Pacific Symposium on Biocomputing 15:157-165(2010

ESTIMATION OF PROTEIN AND DOMAIN INTERACTIONS IN THE SWITCHING MOTILITY SYSTEM OF *MYXOCOCCUS XANTHUS*

F. MORCOS[1], M. SIKORA[2], M. ALBER[3], D. KAISER[4] and J. A. IZAGUIRRE[1]

[1]*Department of Computer Science & Engineering,*
[2]*Department of Electrical Engineering,* [3]*Department of Mathematics,*
University of Notre Dame, Notre Dame, IN 46556, USA
E-mail: amorcosg@nd.edu, msikora@ieee.org, malber@nd.edu, izaguirr@nd.edu

[4]*Departments of Biochemistry and Developmental Biology,*
Stanford University, Stanford, CA
E-mail: adkaiser@stanford.edu

The gram-negative myxobacterium *Myxococcus xanthus* is equipped with an interesting motility system that allows it to reverse direction on average every 8 minutes by switching the construction of two motility engines at the ends of this rod-shaped bacterium. While the mechanisms responsible for timing and engine construction/deconstruction are relatively well understood, there are several competing hypotheses as to how they are coupled together. In this paper we examine the evidence for protein interactions underlying these possible couplings using a novel framework consisting of a probabilistic model describing protein and domain interactions and a belief propagation inference algorithm. When provided with large amount of indirect pieces of information, such as high-throughput experiment results, and protein structures, we can reliably determine the relative likelihoods of these hypotheses, even though each individual piece of evidence by itself has very limited reliability. The same framework can be used to map large protein and domain interaction networks in myxobacteria and other organisms.

Keywords: PPI; DDI; M. xanthus; motility reversals; inference; Sum-Product Algorithm

1. Introduction

The availability of large quantities of experimental data, protein information, crystal structures of proteins and domains stimulated research on quantitative methods to estimate protein-protein (PPI) and domain-domain (DDI) interaction likelihoods. At the present time, most of these efforts have been focused on proving the effectiveness of these algorithms on known protein datasets. However, fewer efforts have been put into using such predictive tools to understand and study the dynamics of protein networks of actual interest in biology from an *in silico* perspective. In this work, we aim to do this by concentrating in a particular pathway related to motility in myxobacteria. Specifically, our study focuses on the study of *Myxococcus xanthus*, a gram-negative myxobacterium which under extreme conditions of starvation stops swarming and starts a new developmental phase to coordinate their motion cooperatively to aggregate and ultimately form fruiting bodies.[1] This multicellular cooperation in an unicellular organism makes *Myxococcus xanthus* an important model to understand regulatory pathways where many cells work together to achieve a common functional role.

To achieve this we use a PPI/DDI inference framework to provide a set of possible interactions underlying signaling pathways in *M. xanthus*. *M. xanthus* cells reverse the direction of gliding, which leads to efficient swarming, by controlling the assembly and disassembly of motility engines. For these reasons, *M. xanthus* is an interesting system to perform PPI predictions, since many of its proteins and domains have homology with chemosensory and other regulatory proteins in other well studied proteobacteria like *Escherichia coli*, *Salmonella typhimurium* and other sequenced δ-proteobacteria. Structures of domains in *E. coli* proteins are available, a feature required for the structural scoring. We also include structural complexes of domains obtained from the iPfam database to enhance the predictive nature of our methodology.

Pacific Symposium on Biocomputing 15:157-165(2010

2. Methods

2.1. *Sum-Product Algorithm for DDI/PPI Inference*

Posterior probabilities for domain and protein interaction pairs can be obtained by calculating the marginals of a joint probability distribution that is a function of protein interaction experiments, domain interaction evidence and the relationship between domain composition of proteins and the fact that individual independent domains establish physical contacts leading to protein interactions. We calculate this using the Sum-Product Algorithm (SPA),[2] an efficient method to compute marginals of multivariate functions that factor into products of simpler functions. The SPA uses a graph representation of this joint probability distribution, called the *factor graph*, and obtains marginal values by iteratively exchanging messages along the graph edges. One of the advantages of SPA is the ability to improve prediction accuracy against other known methods like maximum likelihood estimation, while at the same time correct experimental errors. SPA will not only estimate the probability of potential interaction but also will re-score the experimental input data and correct putative experimental errors. We present a more in-depth derivation and validation of the use of SPA for PPI and DDI inference in.[3] In this work, we include a brief overview of the method and focus on its application to protein networks in *M. xanthus*.

