

**Supplementary material for:**

**The role of DNA-binding specificity in the evolution  
of bacterial regulatory networks**

Irma Lozada-Chávez, Vladimir Espinosa Angarica, Julio Collado-Vides and Bruno Contreras-  
Moreira

## References for Table 1 in main text

1. Kiley, P. J. & Beinert, H. (1998). Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol Rev* **22**, 341-52.
2. Becker, S., Holighaus, G., Gabrielczyk, T. & Unden, G. (1996). O<sub>2</sub> as the regulatory signal for FNR-dependent gene regulation in *Escherichia coli*. *J Bacteriol* **178**, 4515-21.
3. Khoroshilova, N., Popescu, C., Munck, E., Beinert, H. & Kiley, P. J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O<sub>2</sub>: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci U S A* **94**, 6087-92.
4. Reents, H., Munch, R., Dammeyer, T., Jahn, D. & Hartig, E. (2006). The Fnr regulon of *Bacillus subtilis*. *J Bacteriol* **188**, 1103-12.
5. Newman, E. B. & Lin, R. (1995). Leucine-responsive regulatory protein: a global regulator of gene expression in *E. coli*. *Annu Rev Microbiol* **49**, 747-75.
6. Willins, D. A., Ryan, C. W., Platko, J. V. & Calvo, J. M. (1991). Characterization of Lrp, and *Escherichia coli* regulatory protein that mediates a global response to leucine. *J Biol Chem* **266**, 10768-74.
7. Belitsky, B. R., Gustafsson, M. C., Sonenshein, A. L. & Von Wachenfeldt, C. (1997). An lrp-like gene of *Bacillus subtilis* involved in branched-chain amino acid transport. *J Bacteriol* **179**, 5448-57.
8. Pedersen, H. & Valentin-Hansen, P. (1997). Protein-induced fit: the CRP activator protein changes sequence-specific DNA recognition by the CytR repressor, a highly flexible LacI member. *Embo J* **16**, 2108-18.
9. Thomsen, L. E., Pedersen, M., Norregaard-Madsen, M., Valentin-Hansen, P. & Kallipolitis, B. H. (1999). Protein-ligand interaction: grafting of the uridine-specific determinants from the CytR regulator of *Salmonella typhimurium* to *Escherichia coli* CytR. *J Mol Biol* **288**, 165-75.
10. Lulko, A. T., Buist, G., Kok, J. & Kuipers, O. P. (2007). Transcriptome analysis of temporal regulation of carbon metabolism by CcpA in *Bacillus subtilis* reveals additional target genes. *J Mol Microbiol Biotechnol* **12**, 82-95.
11. Moreno, M. S., Schneider, B. L., Maile, R. R., Weyler, W. & Saier, M. H., Jr. (2001). Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol* **39**, 1366-81.
12. Lorca, G. L., Chung, Y. J., Barabote, R. D., Weyler, W., Schilling, C. H. & Saier, M. H., Jr. (2005). Catabolite repression and activation in *Bacillus subtilis*: dependency on CcpA, HPr, and HprK. *J Bacteriol* **187**, 7826-39.
13. Lopez, J. M. & Thoms, B. (1977). Role of sugar uptake and metabolic intermediates on catabolite repression in *Bacillus subtilis*. *J Bacteriol* **129**, 217-24.
14. Bagg, A. & Neilands, J. B. (1987). Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry* **26**, 5471-7.
15. Baichoo, N., Wang, T., Ye, R. & Helmann, J. D. (2002). Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. *Mol Microbiol* **45**, 1613-29.
16. Ollinger, J., Song, K. B., Antelmann, H., Hecker, M. & Helmann, J. D. (2006). Role of the Fur regulon in iron transport in *Bacillus subtilis*. *J Bacteriol* **188**, 3664-73.
17. Baichoo, N. & Helmann, J. D. (2002). Recognition of DNA by Fur: a reinterpretation of the Fur box consensus sequence. *J Bacteriol* **184**, 5826-32.

18. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W. & Strynadka, N. C. (2001). Crystal structure of LexA: a conformational switch for regulation of self-cleavage. *Cell* **106**, 585-94.
19. Miller, M. C., Resnick, J. B., Smith, B. T. & Lovett, C. M., Jr. (1996). The bacillus subtilis dinR gene codes for the analogue of Escherichia coli LexA. Purification and characterization of the DinR protein. *J Biol Chem* **271**, 33502-8.
20. Mizushima, T., Nishida, S., Kurokawa, K., Katayama, T., Miki, T. & Sekimizu, K. (1997). Negative control of DNA replication by hydrolysis of ATP bound to DnaA protein, the initiator of chromosomal DNA replication in Escherichia coli. *Embo J* **16**, 3724-30.
21. Sekimizu, K., Bramhill, D. & Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E. coli chromosome. *Cell* **50**, 259-65.
22. Fukuoka, T., Moriya, S., Yoshikawa, H. & Ogasawara, N. (1990). Purification and characterization of an initiation protein for chromosomal replication, DnaA, in Bacillus subtilis. *J Biochem (Tokyo)* **107**, 732-9.
23. Yamamoto, K. & Ishihama, A. (2006). Characterization of copper-inducible promoters regulated by CpxA/CpxR in Escherichia coli. *Biosci Biotechnol Biochem* **70**, 1688-95.
24. Howell, A., Dubrac, S., Noone, D., Varughese, K. I. & Devine, K. (2006). Interactions between the YycFG and PhoPR two-component systems in Bacillus subtilis: the PhoR kinase phosphorylates the non-cognate YycF response regulator upon phosphate limitation. *Mol Microbiol* **59**, 1199-215.
25. Makino, K., Shinagawa, H., Amemura, M., Kawamoto, T., Yamada, M. & Nakata, A. (1989). Signal transduction in the phosphate regulon of Escherichia coli involves phosphotransfer between PhoR and PhoB proteins. *J Mol Biol* **210**, 551-9.
26. Baruah, A., Lindsey, B., Zhu, Y. & Nakano, M. M. (2004). Mutational analysis of the signal-sensing domain of ResE histidine kinase from Bacillus subtilis. *J Bacteriol* **186**, 1694-704.
27. Martinez-Antonio, A. & Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* **6**, 482-9.
28. Moreno-Campuzano, S., Janga, S. C. & Perez-Rueda, E. (2006). Identification and analysis of DNA-binding transcription factors in Bacillus subtilis and other Firmicutes--a genomic approach. *BMC Genomics* **7**, 147.
29. Contreras-Moreira, B., Branger, P. A. & Collado-Vides, J. (2007). TFmodeller: comparative modelling of protein-DNA complexes. *Bioinformatics*.
30. DiGiuseppe, P. A. & Silhavy, T. J. (2003). Signal detection and target gene induction by the CpxRA two-component system. *J Bacteriol* **185**, 2432-40.

## Supplementary methods

### M1. Estimation of contact-based specificity

These are the steps followed to estimate the *a priori* binding specificity of a given transcription factor  $T$  based on its contact pattern:

- 1) The most similar protein-DNA complex  $C$  in the Protein Data Bank (PDB) [28] is identified using PSI-BLAST[29] and the set of residues contacting nitrogen bases is defined, using the TFmodeller program with default parameters[30]. These contacting residues are considered to be the specific subset  $S$  of the complete protein-DNA interface.
- 2)  $C$  is used to scan a 40% non-redundant library of protein-DNA complexes using MAMMOTH[31] with a log (e-value) cut-off of 4.5 in order to get a list  $L$  of structurally similar complexes.
- 3) For every complex in  $L$ , residues in contact with DNA nitrogen bases are mapped to residues in  $T$  through MAMMOTH alignments. The complete set of  $L$  mapped residues is considered to be the non-specific subset  $NS$  of the protein-DNA interface in  $T$ . This step is very sensitive to the size of  $L$ , so we skipped cases in which the final  $NS$  set included fewer contacts than the mean minus one standard deviation, to exclude cases where the number of available similar complexes was not sufficient (12.9 contacts for *E. coli* and 12.4 for *B. subtilis*).
- 4) We estimate the specificity of  $T$  as the ratio  $S/(NS+S)$ , such as in the work of Luscombe and Thornton [32].

Following these steps we were able to estimate specificities for 49 TFs in *E. coli* and for 33 TFs in *B. subtilis*.

## Supplementary tables

**Table S1. Transcriptional regulatory interactions of *Escherichia coli* K12 from RegulonDB release version 5.5 (<http://regulondb.ccg.unam.mx/>) and for *Bacillus subtilis* from DBTBS release version 4.1 (<http://dbtbs.hgc.jp/>).**

**k:** Total number of regulated genes by a transcription factor (TF) with known DNA binding sites.

**n:** Total number of DNA binding sites that have been characterized for a TF.

<b><i>Escherichia coli</i></b> <b>RegulonDB version 5.5</b>				
Transcription Factor			k # genes	n # bsDNA
crp	16131236	b3357	276	207
fis	16131149	b3261	76	121
ihf	1788005	b1712	160	78
arcA	16132218	b4401	92	77
narL	16129184	b1221	81	76
fnr	16129295	b1334	135	75
lrp	16128856	b0889	59	54
fur	16128659	b0683	36	47
h-ns	1787489	b1237	45	34
cpxR	16131752	b3912	10	33
lexA	16131869	b4043	14	23
metJ	16131776	b3938	10	23
ompR	16131282	b3405	14	20
argR	16131127	b3237	22	18
flhD	49176166	b1892	26	18
glpR	16131297	b3423	8	18
phoP	16129093	b1130	2	18
soxS	16131888	b4062	20	18
phoB	16128384	b0399	31	17
tyrR	16129284	b1323	10	17
marA	16129490	b1531	14	16
purR	16129616	b1658	27	16
malT	16131294	b3418	10	15
narP	16130130	b2193	21	14
araC	16128058	b0064	7	13
fruR	16128073	b0080	25	12
cytR	16131772	b3934	10	10
fadR	16129150	b1187	7	10
gntR	49176356	b3438	9	10
iclR	16131844	b4018	3	10
nac	16129930	b1988	10	10
nagC	16128652	b0676	9	10
trpR	16132210	b4393	11	10
oxyR	16131799	b3961	12	9
cysB	16129236	b1275	17	8
dnaA	16131570	b3702	8	8
iscR	16130456	b2531	3	8
deoR	16128808	b0840	6	7
galR	16130741	b2837	5	7
galS	16130089	b2151	3	7
mlc	16129552	b1594	8	7

