-
	-	BB0414-BB0182	-	-	-
cheR-TP0121	-	-	b1884-b4146	-	STM1918-STM4333
cheR-recR	-	-	b1884-b0472	-	STM1918-STM0486

cheR-dnaH	-	BB0040-BB0765	b1884-b0470	-	STM1918-STM0484
	-	BB0040-BB0461	-	-	-
	-	BB0670-BB0765	-	-	-
	-	BB0670-BB0461	-	-	-
	-	BB0414-BB0765	-	-	-
	-	BB0414-BB0461	-	-	-
cheR-fliG-1	-	BB0040-BB0290	b1884-b1939	-	STM1918-STM1970
	-	BB0040-BB0221	-	-	-
	-	BB0670-BB0290	-	-	-
	-	BB0670-BB0221	-	-	-
	-	BB0414-BB0290	-	-	-
	-	BB0414-BB0221	-	-	-
mcp4-TP0064	-	-	-	-	-
mcp4-TP0088	-	-	-	-	-
mcp4-flgC	BSU31240-BSU16190	BB0597-BB0293	b4355-b1074	HP0599-HP1558	STM1919-STM1175
	BSU13950-BSU16190	BB0680-BB0293	b1421-b1074	HP0082-HP1558	STM3138-STM1175
	BSU03440-BSU16190	BB0596-BB0293	b1886-b1074	HP0103-HP1558	STM3152-STM1175
	BSU10380-BSU16190	BB0681-BB0293	b3072-b1074	HP0099-HP1558	STM1626-STM1175
	BSU33690-BSU16190	BB0578-BB0293	b1885-b1074	-	STM3577-STM1175
	BSU07360-BSU16190	-	-	-	STM1657-STM1175
	BSU31260-BSU16190	-	-	-	STM4533-STM1175
	BSU31230-BSU16190	-	-	-	STM3216-STM1175
	BSU18610-BSU16190	-	-	-	-
	BSU31250-BSU16190	-	-	-	-
csrA-TP0226	-	-	-	-	-
csrA-TP0518	BSU35370-BSU15800	-	-	HP1442-HP1291	-
TP0658-tprI	-	-	-	-	-
TP0658-argS	BSU35380-BSU37330	BB0183-BB0594	-	HP1154-HP0319	-
	-	-	-	HP1377-HP0319	-
TP0658-hydG	BSU35380-BSU15530	-	-	-	-
TP0658-TP0941	-	-	-	-	-
TP0658-TP0939	-	-	-	HP1154-HP1111	-
	-	-	-	HP1154-HP0590	-
	-	-	-	HP1377-HP1111	-
	-	-	-	HP1377-HP0590	-
flgL-TP0788	-	-	-	-	-
flgL-orf4	-	BB0147-BB0282	-	-	-
	-	BB0182-BB0282	-	-	-
flgL-dnaN	BSU35360-BSU00020	BB0147-BB0438	b1923-b3701	HP0601-HP0500	STM1959-STM3837
	BSU35400-BSU00020	BB0182-BB0438	b1083-b3701	HP0115-HP0500	STM1184-STM3837
	BSU35150-BSU00020	-	-	HP0295-HP0500	STM2771-STM3837
flgL-recR	BSU35360-BSU00210	-	b1923-b0472	HP0601-HP0925	STM1959-STM0486
	BSU35400-BSU00210	-	b1083-b0472	HP0115-HP0925	STM1184-STM0486
	BSU35150-BSU00210	-	-	HP0295-HP0925	STM2771-STM0486
flgK-TP0661	-	-	-	-	-
flaA-2-TP0233	-	-	-	-	-
flaA-2-amiA	-	-	-	-	-
TP0709-TP0974	-	-	-	-	-
flhF-TP0711	-	-	-	-	-

flhB-TP0561	BSU16380-BSU06200	-	-	-	-
flhB-mgtE	BSU16380-BSU13300	BB0272-BB0380	-	-	-
fliR-TP0561	BSU16370-BSU06200	-	-	-	-
fliR-motA	BSU16370-BSU29730	BB0273-BB0281	b1950-b1890	HP0173-HP0815	STM1981-STM1923
	BSU16370-BSU13690	-	-	-	-
fliR-mgtE	BSU16370-BSU13300	BB0273-BB0380	-	-	-
fliR-mgtC	BSU16370-BSU06650	-	b1950-b3508	-	STM1981-STM3764
fliQ-TP0561	BSU16360-BSU06200	-	-	-	-
fliQ-mgtE	BSU16360-BSU13300	BB0274-BB0380	-	-	-
fliY-TP0788	-	-	-	-	-
fliY-TP0233	-	-	-	-	-
fliY-rpmJ	-	BB0277-BB0499	b1946-b3299	HP1030-HP1297	STM2891-STM3419
	-	-	-	HP0584-HP1297	STM2891-STM0470
	-	-	-	-	STM1418-STM3419
	-	-	-	-	STM1418-STM0470
	-	-	-	-	STM1977-STM3419
	-	-	-	-	STM1977-STM0470
fliY-TP0048	-	BB0277-BB0538	-	HP1030-HP1542	-
	-	BB0277-BB0245	-	HP0584-HP1542	-
	-	BB0277-BB0231	-	-	-
fliY-TP0066	-	-	-	-	-
fliY-TP0700	-	-	-	-	-
fliY-TP0088	-	-	-	-	-
fliY-rpsS	-	BB0277-BB0482	b1946-b3316	HP1030-HP1315	STM2891-STM3436
	-	-	-	HP0584-HP1315	STM1418-STM3436
	-	-	-	-	STM1977-STM3436
fliY-TP0287	-	-	-	-	-
fliY-flgK	-	BB0277-BB0181	b1946-b1082	HP1030-HP1119	STM2891-STM1183
	-	-	-	HP0584-HP1119	STM1418-STM1183