We denote by $A_{i,j} = 1$ a hypothesis that proteins i and j interact, by $B_{x,y} = 1$ a hypothesis that domains x and y interact, and use $M_{i,j}$ and $N_{x,y}$ to quantify the results of interaction measurements or experimental evidence performed on the protein pair (i, j) and domain pair (x, y), respectively. According to the model, the joint probability distribution $P(\mathbf{A}, \mathbf{B}, \mathbf{M}, \mathbf{N}|\mathbf{H})$, where $\mathbf{H}$ denotes the set of the domain architectures of proteins, factors into a product

$$P(\mathbf{A}, \mathbf{B}, \mathbf{M}, \mathbf{N}|\mathbf{H}) = \prod_{(i,j)} P(M_{i,j}|A_{i,j}) \prod_{(x,y)} P(N_{x,y}|B_{x,y}) P(B_{x,y}) P(\mathbf{A}|\mathbf{B}, \mathbf{H}). \quad (1)$$

The terms $P(M_{i,j}|A_{i,j})$ and $P(N_{x,y}|B_{x,y})$ represent the probability that an experiment produces a positive result given that proteins i and j actually interact and the probability of a positive measurement given that domain x interacts with y, respectively. These terms serve as the main information input points of the model. The term $P(B_{x,y})$ represents the a priori probability of a domain-domain interactions and is set for all x and y to an estimated probability that two randomly selected domains would interact. The central part of our model is described through deterministic relations of the form

$$A_{i,j} = \bigvee_{(x,y) \in \mathcal{B}_{i,j}(\mathbf{H})} B_{x,y} \text{ for all } (i, j), \quad (2)$$

where $\mathcal{B}_{i,j}(\mathbf{H})$ is the set of domain pairs, such that one domain is present in protein i and the other in j. Relations (2) state that a protein pair interacts if and only if at least one of its domain pairs interacts. This set of equations must be satisfied by all $A_{i,j}$ and $B_{x,y}$. Consequently, the probability $P(\mathbf{A}|\mathbf{B}, \mathbf{H})$ can be factored into a product of individual indicator functions for each (i, j). The factor graph illustrating the

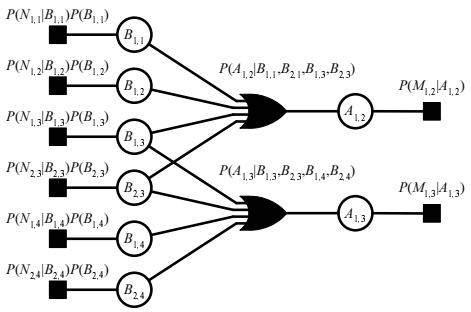

Fig. 1. **Factor graph representation of the joint probability distribution** $P(\mathbf{A}, \mathbf{B}, \mathbf{M}, \mathbf{N}|\mathbf{H})$.

complete factorization is presented in Figure 1. The SPA iteratively recomputes messages along each edge in

Pacific Symposium on Biocomputing 15:157-165(2010

the graph according to equations specific to the type of the node from which the message originates. After completing a predefined number of iterations, the algorithm obtains the *a posteriori* probabilities of PPI, $P(A_{i,j} = 1|\mathbf{M}, \mathbf{N}, \mathbf{H})$, and DDI, $P(B_{x,y} = 1|\mathbf{M}, \mathbf{N}, \mathbf{H})$, which are the final result of the inference task. They are obtained from all messages incoming into their corresponding variable nodes, i.e.,

$$P(A_{i,j} = 1|\mathbf{M}, \mathbf{N}, \mathbf{H}) = \left(1 + \prod_{k=1}^{K^{A_{i,j}}} e^{\alpha_k^{A_{i,j}}}\right)^{-1}, \tag{3}$$