<b><i>Bacillus subtilis</i></b> <b>DBTBS version 4.1</b>				
Transcription Factor			k # genes	n # bsDNA
comK	16078106	BG11059	47	120
phoP	16079963	BG10363	35	65
ccpA	16080026	BG10376	97	48
mtrB	16079334	BG10278	4	41
spoIIID	16080695	BG10408	11	39
spo0A	16079478	BG10765	12	38
degU	16080602	BG10393	19	37
tnrA	16078396	BG11805	31	37
abrB	16077105	BG10100	72	34
codY	16078680	BG10968	29	34
gerE	16079893	BG10355	20	31
fur	16079409	BG11766	34	26
purR	16077115	BG10110	22	24
dnaA	16077069	BG10065	2	16
spoVT	16077124	BG10119	15	16
hpr	16078063	BG10659	14	14
perR	16081061	BG12227	12	13
cssR	16080354	BG14131	3	12
glnR	16078808	BG10424	8	12
resD	16079369	BG10534	13	12
sinR	16079517	BG10754	6	12
ylbO	16078573	BG13367	2	12
araR	16080450	BG11913	12	11
rocR	16081087	BG10723	3	11
comA	16080219	BG10381	10	10
ctsR	16077151	BG10145	11	10
glpP	16077992	BG10185	5	10
lexA	16078848	BG10678	4	10
rok	16078488	BG13307	1	10
ccpC	16078478	BG13297	2	9
citT	16077826	BG12577	1	8
fnr	16080784	BG11343	8	8
yycF	16081093	BG10001	4	8
pucR	16080295	BG13983	15	7
deoR	16080994	BG10982	1	6
ydiH	16077664	BG12205	4	6
ahrC	16079481	BG10309	7	5
cggR	16080448	BG14085	1	5
iolR	16081028	BG11364	3	5
mta	16080713	BG12482	4	5
citR	16078008	BG10853	1	4

melR	16131944	b4118	2	6
modE	16128729	b0761	35	6
rob	16132213	b4396	8	6
torR	16128961	b0995	10	6
caiF	49175992	b0034	10	4
fhlA	16130638	b2731	26	4
gcvA	16130715	b2808	3	4
metR	16131677	b3828	3	4
rhaS	16131745	b3905	4	4
xylR	16131440	b3569	5	4
ada	16130150	b2213	4	3
dcuR	16131950	b4124	7	3
idnR	16132086	b4264	4	3
lacI	49176012	b0345	3	3
pspF	49176097	b1303	6	3
uxuR	16132145	b4324	2	3
cbl	16129929	b1987	4	2
cynR	16128323	b0338	3	2
exuR	16130989	b3094	3	2
ilvY	16131631	b3773	1	2
malI	16129578	b1620	2	2
paaX	16129360	b1399	12	2
rhaR	16131746	b3906	1	2
slyA	16129600	b1642	1	2
treR	16132063	b4241	2	2
uhpA	16131539	b3669	1	2
xapR	16130331	b2405	2	2
betI	49176011	b0313	3	1
birA	16131807	b3973	5	1
cusR	16128554	b0571	4	1
evgA	16130301	b2369	1	1
fecI	16132114	b4293	6	1
glcC	16130880	b2980	5	1
hyfR	49176231	b2491	10	1
kdpE	16128670	b0694	3	1
lrhA	16130224	b2289	2	1
mhpR	16128331	b0346	6	1
nhaR	16128014	b0020	2	1
rbsR	16131621	b3753	5	1
soxR	16131889	b4063	1	1
tdcA	16131011	b3118	6	1
tdcR	16131012	b3119	7	1
yiaJ	16131445	b3574	9	1
zraR	16131834	b4004	2	1
<b>85 TFs</b>			<b>1593</b>	<b>1314</b>

gltR	16079720	BG11942	1	4
hrcA	16079603	BG10662	2	4
paiA	16080268	BG10695	1	4
paiB	16080267	BG10696	1	4
pyrR	16078611	BG10712	1	4
tenA	16078230	BG10791	4	4
xre	16078316	BG10994	4	4
ykmA	16078380	BG13239	1	4
adaA	16077249	BG10166	2	3
bkdR	16079466	BG11721	7	3
gltC	16078907	BG10810	2	3
levR	16079762	BG10677	1	3
yufM	16080205	BG12348	3	3
zur	16079565	BG11668	3	3
acoR	16077877	BG11790	4	2
ccpB	16081139	BG10045	2	2
fruR	16078502	BG12589	2	2
gabR	16077457	BG12042	2	2
gutR	16077681	BG10178	2	2
licT	16080959	BG10474	2	2
lytR	16080618	BG10404	3	2
mntR	16079508	BG11702	5	2
sacT	16080858	BG10593	4	2
sacY	16080893	BG10559	1	2
treR	16077849	BG11011	1	2
xylR	16078822	BG11986	2	2
yfhP	16077928	BG12890	1	2
yocG	16078980	BG13520	1	2
ywfK	16080817	BG10635	2	2
arsR	16079634	BG11301	3	1
azlB	16079725	BG11914	4	1
birA	16079301	BG11206	7	1
bltR	16079711	BG10904	2	1
bmrR	50812267	BG10304	1	1
exuR	16078302	BG13211	1	1
glcT	16078452	BG12593	3	1
hxlR	16077416	BG11184	2	1
kdgR	16079270	BG11398	3	1
kipR	16077477	BG11214	6	1
lacR	16080470	BG12435	5	1
licR	16080911	BG11346	1	1
lmrA	16077337	BG12612	1	1
lrpA	16077572	BG12122	1	1
lrpB	16077573	BG12123	1	1
med	50812218	BG13126	1	1
tenI	16078231	BG10792	1	1
xpf	16078321	BG10998	10	1
yfiA	16077886	BG11847	1	1
<b>89 TFs</b>			<b>726</b>	<b>940</b>

**Table S2. Paralogous groups after assigning SUPERFAMILY and PFAM domains (see Materials and Methods), including 151 transcription factors of *Escherichia coli*.**

- † Gene(s) that belongs to other PFAM group [not shown].  
 \* PFAM group that include more gene members than their SFAM family [not shown].  
 † PFAM group that is shared with other SFAM family.

**Transcription factor genes**

SFAM domain-#	PFAM domain-#	# genes	# TFs	Members
<b>Groups of orthologous TFs formed by congruence in domains and total members</b>				
46689-1	PF00165.12-2	1	1	16130150 b2213 ada
46767-1	PF01035.11-1			
53155-1	PF02805.5-1			
57884-1	PF02870.6-1			
46689-1	PF00158.16-1	1	1	16129284 b1323 tyrR
52540-1	PF01842.14-1			
55021-1				
55785-1				
46689-1	PF00158.16-1	1	1	16130638 b2731 fh1A
52540-1	PF01590.15-1			
55781-2	PF02954.9-1			
	PF07728.4-1			
46785-1	PF02237.6-1	1	1	16131807 b3973 birA
50037-1	PF03099.9-1			
55681-1	PF08279.1-1			
46785-1	PF00392.11-1	1	1	16129150 b1187 fadR
48008-1	PF07840.2-1			
46785-1	PF00126.16-1	1	1	16128729 b0761 modE
50331-2	PF03459.7-2			
46785-1	PF00717.13-1	1	1	16131869 b4043 lexA
51306-1	PF01726.7-1			
46785-1	PF01316.12-1	1	1	16131127 b3237 argR
55252-1	PF02863.9-1			
48295-1	PF00308.8-1	1	1	16131570 b3702 dnaA
52540-1	PF08299.1-1			
47598-1	PF01340.10-1	1	1	16131776 b3938 metJ
48295-1	PF01371.10-1	1	1	16132210 b4393 trpR
63592-1	PF05247.3-1	1	1	49176166 b1892 flhD
<b>Groups of paralogous TFs formed by congruence in domains and total members</b>				
47413-1	PF00356.11-1	14	11	16128073 b0080 fruR 16129281 b1320 ycjW 16129578 b1620 malI
53822-1	PF00532.11-1			16129616 b1658 purR 16130089 b2151 gals 16130621 b2714 ascG
				16130741 b2837 galR 16130970 b3075 ebgR 16131621 b3753 rbsR
				16131772 b3934 cytR 16132063 b4241 treR 16132086 b4264 idnR
				49176012 b0345 lacI 49176356 b3438 gntR
46785-1	PF01614.9-1	8	3	16128257 b0272 yagI 16128331 b0346 mhpR 16128490 b0506 allR
55781-1				16129781 b1827 kdgR 16130183 b2248 yfaX 16131445 b3574 yiaJ
				16132120 b4299 yjhI 90111675 b4018 iclR
46785-1	PF01037.10-1	3	1	16128856 b0889 lrp 16131611 b3743 asnC 90111137 b0447 ybaO
54909-1				
46785-1	PF00027.18-1	2	2	16129295 b1334 fnr 16131236 b3357 crp
51206-1	PF00325.11-1			

47729-1	PF00216.11-1	4	1	16128425 b0440 hupB <b>16128879 b0912 ihfB 16129668 b1712 ihfA</b> 16131830 b4000 hupA
81273-2	PF00816.10-1	2	1	<b>16129198 b1237 hns</b> 16130583 b2669 stpA
<b>Groups of paralogous TFs formed by congruence in domains and TF members</b>				
46689-1 52172-1 52540-1	PF00072.13-1 PF00158.16-1 PF02954.9-1 PF07728.4-1	3	1	16131708 b3868 glnG <b>16131834 b4004 zraR</b>  †16130157 b2220 atoC
46785-1 53850-1	PF00126.16-1 PF03466.9-1	45	12	<b>16128014 b0020 nhaR</b> 16128195 b0208 yafC 16128239 b0254 perR 16128301 b0316 yahB <b>16128323 b0338 cynR</b> 16128488 b0504 ybbS 16128867 b0900 ycaN <b>16129236 b1275 cysB</b> 16129289 b1328 ycjZ 16129300 b1339 abgR 16129485 b1526 yneJ 16129553 b1595 ynfL 16129617 b1659 ydhB <b>16129929 b1987 cbl 16129930 b1988 nac</b> 16130095 b2157 yeiE <b>16130224 b2289 lrhA</b> 16130296 b2364 dsdC <b>16130331 b2405 xapR</b> 16130335 b2409 yfeR 16130462 b2537 hcaR <b>16130715 b2808 gcvA</b> 16130743 b2839 lysR 16130817 b2916 argP 16130956 b3060 ygiP 16131000 b3105 yhaJ <b>16131011 b3118 tdcA</b> 16131133 b3243 aaeR 16131393 b3521 yhjC 16131456 b3585 yiaU 16131579 b3711 yidZ <b>16131631 b3773 ilvY 16131677 b3828 metR</b> <b>16131799 b3961 oxyR</b> 16132148 b4327 yjiE 49176046 b0768 ybhD 49176402 b4480 hdfR 90111083 b0076 leuO 90111155 b0629 ybeF 90111269 b1422 ydcI 90111335 b1799 yeaT 90111371 b2015 yeeY 90111462 b2577 yfiE 90111513 b2921 ygfI  †16128586 b0603 ybdO
88659-1 88946-1	PF04542.4-1 PF04545.5-1 PF08281.1-1	3	1	<b>16132114 b4293 fecI</b>  †16130498 b2573 rpoE 16131333 b3461 rpoH
46689-2 55136-1	PF00165.12-2 PF06445.5-1	2	1	<b>16132213 b4396 rob</b>  †90111109 b0300 ykgA
46689-1 52540-1	PF00158.16-1 PF02954.9-1 PF07728.4-1	2	1	<b>90111246 b1303 pspF</b>  †16128315 b0330 prpR
52540-1 55781-1	PF00158.16-1 PF01590.15-1 PF02954.9-1	2	1	<b>90111447 b2491 hyfR</b>  †90111481 b2709 norR
100950-1 46785-1	PF00455.13-1 PF08220.3-1 PF08279.1-1	9	1	16129245 b1284 yciT 16130614 b2707 srlR 16130642 b2735 ygbI 16130712 b2805 fucR 16131023 b3131 agaR <b>16131297 b3423 glpR</b> 90111660 b3884 yihW  †16132013 b4191 ulaR 16129724 b1770 ydjF
46894-1 48452-1	PF00196.10-1	2	1	<b>16131294 b3418 malT</b>  †16130756 b2852 ygeH
100950-1	PF00455.13-1 PF08220.3-1	9	1	* <b>16128808 b0840 deoR</b> 16132121 b4300 sgcR  †16128598 b0615 citF 16128654 b0678 nagB 16130158 b2221 atoD 16130159 b2222 atoA 16131033 b3141 agaI 49176282 b2912 ygfA 49176397 b3718 yieK
46955-1	PF00376.13-1	10	1	16128471 b0487 cueR 16129125 b1162 ycgE 16130065 b2127 mlrA 16131171 b3292 zntR <b>16131889 b4063 soxR</b>