	-	-	-	-	STM1977-STM1183
fliY-TP0629	-	-	-	-	-
fliY-TP0833	-	-	-	-	-
fliY-flgC	-	BB0277-BB0293	b1946-b1074	HP1030-HP1558	STM2891-STM1175
	-	-	-	HP0584-HP1558	STM1418-STM1175
	-	-	-	-	STM1977-STM1175
fliY-TP0843	-	BB0277-BB0602	b1946-b1000	HP1030-HP1336	STM2891-STM1112
	-	BB0277-BB0655	-	HP1030-HP1024	STM1418-STM1112
	-	-	-	HP0584-HP1336	STM1977-STM1112
	-	-	-	HP0584-HP1024	-
fliY-TP0100	-	BB0277-BB0061	b1946-b3860	HP1030-HP0824	STM2891-STM3915
	-	-	b1946-b3781	HP1030-HP1458	STM2891-STM2248
	-	-	b1946-b2582	HP0584-HP0824	STM2891-STM3997
	-	-	b1946-b2195	HP0584-HP1458	STM2891-STM3193
	-	-	-	-	STM2891-STM1116
	-	-	-	-	STM2891-STM3813
	-	-	-	-	STM2891-STM2649
	-	-	-	-	STM1418-STM3915
	-	-	-	-	STM1418-STM2248
	-	-	-	-	STM1418-STM3997

	-	-	-	-	STM1418-STM3193
	-	-	-	-	STM1418-STM1116
	-	-	-	-	STM1418-STM3813
	-	-	-	-	STM1418-STM2649
	-	-	-	-	STM1977-STM3915
	-	-	-	-	STM1977-STM2248
	-	-	-	-	STM1977-STM3997
	-	-	-	-	STM1977-STM3193
	-	-	-	-	STM1977-STM1116
	-	-	-	-	STM1977-STM3813
	-	-	-	-	STM1977-STM2649
fliY-TP0772	-	-	-	-	-
fliY-TP0579	-	-	-	-	-
fliY-flaB3	-	BB0277-BB0147	b1946-b1923	HP1030-HP0601	STM2891-STM1959
	-	BB0277-BB0182	b1946-b1083	HP1030-HP0115	STM2891-STM1184
	-	-	-	HP1030-HP0295	STM2891-STM2771
	-	-	-	HP0584-HP0601	STM1418-STM1959
	-	-	-	HP0584-HP0115	STM1418-STM1184
	-	-	-	HP0584-HP0295	STM1418-STM2771
	-	-	-	-	STM1977-STM1959
	-	-	-	-	STM1977-STM1184
	-	-	-	-	STM1977-STM2771
fliY-TP0782	-	BB0277-BB0262	b1946-b1856	HP1030-HP0750	STM2891-STM3038
	-	BB0277-BB0255	b1946-b2742	HP1030-HP1544	STM2891-STM2925
	-	BB0277-BB0761	-	HP1030-HP1543	STM2891-STM1890
	-	BB0277-BB0246	-	HP1030-HP0506	STM1418-STM3038
	-	-	-	HP1030-HP1054	STM1418-STM2925
	-	-	-	HP0584-HP0750	STM1418-STM1890
	-	-	-	HP0584-HP1544	STM1977-STM3038
	-	-	-	HP0584-HP1543	STM1977-STM2925
	-	-	-	HP0584-HP0506	STM1977-STM1890
	-	-	-	HP0584-HP1054	-
fliY-dnaH	-	BB0277-BB0765	b1946-b0470	HP1030-HP0717	STM2891-STM0484
	-	BB0277-BB0461	-	HP0584-HP0717	STM1418-STM0484
	-	-	-	-	STM1977-STM0484
fliY-fliG-1	-	BB0277-BB0290	b1946-b1939	HP1030-HP0352	STM2891-STM1970
	-	BB0277-BB0221	-	HP0584-HP0352	STM1418-STM1970
	-	-	-	-	STM1977-STM1970
fliM-TP0066	-	-	-	-	-
fliM-rpoE	BSU16310-BSU27120	-	b1945-b2573	-	STM1976-STM2640
	BSU16310-BSU12560	-	b1945-b4293	-	-
	BSU16310-BSU38700	-	-	-	-
	BSU16310-BSU14730	-	-	-	-
	BSU16310-BSU00980	-	-	-	-
	BSU16310-BSU26840	-	-	-	-
	BSU16310-BSU09520	-	-	-	-
	BSU16310-BSU23100	-	-	-	-
	BSU16310-BSU01730	-	-	-	-
fliM-fliG-1	BSU16310-BSU16220	BB0278-BB0290	b1945-b1939	HP1031-HP0352	STM1976-STM1970

	-	BB0278-BB0221	-	-	-
fliL-TP0260	-	-	-	-	-
fliL-TP0561	BSU16300-BSU06200	-	-	-	-
motA-motA	BSU29730-BSU29730	BB0281-BB0281	b1890-b1890	HP0815-HP0815	STM1923-STM1923
	BSU29730-BSU13690	-	-	-	-
	BSU13690-BSU29730	-	-	-	-
	BSU13690-BSU13690	-	-	-	-
motA-mgtE	BSU29730-BSU13300	BB0281-BB0380	-	-	-
	BSU13690-BSU13300	-	-	-	-
flgE-rpmJ	-	BB0283-BB0499	b1076-b3299	HP0908-HP1297	STM1177-STM3419
	-	-	-	HP0870-HP1297	STM1177-STM0470
flgE-tpn50	-	BB0283-BB0167	b1076-b0957	HP0908-HP1125	STM1177-STM0749
	-	-	b1076-b2605	HP0870-HP1125	STM1177-STM3645
	-	-	b1076-b0741	-	STM1177-STM1070