$$P(B_{x,y} = 1|\mathbf{M}, \mathbf{N}, \mathbf{H}) = \left(1 + \prod_{k=1}^{K^{B_{x,y}}} e^{\alpha_k^{B_{x,y}}}\right)^{-1}, \tag{4}$$

generated in the final iteration. Where $\alpha_k^{A_{i,j}}$ is the log-ratio message entering node $A_{i,j}$ along the branch k in the factor graph of Figure 1. In this model, the domain pair architecture as well as potential protein pairs are connected via a factor graph. Hence, protein and domain pair probability estimates depend on the shared domain pair architecture. Since protein domain architectures define the topology of the factor graphs they directly affect the computation of probabilities via message passing.

2.2. *Data Sets*

Several sources were used to extract relevant information for PPI inference. Protein information was retrieved from Uniprot[4] and 3-D protein structures from the Protein Data Bank.[5] Domain composition of each protein was extracted from Pfam 22.[6] PPI data were obtained from two main sources: Database of Interacting Protein (DIP),[7] where most of the interaction pairs were obtained using yeast two-hybrid assays and IntAct,[8] which contains a higher number of interactions with different reliability levels. In[3] we use a high quality binary dataset of PPI in *Saccharomyces cerevisiae*, compiled by Yu et al.[9] for performance evaluation and validation. The algorithm also uses 6,081 iPfam domain interaction pairs as input.[10]

Table 1. Data sets used by our inference *in silico* methodology

Data Source	**Description**
Uniprot	Protein general information and sequence
DIP Database and literature (May 2007)	**33,234** protein interactions
IntAct	**150,876** protein interactions
Yeast Golden Set (Yu et al.)	**2,581** protein interactions
Pfam version 22	Domain composition of proteins
iPfam	**6,081** Domain-domain interactions with complex
PDB	Protein and Domain three dimensional structure

For the *M. xanthus* system we used a list of 23 proteins that are known to be related to the Frz system. Out of 253 possible pairings, SPA found a list of 66 potential PPI with a probability larger than 0. These interactions contained 18 unique proteins. We discarded protein interactions that do not share the same cellular location to obtain Figure 2. Finally, Table 2 shows 10 protein pairs that either have experimentally confirmed interactions or that contain potential interactions that seem biologically meaningful for the motility reversal model in *M. xanthus*. The discarded protein interactions have a mean score of 0.6634 while the 10 selected interactions have a mean score of 0.8064.

2.3. *SPA run time and scalability*

Run time of this algorithm depends on the size of the factor graph and the number of iterations defined in the message passing algorithm. The factor graph scales quadratically with the number of proteins and domains

Pacific Symposium on Biocomputing 15:157-165(2010

analyzed. However, given the nature of the input data, where domains are not shared among all the proteins and the fact that we only have measurements for a smaller subset of protein pairs, a series of disconnected factor graphs are created for the input protein and domain datasets. We process these disconnected factor graphs independently, which allows parallelization and scalability of the inference task. We do predictions using the high performance computing cluster (HPCC) of the University of Notre Dame. Using 2.29 GHz Quad-Core AMD Opteron processors we run SPA with 19 iterations in less than two hours including the largest connected subgraph.