				†16128543 b0560 nohB 16129104 b1141 ymfG 16129507 b1548 nohA 16130329 b2399 yfeD 90111429 b2398 yfeC
53067-2	PF00480.10-1	25	2	<b>16128652 b0676 nagC 16129552 b1594 dgsA</b> 16131910 b4084 alsK 90111457 b2550 yphH 90111558 b3222 nanK  †16128057 b0063 araB 16128087 b0094 ftsA 16129082 b1119 nagK 16129761 b1807 yeaZ 16130231 b2296 ackA 16130320 b2388 glk 16130379 b2454 eutJ 16130427 b2502 ppx 16130683 b2776 ygcE 16130710 b2803 fucK 16131008 b3115 tdcD 16131435 b3564 xylB 16131451 b3580 lyxK 16131744 b3904 rhaB 16131764 b3926 glpK 49176118 b1511 lsrK 49176408 b3779 gpp 90111128 b0394 mak 90111380 b2069 yegD 90111564 b3251 mreB
<b>Groups of paralogous TFs formed by groups of domains and same TF members</b>				
46894-1 52172-1	PF00072.13-1 PF00196.10-1 PF04545.5-1 PF08281.1-1  PF00072.13-1 PF00196.10-1	10	4	<b>16129184 b1221 narL</b>  16128519 b0535 fimZ 16129861 b1914 uvrY <b>16130130 b2193 narP</b> 16130154 b2217 rCSB <b>16130301 b2369 evgA 16131539 b3669 uhpA</b>  †16129566 b1608 rstA 16132215 b4398 creB 90111499 b2855 ygeK
52172-1	PF00072.13-1 PF00486.18-1  PF00072.13-1	18	9	<b>*ψ16128384 b0399 phoB 16128554 b0571 cusR</b> <b>16128670 b0694 kdpE 16128961 b0995 torR 16129093 b1130 phoP</b> 16130019 b2079 baeR 16130313 b2381 ypdB 16130921 b3025 qseB <b>16131282 b3405 ompR 16131752 b3912 cpxR</b> 16131939 b4113 basR <b>16132218 b4401 arcA</b> 90111362 b1969 yedW 90111392 b2125 yehT  16128603 b0620 citB 16129196 b1235 rssB 16129834 b1882 cheY <b>16131950 b4124 dcuR</b>
46785-1	PF01475.9-1  PF01047.13-1  PF02082.9-1  PF00392.11-1 PF07729.1-1	30	6	<b>16128659 b0683 fur</b> 90111679 b4046 zur  16130596 b2684 mprA 90111289 b1530 marR <b>90111310 b1642 slyA</b>  <b>16130456 b2531 iscR</b> 16132000 b4178 yjeB  16128106 b0113 pdhR 16129499 b1540 ydfH <b>16130880 b2980 glcC</b> 16131475 b3604 lldR <b>16132145 b4324 uxuR</b> 49176329 b3226 nanR 49176392 b4479 dgoR 49176401 b3755 yieP 90111276 b1450 yncC 90111478 b2664 csiR <b>90111537 b3094 exuR</b>  †16129112 b1149 ymfN 16130581 b2667 ygaV 16130966 b3071 yqjI 16131373 b3501 arsR 49176497 b4357 yjjM 90111706 b4212 ytfH 16128705 b0730 mngR 16130039 b2101 yegW 16131552 b3684 yidP 16131712 b3872 yihL 16131928 b4102 phnF 90111578 b3375 frlR
46689-1	PF02954.9-1  PF00440.13-1	18	2	<b>16131149 b3261 fis</b>  16128814 b0846 ybjK 16131801 b3963 fabR <b>49176011 b0313 betI</b> 90111712 b4251 yJgJ  †16128529 b0546 ybcM 16128547 b0564 appY 16129345 b1384 feaR 16131387 b3515 gadW 16128549 b0566 envY 16129458 b1499 ydeO 16130362 b2437 yfeG 16131388 b3516 gadX 16131428 b3557 insJ 16131792 b3954 yijO 16131942 b4116 adiY 90111143 b0502 yIbG 90111689 b4135 yJdC

<b>Groups of paralogous TFs formed by preference in SFAM domains and TF members</b>				
46689-1 47454-1 51182-1	PF00165.12-2	1	1	<b>*ψ16131944 b4118 me1R</b>
46689-1 53822-1	PF00165.12-2	1	1	<b>*ψ16131440 b3569 xylR</b>
46689-2 51182-1	PF00165.12-2 PF02311.9-1	2	1	<b>*ψ16131746 b3906 rhaR</b> †16129652 b1696 ydiP
46689-2 51215-1	PF00165.12-2 PF02311.9-1	2	2	<b>*ψ16128058 b0064 araC 16131745 b3905 rhaS</b>
46689-2	PF00165.12-2	6	2	<b>*ψ16128290 b0305 ykgD 16130314 b2382 ypdC</b> 16131550 b3680 yidL <b>16131888 b4062 soxS 90111290 b1531 marA</b> †90111526 b3010 yqhC
<b>Groups of paralogous TFs only formed by PFAM domains</b>				
	PF07848.2-1 PF08223.1-1	1	1	<b>16129360 b1399 paaX</b>

**Table S3. Paralogous groups after assigning SUPERFAMILY and PFAM domains (see Materials and Methods), including 93 transcription factors of *Bacillus subtilis*.**

† Gene(s) that belongs to other PFAM group [not shown].

\* PFAM group that include more gene members than their SFAM family [not shown].

ψ PFAM group that is shared with other SFAM family.

**Transcription factor genes**

SFAM domain-#	PFAM domain-#	# genes	# TFs	Members
<b>Groups of orthologous TFs formed by congruence in domains and total members</b>				
46689-1	PF00158.16-1	1	1	16079466 BG11721 bkdR
51735-1	PF00989.13-2			
52540-1	PF02954.9-1			
55785-2	PF08448.1-2			
46785-1	PF00004.19-1	1	1	16079762 BG10677 levR
52540-1	PF00158.16-1			
53062-1	PF00874.9-2			
63520-2	PF03610.5-1			
46785-2	PF00359.11-1	1	1	16080911 BG11346 licR
52794-1	PF00874.9-2			
55804-1	PF05043.3-1			
63520-2	PF08279.1-2			
46689-1	PF00158.16-1	1	1	16077877 BG11790 acoR
52540-1	PF01590.15-1			
55781-1	PF02954.9-1			
	PF07728.4-1			
46689-1	PF00158.16-1	1	1	16081087 BG10723 rocR
52540-1	PF02954.9-1			
55785-1	PF07726.1-1			
	PF07728.4-1			
46689-2	PF00165.12-2	1	1	16077249 BG10166 adaA
57884-1	PF02805.5-1			
46785-1	PF01325.9-1	1	1	16079508 BG11702 mntR
47979-1	PF02742.5-1			
46785-1	PF00027.18-1	1	1	16080784 BG11343 fnr
51206-1	PF00325.11-1			
46785-1	PF00717.13-1	1	1	16078848 BG10678 lexA
51306-1	PF01726.7-1			
46785-1	PF02629.8-1	1	1	16077664 BG12205 ydiH
51735-1	PF06971.3-1			
46785-1	PF00392.11-1	1	1	16080450 BG11913 araR
53822-1	PF00532.11-1			
46785-1	PF01316.12-1	1	1	16079481 BG10309 ahrC
55252-1	PF02863.9-1			
46785-1	PF02237.6-1	1	1	16079301 BG11206 birA
55681-1	PF03099.9-1			
	PF08279.1-1			
46955-1	PF00376.13-1	1	1	16080713 BG12482 mta
89082-1	PF07739.3-1			
48295-1	PF00308.8-1	1	1	16077069 BG10065 dnaA
52540-1	PF08299.1-1			

46894-1	PF00196.10-1 PF04545.5-1 PF08281.1-1	1	1	<b>16079893 BG10355 gerE</b>
48613-1	PF03070.5-1	1	1	<b>16078230 BG10791 tenA</b>
51219-1	PF02081.5-1	1	1	<b>16079334 BG10278 mtrB</b>
110391-1	PF04309.2-1	1	1	<b>16077992 BG10185 glpP</b>
<b>Groups of paralogous TFs formed by congruence in domains and total members</b>				
46785-1 53850-1	PF00126.16-1 PF03466.9-1	19	6	16077386 BG12009 ycgK 16077430 BG12022 yclA <b>16078008 BG10853 citr</b> 16078478 BG13297 ccpC 16078903 BG13552 yofA 16078907 BG10810 gltC 16078936 BG13491 yofA 16079716 BG12288 yrdQ <b>16079720 BG11942 gltR</b> 16079740 BG12278 yraN 16079992 BG13877 ytlI 16080344 BG14032 yusT 16080452 BG14087 yvbU <b>16080655 BG10470 alsR</b> 16080669 BG12517 ywqM <b>16080817 BG10635 ywfK</b> 16080882 BG10570 ywbI 16080939 BG11898 yxjO 16081119 BG10026 yybE
50151-1 63520-2	PF00874.9-2 PF03123.5-1	4	4	<b>16078452 BG12593 glcT</b> 16080858 BG10593 sacT <b>16080893 BG10559 sacY</b> 16080959 BG10474 licT
51391-1	PF02581.6-1	2	1	<b>16078231 BG10792 tenI</b> 16080880 BG10572 thiE
<b>Groups of paralogous TFs formed by congruence in domains and TF members</b>				
46689-1 48498-1	PF00440.13-1	10	1	* <b>16077337 BG12612 lmrA</b> 16077599 BG12146 ydeS 16077625 BG12170 ydgC 16077904 BG12894 yfiR 16078079 BG10432 yhgD 16080411 BG14062 yvaF 16080501 BG12428 yvdT 16081051 BG11108 yxaF 50812197 BG14194 yezE †16079907 BG12330 ysiA
46785-1 53383-1	PF00155.11-1 PF00392.11-1	6	1	* <b>16077457 BG12042 gabR</b> 16077585 BG12133 ydeF 16077591 BG12139 ydeL 16077604 BG12151 ydfD 16078013 BG13015 yhdI 50812213 BG13101 yisV
55729-1	PF00583.14-1	48	1	16077488 BG12054 ydaF 16077602 BG12149 ydfB 16077627 BG12172 ydgE 16077644 BG12186 ydhI 16077660 BG12201 ydiD 16077759 BG11809 yesJ 16077811 BG12962 yfmK 16078014 BG13016 yhdJ 16078095 BG13059 yhfO 16078163 BG13113 yitH 16078164 BG13114 yitI 16078249 BG13159 yjcF 16078254 BG13164 yjck 16078269 BG13179 yjdG 16078373 BG11421 ykkB 16078574 BG13368 yldbP 16078815 BG12256 ynaD 16078914 BG13472 yofA 16078929 BG13486 yofP 16078966 BG13509 yobR 16079033 BG11197 cgeE 16079214 BG13579 yokL 16079381 BG10522 ribT 16079426 BG11754 yqjY 16079698 BG11780 yrkN 16079713 BG10906 bltD 16079885 BG12336 ysnE 16079991 BG13883 ytmI 16080021 BG10369 acuA <b>16080268 BG10695 paiA</b> 16080442 BG14084 yvbK 16080709 BG12486 ywnH 16080839 BG10612 spsD 16081002 BG11888 yxeL 16081038 BG11354 yxbd 16081081 BG11467 yycN 16081120 BG10027 yybD 16081124 BG10031 yyaT 16081126 BG10033 yyaR  †16077239 BG11571 ybbJ 16078214 BG13132 yjbC 16078458 BG12252 ykwB 16078532 BG13331 ykcC 16078976 BG13516 yocC 16079028 BG13543 yodP 16079424 BG11756 yqkA 16080152 BG13948 yuaI 16080749 BG10938 ywlB