flgD-TP0788	-	-	-	-	-
flgD-gph-2	BSU16280-BSU34970	BB0284-BB0676	b1075-b3385	HP0907-HP1207	STM1176-STM3482
flgD-orf4	-	BB0284-BB0282	-	-	-
flgD-TP0661	-	-	-	-	-
flgD-TP0629	-	-	-	-	-
flgD-flgE	-	BB0284-BB0283	b1075-b1076	HP0907-HP0908	STM1176-STM1177
	-	-	-	HP0907-HP0870	-
flgD-flgB	BSU16280-BSU16180	BB0284-BB0294	b1075-b1073	HP0907-HP1559	STM1176-STM1174
flgD-fliE	BSU16280-BSU16200	BB0284-BB0292	b1075-b1937	HP0907-HP1557	STM1176-STM1968
flgD-TP0100	BSU16280-BSU18010	BB0284-BB0061	b1075-b3860	HP0907-HP0824	STM1176-STM3915
	BSU16280-BSU28500	-	b1075-b3781	HP0907-HP1458	STM1176-STM2248
	BSU16280-BSU05510	-	b1075-b2582	-	STM1176-STM3997
	BSU16280-BSU04550	-	b1075-b2195	-	STM1176-STM3193
	BSU16280-BSU14230	-	-	-	STM1176-STM1116
	BSU16280-BSU20030	-	-	-	STM1176-STM3813
	BSU16280-BSU32770	-	-	-	STM1176-STM2649
	BSU16280-BSU23150	-	-	-	-
	BSU16280-BSU29840	-	-	-	-
	BSU16280-BSU13840	-	-	-	-
	BSU16280-BSU21460	-	-	-	-
	BSU16280-BSU01980	-	-	-	-
flgD-gyrA	BSU16280-BSU00070	BB0284-BB0435	b1075-b3019	HP0907-HP0701	STM1176-STM3174
	BSU16280-BSU18100	BB0284-BB0035	b1075-b2231	-	STM1176-STM2272
flgD-leuS	BSU16280-BSU30320	BB0284-BB0251	b1075-b0642	HP0907-HP1547	STM1176-STM0648
flgD-TP0579	-	-	-	-	-
flgD-flaB3	BSU16280-BSU35360	BB0284-BB0147	b1075-b1923	HP0907-HP0601	STM1176-STM1959
	BSU16280-BSU35400	BB0284-BB0182	b1075-b1083	HP0907-HP0115	STM1176-STM1184
	BSU16280-BSU35150	-	-	HP0907-HP0295	STM1176-STM2771
flgD-TP0050	-	BB0284-BB0103	-	HP0907-HP0735	-
flaB2-TP0064	-	-	-	-	-
flaB2-TP0658	BSU35360-BSU35380	BB0147-BB0183	-	HP0601-HP1154	-
	BSU35400-BSU35380	BB0182-BB0183	-	HP0601-HP1377	-
	BSU35150-BSU35380	-	-	HP0115-HP1154	-
	-	-	-	HP0115-HP1377	-
	-	-	-	HP0295-HP1154	-

	-	-	-	HP0295-HP1377	-
metK-dnaN	BSU30550-BSU00020	BB0376-BB0438	b2942-b3701	HP0197-HP0500	STM3090-STM3837
metK-TP0955	-	-	-	-	-
flaB1-TP0658	BSU35360-BSU35380	BB0147-BB0183	-	HP0601-HP1154	-
	BSU35400-BSU35380	BB0182-BB0183	-	HP0601-HP1377	-
	BSU35150-BSU35380	-	-	HP0115-HP1154	-
	-	-	-	HP0115-HP1377	-
	-	-	-	HP0295-HP1154	-
	-	-	-	HP0295-HP1377	-
flaB1-TP0050	-	BB0147-BB0103	-	HP0601-HP0735	-
	-	BB0182-BB0103	-	HP0115-HP0735	-
	-	-	-	HP0295-HP0735	-
flaB3-TP0788	-	-	-	-	-
flaB3-TP0066	-	-	-	-	-
flaB3-TP0174	-	-	-	-	-
flaB3-TP0702	BSU35360-BSU32340	BB0147-BB0262	b1923-b1856	HP0601-HP0750	STM1959-STM3038
	BSU35360-BSU36550	BB0147-BB0255	b1923-b2742	HP0601-HP1544	STM1959-STM2925
	BSU35360-BSU27980	BB0147-BB0761	b1083-b1856	HP0601-HP1543	STM1959-STM1890
	BSU35400-BSU32340	BB0147-BB0246	b1083-b2742	HP0601-HP0506	STM1184-STM3038
	BSU35400-BSU36550	BB0182-BB0262	-	HP0601-HP1054	STM1184-STM2925
	BSU35400-BSU27980	BB0182-BB0255	-	HP0115-HP0750	STM1184-STM1890
	BSU35150-BSU32340	BB0182-BB0761	-	HP0115-HP1544	STM2771-STM3038
	BSU35150-BSU36550	BB0182-BB0246	-	HP0115-HP1543	STM2771-STM2925
	BSU35150-BSU27980	-	-	HP0115-HP0506	STM2771-STM1890
	-	-	-	HP0115-HP1054	-
	-	-	-	HP0295-HP0750	-
	-	-	-	HP0295-HP1544	-
	-	-	-	HP0295-HP1543	-
	-	-	-	HP0295-HP0506	-
	-	-	-	HP0295-HP1054	-
flaB3-cheR	-	BB0147-BB0040	b1923-b1884	-	STM1959-STM1918
	-	BB0147-BB0670	b1083-b1884	-	STM1184-STM1918
