Data Article

Data for the qualitative modeling of the osmotic stress response to NaCl in *Escherichia coli*

Delphine Ropers [a],*, Aline Métris [b]

[a] *Inria, Grenoble – Rhône-Alpes Research Center, Saint Ismier, France*
[b] *Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, UK*

A R T I C L E   I N F O

A B S T R A C T

Qualitative modeling approaches allow to provide a coarse-grained description of the functioning of cellular networks when experimental data are scarce and heterogeneous. We translate the primary literature data on the response of *Escherichia coli* to hyperosmotic stress caused by NaCl addition into a piecewise linear (PL) model. We provide a data file of the qualitative model, which can be used for simulation of changes of protein concentrations and of DNA coiling during the physiological response of the bacterium to the stress. The qualitative model predictions are directly comparable to the available experimental data. This data is related to the research article entitled "Piecewise linear approximations to model the dynamics of adaptation to osmotic stress by food-borne pathogens" (Metris et al., 2016) [1].

Specifications Table

Subject area	*Computational biology*
More specific subject area	*Qualitative modeling of the dynamics of a gene regulatory network*
Type of data	*Graph, figures, table, model equations, model files (GNAML and SBML formats)*

How data was acquired	*Simulations were performed by means of the publicly available tool Genetic Network Analyzer*
Data format	*Equations, GNAML and SBML files*
Experimental factors	*Primary literature data*
Experimental features	*Gene expression and transcription factor binding sites of E. coli and Salmonella during osmotic stress response*
Data source location	*Inria, Saint Ismier, France*
Data accessibility	*The data is with this article*

Value of the data

- The reconstruction of the osmotic stress response network of *E. coli* provides a compilation of current knowledge on this process.
- The piecewise-linear model of the network is useful to exploit the heterogeneous and scarce experimental data on hyper-osmotic stress: their comparison with the model predictions allow to verify if we have a good understanding of the network functioning or if additional hypotheses should be formulated to reconcile potential discrepancies.
- The model can be easily extended to describe the response of *E. coli* to alternative osmotic stresses, caused for instance by other humectants and low moisture.

1. Data

The data provided in this article include a reconstruction of the hyper-osmotic stress response network of *Escherichia coli* and its translation into a piece-wise linear model. Files with the model equations in tow formats (GNAML and SBML) are also given, for computer simulation of the physiological response of *E. coli* to the presence or the absence of salt.

2. Experimental design, materials and methods

The PL modeling of the osmotic stress response network is briefly described below (see Batt et al. [2] and references therein for more information) and it is illustrated with a simple example in Fig. 1. Four steps were necessary before we could generate predictions on the network behavior that could be compared to experimental data.

2.1. Reconstruction of the osmotic stress response network

Our analysis of the physiological response of *E. coli* to hyper-osmotic stress is centered around proteins and markers known to play a key role in this process. Based on an extensive search of the literature and previous work [4,5], we have reconstructed a network of eight genes: the sigma factor RpoS, the transcription factors Fis, IHF, and CRP, the symporter ProP, the trehalose synthase OtsAB, and markers of the osmotic stress response, OsmY, and of cellular growth, the stable RNAs. The assumptions made to reconstruct the network and the role of the different network components and their interactions are summarized in Table 1. The reader is referred to Metris et al. [1] for additional information. An example is provided in Fig. 1A, in the case of the genes involved in the synthesis of the osmoprotectant trehalose, *otsA* and *otsB*. Expression of these genes is both osmotically and growth-phase regulated in a RpoS-dependent manner as determined by gene expression of mutants in Hengge-Aronis et al. [3]. Since the two genes are organized in an operon and share the same regulation [6], we consider OtsAB to be the product of a single gene *otsAB*, whose promoter P is recognized by RpoS.
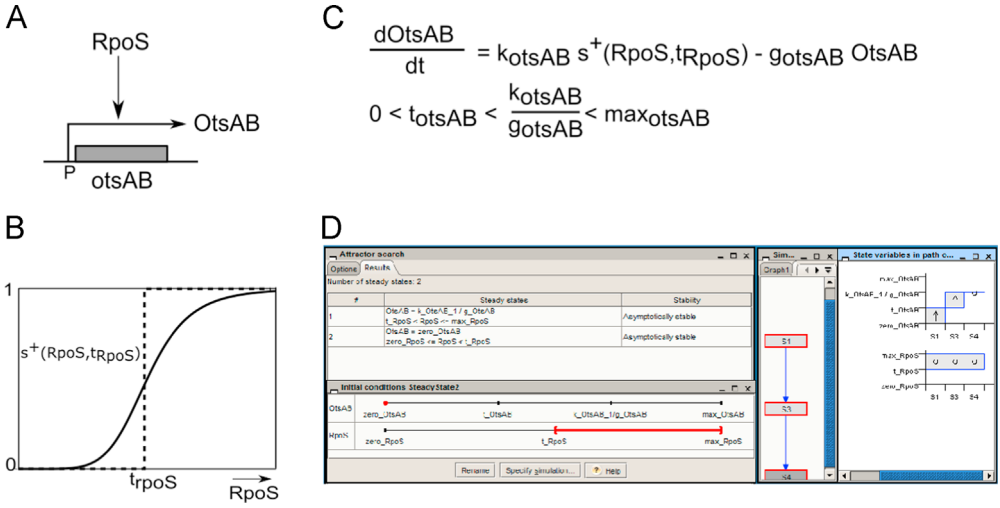
Fig. 1. PL modeling of a simple network. (A) Network reconstruction from literature data. The expression of the operon genes *otsA* and *otsB* is RpoS-dependent [3]. Trehalose synthesis is considered to be under the control of OtsAB, the product of a single gene *otsAB*, whose promoter P is recognized by RpoS. (B) Positive step function (dashed line) used to approximate the sigmoidal Hill function (continuous line) often found in the regulation of gene expression. (C) PL model of the simple network and ordering of the threshold and focal parameters. (D) Screenshot of the initial conditions for simulation and of the state transition graph, attractor search and simulation results returned by Genetic Network Analyzer.