48452-1	PF00515.17-1 PF00931.12-1 PF05729.2-1 PF07719.6-1	16	1	<b>16077681 BG10178 gutR</b>  †16077351 BG12663 rapJ 16077568 BG12119 rapI 16077751 BG11031 rapH 16078308 BG10652 rapA 16078951 BG12664 rapK 16079316 BG11496 ypiA 16079543 BG11683 yqgP 16079636 BG11299 rapE 16079802 BG13788 yrrB 16079876 BG12338 ysoA 16080534 BG12396 yvcD 16080691 BG11967 rapD 16080722 BG11965 rapB 16080798 BG11968 rapF 16081082 BG11466 rapG
53067-2	PF00480.10-1	16	1	16077653 BG12195 ydhR <b>16078822 BG11986 xylR</b> 16079541 BG11685 glcK  †16077659 BG12200 ydiC 16077994 BG10187 glpK 16078511 BG10987 mreBH 16078592 BG10231 ftsA 16078824 BG10807 xylB 16078916 BG13474 yocA 16079463 BG11724 buk 16079855 BG10326 mreB 16079931 BG11905 araB 16079999 BG10813 ackA 16080172 BG12385 yulC 16081058 BG10649 gntK 50812298 BG10916 mbI
53697-1	PF01380.12-1 PF01418.8-1	6	1	16077237 BG11569 ybbH <b>16077886 BG11847 yfiA</b>  †16077238 BG11570 ybbI 16077414 BG11182 hxIb 16080187 BG12366 pgi 16080314 BG14002 yurP
88659-1	PF04545.5-1 PF08281.1-1	5	1	<b>16078321 BG10998 xpf</b> 16080376 BG14147 yvrI  †16078660 BG10829 ylxM 16079692 BG10919 spoIIIC 16078258 BG13168 yjcO
89447-1	PF04014.8-1	4	2	<b>16077105 BG10100 abrB 16077124 BG10119 spoVT</b> 16078512 BG10988 abh  †16078577 BG11425 yIIb
53822-1	PF02608.4-1	3	1	16080206 BG12349 yufN <b>50812218 BG13126 med</b>  †16080649 BG10881 rbsB
<b>Groups of paralogous TFs formed by groups of domains and same TF members</b>				
46894-1 52172-1	PF00072.13-1 PF00196.10-1  PF00072.13-1 PF00196.10-1 PF08281.1-1  PF00072.13-1  PF00072.13-1 PF00486.18-1	12	5	16077269 BG12723 ybdJ <b>16080219 BG10381 comA</b> 16080361 BG14133 yvqC 16080459 BG12446 yvfU <b>16080602 BG10393 degU</b> 16080942 BG11895 yxjL  16077998 BG13006 yhcZ <b>16078980 BG13520 yocG</b>  <b>*Ψ16079478 BG10765 spo0A</b>  <b>*Ψ16080354 BG14131 cssR</b>  †16077609 BG12156 ydfI 16077897 BG12216 yfiK
47413-1 53822-1	PF00356.11-1 PF00532.11-1  PF00356.11-1	11	5	16078120 BG13079 yhjM 16078147 BG10847 degA <b>16078302 BG13211 exuR</b> 16078451 BG13327 ykvZ <b>16079270 BG11398 kdgR 16080026 BG10376 ccpA</b> 16080078 BG12626 msmR 16080516 BG12414 yvdE 16080644 BG10876 rbsR <b>16081139 BG10045 ccpB</b>  <b>16080470 BG12435 lacR</b>

46785-1 55781-1	PF01614.9-1 PF01628.11-1	2	2	<b>16077477 BG11214 kipR</b> <b>16079603 BG10662 hrcA</b>
100950-1 46785-1	PF00455.13-1 PF08220.3-1 PF00455.13-1 PF08220.3-1 PF08279.1-1	4	2	<b>16078502 BG12589 fruR</b> 16080173 BG12384 yulB 16080683 BG12503 glcR <b>16081028 BG11364 iolR</b>
52172-1	PF00072.13-1 PF00486.18-1 PF00072.13-1	20	5	<b>*Ψ16077324 BG11167 ycbL</b> 16077443 BG12031 yclJ 16078390 BG13252 ykoG <b>16079369 BG10534 resD</b> 16079696 BG11782 yrkP <b>16079963 BG10363 phoP</b> 16080092 BG12340 ytsA 16080375 BG12452 yvrH 16080525 BG12405 yvcP 16081017 BG11126 yxdJ <b>16081093 BG10001 yycF</b> <b>*Ψ16077513 BG12074 dctR</b> <b>16077826 BG12577 citT</b> 16078696 BG10258 cheY 16078857 BG11250 yneI <b>16080205 BG12348 yufM</b> 16080766 BG10411 spo0F †16077314 BG10872 ycbB 16077343 BG11998 ycch 16079944 BG11953 lytT
46785-1	PF01047.13-1 PF01638.8-1 PF00392.11-1 PF07702.3-1 PF01022.10-1 PF01475.9-1 PF01037.10-1 PF04198.2-1	68	10	16077377 BG12003 ycgE 16077544 BG12095 ydcH 16077631 BG12174 ydgG 16077634 BG12177 ydgJ 16077789 BG12868 yetL 16077908 BG12898 yfiV 16077964 BG13003 yhbI <b>16078063 BG10659 hpr</b> 16078115 BG13074 yhjH <b>16078380 BG13239 ykmA</b> 16078399 BG13258 ykoM 16078431 BG13310 ykvE 16079229 BG11626 ypoP 16079892 BG12332 ysmB 16080339 BG14027 yusO 16080558 BG14122 yvnA 16080561 BG11018 yvmB 16080697 BG12495 ywoH 16080807 BG12455 ywhA 16080896 BG10556 ywaE 16081053 BG11106 yxaD 16081123 BG10030 yybA <b>16077416 BG11184 hxlR</b> 16077594 BG12809 ydzF 16077596 BG12143 ydeP 16078440 BG13316 ykvN 16079012 BG13531 yodB 16079955 BG13831 ytcD 16080421 BG14071 yvaP 16081106 BG10013 yybR 16077306 BG12747 ybgA 16077652 BG12194 ydhQ <b>16077849 BG11011 treR</b> 16078744 BG13423 ymfC 16080309 BG13997 yurK 16080556 BG14124 yvoA 16081065 BG11484 yydK 16077456 BG12781 yczG 16077600 BG12147 ydeT 16077606 BG12153 ydfF 16078972 BG13748 yoza <b>16079634 BG11301 arsR</b> 16080432 BG14078 yvbA <b>16077938 BG12227 perR</b> <b>16079409 BG11766 fur</b> <b>16079565 BG11668 zur</b> <b>*Ψ16077572 BG12122 lrpA</b> <b>*16080448 BG14085 cggR</b> †16077971 BG11584 yhcf 16078003 BG13011 yhdE 16078910 BG10738 rtp 16078970 BG13513 yobV 16079765 BG12303 yrhO 16079806 BG13813 yrzC 16080054 BG13939 ytzE 16080098 BG13911 ytrA

				16080159 BG12342 yuaC 16080427 BG14073 yvaV 16080437 BG14079 yvbF 16080716 BG12479 ywnA 16080810 BG12454 ywgB 16081057 BG10648 gntR 16077319 BG11162 ycbG 16077366 BG12775 yceK 16077464 BG12047 ycnK 16077637 BG12180 ydhC 16077722 BG12874 yezC 16077901 BG12220 padR
<b>Groups of paralogous TFs formed by preference in SFAM domains and same TFs</b>				
46785-1 54909-1	PF01037.10-1	5	2	* $\Psi$ 16077492 BG12056 lrpC <b>16077573 BG12123 lrpB</b> <b>16079725 BG11914 azlB</b> 16080193 BG12361 alaR  †16080664 BG12522 ywrC
46955-1 55136-1	PF00376.13-1 PF06445.5-1 PF00376.13-1	3	2	16077613 BG12159 ydfL <b>50812267 BG10304 bmrR</b>  * $\Psi$ 16079711 BG10904 bltR
47406-1 47413-1	PF01381.12-1	1	1	* $\Psi$ 16079517 BG10754 sinR
46785-1 53271-1	PF00156.16-1	1	1	* $\Psi$ 16077115 BG10110 purR
46955-1	PF00376.13-1	7	2	* $\Psi$ 16077806 BG12967 yfmP 16078021 BG13023 yhdQ <b>16078396 BG11805 tnrA 16078808 BG10424 glnR</b> 16079754 BG13777 yraB 16081132 BG10038 yyaN  †16080756 BG11314 ywkC
53271-1	PF00156.16-1	8	1	* $\Psi$ 16078611 BG10712 pyrR 16077119 BG10114 prs 16077136 BG10131 hprT 16078620 BG10720 pyrE 16079265 BG11079 xpt 16079815 BG12568 apt 16080742 BG10945 upp  †16080598 BG10397 comFC
47413-1	PF01381.12-1	17	2	* $\Psi$ 16077148 BG12701 yazB 16077549 BG12101 ydcN <b>16078316 BG10994 xre</b> 16078934 BG13754 yozG 16078945 BG13497 yobD 16079137 BG13652 yopS 16079141 BG13648 yopO 16079161 BG13628 yonR <b>16079416 BG10299 ansR</b> 16079685 BG11258 yqaG 16079687 BG11257 yqaF 16079688 BG11256 yqaE 16080418 BG14159 yvzC 16080419 BG14069 yvaN 16080420 BG14070 yvaO 16080491 BG11858 slr  †16079042 BG13745 yotL
<b>Groups of paralogous TFs only formed by PFAM domains</b>				
	PF03816.4-1	3	1	16080605 BG12449 yvhJ <b>16080618 BG10404 lytR</b> 16080637 BG12537 ywtF
	PF04307.4-1	2	1	<b>16077928 BG12890 yfhP</b> 16080388 BG14153 yvsG
	PF04299.2-1	1	1	<b>16080267 BG10696 paiB</b>
	PF05848.1-1	1	1	<b>16077151 BG10145 ctsR</b>
	PF06018.3-1 PF08222.1-1	1	1	<b>16078680 BG10968 codY</b>
	PF06338.1-1	1	1	<b>16078106 BG11059 comK</b>
	PF07905.1-1	1	1	<b>16080295 BG13983 pucR</b>

**Table S4. Expected and observed co-regulated target genes for all *E. coli* transcription factors with more than 10 expected targets.**  
 Shadowed cells highlight cases in which there are more co-regulation events than theoretically expected.