	-	BB0147-BB0414	-	-	STM2771-STM1918
	-	BB0182-BB0040	-	-	-
	-	BB0182-BB0670	-	-	-
	-	BB0182-BB0414	-	-	-
flaB3-TP0661	-	-	-	-	-
flaB3-TP0629	-	-	-	-	-
flaB3-flgB	BSU35360-BSU16180	BB0147-BB0294	b1923-b1073	HP0601-HP1559	STM1959-STM1174
	BSU35400-BSU16180	BB0182-BB0294	b1083-b1073	HP0115-HP1559	STM1184-STM1174
	BSU35150-BSU16180	-	-	HP0295-HP1559	STM2771-STM1174
flaB3-TP0658	BSU35360-BSU35380	BB0147-BB0183	-	HP0601-HP1154	-
	BSU35400-BSU35380	BB0182-BB0183	-	HP0601-HP1377	-
	BSU35150-BSU35380	-	-	HP0115-HP1154	-
	-	-	-	HP0115-HP1377	-
	-	-	-	HP0295-HP1154	-
	-	-	-	HP0295-HP1377	-
flaB3-gyrA	BSU35360-BSU00070	BB0147-BB0435	b1923-b3019	HP0601-HP0701	STM1959-STM3174
	BSU35360-BSU18100	BB0147-BB0035	b1923-b2231	HP0115-HP0701	STM1959-STM2272

	BSU35400-BSU00070	BB0182-BB0435	b1083-b3019	HP0295-HP0701	STM1184-STM3174
	BSU35400-BSU18100	BB0182-BB0035	b1083-b2231	-	STM1184-STM2272
	BSU35150-BSU00070	-	-	-	STM2771-STM3174
	BSU35150-BSU18100	-	-	-	STM2771-STM2272
flaB3-flaB3	BSU35360-BSU35360	BB0147-BB0147	b1923-b1923	HP0601-HP0601	STM1959-STM1959
	BSU35360-BSU35400	BB0147-BB0182	b1923-b1083	HP0601-HP0115	STM1959-STM1184
	BSU35360-BSU35150	BB0182-BB0147	b1083-b1923	HP0601-HP0295	STM1959-STM2771
	BSU35400-BSU35360	BB0182-BB0182	b1083-b1083	HP0115-HP0601	STM1184-STM1959
	BSU35400-BSU35400	-	-	HP0115-HP0115	STM1184-STM1184
	BSU35400-BSU35150	-	-	HP0115-HP0295	STM1184-STM2771
	BSU35150-BSU35360	-	-	HP0295-HP0601	STM2771-STM1959
	BSU35150-BSU35400	-	-	HP0295-HP0115	STM2771-STM1184
	BSU35150-BSU35150	-	-	HP0295-HP0295	STM2771-STM2771
flaB3-TP0873	-	-	-	-	-
flaB3-TP0050	-	BB0147-BB0103	-	HP0601-HP0735	-
	-	BB0182-BB0103	-	HP0115-HP0735	-
	-	-	-	HP0295-HP0735	-
fliS-TP0064	-	-	-	-	-
fliS-TP0333	BSU35330-BSU04630	BB0550-BB0346	b1925-b0891	HP0753-HP0785	STM1961-STM0961
fliS-TP0260	-	-	-	-	-
fliS-TP0066	-	-	-	-	-
fliS-ruvC	-	-	b1925-b1863	HP0753-HP0877	STM1961-STM1898
fliS-TP0088	-	-	-	-	-
fliS-TP0518	BSU35330-BSU15800	-	-	HP0753-HP1291	-
fliS-TP0268	BSU35330-BSU12430	BB0550-BB0542	b1925-b0568	HP0753-HP0275	STM1961-STM3616
	BSU35330-BSU03770	BB0550-BB0106	b1925-b2851	HP0753-HP1479	STM1961-STM2886
	BSU35330-BSU27490	BB0550-BB0133	-	-	STM1961-STM1403
	BSU35330-BSU28240	BB0550-BB0459	-	-	STM1961-STM1228
	BSU35330-BSU25830	BB0550-BB0261	-	-	STM1961-STM1399
	BSU35330-BSU05010	BB0550-BB0171	-	-	-
	BSU35330-BSU02820	BB0550-BB0537	-	-	-
	BSU35330-BSU36380	BB0550-BB0058	-	-	-
	BSU35330-BSU18910	BB0550-BB0210	-	-	-
	BSU35330-BSU22590	BB0550-BB0170	-	-	-
	BSU35330-BSU37460	BB0550-BB0326	-	-	-
	BSU35330-BSU36690	BB0550-BB0195	-	-	-
	BSU35330-BSU34810	BB0550-BB0714	-	-	-
	BSU35330-BSU06140	BB0550-BB0209	-	-	-
	BSU35330-BSU06830	BB0550-BB0298	-	-	-
	BSU35330-BSU40300	-	-	-	-
fliS-TP0287	-	-	-	-	-
fliS-flgK	BSU35330-BSU35410	BB0550-BB0181	b1925-b1082	HP0753-HP1119	STM1961-STM1183
fliS-TP0661	-	-	-	-	-
fliS-TP0629	-	-	-	-	-
fliS-flgB	BSU35330-BSU16180	BB0550-BB0294	b1925-b1073	HP0753-HP1559	STM1961-STM1174
fliS-TP0843	-	BB0550-BB0602	b1925-b1000	HP0753-HP1336	STM1961-STM1112
	-	BB0550-BB0655	-	HP0753-HP1024	-
fliS-prfB	BSU35330-BSU35290	BB0550-BB0074	b1925-b0191	HP0753-HP0171	STM1961-STM0315