2.2. Translation of the gene regulatory network into a PL model

We consider the network as composed of four different modules, each one accomplishing a specific task, that of setting (1) the concentration of the potassium salt of glutamate and the DNA supercoiling level; (2) the concentration of the general stress response factor RpoS; (3) the concentration of the complex CRP–cAMP; (4) the concentration of OsmY and the growth rate. We do not model explicitly the concentration of RNA polymerase nor the concentration of the σ^{70} factor, as they are not known to vary in response to osmotic stress. We indicate in the table below the model equations and corresponding parameter ordering. The notations will follow the convention used above, k, representing protein synthesis rates, g, degradation rates, and t, threshold parameters.

Regulatory influences are described by means of step functions that change value at a threshold concentration. These functions simplify the sigmoidal Hill functions generally used to describe cooperative processes involved in the regulation of gene expression. For instance, the positive influence of RpoS on *otsAB* expression is described by a positive step function (see Panel B of Fig. 1). It evaluates to 1 when RpoS is above its threshold concentration t_{Rpos}, and to 0 otherwise. Negative step functions are used to describe cooperative inhibition of gene expression. Hence, the auto-inhibition of IHF expression is described by the step function $s^- \left(IHF, t_{ihf}^2 \right)$, which equals to 1 when IHF is below its threshold value t_{ihf}^2 and to 0 otherwise (see Table 1).

PL equations describe the rate of change of protein or RNA concentrations as the difference between their synthesis rate and their degradation rate. For instance, the PL equation for OtsAB in Panel C of Fig. 1 states that OtsAB is synthesized at a rate k_{otsAB} when RpoS is above its threshold concentration, while the synthesis rate is null in the absence of RpoS. OtsAB is degraded at a basal rate g_{otsAB} OtsAB. The concentration to which OtsAB tends when it is synthesized, $\frac{k_{otsAB}}{g_{otsAB}}$, should be above its threshold level: $\frac{k_{otsAB}}{g_{otsAB}} > t_{otsAB}$; otherwise the protein would never reach a level above which it is able to produce trehalose and, indirectly, to affect the efflux of potassium.

We represent the input signal, i.e. the application of an osmotic stress to the system, by a step function $s^+(u, t_u)$ which equates to 1 upon osmotic shock. It captures the molecular changes induced

Table 1
Equations of the PL model with corresponding assumptions and parameter ordering.