<b>TF name</b>	<b>exp_TGs</b>	<b>obs_TGs</b>
CRP	284	330
FNR	188	196
IHF	136	179
Fis	116	80
ArcA	103	122
NarL	70	96
H-NS	67	76
Lrp	65	69
FliHDC	55	29
Fur	55	41
ModE	33	46
NtrC	31	13
CpxR	31	10
NarP	28	40
PhoB	25	12
FruR	23	24
PhoP	22	11
PurR	21	13
FliA	21	27
LexA	21	2
GadE	20	19
IscR	18	20
SoxS	18	17
ArgR	18	17
CysB	16	9
MarA	16	15
RcsAB	13	7
NagC	12	15
GadX	12	12
OxyR	11	11



OmpR	11	11
Rob	10	11

**Table S5. Summaries of the roles of global regulators in *E. coli* and *B. subtilis*.**

<i>Escherichia coli</i>		<i>Bacillus subtilis</i>
<b>CRP</b>		<b>CcpA</b>
<b>Function</b>	<b>cAMP Receptor Protein<sup>1;2</sup>:</b> Major global regulator of catabolite repression, genes required for energy production, amino acid and nucleotide metabolism and ion transport systems.	<b>Catabolite Control Protein A<sup>3;4</sup>:</b> Central regulator of carbon catabolite repression in response to intracellular metabolite levels, also genes involved in membrane bioenergetics, transport and stress-response.
<b>Effector</b>	Cyclic AMP <sup>5;6</sup>	Ser46-phosphorylated forms of HPr and of Crh, Fructose-1,6-bisphosphate and Glucose 6-phosphate <sup>7</sup> .
<b>FNR</b>		<b>TnrA</b>
<b>Function</b>	<b>Fumarate and Nitrate Reduction<sup>8;9</sup>:</b> Stimulates genes required for fermentation and anaerobic respiration, repressing some genes that function only during aerobic growth.	<b>Transcriptional Nitrogen Regulator protein A<sup>10;11</sup>:</b> Regulates genes involved in the transport and catabolism of nitrogen containing compounds during a nitrogen-limited growth.
<b>Effector</b>	O <sub>2</sub> through a [4Fe-4S] <sup>2+</sup> redox-sensitive iron-sulphur cluster <sup>12;13</sup>	Glutamine synthetase <sup>14;15</sup> .
<b>LRP</b>		<b>CodY</b>
<b>Function</b>	<b>Leucine-Responsive regulatory Protein<sup>16;17;18</sup>:</b> Controls genes involved in branched-chain amino acid transport, degradation and biosynthesis. LRP is a mediator of the leucine response and regulates genes involved in the production of pili, porins, sugar transporters and nucleotide transhydrogenases.	<b>GTP-sensing transcriptional pleiotropic repressor<sup>19;20;21</sup>:</b> Involved in the transition from exponential growth to stationary phase and early sporulation, controlling genes involved in genetic competence, N and acetate metabolism, nutrient limitation, extracellular enzymes, transport systems, intracellular catabolic pathways, chemotaxis and antibiotic synthesis.
<b>Effector</b>	Leucine <sup>18</sup>	GTP, isoleucine and valine <sup>22;23</sup> .

<b>ArcA</b>		<b>spo0A</b>	
<b>Function</b>	<b>Aerobic Respiration Control compound A</b> <sup>24; 25</sup> : Regulates the aerobic/anaerobic expression of genes involved in respiration and fermentation, also required for conjugative transfer of plasmids.	<b>Stage 0 SPorulation compound A</b> <sup>26; 27</sup> : Spo0A regulates the initiation of sporulation and various other stationary phase processes. It also prevents chromosomal replication.	
<b>Effector</b>	Phosphorylated by ArcB <sup>28</sup> , which in turn senses pyruvate, NADH, lactate and acetate	Phosphorylated by Spo0A <sup>29</sup> , which in turn senses glucose <sup>30</sup> .	
<b>FIS</b>		<b>AbrB</b>	
<b>Function</b>	<b>Factor-for-Inversion Stimulation protein</b> <sup>31; 32</sup> : Stimulates initiation of chromosome replication and the expression of rRNA and tRNA genes, involved in lambda site-specific recombination in vivo, enhancing integrative and excisive recombination. It also couples DNA topology with physiological changes.	<b>AntiBiotic Resistance protein B</b> <sup>33</sup> : Regulates the transition between vegetative growth (biofilm formation) and the onset of stationary phase and sporulation, controlling the production of antibiotics and extracellular enzymes.	
<b>Effector</b>	Superhelical density of DNA <sup>34</sup>	Superhelical density of DNA? <sup>33; 35; 36</sup>	
<b>IHF</b>		<b>ComK</b>	
<b>Function</b>	<b>Integration Host Factor</b> <sup>37; 38</sup> : Introduces sharp bends in the DNA, being required for site-specific recombination, transposition, repair, replication and transcriptional control. IHF is essential in the lysogenic life cycle of bacteriophage lambda.	<b>COMPetence regulatory Protein K</b> <sup>39; 40</sup> : ComK affects homologous recombination and is required for the expression of late competence genes, which are required for binding, fragmentation, and uptake of DNA.	
<b>Effector</b>	Curved DNA <sup>41</sup> .	Possibly DNA fragments <sup>39</sup>	
<b>H-NS</b>			
<b>Function</b>	<b>Heat-stable Nucleoid-Structuring protein</b> <sup>42; 43</sup> : Pleiotropic regulator that controls compaction of chromosomal DNA, directly or indirectly controlling adaptation to environment. Represses horizontally acquired genes under normal growth conditions.		
<b>Effector</b>	Intrinsically curved DNA <sup>44</sup>		

**Table S6. Hydrogen bond scoring matrix.**

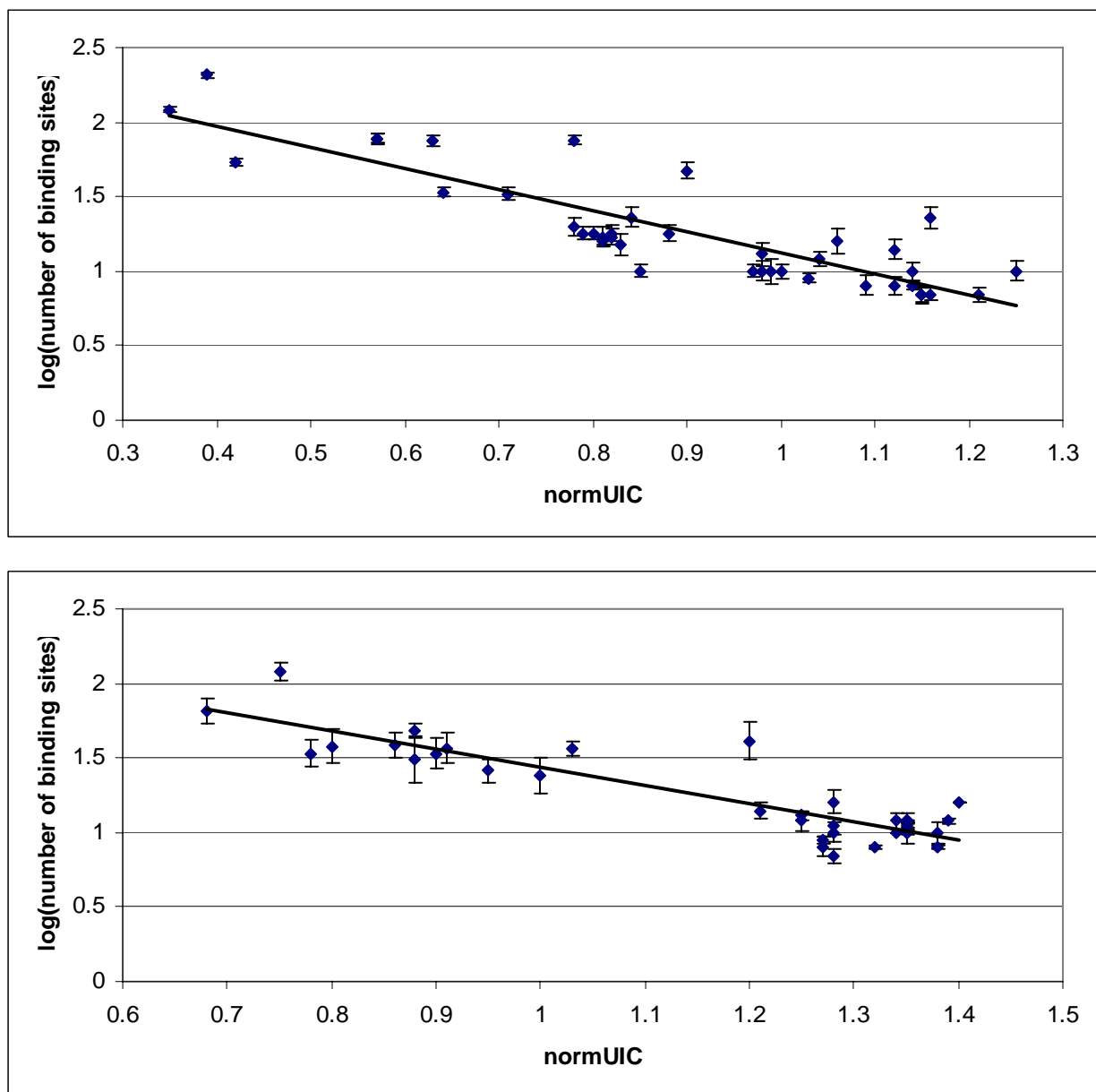
AMINOACIDS		THYMINE				CYTOSINE				ADENINE				GUANINE								
		N1	O2	N3	O4	N1	O2	N3	N4	N9	N7	N6	N1	N3	N9	N7	O6	N1	N2	N3		
ARG	R	NE	-7.343	1.410	-7.343	1.005	-7.343	1.337	-0.859	-7.343	-7.343	-7.343	-7.343	-0.545	-7.343	2.044	2.423	-7.343	-7.343	-0.520	NE	
		NH1	-7.343	2.509	-7.343	0.717	-7.343	1.624	1.337	-7.343	-7.343	0.552	-7.343	-0.545	1.938	-7.343	3.370	3.192	-7.343	-7.343	0.577	NH1
		NH2	-7.343	2.451	-7.343	1.816	-7.343	2.084	0.931	-7.343	-7.343	-7.343	-7.343	-0.545	1.651	-7.343	3.713	3.967	-7.343	-7.343	0.577	NH2
LYS	K	NZ	-6.715	1.843	-6.715	1.207	-6.715	1.096	-1.464	-6.715	-6.715	0.925	-6.715	-1.151	0.925	-6.715	2.720	2.930	-6.715	-6.715	-0.030	NZ
		OG	-6.277	1.125	-0.259	0.433	-0.045	1.339	0.647	0.870	-6.277	0.961	0.961	1.184	1.184	-6.277	1.902	1.679	-6.277	2.164	0.986	OG
ASN	N	OG1	-6.028	0.814	0.122	1.730	-0.356	0.336	-0.356	1.944	-6.028	1.054	1.054	-6.028	1.054	-6.028	0.674	1.590	-0.018	0.674	-6.028	OG1
		OD1	-5.961	-5.961	1.099	-5.961	0.621	-5.961	-5.961	3.104	-5.961	-5.961	4.069	-5.961	-5.961	-5.961	-5.961	-5.961	0.960	3.038	-5.961	OD1
GLN	Q	ND2	-5.961	3.178	-5.961	3.806	-5.961	2.565	-5.961	-5.961	-5.961	3.878	-5.961	1.627	3.013	-5.961	2.345	3.357	-5.961	-5.961	2.750	ND2
		OE1	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	2.756	-5.613	-5.613	4.223	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	4.193	-5.613	OE1
HIS	H	NE2	-5.613	1.444	-5.613	2.542	-5.613	2.910	0.965	-5.613	-5.613	3.763	-5.613	1.972	2.377	-5.613	1.997	3.383	-5.613	-5.613	3.868	NE2
		ND1	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	2.055	-5.375	-5.375	-5.375	2.773	3.466	-5.375	-5.375	-5.375	ND1
TYR	Y	NE2	-5.375	2.913	-5.375	-5.375	-5.375	3.127	2.434	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	2.055	-5.375	3.466	4.477	-5.375	2.080	NE2
		OH	-6.190	0.312	0.312	1.697	-6.190	-0.166	-6.190	-0.166	-6.190	0.148	1.533	-6.190	1.245	-6.190	0.865	0.173	-6.190	1.963	0.173	OH
GLU	E	OE1	-5.136	-5.136	1.290	-5.136	-5.136	-5.136	3.854	-5.136	-5.136	1.126	-5.136	-5.136	-5.136	-5.136	-5.136	2.536	2.941	-5.136	OE1	
		OE2	-5.136	-5.136	-5.136	-5.136	1.504	-5.136	-5.136	3.295	-5.136	-5.136	2.223	-5.136	-5.136	-5.136	-5.136	-5.136	2.941	3.452	-5.136	OE2
ASP	D	OD1	-5.094	-5.094	-5.094	-5.094	-5.094	-5.094	3.606	-5.094	-5.094	2.216	-5.094	-5.094	-5.094	-5.094	-5.094	2.646	3.744	-5.094	OD1	
		OD2	-5.094	-5.094	2.380	-5.094	-5.094	-5.094	-5.094	3.606	-5.094	-5.094	-5.094	-5.094	-5.094	-5.094	-5.094	-5.094	3.850	4.112	-5.094	OD2
CYS	C	SG	-3.434	-3.434	-3.434	4.167	3.689	-3.434	-3.434	-3.434	-3.434	-3.434	4.696	-3.434	-3.434	-	-	-	-	-	-	SG
		SD	-	-	-	-	-	-	-	-	-	-	-4.025	-4.025	-4.025	-4.025	-4.025	-	-	-	-	SD
TRP	W	NE1	-	-	-	-	-	-	-	-	-	-4.820	-4.820	-4.820	-4.820	2.444	-	-	-	-	-	NE1