	-	-	b1925-b0236	-	STM1961-STM0240

	-	-	b1925-b2891	-	-
fliS-def	BSU35330-BSU15720	BB0550-BB0065	b1925-b3287	HP0753-HP0793	STM1961-STM3406
	BSU35330-BSU14560	-	-	-	-
fliS-TP0579	-	-	-	-	-
fliS-TP0121	BSU35330-BSU19690	-	b1925-b4146	-	STM1961-STM4333
fliS-TP0050	-	BB0550-BB0103	-	HP0753-HP0735	-
fliS-ruvB	-	BB0550-BB0022	b1925-b1860	HP0753-HP1059	STM1961-STM1894
fliS-TP0979	BSU35330-BSU00390	BB0550-BB0194	b1925-b1100	HP0753-HP1573	STM1961-STM1202
	-	-	-	-	STM1961-STM4564
fliS-TP0700	-	-	-	-	-
fliS-TP0702	BSU35330-BSU32340	BB0550-BB0262	b1925-b1856	HP0753-HP0750	STM1961-STM3038
	BSU35330-BSU36550	BB0550-BB0255	b1925-b2742	HP0753-HP1544	STM1961-STM2925
	BSU35330-BSU27980	BB0550-BB0761	-	HP0753-HP1543	STM1961-STM1890
	-	BB0550-BB0246	-	HP0753-HP0506	-
	-	-	-	HP0753-HP1054	-
fliS-mraZ	BSU35330-BSU15130	-	b1925-b0081	-	STM1961-STM0119
fliS-fliK	-	-	-	-	-
fliS-gyrA	BSU35330-BSU00070	BB0550-BB0435	b1925-b3019	HP0753-HP0701	STM1961-STM3174
	BSU35330-BSU18100	BB0550-BB0035	b1925-b2231	-	STM1961-STM2272
fliS-TP0782	BSU35330-BSU32340	BB0550-BB0262	b1925-b1856	HP0753-HP0750	STM1961-STM3038
	BSU35330-BSU36550	BB0550-BB0255	b1925-b2742	HP0753-HP1544	STM1961-STM2925
	BSU35330-BSU27980	BB0550-BB0761	-	HP0753-HP1543	STM1961-STM1890
	-	BB0550-BB0246	-	HP0753-HP0506	-
	-	-	-	HP0753-HP1054	-
fliS-TP0929	-	-	-	-	-
fliS-TP0310	BSU35330-BSU40900	BB0550-BB0114	b1925-b4059	HP0753-HP1245	STM1961-STM4256
	BSU35330-BSU36310	-	-	-	-
fliS-rplI	BSU35330-BSU40500	BB0550-BB0112	b1925-b4203	HP0753-HP0514	STM1961-STM4394
TP0959-TP0260	-	-	-	-	-
TP0959-TP0088	-	-	-	-	-
TP0959-TP0467	-	-	-	-	-
TP0959-ruvB	-	-	-	-	STM1182-STM1894
TP0959-gyrA	-	-	-	-	STM1182-STM3174
	-	-	-	-	STM1182-STM2272
flgG-2-TP0877	-	BB0774-BB0374	-	-	-
	-	BB0775-BB0374	-	-	-
flgG-2-TP0788	-	-	-	-	-
flgG-2-TP0314	-	-	-	-	-
flgG-2-TP0287	-	-	-	-	-
flgG-2-TP0911	BSU36400-BSU16080	-	-	HP1092-HP1575	-
	BSU36390-BSU16080	-	-	HP1585-HP1575	-
	BSU16290-BSU16080	-	-	-	-
flgG-2-TP0751	-	-	-	-	-
flgG-2-flaA-2	-	-	-	-	-
flgG-2-flgC	BSU36400-BSU16190	BB0774-BB0293	b1078-b1074	HP1092-HP1558	STM1179-STM1175
	BSU36390-BSU16190	BB0775-BB0293	-	HP1585-HP1558	-
	BSU16290-BSU16190	-	-	-	-
flgG-2-TP0772	-	-	-	-	-
flgG-2-htrA-1	BSU36400-BSU35220	BB0774-BB0104	b1078-b0161	HP1092-HP1019	STM1179-STM0209

	BSU36400-BSU12900	BB0775-BB0104	b1078-b3234	HP1585-HP1019	STM1179-STM3349
	BSU36400-BSU40360	-	b1078-b3235	-	STM1179-STM3348
	BSU36390-BSU35220	-	-	-	-
	BSU36390-BSU12900	-	-	-	-
	BSU36390-BSU40360	-	-	-	-
	BSU16290-BSU35220	-	-	-	-
	BSU16290-BSU12900	-	-	-	-
	BSU16290-BSU40360	-	-	-	-
flgG-2-def	BSU36400-BSU15720	BB0774-BB0065	b1078-b3287	HP1092-HP0793	STM1179-STM3406
	BSU36400-BSU14560	BB0775-BB0065	-	HP1585-HP0793	-
	BSU36390-BSU15720	-	-	-	-
	BSU36390-BSU14560	-	-	-	-
	BSU16290-BSU15720	-	-	-	-
	BSU16290-BSU14560	-	-	-	-
flgG-2-TP0050	-	BB0774-BB0103	-	HP1092-HP0735	-
	-	BB0775-BB0103	-	HP1585-HP0735	-
flgG-2-gatC	BSU36400-BSU06670	BB0774-BB0343	-	HP1092-HP0975	-
	BSU36390-BSU06670	BB0775-BB0343	-	HP1585-HP0975	-
	BSU16290-BSU06670	-	-	-	-
flgG-2-murC	BSU36400-BSU29790	BB0774-BB0817	b1078-b0091	HP1092-HP0623	STM1179-STM0129