3. Predicting interactors in *M. xanthus* reversal system

We studied how the control of reversals in gliding direction of *M. xanthus* takes place. *M. xanthus* has two different multi-protein engines: S-motility and A-motility; both involve social interactions. S-motility depends on Type IV Pili motors.[11] These pili attach to fibrils surrounding other *M. xanthus* cells and by means of retraction pull the cell forward. A-motility, is associated with slime secretion. One hypothesis states that slime coming out of several molecular nozzles at polar ends of the cell pushes the cell forward[12].[13] Pili should be present in the leading part of the cell while slime nozzles be part of the back of the cell. This configuration allows for pilus retraction and slime propulsion at the same time. A second hypothesis relates A-motility with non-polar cell surface adhesion complexes.[14] *M. xanthus* cells are able to move over surfaces by gliding, they reverse their direction of movement with an average period of approximately 7.2 minutes.[15] This period seems to be optimal to achieve higher swarming rates and cell flux.[16] These reversals occur by switching motility engines from one end of the cell to the other. Under the slime propulsion model, a process of nozzle inactivation takes place when secretion is switched from one pole to the other, and pili are inactivated when they switch from one end to the other. A control circuit is needed for engine switching that would lead to cell reversals. That circuit includes the Frz (frizzy) system which is deemed responsible of controlling the frequency of reversals in *M. xanthus*.[17] This circuit is based on a two-component system that clocks the time to exchanging motility engines in the ends of the cell. Frizzy proteins that control the reversal frequency include FrzCD (uniprot accession: P43500), a methyl-accepting chemosensory protein having similarities with MCP proteins used for flagellar swarming in *E. coli*, but which is cytoplasmic in *M. xanthus* and not a membrane protein as in *E. coli*.[15] Also important is FrzE (P18769), a histidine kinase composed of two domains similar to CheY (P0AE68) and CheA (P07363) proteins in *E. coli*.[18] FrzE has autophosphorylating function induced by a phosphate transfer from the HK domain (CheA) to its response regulator (CheY). It is suggested that FrzE∼P is responsible for signaling reversal of polarity to the motility engine. FrzG (P31758) is a methyltransferase similar to CheB (P07330), which is involved in the methylation of FrzCD, for adaptation. During fruiting body development, the FruA (Q1D7Q3) response regulator triggers FrzCD methylation. During growth there is no FruA, and FrzCD is spontaneously methylated at a low rate.[19]

Igoshin et al.[20] constructed a dynamical model of the Frz system. During reversals the levels of methylated FrzCD and FrzE∼P oscillated consistently out of phase, suggesting that these proteins form a negative feedback loop. The loop describes accurately the frequency of reversals observed in *M. xanthus* when nutrients are scarce and development towards fruiting body formation is in its initial phase. For the model of Igoshin et al. to be valid, it requires one of two potential protein pairs. These hypotheses still remain to be tested experimentally. The two potential interactions are the following: FrzE-FrzF and FrzE-FrzG. To test these hypotheses, we used our predictive methodology on a list of proteins involved in the FruA-Frz network. Since several interactions forming part of the reversal mechanism are understood and their existence has been established, we have a way to assess our predictions. Reproducing interactions that are already studied provides support to our methodology and at the same time shows the potential of *in silico* methodologies to reproduce scientific efforts to elucidate protein interactions in a given pathway. On the other hand, predicting protein interactions in proteins of and related to the Frz-Mgl network let us investigate the plausibility of these competing hypotheses. The prediction algorithm was run using the data sets presented in Methods. Figure 2 presents a PPI network that was the outcome of the algorithm.

In Figure 2, nodes represent signaling proteins while the edge represent the physical interaction pre-

Pacific Symposium on Biocomputing 15:157-165(2010

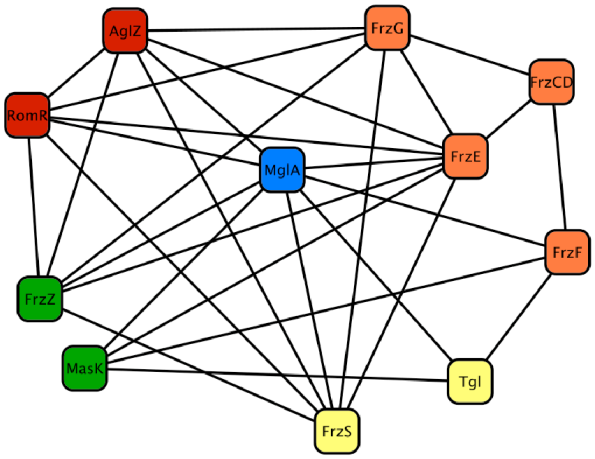

Fig. 2. **Predicted network of interactions in the Frz system of *M. xanthus*.**

diction. Self interactions have been subtracted to facilitate readability. This network has also a color code classification: orange nodes represent proteins in the negative feedback loop of Igoshin's model; red proteins are related to A-motility; yellow nodes play a role in S-motility; green nodes are being identified to have a role in both A and S motility systems and finally MglA is represented by a blue node. This node plays an important role in the connection of the reversal control and the actual motility mechanisms.