Module	State variable	Assumptions	Equation	Ordering
Potassium glutamate/DNA supercoiling module	ProP, transporter of proline and glycine betaine	2 promoters [7], the σ70 dependent promoter P1, inhibited by the complex CRP–cAMP [8] and the σS-dependent promoter P2 stimulated by proteins Fis and CRP–cAMP [9,10]	$\frac{d}{dt}ProP = k^1_{proP}\left(1 - s^+\left(CRP, t^3_{crp}\right)s^+(u, t_u)\right) + k^2_{proP}s^+\left(CRP, t^3_{crp}\right)s^+(u, t_u)$ $\times\, s^+\left(Fis, t^3_{fis}\right)s^+(RpoS, t_{rpoS}) - g_{proP}ProP$ k^1_{proP} and k^2_{proP}, synthesis rates for each promoter	$0 < t_{proP} < \frac{k^1_{proP}}{g_{proP}} < \frac{k^2_{proP}}{g_{proP}} < max_{proP}$
	OtsAB, trehalose synthase	σS-dependent [3]	$\frac{d}{dt}OtsAB = k_{otsAB}s^+(RpoS, t_{rpoS}) - g_{otsAB}OtsAB$	$0 < t_{otsAB} < \frac{k_{otsAB}}{g_{otsAB}} < max_{otsAB}$
RpoS module	RpoS, general stress factor	intracellular concentration assumed to come from cellular stabilization (see [1] for additional information)	$\frac{d}{dt}RpoS = k^1_{rpoS} - \left(g^1_{rpoS} + g^2_{rpoS}s^-(u, t_u)\right)RpoS$ g^1_{rpoS} and g^2_{rpoS}, degradation rate constants, one unregulated, and one modulated by osmotic stress	$0 < \frac{k_{rpoS}}{g^1_{rpoS}+g^2_{rpoS}} < t_{rpoS} < \frac{k_{rpoS}}{g^1_{rpoS}} < max_{rpoS}$
CRP–cAMP module	CRP, global regulator	Active in transcription when bound to the cAMP [11]. Transcription of P1 is inhibited by Fis [12]	$\frac{d}{dt}CRP = k^1_{crp}s^+\left(CRP, t^1_{crp}\right)s^+(u, t_u)s^-\left(Fis, t^2_{fis}\right) + k^2_{crp} - g_{crp}CRP$ k^1_{crp} and k^2_{crp}, synthesis rates for each promoter	$0 < t^1_{crp} < t^2_{crp} < \frac{k^2_{crp}}{g_{crp}} < t^3_{crp} < \frac{k^1_{crp}+k^2_{crp}}{g_{crp}}$ $< max_{crp}$
Output module	rrn, stable RNAs	P1 is stimulated by Fis. Both promoters are assumed to be repressed by osmoprotectants [1,13,14].	$\frac{d}{dt}rrn = k^1_{rrn}s^+\left(Fis, t^1_{fis}\right)\left(1 - s^+(u, t_u)s^-(OtsAB, t_{otsAB})s^-(ProP, t_{proP})\right)$ $+ k^2_{rrn}\left(1 - s^+(u, t_u)s^-(OtsAB, t_{otsAB})s^-(ProP, t_{proP})\right) - g_{rrn}rrn$ k^1_{rrn} and k^2_{rrn}, synthesis rates for each promoter	$0 < \frac{k^2_{rrn}}{g_{rrn}} < \frac{k^1_{rrn}+k^2_{rrn}}{g_{rrn}} < max_{rrn}$
	Fis, factor for inversion stimulation	Expression is stimulated by IHF and repressed by both CRP–cAMP and Fis itself [1,15,16].	$\frac{d}{dt}Fis = k_{fis}s^+\left(IHF, t^1_{ihf}\right)\left(1 - s^+\left(CRP, t^2_{crp}\right)s^+(u, t_u)\right)s^-\left(Fis, t^4_{fis}\right) - g_{fis}Fis$	$0 < t^1_{fis} < t^2_{fis} < t^3_{fis} < t^4_{fis} < \frac{k_{fis}}{g_{fis}} < max_{fis}$
	IHF, integration host factor	σS-dependent and inhibits itself	$\frac{d}{dt}IHF = k^1_{ihf} + k^2_{ihf}s^+(RpoS, t_{rpoS})s^-\left(IHF, t^2_{ihf}\right) - g_{ihf}IHF$ k^1_{ihf} and k^2_{ihf} correspond to IHF synthesis mediated by σ70 and σS respectively	$0 < t^1_{ihf} < \frac{k^1_{ihf}}{g_{ihf}} < t^2_{ihf} < \frac{k^1_{ihf}+k^2_{ihf}}{g_{ihf}} < max_{ihf}$
	OsmY, osmotically induced membrane protein	σS responsible for osmY expression under osmotic stress [17–19]	$\frac{d}{dt}OsmY = k^1_{osmY}s^+(u, t_u)s^-(OtsAB, t_{otsAB})s^-(ProP, t_{proP})$ $+ k^2_{osmY}s^+(RpoS, t_{rpoS})s^+(u, t_u) \times s^-(OtsAB, t_{otsAB})s^-(ProP, t_{proP})$ $- g_{osmY}OsmY$ k^1_{osmY} and k^2_{osmY} correspond to OsmY synthesis mediated by σ70 and σS respectively	$0 < \frac{k^1_{osmY}}{g_{osmY}} < \frac{k^1_{osmY}+k^2_{osmY}}{g_{osmY}} < max_{osmY}$

Table 2
Literature survey of the dynamic response of *otsAB* operon following a hyper-osmotic stress.