	BSU36390-BSU29790	BB0775-BB0817	b1078-b4233	HP1585-HP0623	STM1179-STM4416
	BSU16290-BSU29790	-	-	-	-
flgG-2-nusG	BSU36400-BSU01010	BB0774-BB0394	b1078-b3982	HP1092-HP1203	STM1179-STM3977
	BSU36390-BSU01010	BB0775-BB0394	b1078-b3842	HP1585-HP1203	STM1179-STM4148
	BSU16290-BSU01010	-	-	-	-
flgG-2-orf4	-	BB0774-BB0282	-	-	-
	-	BB0775-BB0282	-	-	-
flgG-2-gyrA	BSU36400-BSU00070	BB0774-BB0435	b1078-b3019	HP1092-HP0701	STM1179-STM3174
	BSU36400-BSU18100	BB0774-BB0035	b1078-b2231	HP1585-HP0701	STM1179-STM2272
	BSU36390-BSU00070	BB0775-BB0435	-	-	-
	BSU36390-BSU18100	BB0775-BB0035	-	-	-
	BSU16290-BSU00070	-	-	-	-
	BSU16290-BSU18100	-	-	-	-
flgG-2-flaB3	BSU36400-BSU35360	BB0774-BB0147	b1078-b1923	HP1092-HP0601	STM1179-STM1959
	BSU36400-BSU35400	BB0774-BB0182	b1078-b1083	HP1092-HP0115	STM1179-STM1184
	BSU36400-BSU35150	BB0775-BB0147	-	HP1092-HP0295	STM1179-STM2771
	BSU36390-BSU35360	BB0775-BB0182	-	HP1585-HP0601	-
	BSU36390-BSU35400	-	-	HP1585-HP0115	-
	BSU36390-BSU35150	-	-	HP1585-HP0295	-
	BSU16290-BSU35360	-	-	-	-
	BSU16290-BSU35400	-	-	-	-
	BSU16290-BSU35150	-	-	-	-
flgG-2-TP0782	BSU36400-BSU32340	BB0774-BB0262	b1078-b1856	HP1092-HP0750	STM1179-STM3038
	BSU36400-BSU36550	BB0774-BB0255	b1078-b2742	HP1092-HP1544	STM1179-STM2925
	BSU36400-BSU27980	BB0774-BB0761	-	HP1092-HP1543	STM1179-STM1890
	BSU36390-BSU32340	BB0774-BB0246	-	HP1092-HP0506	-
	BSU36390-BSU36550	BB0775-BB0262	-	HP1092-HP1054	-
	BSU36390-BSU27980	BB0775-BB0255	-	HP1585-HP0750	-
	BSU16290-BSU32340	BB0775-BB0761	-	HP1585-HP1544	-

	BSU16290-BSU36550	BB0775-BB0246	-	HP1585-HP1543	-
	BSU16290-BSU27980	-	-	HP1585-HP0506	-
	-	-	-	HP1585-HP1054	-
flgG-2-TP0024	BSU36400-BSU11640	BB0774-BB0725	b1078-b3290	-	STM1179-STM3409
	BSU36400-BSU31090	BB0775-BB0725	-	-	-
	BSU36400-BSU14510	-	-	-	-
	BSU36390-BSU11640	-	-	-	-
	BSU36390-BSU31090	-	-	-	-
	BSU36390-BSU14510	-	-	-	-
	BSU16290-BSU11640	-	-	-	-
	BSU16290-BSU31090	-	-	-	-
	BSU16290-BSU14510	-	-	-	-
flgG-2-fliG-1	BSU36400-BSU16220	BB0774-BB0290	b1078-b1939	HP1092-HP0352	STM1179-STM1970
	BSU36390-BSU16220	BB0774-BB0221	-	HP1585-HP0352	-
	BSU16290-BSU16220	BB0775-BB0290	-	-	-
	-	BB0775-BB0221	-	-	-
TP0965-uvrA	BSU14350-BSU35160	BB0141-BB0837	b3265-b4058	HP1328-HP0705	STM0941-STM4254
	BSU33280-BSU35160	-	b3513-b4058	HP0606-HP0705	STM2692-STM4254
	-	-	b3487-b4058	HP1488-HP0705	STM3587-STM4254
	-	-	b0463-b4058	HP0970-HP0705	STM3390-STM4254
	-	-	b0574-b4058	-	STM4260-STM4254
	-	-	b0795-b4058	-	STM2126-STM4254
	-	-	b2074-b4058	-	STM0476-STM4254
	-	-	-	-	STM0818-STM4254
TP0981-rpsR	-	-	b1489-b4202	-	STM3615-STM4393
	-	-	b1025-b4202	-	STM2503-STM4393
	-	-	b3252-b4202	-	STM1283-STM4393
	-	-	b0834-b4202	-	STM2672-STM4393
	-	-	b3529-b4202	-	STM1987-STM4393
	-	-	b2604-b4202	-	STM1703-STM4393
	-	-	b1956-b4202	-	STM0385-STM4393
	-	-	b1490-b4202	-	STM4551-STM4393
	-	-	b1341-b4202	-	STM3375-STM4393
	-	-	b1786-b4202	-	-
	-	-	b1522-b4202	-	-
	-	-	b2067-b4202	-	-
	-	-	b1535-b4202	-	-
	-	-	b0385-b4202	-	-
	-	-	b1785-b4202	-	-
TP0981-TP0877	-	-	-	-	-
TP0981-TP0064	-	-	-	-	-
TP0981-TP0530	-	-	-	-	-
TP0981-TP0089	BSU29650-BSU37310	-	b1489-b3357	-	STM3615-STM3466