**Table S7. Van der Waals contacts scoring matrix.**

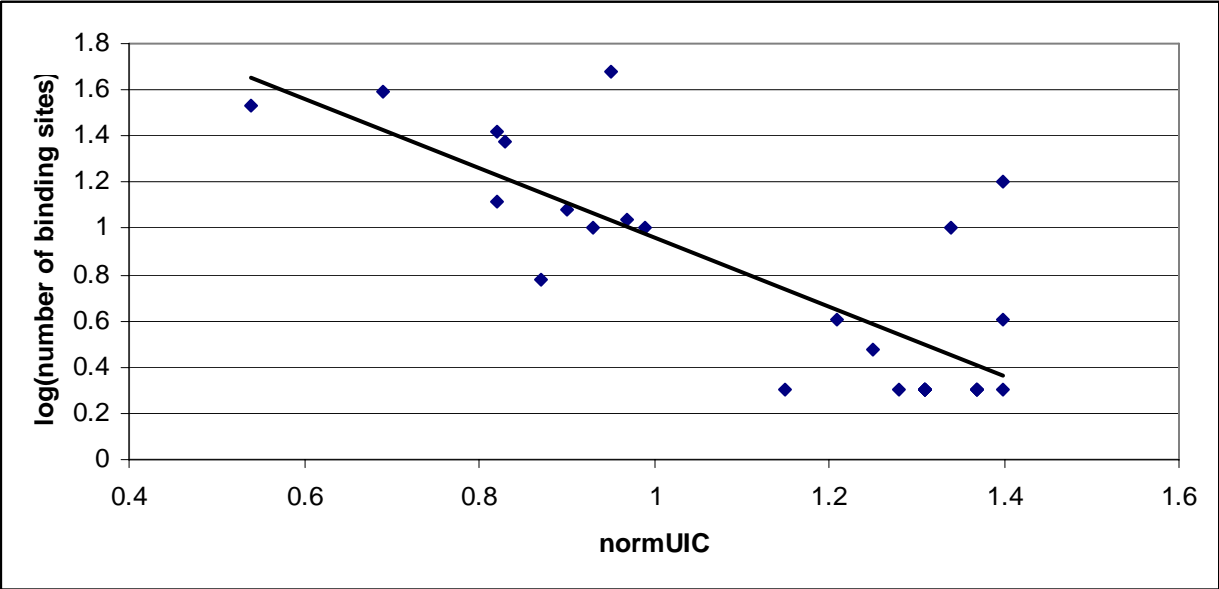
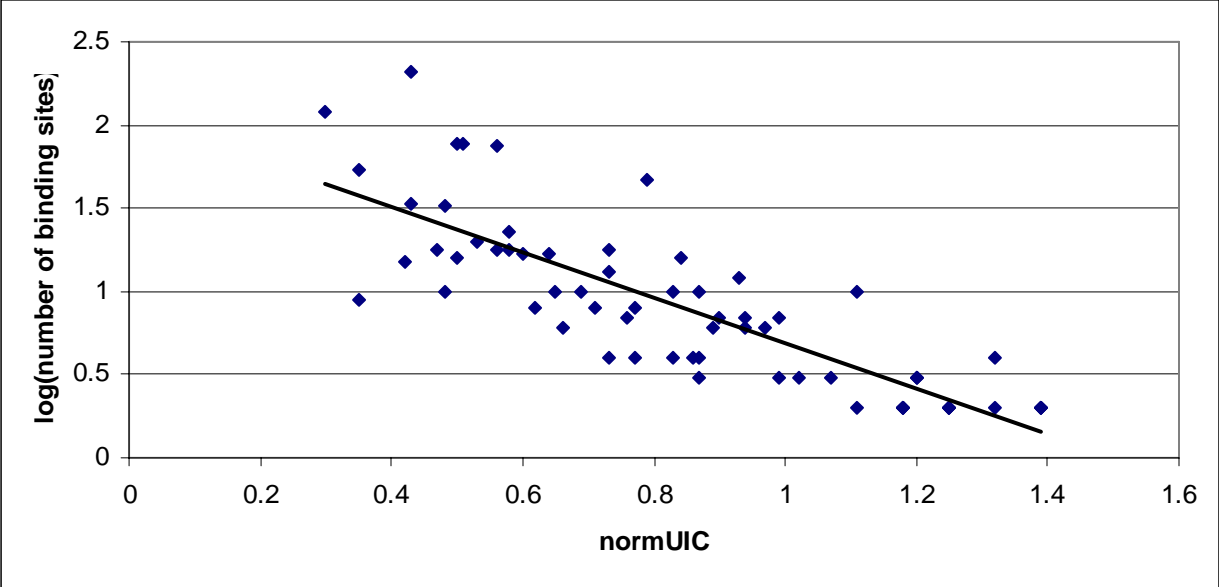
AMINOACIDS		THYMINE					CYTOSINE				ADENINE				GUANINE								
		C2	C4	C5	C5M	C6	C2	C4	C5	C6	C8	C5	C6	C2	C4	C8	C5	C6	C2	C4			
ARG	R	CB	-7.343	-7.343	-7.343	1.787	-7.343	-0.859	-0.167	0.526	-7.343	-0.545	-0.545	-0.545	0.552	-7.343	0.683	-7.343	-0.703	-7.343	CB		
		CG	-7.343	-0.158	-0.158	1.633	-7.343	0.526	-0.167	-7.343	-7.343	0.552	0.147	0.552	1.400	0.147	0.683	0.683	-0.010	-0.703	-7.343	CG	
		CZ	-0.158	0.535	2.614	3.670	-7.343	0.931	1.537	2.723	2.184	2.018	2.018	1.651	2.018	1.651	2.241	1.781	1.494	0.683	1.599	CZ	
LYS	K	CB	-6.715	-6.715	-6.715	1.430	-6.715	-0.369	1.016	0.728	-6.715	-0.055	-0.055	0.637	1.553	-0.055	-6.715	-6.715	-0.212	0.480	-6.715	CB	
		CG	0.332	1.025	0.332	2.123	-6.715	-0.369	-6.715	-0.369	-6.715	0.637	-0.055	0.637	1.042	-6.715	-0.212	1.173	0.885	1.173	0.480	CG	
		CE	-6.715	0.332	0.332	2.411	-6.715	0.728	0.323	1.239	-0.369	1.042	0.637	-6.715	2.023	-6.715	0.480	1.578	2.271	0.480	0.480	CE	
SER	S	CB	0.656	-6.277	-0.036	2.447	-6.277	-0.045	0.359	1.206	0.870	-6.277	-6.277	-0.423	1.366	-0.423	0.804	0.111	0.516	0.516	-0.580	CB	
	T	CG	-6.028	-6.028	-6.028	2.828	-6.028	1.028	0.336	0.336	1.028	-6.028	-6.028	-6.028	0.649	-6.028	0.492	-6.028	-6.028	-6.028	-6.028	CB	
		CG2	1.037	1.730	1.730	3.744	-6.028	1.028	0.336	1.433	1.433	-6.028	-6.028	-6.028	1.747	0.649	2.283	0.492	0.492	0.492	-6.028	CG2	
ASN	N	CB	-5.961	-5.961	-5.961	2.708	-5.961	-5.961	0.621	1.313	0.621	2.032	-5.961	-5.961	2.320	-5.961	1.470	-5.961	1.875	0.777	0.777	CB	
		CG	-5.961	1.322	-5.961	3.401	-5.961	1.313	0.621	2.565	2.006	2.320	-5.961	-5.961	0.935	-5.961	2.568	-5.961	0.777	-5.961	-5.961	CG	
GLN	Q	CB	-5.613	-5.613	-5.613	2.765	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	-5.613	1.122	1.122	1.122	-5.613	CB	
		CG	-5.613	-5.613	-5.613	3.276	-5.613	-5.613	-5.613	-5.613	0.965	-5.613	-5.613	-5.613	1.279	-5.613	1.814	2.220	1.122	1.814	1.814	CG	
HIS	H	CB	-5.375	3.541	-5.375	3.541	-5.375	-5.375	3.245	2.839	-5.375	2.460	3.153	3.559	3.153	3.153	2.996	-5.375	2.303	-5.375	-5.375	CB	
		CG	-5.375	3.541	3.947	4.457	-5.375	-5.375	2.839	2.839	-5.375	3.153	3.846	4.252	3.559	3.559	2.996	3.401	2.303	-5.375	2.996	CG	
		CD2	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	-5.375	CD2	
TYR	Y	CB	2.144	2.144	-6.190	3.242	-6.190	-6.190	1.442	1.442	-6.190	3.142	-6.190	3.142	2.854	-6.190	-6.190	-6.190	1.599	2.697	2.697	-6.190	CB
		CG	3.530	3.935	2.144	3.935	-6.190	-6.190	2.135	2.828	1.442	2.854	3.365	3.142	-6.190	2.854	2.985	2.697	2.697	2.697	1.599	CG	
		CD1	3.935	3.530	3.530	4.223	-6.190	-6.190	1.442	3.051	1.442	1.756	2.854	2.449	2.449	2.854	3.390	3.390	2.291	2.985	3.390	CD1	
		CD2	3.242	3.753	-6.190	3.753	-6.190	2.540	2.540	3.233	-6.190	3.142	3.365	3.365	1.756	2.854	3.390	2.985	3.208	2.291	1.599	CD2	
		CZ	3.753	3.935	4.089	4.446	-6.190	2.135	2.135	3.744	3.638	3.835	1.756	1.756	1.756	-6.190	4.083	3.208	2.291	1.599	3.544	CZ	
ALA	A	CB	1.705	2.398	2.803	3.844	-4.977	1.004	-4.977	0.314	1.004	2.010	0.627	0.627	1.723	2.010	1.161	0.470	0.470	1.853	0.470	CB	
	E	CG	-5.136	-5.136	-5.136	2.611	-5.136	-5.136	0.812	-5.136	-5.136	1.818	-5.136	-5.136	-5.136	-5.136	-5.136	-5.136	-5.136	-5.136	-5.136	CB	
		CG	-5.136	-5.136	-5.136	2.611	-5.136	-5.136	1.504	-5.136	-5.136	1.126	-5.136	-5.136	-5.136	-5.136	-5.136	0.969	-5.136	-5.136	-5.136	CG	
ILE	I	CB	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	2.461	2.461	-4.407	-4.407	2.775	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	CB	
		CG1	-4.407	-4.407	-4.407	3.163	-4.407	3.559	2.461	-4.407	2.461	-4.407	-4.407	2.775	3.873	-4.407	2.618	3.311	3.311	2.618	3.311	CG1	
		CG2	3.163	-4.407	-4.407	3.856	-4.407	2.461	2.461	3.154	2.461	4.161	2.775	-4.407	3.468	3.468	3.311	2.618	-4.407	2.618	2.618	CG2	
		CD1	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	-4.407	CD1
ASP	D	CB	-5.094	1.911	-5.094	1.911	-5.094	-5.094	-5.094	2.595	1.210	1.523	-5.094	2.621	1.523	-5.094	-5.094	-5.094	-5.094	-5.094	1.366	CB	
		CG	-5.094	-5.094	-5.094	3.296	-5.094	1.210	1.902	3.511	2.307	-5.094	-5.094	-5.094	2.216	-5.094	-5.094	-5.094	1.366	2.751	-5.094	CG	
VAL	V	CB	-4.615	-4.615	-4.615	2.572	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	-4.615	2.027	-4.615	CB	
		CG1	2.572	-4.615	-4.615	4.181	-4.615	1.870	2.563	1.870	-4.615	2.877	3.282	2.877	3.975	2.184	2.720	-4.615	2.720	3.125	-4.615	CG1	
		CG2	-4.615	3.265	-4.615	4.768	-4.615	1.870	-4.615	1.870	-4.615	-4.615	-4.615	2.184	4.129	-4.615	2.027	2.027	2.027	2.027	2.027	CG2	
PHE	F	CB	-3.434	-3.434	-3.434	4.391	-3.434	-3.434	-3.434	3.689	3.689	4.003	-3.434	-3.434	-3.434	-3.434	-	-	-	-	-	CB	
		CG	3.481	3.481	-5.421	-5.421	-5.421	-5.421	2.086	3.185	2.779	2.400	2.400	3.499	4.010	-5.421	-5.421	2.243	2.243	2.243	-5.421	CB	
		CG1	3.481	3.481	2.788	3.481	-5.421	-5.421	3.472	3.472	2.779	2.400	3.786	2.400	3.093	3.093	2.243	-5.421	2.243	2.243	3.341	CG1	
		CD1	4.397	3.887	3.481	3.887	-5.421	-5.421	3.185	2.086	2.086	3.499	3.786	3.093	3.093	3.499	2.936	2.936	3.629	3.341	-5.421	CD1	
		CD2	2.788	3.481	2.788	3.887	-5.421	-5.421	3.185	3.472	3.185	3.499	4.192	3.499	3.499	3.786	2.243	2.243	2.243	3.341	2.243	CD2	
	CZ	4.734	3.481	3.887	3.481	-5.421	4.032	3.472	3.472	2.779	4.010	3.499	-5.421	4.346	3.786	2.243	2.936	2.936	3.629	3.341	CZ		
LEU	L	CB	-5.069	-5.069	-5.069	-5.069	-5.069	1.328	-5.069	-5.069	1.328	1.641	1.641	-5.069	-5.069	1.641	1.484	-5.069	-5.069	-5.069	1.484	CB	
		CG	-5.069	2.029	-5.069	-5.069	-5.069	-5.069	1.328	-5.069	-5.069	-5.069	2.334	2.334	-5.069	-5.069	-5.069	1.484	-5.069	-5.069	1.484	CG	
		CD1	2.722	2.722	2.029	2.722	-5.069	2.936	2.020	2.020	2.713	3.027	3.432	3.250	3.838	2.739	2.177	1.484	2.177	1.484	1.484	CD1	
		CD2	-5.069	-5.069	2.029	3.127	-5.069	2.425	1.328	2.020	1.328	2.739	1.641	2.739	3.250	1.641	3.275	2.870	2.582	1.484	2.870	CD2	
MET	M	CB	-4.025	-4.025	-4.025	-4.025	-4.025	3.636	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	3.950	-4.025	3.793	-4.025	-4.025	-4.025	-4.025	CB	
		CG	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	-4.025	4.643	-4.025	4.486	-4.025	-4.025	3.793	CG	
		CE	-4.025	-4.025	4.338	4.338	-4.025	-4.025	-4.025	3.636	-4.025	3.950	4.643	4.643	6.030	4.643	3.793	3.793	-4.025	-4.025	3.793	CE	
TRP	W	CB	-4.820	4.217	4.217	4.217	-4.820	-4.820	3.515	4.901	-4.820	-4.820	-4.820	3.830	3.830	-4.820	-4.820	3.672	4.365	3.672	-4.820	CB	
		CG	4.217	4.910	4.910	4.217	-4.820	-4.820	3.515	4.901	3.515	-4.820	-4.820	-4.820	3.830	3.830	-4.820	3.672	3.672	4.365	-4.820	CG	
		CD1	4.217	4.910	4.217	4.217	-4.820	-4.820	4.208	3.515	4.208	-4.820	-4.820	-4.820	3.830	3.830	-4.820	-4.820	-4.820	3.672	-4.820	CD1	
		CD2	-4.820	5.316	5.604	4.910	-4.820	-4.820	4.208	5.307	4.614	3.830	3.830	-4.820	3.830	3.830	-4.820	-4.820	4.365	4.365	5.058	4.365	CD2
		CG	3.575	3.575	2.882	3.980	-4.304	2.181	-4.304	-4.304													