Table 2 presents a summary of the most important predicted interactions in the signaling network of *M. xanthus*. It includes literature references and SPA scores to those interactions discuss here, illustrating the correct predictive capabilities of our methodology. The bottom part of the table encompasses those PPI for which we have not found literature describing them, however, our predictions present computational evidence about their interaction and our analysis shows that their biological role might be of importance and thus can be further studied.

Table 2. Predicted Interactions found in Literature and novel interactions for *M. xanthus*

Predicted Interaction	Reference	SPA score
FrzE - FrzG	[20] Hypothesis (high likelihood)	0.96
FrzE - FrzCD	[21]	0.99
FrzE - FrzZ	[22] (suggested)	1.00
FrzF - FrzCD	[21]	0.46
MglA - MasK (Q1DB00)	[23]	1.00
MglA - AglZ	[24]	0.56
MglA - FrzE	Predicted , (motility bridge)	0.639
MglA - FrzF	Predicted	0.95
MglA - RomR	Predicted (A-motility)	0.56
MglA - Tgl	Predicted (S-motility)	0.95

3.1. *Analysis of hypotheses in Igoshin's model*

We turned our discussion to those potential interactions that make possible a negative feedback loop yielding a biochemical oscillator in the signaling circuit of *M. xanthus*. Based on these results, we provide computational support for one of these hypothesis. The hypothesis FrzE-FrzG, was predicted with high score (see

Pacific Symposium on Biocomputing 15:157-165(2010

Table 2). The PDB `1a2o` provides structural evidence that proteins with these domains might create interacting complexes. Djordjevicj et al.[25] determined the crystal structure of CheB. It has two domains, a response regulator and a CheB methylesterase joined by a linker. They found that in its unphosphorylated state, the RR would bind the CheB_methylest domain and impede its activity. The crystal in `1a2o` shows how an unphosphorylated response regulator interacts with the effector domain CheB_methylest with a surface area of 1000 $Å^2$. They also argued that phosphorylation of the response regulator would induce a conformational change that takes both domains apart and free the methylesterase to perform its catalytic function. Phosphorylation induces minor reorientation of regulator helices $\alpha 4$ and $\alpha 5$, resulting in disruption of the inter-domain interaction. Protein CheA phosphorylates the response regulator in CheB.[26] In the regulatory circuit of *M. xanthus*, the FrzG-FrzE interaction is different because FrzG only has the CheB_methylest domain, hence this domain could be active in the absence of a regulator. We hypothesize that the unphosphorylated response regulator in FrzE could help inhibit FrzG function, while an autophosphorylated FrzE could not block FrzG methylesterase and activating it as a consequence. This hypothesis is plausible but unconfirmed. Our results suggest that the biochemical oscillator depends on the FrzE-FrzG interaction.

3.2. *Searching for interactions with MglA, the proposed switch protein*

Although the mechanism to control the frequency of reversals in *M. xanthus* has been studied extensively through the Frz system, and other two-component systems,[27] there are still some missing links needed to uncover the connection between the MglA switch and the actual A and S engines. MglA has been shown experimentally to have an influence in both A and S engines.[28] Cells with mutant *mglA* genes fail to reverse and fail to swarm. They show simultaneous secretion of slime from both ends of the cell;[13] it is suggested, but not experimentally confirmed, that pili are also present from both ends of the cell. Thus MglA may serve as a switch to start disassembly of engines at both ends of the cell.[16] MglA-GTP might pick the end that should lose pili by losing Tgl (P95324) and would also select the end to inactivate slime propulsion by potentially interacting with CglB (O31191).[29]

We included MglA, with a conserved Ras domain, in our list of interactors. We were interested in the interaction patterns of MglA with respect to proteins in the Frz system and other proteins related to motility. We found that MglA has a potential central role in both motility systems and the Frz system given its predicted interactions with different members of these systems. Figure 2 depicts MglA in blue, and presents edges with several proteins. We start our discussion by analyzing the inferred interaction MglA-FrzE, since this interaction could potentially represent the link between the Frz system and the two motility systems in *M. xanthus*.