	BSU04280-BSU37310	-	b1489-b1334	-	STM2503-STM3466
	BSU09120-BSU37310	-	b1025-b3357	-	STM1283-STM3466
	-	-	b1025-b1334	-	STM2672-STM3466
	-	-	b3252-b3357	-	STM1987-STM3466
	-	-	b3252-b1334	-	STM1703-STM3466
	-	-	b0834-b3357	-	STM0385-STM3466

	-	-	b0834-b1334	-	STM4551-STM3466
	-	-	b3529-b3357	-	STM3375-STM3466
	-	-	b3529-b1334	-	-
	-	-	b2604-b3357	-	-
	-	-	b2604-b1334	-	-
	-	-	b1956-b3357	-	-
	-	-	b1956-b1334	-	-
	-	-	b1490-b3357	-	-
	-	-	b1490-b1334	-	-
	-	-	b1341-b3357	-	-
	-	-	b1341-b1334	-	-
	-	-	b1786-b3357	-	-
	-	-	b1786-b1334	-	-
	-	-	b1522-b3357	-	-
	-	-	b1522-b1334	-	-
	-	-	b2067-b3357	-	-
	-	-	b2067-b1334	-	-
	-	-	b1535-b3357	-	-
	-	-	b1535-b1334	-	-
	-	-	b0385-b3357	-	-
	-	-	b0385-b1334	-	-
	-	-	b1785-b3357	-	-
	-	-	b1785-b1334	-	-
TP0981-TP0661	-	-	-	-	-
TP0981-flgB	BSU29650-BSU16180	-	b1489-b1073	-	STM3615-STM1174
	BSU04280-BSU16180	-	b1025-b1073	-	STM2503-STM1174
	BSU09120-BSU16180	-	b3252-b1073	-	STM1283-STM1174
	-	-	b0834-b1073	-	STM2672-STM1174
	-	-	b3529-b1073	-	STM1987-STM1174
	-	-	b2604-b1073	-	STM1703-STM1174
	-	-	b1956-b1073	-	STM0385-STM1174
	-	-	b1490-b1073	-	STM4551-STM1174
	-	-	b1341-b1073	-	STM3375-STM1174
	-	-	b1786-b1073	-	-
	-	-	b1522-b1073	-	-
	-	-	b2067-b1073	-	-
	-	-	b1535-b1073	-	-
	-	-	b0385-b1073	-	-
	-	-	b1785-b1073	-	-
TP0981-TP0843	-	-	b1489-b1000	-	STM3615-STM1112
	-	-	b1025-b1000	-	STM2503-STM1112
	-	-	b3252-b1000	-	STM1283-STM1112
	-	-	b0834-b1000	-	STM2672-STM1112
	-	-	b3529-b1000	-	STM1987-STM1112
	-	-	b2604-b1000	-	STM1703-STM1112
	-	-	b1956-b1000	-	STM0385-STM1112
	-	-	b1490-b1000	-	STM4551-STM1112
	-	-	b1341-b1000	-	STM3375-STM1112
	-	-	b1786-b1000	-	-

	-	-	b1522-b1000	-	-
	-	-	b2067-b1000	-	-
	-	-	b1535-b1000	-	-
	-	-	b0385-b1000	-	-
	-	-	b1785-b1000	-	-
TP0981-gatC	BSU29650-BSU06670	-	-	-	-
	BSU04280-BSU06670	-	-	-	-
	BSU09120-BSU06670	-	-	-	-
TP0981-rpoE	BSU29650-BSU27120	-	b1489-b2573	-	STM3615-STM2640
	BSU29650-BSU12560	-	b1489-b4293	-	STM2503-STM2640
	BSU29650-BSU38700	-	b1025-b2573	-	STM1283-STM2640
	BSU29650-BSU14730	-	b1025-b4293	-	STM2672-STM2640
	BSU29650-BSU00980	-	b3252-b2573	-	STM1987-STM2640
	BSU29650-BSU26840	-	b3252-b4293	-	STM1703-STM2640
	BSU29650-BSU09520	-	b0834-b2573	-	STM0385-STM2640
	BSU29650-BSU23100	-	b0834-b4293	-	STM4551-STM2640
	BSU29650-BSU01730	-	b3529-b2573	-	STM3375-STM2640
	BSU04280-BSU27120	-	b3529-b4293	-	-
	BSU04280-BSU12560	-	b2604-b2573	-	-
	BSU04280-BSU38700	-	b2604-b4293	-	-
	BSU04280-BSU14730	-	b1956-b2573	-	-
	BSU04280-BSU00980	-	b1956-b4293	-	-
	BSU04280-BSU26840	-	b1490-b2573	-	-
	BSU04280-BSU09520	-	b1490-b4293	-	-
	BSU04280-BSU23100	-	b1341-b2573	-	-
	BSU04280-BSU01730	-	b1341-b4293	-	-
	BSU09120-BSU27120	-	b1786-b2573	-	-
	BSU09120-BSU12560	-	b1786-b4293	-	-
	BSU09120-BSU38700	-	b1522-b2573	-	-
	BSU09120-BSU14730	-	b1522-b4293	-	-
	BSU09120-BSU00980	-	b2067-b2573	-	-
	BSU09120-BSU26840	-	b2067-b4293	-	-
	BSU09120-BSU09520	-	b1535-b2573	-	-
	BSU09120-BSU23100	-	b1535-b4293	-	-
	BSU09120-BSU01730	-	b0385-b2573	-	-
	-	-	b0385-b4293	-	-
	-	-	b1785-b2573	-	-
	-	-	b1785-b4293	-	-
TP0981-gyrA	BSU29650-BSU00070	-	b1489-b3019	-	STM3615-STM3174