## Supplementary figures

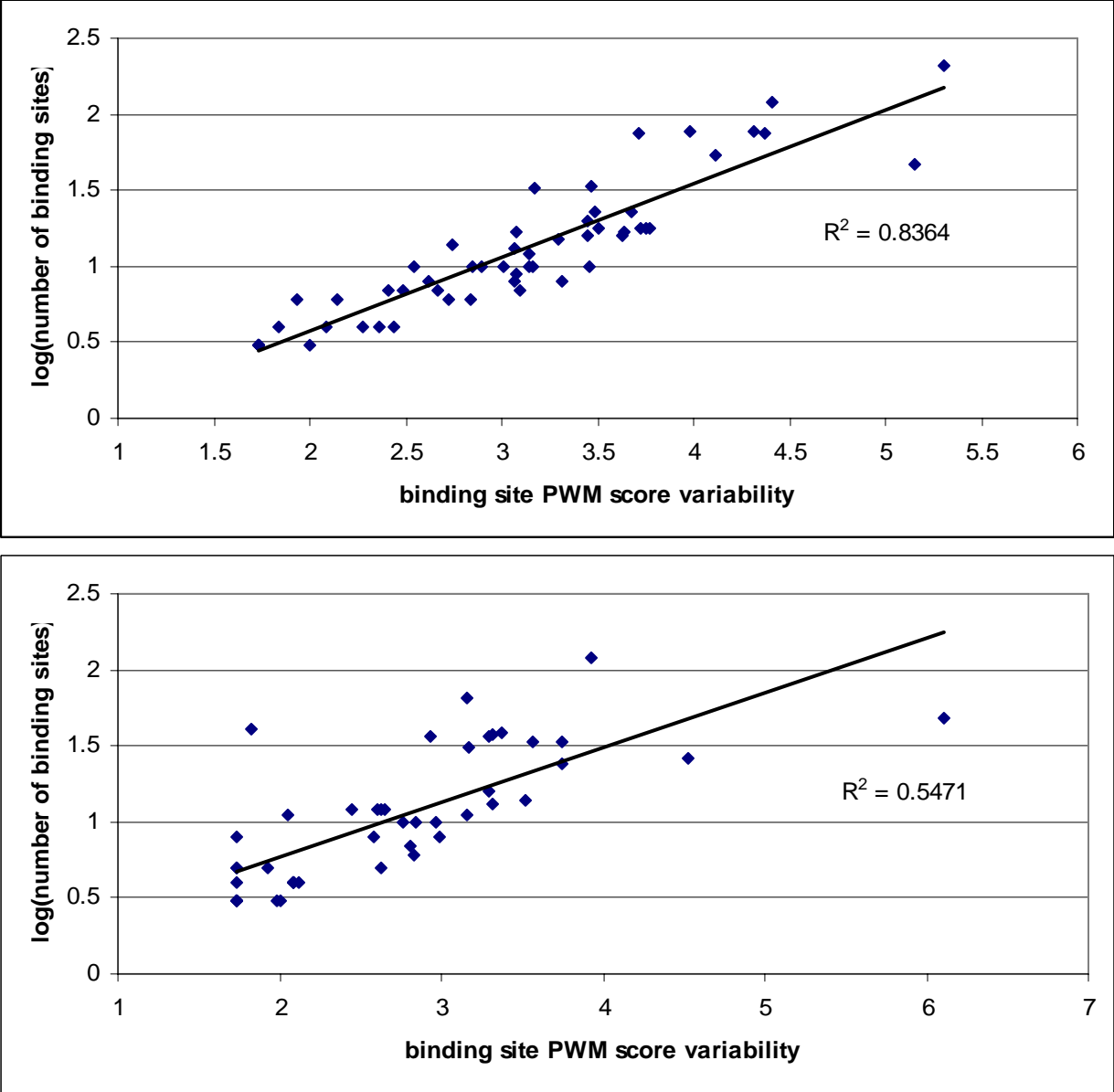
**Figure S1. Scatter plot of mean normalized information content of 100 random samples versus number of binding sites in *E. coli* (top) and *B. subtilis* (bottom).** Standard deviation bars are depicted around each mean IC value. A linear fit is also plotted to illustrate the observed correlation coefficients of -0.86 and -0.89. Samples included 30% of the available sites for each TF.