Our estimation algorithm produced a score for Ras-Response_reg, supporting the plausibility that a potential interaction between MglA and FrzE results from a physical contact between the Ras domain in MglA and the phosphorylated Response_reg in FrzE. The domain pair Ras-Response_reg-P which underlies the FrzE-MglA interaction leads to the prediction of other interactions that seem important to our study. The protein pairs MglA-FrzZ (Q7BU54) and MglA-FrzS (Q1D4U9) are also plausible interactions between a Ras domain pair and response regulators. These results suggest how MglA has an important role for the switching of motility in the Frz system of *M. xanthus*.

3.3. *MglA interacts with Tgl and RomR*

We also investigated a possible interaction between MglA and engine proteins. Tgl is a lipoprotein needed for the assembly of the PilQ (Q9ZFG1) secretin and thus for S-motility. Tgl is a stimulatable protein found in membrane. It allows PilQ to assemble with the help of six TPR (Tetratrico peptide repeat superfamily) repeats. One hypothesis is that three of these repeats interact with one monomer of PilQ, while the three remaining repeats interact with the adjacent PilQ monomer.[29] We investigated if MglA interacts with Tgl. Our prediction algorithm infered that TPR_1, one of the domains present in Tgl, interacts with high probability with the Ras domain in MglA. This domain interaction would potentially trigger the destruction of

Pacific Symposium on Biocomputing 15:157-165(2010

Tgl or its removal from PilQ.[29] The inclusion of iPfam evidence in our algorithm allowed us to identify a crystal structure (PDB: `1e96`) showing the Ras domain in complex with TPR_1. When this structure was obtained, Rittinger et al. suggested that this interaction is important for the assembly of protein (enzyme) complexes.[30] This proposition provides more plausibility to the hypothesis that MglA-GTP causes PilQ to disassemble. It is possible that through this interaction, MglA opens Tgl in such a way that it makes it accessible to a protease that would then contribute to the disassemble of PilQ at one end of the rod shaped cell.[31]

RomR is an essential protein for slime engine function. Its switching from the extremes of the cell depends on MglA. Thus, we are interested in the possible interaction MglA-RomR. Søgaard-Andersen et al.[32] report that RomR has a Response_reg (receiver) domain in the N terminal and an output domain in the C terminal of the protein. Søgaard-Andersen et al. proposed that correct RomR polarity depends on the small GTPase MglA. They also state the importance of RomR for A-motility and how this protein relocates synchronously with another Frz protein: FrzS. They concluded that the Receiver domain is involved in dynamic RomR localization and that this is required for reversals, while the output domain is a polar targeting determinant. Their studies also showed that when the receiver domain is not phosphorylated then no reversals are observed. On the other hand, if the receiver domain is always phosphorylated then there is a 1.5 fold increase in reversal frequency. However, the RomR kinase has yet to be identified. With this information and based on our prediction of interaction between Ras-Response_reg-P, we provide computational evidence that indeed MglA could interact with RomR.

4. Discussion

In this work, PPI/DDI prediction methods are combined with the biological understanding of the reversal system of *M. xanthus* as an attempt to improve present knowledge of the players involved in this process. A contribution of this work is to illustrate how this framework is used to study small networks that perform important functions in an organism. The focus in *M. xanthus*' reversal system is due to its importance in development and swarming. Support for this approach is provided by predicting interactions previously reported in the literature. This study provides an assessment of the contribution of interactions FrzE-FrzF and FrzE-FrzG as being part of the biochemical oscillator that controls the reversal frequency. Results suggest that FrzE-FrzG plays an important role in negative feedback circuit. Furthermore, this framework let us reach a more detailed explanation of how the response regulator domain in FrzE possibly interacts with the CheB methylesterase domain in FrzG. This interaction possibly inhibits FrzG demethylation activity. When the response regulator is phosphorylated this interaction is broken allowing FrzG to demethylate FrzCD. This mechanism is needed to support the negative feedback model in the "Frizzilator".

We investigated the role of MglA as a switch to control the construction of motility engines in *M. xanthus*. Predictions showed the central role of this protein, first as a bridge with the negative feedback system by interacting with FrzE and how its interactions with RomR, Tgl, FrzS and AglZ (Q1D823) are important for both A and S engines. With these predictions, along with the present understanding of the reversal switch it is possible to expand the model of this system. This model is illustrated in Figure 3.