	BSU29650-BSU18100	-	b1489-b2231	-	STM3615-STM2272
	BSU04280-BSU00070	-	b1025-b3019	-	STM2503-STM3174
	BSU04280-BSU18100	-	b1025-b2231	-	STM2503-STM2272
	BSU09120-BSU00070	-	b3252-b3019	-	STM1283-STM3174
	BSU09120-BSU18100	-	b3252-b2231	-	STM1283-STM2272
	-	-	b0834-b3019	-	STM2672-STM3174
	-	-	b0834-b2231	-	STM2672-STM2272
	-	-	b3529-b3019	-	STM1987-STM3174
	-	-	b3529-b2231	-	STM1987-STM2272
	-	-	b2604-b3019	-	STM1703-STM3174

	-	-	b2604-b2231	-	STM1703-STM2272
	-	-	b1956-b3019	-	STM0385-STM3174
	-	-	b1956-b2231	-	STM0385-STM2272
	-	-	b1490-b3019	-	STM4551-STM3174
	-	-	b1490-b2231	-	STM4551-STM2272
	-	-	b1341-b3019	-	STM3375-STM3174
	-	-	b1341-b2231	-	STM3375-STM2272
	-	-	b1786-b3019	-	-
	-	-	b1786-b2231	-	-
	-	-	b1522-b3019	-	-
	-	-	b1522-b2231	-	-
	-	-	b2067-b3019	-	-
	-	-	b2067-b2231	-	-
	-	-	b1535-b3019	-	-
	-	-	b1535-b2231	-	-
	-	-	b0385-b3019	-	-
	-	-	b0385-b2231	-	-
	-	-	b1785-b3019	-	-
	-	-	b1785-b2231	-	-
TP0981-flaB3	BSU29650-BSU35360	-	b1489-b1923	-	STM3615-STM1959
	BSU29650-BSU35400	-	b1489-b1083	-	STM3615-STM1184
	BSU29650-BSU35150	-	b1025-b1923	-	STM3615-STM2771
	BSU04280-BSU35360	-	b1025-b1083	-	STM2503-STM1959
	BSU04280-BSU35400	-	b3252-b1923	-	STM2503-STM1184
	BSU04280-BSU35150	-	b3252-b1083	-	STM2503-STM2771
	BSU09120-BSU35360	-	b0834-b1923	-	STM1283-STM1959
	BSU09120-BSU35400	-	b0834-b1083	-	STM1283-STM1184
	BSU09120-BSU35150	-	b3529-b1923	-	STM1283-STM2771
	-	-	b3529-b1083	-	STM2672-STM1959
	-	-	b2604-b1923	-	STM2672-STM1184
	-	-	b2604-b1083	-	STM2672-STM2771
	-	-	b1956-b1923	-	STM1987-STM1959
	-	-	b1956-b1083	-	STM1987-STM1184
	-	-	b1490-b1923	-	STM1987-STM2771
	-	-	b1490-b1083	-	STM1703-STM1959
	-	-	b1341-b1923	-	STM1703-STM1184
	-	-	b1341-b1083	-	STM1703-STM2771
	-	-	b1786-b1923	-	STM0385-STM1959
	-	-	b1786-b1083	-	STM0385-STM1184
	-	-	b1522-b1923	-	STM0385-STM2771
	-	-	b1522-b1083	-	STM4551-STM1959
	-	-	b2067-b1923	-	STM4551-STM1184
	-	-	b2067-b1083	-	STM4551-STM2771
	-	-	b1535-b1923	-	STM3375-STM1959
	-	-	b1535-b1083	-	STM3375-STM1184
	-	-	b0385-b1923	-	STM3375-STM2771
	-	-	b0385-b1083	-	-
	-	-	b1785-b1923	-	-
	-	-	b1785-b1083	-	-

TP0981-recR	BSU29650-BSU00210	-	b1489-b0472	-	STM3615-STM0486
	BSU04280-BSU00210	-	b1025-b0472	-	STM2503-STM0486
	BSU09120-BSU00210	-	b3252-b0472	-	STM1283-STM0486
	-	-	b0834-b0472	-	STM2672-STM0486
	-	-	b3529-b0472	-	STM1987-STM0486
	-	-	b2604-b0472	-	STM1703-STM0486
	-	-	b1956-b0472	-	STM0385-STM0486
	-	-	b1490-b0472	-	STM4551-STM0486
	-	-	b1341-b0472	-	STM3375-STM0486
	-	-	b1786-b0472	-	-
	-	-	b1522-b0472	-	-
	-	-	b2067-b0472	-	-
	-	-	b1535-b0472	-	-
	-	-	b0385-b0472	-	-
	-	-	b1785-b0472	-	-
TP0981-dnaH	BSU29650-BSU00190	-	b1489-b0470	-	STM3615-STM0484
	BSU04280-BSU00190	-	b1025-b0470	-	STM2503-STM0484
	BSU09120-BSU00190	-	b3252-b0470	-	STM1283-STM0484
	-	-	b0834-b0470	-	STM2672-STM0484
	-	-	b3529-b0470	-	STM1987-STM0484
	-	-	b2604-b0470	-	STM1703-STM0484
	-	-	b1956-b0470	-	STM0385-STM0484
	-	-	b1490-b0470	-	STM4551-STM0484
	-	-	b1341-b0470	-	STM3375-STM0484
	-	-	b1786-b0470	-	-
	-	-	b1522-b0470	-	-
	-	-	b2067-b0470	-	-
	-	-	b1535-b0470	-	-
	-	-	b0385-b0470	-	-
	-	-	b1785-b0470	-	-