Protein (reference)	Conditions	Method	Strains	Time			
OtsA [22]	LB with 5% NaCl compared to LB with no salt	Decimal logarithm of the gene transcript relative to 16S gene measured by qRT-PCR		15 min.		60 min.	
			E. coli K303	< 0.3 (log(2))		< 1.5	
			E. coli K356	> −1		< 1	
			E. coli K331-4	< 0.3 (log(2))		< 1	
			E. coli N09-1298	> −1		> 1	
			E. coli FAM21843	> 0.3 (log(2))		> 1	
OtsB [23]	glucose minimal medium with 0.4 M NaCl in aerobic conditions	Ratio of protein induction measured by two dimensional gel electrophoresis	E. coli MG1655	0 min 0	10 min 0	30 min > 0	60 min 4.5
OtsA [24]	glucose minimal medium with 0.4 M NaCl in aerobic conditions	Ratio of gene expression as measured by micro-array as compared to no added NaCl in the medium	E. coli MG1655	9 min 1.5			
		Ratio obtained by Northern blot analysis		8.1			
	LB with 6% NaCl	Fold change in gene expression as measured by micro-array as compared to no added NaCl in the medium	S. Typhimurium 4/74	0 h	6 h	24 h	
OtsA [25]				< 3	6.6	< 3	
OtsB [25]				< 3	8.1	< 3	
OtsB [26]	defined medium+0.3 M NaCl	mRNA level normalized to 16S as measured by qRT-PCR	S. Typhimurium LT2	0 min 0	20 min > 1	60 min 3	

by osmotic stress such as supercoiling and accumulation of glutamate as explained in Section 2.1 in Metris et al. [1].

2.3. Qualitative simulation of the PL model

The model in Table 1 has been implemented in and simulated with the publicly available software tool Genetic Network Analyzer (GNA 8.4, Genostar, http://userclub.genostar.com/en/genostar-software/gnasim.html). The corresponding model file is given in the supplementary data in GNAML and qualitative SBML formats [20]. An example of qualitative simulation with the simple network model is given in Panel D of Fig. 1.

Running the attractor search functionality of GNA shows that there are two stable steady states for the osmotic stress response model, one characteristic for normal growth in the absence of osmotic stress and one reached in case of osmotic stress. We can simulate how the system responds to salt addition and leaves the non-stressed state for the stressed one, by perturbing the state of normal growth with the signal of stress switched on ($u > t_u$). The simulation returns a state transition graph composed of 192 states, showing all the possible dynamic behaviors of the system going from the initial state of normal growth to the stressed state.

2.4. Comparison of model predictions with experimental data

Each path in the state transition graph describes the evolution of protein and RNA concentrations, which can be confronted to the experimental data [1]. For instance, literature data summarized in Table 2 suggest an accumulation of OtsA and OtsB as trehalose accumulates after a hyper-osmotic stress induced by salt [21].

The expression varies with the conditions and the strains, however the pattern remains similar; a delay before increased expression during adaptation. The trend is also similar for *E. coli* and *Salmonella typhimurium* , so, in this case, the same equation may be used for both species. In the simple example of Fig. 1, the concentration of protein OtsAB is predicted to increase once RpoS accumulates in the cell. This prediction is hence consistent with experimental data. This type of analysis may be carried out for the different osmotic genes with more complex regulation as shown in Metris et al. [1].

Acknowledgments

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.mssp.2016.09.018.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.mssp.2016.09.018.