The model shows how MglA serves as a link between the signaling pathway involving Frz proteins and the motility engines. MglA receives a signaling message from FrzE that triggers the destruction of old engines in the ends of the cell. Potential interactions of MglA with Tgl and FrzS have an impact in the switching of S-motility engines. Parallel to these interactions, MglA is predicted to be interacting with RomR and AglZ, showing a connection with the A-motility engines. This expanded model could help devise testable hypotheses in the motility system of *M. xanthus*.

Acknowledgments

This work was supported in part by NSF grant CCF-0622940.

Pacific Symposium on Biocomputing 15:157-165(2010

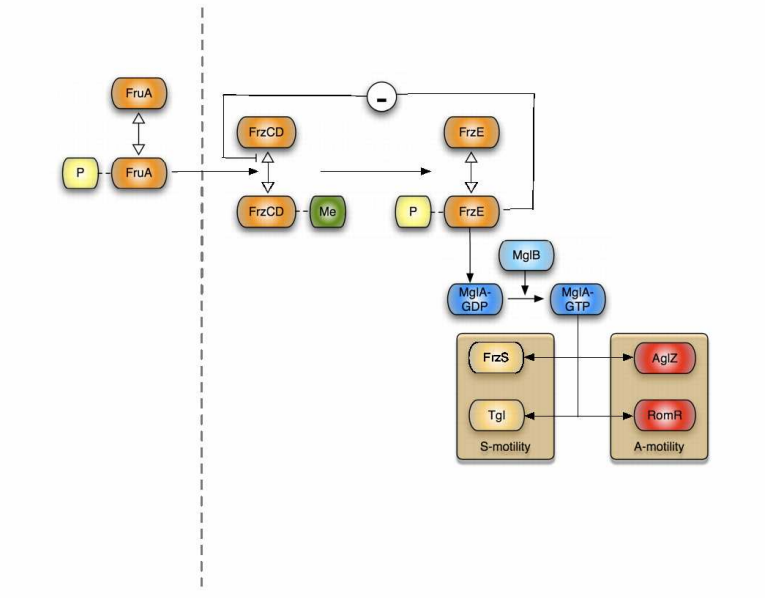

Fig. 3. **Expanded model of switching motility in *M. xanthus***. Switching is triggered by FrzE-MglA. MglA-GTP interacts with A and S motility systems.