References

[1] A. Metris, S. George, D. Ropers, Piecewise linear approximations to model the dynamics of adaptation to osmotic stress by food-borne pathogens, Int. J. Food Microbiol. (2016), http://dx.doi.org/10.1016/j.ijfoodmicro.2016.06.022.
[2] G. Batt, et al., Genetic network analyzer: a tool for the qualitative modeling and simulation of bacterial regulatory networks, Methods Mol. Biol. 804 (2012) 439–462.
[3] R. Hengge-Aronis, et al., Trehalose synthesis genes are controlled by the putative sigma factor encoded by rpoS and are involved in stationary-phase thermotolerance in *Escherichia coli*, J. Bacteriol. 173 (24) (1991) 7918–7924.
[4] D. Ropers, V. Baldazzi, H. de Jong, Model reduction using piecewise-linear approximations preserves dynamic properties of the carbon starvation response in *Escherichia coli*, IEEE/ACM Trans. Comput. Biol. Bioinform. 8 (1) (2011) 166–181.
[5] D. Ropers, et al., Qualitative simulation of the carbon starvation response in *Escherichia coli*, Biosystems 84 (2) (2006) 124–152.
[6] H.M. Giaever, et al., Biochemical and genetic characterization of osmoregulatory trehalose synthesis in *Escherichia coli*, J. Bacteriol. 170 (6) (1988) 2841–2849.
[7] J. Mellies, A. Wise, M. Villarejo, Two different *Escherichia coli* proP promoters respond to osmotic and growth phase signals, J. Bacteriol. 177 (1) (1995) 144–151.
[8] J. Xu, R.C. Johnson, Cyclic AMP receptor protein functions as a repressor of the osmotically inducible promoter proP P1 in *Escherichia coli*, J Bacteriol. 179 (7) (1997) 2410–2417.
[9] J. Xu, R.C. Johnson, Fis activates the RpoS-dependent stationary-phase expression of proP in *Escherichia coli*, J. Bacteriol. 177 (18) (1995) 5222–5231.
[10] J. Xu, R.C. Johnson, Activation of RpoS-dependent proP P2 transcription by the Fis protein in vitro, J. Mol. Biol. 270 (3) (1997) 346–359.
[11] C. Balsalobre, J. Johansson, B.E. Uhlin, Cyclic AMP-dependent osmoregulation of crp gene expression in *Escherichia coli*, J. Bacteriol. 188 (16) (2006) 5935–5944.
[12] H. Ishizuka, et al., Mechanism of the down-regulation of cAMP receptor protein by glucose in *Escherichia coli*: role of autoregulation of the crp gene, EMBO J. 13 (13) (1994) 3077–3082.
[13] J.D. Gralla, D.R. Vargas, Potassium glutamate as a transcriptional inhibitor during bacterial osmoregulation, EMBO J. 25 (7) (2006) 1515–1521.
[14] B.J. Paul, et al., rRNA transcription in *Escherichia coli*, Annu. Rev. Genet. 38 (2004) 749–770.
[15] W. Nasser, et al., CRP modulates fis transcription by alternate formation of activating and repressing nucleoprotein complexes, J. Biol. Chem. 276 (21) (2001) 17878–17886.
[16] T.S. Pratt, et al., Deletion analysis of the fis promoter region in *Escherichia coli*: antagonistic effects of integration host factor and Fis, J. Bacteriol. 179 (20) (1997) 6367–6377.

[17] Y.X. Huo, A.Z. Rosenthal, J.D. Gralla, General stress response signalling: unwrapping transcription complexes by DNA relaxation via the sigma38 C-terminal domain, Mol. Microbiol. 70 (2) (2008) 369–378.

[18] S.J. Lee, J.D. Gralla, Osmo-regulation of bacterial transcription via poised RNA polymerase, Mol. Cell 14 (2) (2004) 153–162.

[19] A.Z. Rosenthal, M. Hu, J.D. Gralla, Osmolyte-induced transcription: −35 region elements and recognition by sigma38 (rpoS), Mol. Microbiol. 59 (3) (2006) 1052–1061.

[20] C. Chaouiya, et al., SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools, BMC Syst. Biol. 7 (2013) 135.

[21] U. Dinnbier, et al., Transient accumulation of potassium glutamate and its replacement by trehalose during adaptation of growing cells of *Escherichia coli* K-12 to elevated sodium chloride concentrations, Arch. Microbiol. 150 (4) (1988) 348–357.

[22] S. Peng, et al., Transcriptional analysis of different stress response genes in *Escherichia coli* strains subjected to sodium chloride and lactic acid stress, FEMS Microbiol. Lett. 361 (2) (2014) 131–137.

[23] A. Weber, S.A. Kogl, K. Jung, Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in *Escherichia coli*, J. Bacteriol. 188 (20) (2006) 7165–7175.

[24] A. Weber, K. Jung, Profiling early osmostress-dependent gene expression in *Escherichia coli* using DNA macroarrays, J. Bacteriol. 184 (19) (2002) 5502–5507.

[25] S. Finn, et al., Exposure of *Salmonella enterica* Serovar Typhimurium to three humectants used in the food industry induces different osmoadaptation systems, Appl. Environ. Microbiol 81 (19) (2015) 6800–6811.

[26] B. Balaji, et al., Timing of induction of osmotically controlled genes in *Salmonella enterica* Serovar Typhimurium, determined with quantitative real-time reverse transcription-PCR, Appl. Environ. Microbiol. 71 (12) (2005) 8273–8283.