References

1. Kaiser, D. *Annu Rev Microbiol* **58**, 75–98 (2004).
2. Kschischang, F., Frey, B., and Loeliger, H.-A. *Information Theory, IEEE Transactions on* **47**(2), 498–519 Feb (2001).
3. Morcos, F., Sikora, M., Alber, M., Kaiser, D., and Izaguirre, J. A. *Information Theory, IEEE Transactions on* **(under review)** (2009).
4. Consortium, U. *Nucleic Acids Res* **35**(Database issue), D193–D197 Jan (2007).
5. Deshpande and et al. *Nucleic Acids Res* **33**(Database issue), D233–D237 Jan (2005).
6. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. L., Studholme, D. J., Yeats, C., and Eddy, S. R. *Nucleic Acids Res* **32**(Database issue), D138–D141 Jan (2004).
7. Xenarios, I., Salwnski, L., Duan, X. J., Higney, P., Kim, S.-M., and Eisenberg, D. *Nucleic Acids Res* **30**(1), 303–305 Jan (2002).
8. Kerrien, S., Alam-Faruque, Y., Aranda, B., Bancarz, I., Bridge, A., Derow, C., Dimmer, E., Feuermann, M., Friedrichsen, A., Huntley, R., Kohler, C., Khadake, J., Leroy, C., Liban, A., Lieftink, C., Montecchi-Palazzi, L., Orchard, S., Risse, J., Robbe, K., Roechert, B., Thorneycroft, D., Zhang, Y., Apweiler, R., and Hermjakob, H. *Nucleic Acids Res* **35**(Database issue), D561–D565 Jan (2007).
9. Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.-F., Dricot, A., Vazquez, A., Murray, R. R., Simon, C., Tardivo, L., Tam, S., Svrzikapa, N., Fan, C., de Smet, A.-S., Motyl, A., Hudson, M. E., Park, J., Xin, X., Cusick, M. E., Moore, T., Boone, C., Snyder, M., Roth, F. P., Barabsi, A.-L., Tavernier, J., Hill, D. E., and Vidal, M. *Science* **322**(5898), 104–110 Oct (2008).
10. Finn, R. D., Marshall, M., and Bateman, A. *Bioinformatics* **21**(3), 410–412 Feb (2005).
11. Spormann, A. M. *Microbiol Mol Biol Rev* **63**(3), 621–641 Sep (1999).
12. Wolgemuth, C., Hoiczyk, E., Kaiser, D., and Oster, G. *Curr Biol* **12**(5), 369–377 Mar (2002).
13. Yu, R. and Kaiser, D. *Mol Microbiol* **63**(2), 454–467 Jan (2007).
14. Sliusarenko, O., Zusman, D. R., and Oster, G. *J Bacteriol* **189**(21), 7920–7921 Nov (2007).
15. Bustamante, V. H., Martnez-Flores, I., Vlamakis, H. C., and Zusman, D. R. *Mol Microbiol* **53**(5), 1501–1513 Sep (2004).
16. Wu, Y., Kaiser, A. D., Jiang, Y., and Alber, M. S. *Proc Natl Acad Sci U S A* **106**(4), 1222–1227 Jan (2009).
17. Zusman, D. R. *J Bacteriol* **150**(3), 1430–1437 Jun (1982).
18. McCleary, W. R. and Zusman, D. R. *J Bacteriol* **172**(12), 6661–6668 Dec (1990).
19. McCleary, W. R., McBride, M. J., and Zusman, D. R. *J Bacteriol* **172**(9), 4877–4887 Sep (1990).
20. Igoshin, O. A., Goldbeter, A., Kaiser, D., and Oster, G. *Proc Natl Acad Sci U S A* **101**(44), 15760–15765 Nov

Pacific Symposium on Biocomputing 15:157-165(2010

(2004).
21. McBride, M. J., Khler, T., and Zusman, D. R. *J Bacteriol* **174**(13), 4246–4257 Jul (1992).
22. Li, Y., Bustamante, V. H., Lux, R., Zusman, D., and Shi, W. *J Bacteriol* **187**(5), 1716–1723 Mar (2005).
23. Thomasson, B., Link, J., Stassinopoulos, A. G., Burke, N., Plamann, L., and Hartzell, P. L. *Mol Microbiol* **46**(5), 1399–1413 Dec (2002).
24. Yang, R., Bartle, S., Otto, R., Stassinopoulos, A., Rogers, M., Plamann, L., and Hartzell, P. *J Bacteriol* **186**(18), 6168–6178 Sep (2004).
25. Djordjevic, S., Goudreau, P. N., Xu, Q., Stock, A. M., and West, A. H. *Proc Natl Acad Sci U S A* **95**(4), 1381–1386 Feb (1998).
26. Hess, J. F., Oosawa, K., Kaplan, N., and Simon, M. I. *Cell* **53**(1), 79–87 Apr (1988).
27. Stock, A. M., Robinson, V. L., and Goudreau, P. N. *Annu Rev Biochem* **69**, 183–215 (2000).
28. Stephens, K., Hartzell, P., and Kaiser, D. *J Bacteriol* **171**(2), 819–830 Feb (1989).
29. Nudleman, E., Wall, D., and Kaiser, D. *Mol Microbiol* **60**(1), 16–29 Apr (2006).
30. Lapouge, K., Smith, S. J., Walker, P. A., Gamblin, S. J., Smerdon, S. J., and Rittinger, K. *Mol Cell* **6**(4), 899–907 Oct (2000).
31. Rodriguez-Soto, J. P. and Kaiser, D. *J Bacteriol* **179**(13), 4372–4381 Jul (1997).
32. Leonardy, S., Freymark, G., Hebener, S., Ellehauge, E., and Sgaard-Andersen, L. *EMBO J* Oct (2007).