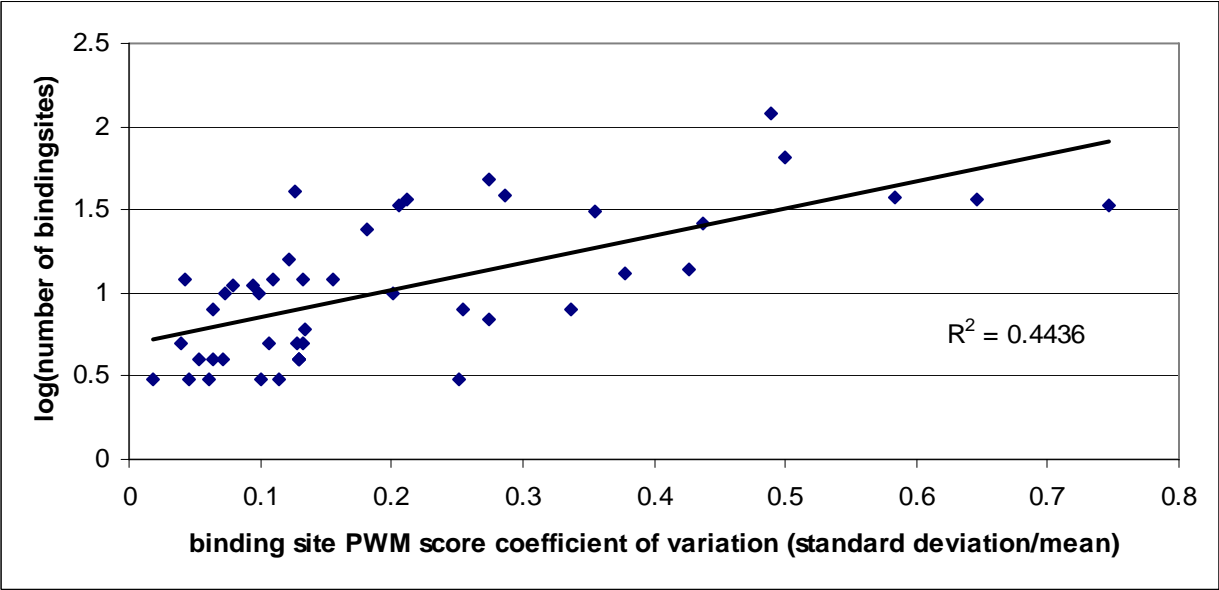
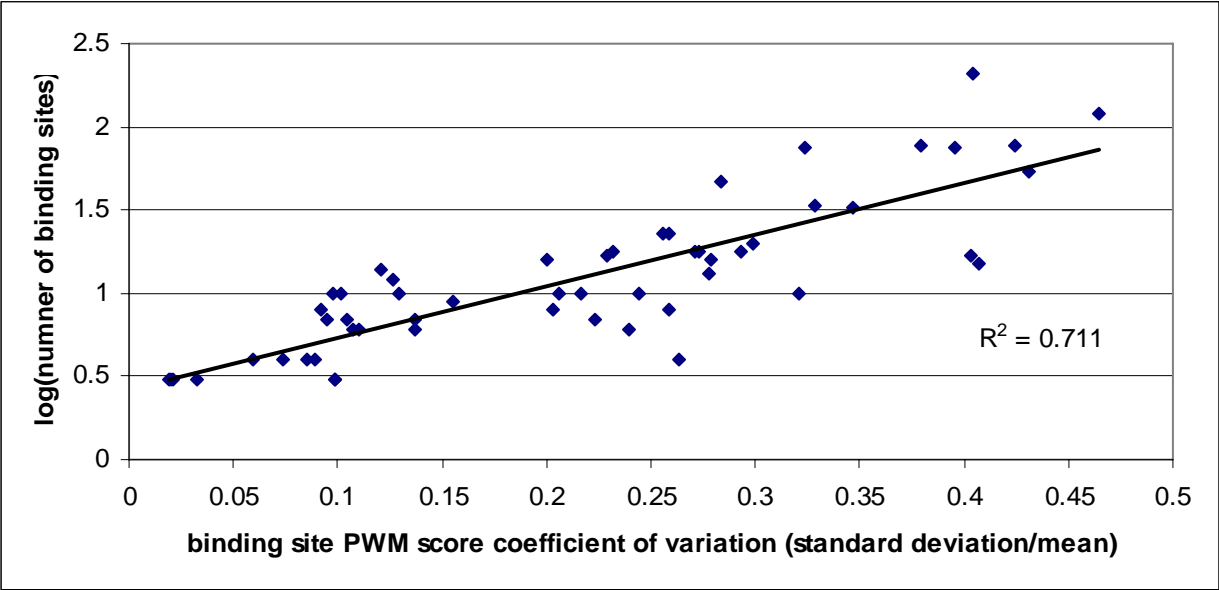
**Figure S2. Scatter plot of normalized information content *versus* number of binding sites in *E. coli* (top) and *B. subtilis* (bottom).** Information content was calculated using the best 10 columns of the motif in order to show that the observed specificity trend is not dependent on the geometry of the site.



**Figure S3. Scatter plot of PWM score variability versus number of binding sites in *E. coli* (top) and *B. subtilis* (bottom).** Here we re-use the set of position-weight matrices (PWM) used in Figure 2 of the manuscript. We took only matrices derived from at least 2 sites and calculated the score of matching each site to the PWM and then applied Equation 1 to calculate the score variability for each transcription factor (see Materials and Methods).

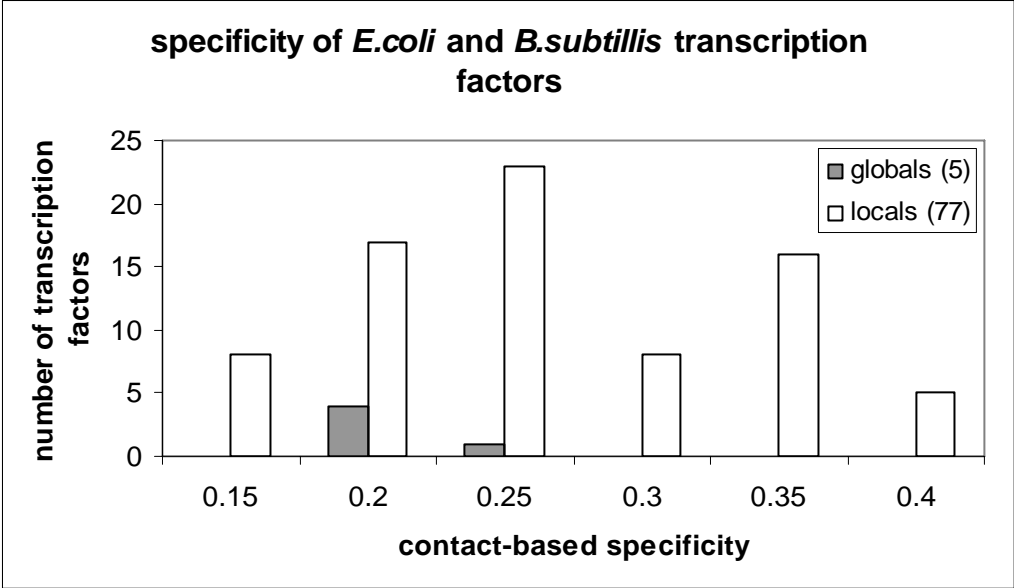


**Figure S4. Scatter plot of PWM score coefficient of variability (standard deviation / mean) versus number of binding sites in *E. coli* (top) and *B. subtilis* (bottom).** Here we re-use the set of position-weight matrices (PWM) used in Figure 2 of the manuscript. We took only matrices derived from at least 2 sites, with standard deviations > 0, and calculated the score of matching each site to the PWM and then applied Equation 1 to calculate the score variability for each transcription factor (see Materials and Methods).

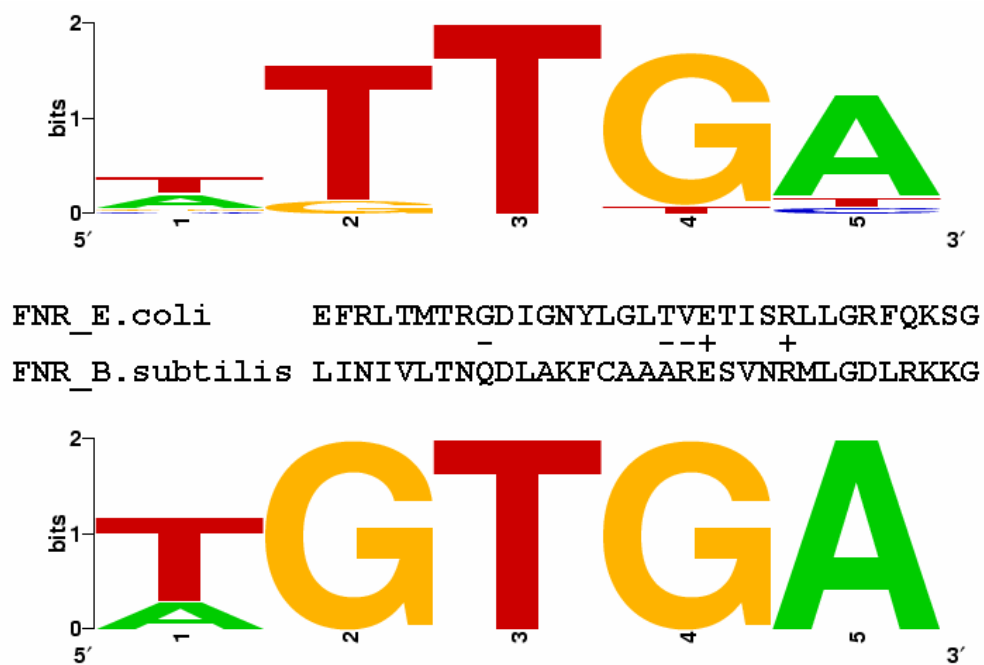




**Figure S5. Histogram transcription factor specificities based on contact-based estimates.** Data includes 49 TFs of *E. coli* and 33 TFs from *B. subtilis*, of which 5 are global regulators. Note that local transcription factors show a wide range of specificities, whilst global TFs are clearly among the lowest specific DNA binders, with values in the range 0.17-0.21. Check Supplementary Methods M1 for details on the calculations.



**Figure S6. Sequence alignment of DNA-binding regions of FNR orthologs in *E. coli* (top) and *B. subtilis* (bottom).** Sequence logos describing the recognized binding sites are shown in the same order. Residues predicted by TFmodeller<sup>45</sup> to contact DNA bases are labeled with + if conserved, otherwise they are marked with -. Sequence logos generated with WEBLOGO<sup>46</sup>.



1. Perrenoud, A. & Sauer, U. (2005). Impact of global transcriptional regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on glucose catabolism in *Escherichia coli*. *J Bacteriol* 187, 3171-9.
2. Zheng, D., Constantinidou, C., Hobman, J. L. & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* 32, 5874-93.
3. Lorca, G. L., Chung, Y. J., Barabote, R. D., Weyler, W., Schilling, C. H. & Saier, M. H., Jr. (2005). Catabolite repression and activation in *Bacillus subtilis*: dependency on CcpA, HPr, and HprK. *J Bacteriol* 187, 7826-39.
4. Moreno, M. S., Schneider, B. L., Maile, R. R., Weyler, W. & Saier, M. H., Jr. (2001). Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol* 39, 1366-81.
5. Saier, M. H., Jr. (1989). Protein phosphorylation and allosteric control of inducer exclusion and catabolite repression by the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Microbiol Rev* 53, 109-20.
6. Zubay, G., Schwartz, D. & Beckwith, J. (1970). Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc Natl Acad Sci U S A* 66, 104-10.
7. Lopez, J. M. & Thoms, B. (1977). Role of sugar uptake and metabolic intermediates on catabolite repression in *Bacillus subtilis*. *J Bacteriol* 129, 217-24.
8. Becker, S., Holighaus, G., Gabrielczyk, T. & Unden, G. (1996). O<sub>2</sub> as the regulatory signal for FNR-dependent gene regulation in *Escherichia coli*. *J Bacteriol* 178, 4515-21.
9. Sawers, G. (1999). The aerobic/anaerobic interface. *Curr Opin Microbiol* 2, 181-7.
10. Wray, L. V., Jr., Ferson, A. E., Rohrer, K. & Fisher, S. H. (1996). TnrA, a transcription factor required for global nitrogen regulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 93, 8841-5.
11. Wray, L. V., Jr., Zalieckas, J. M. & Fisher, S. H. (2000). Purification and in vitro activities of the *Bacillus subtilis* TnrA transcription factor. *J Mol Biol* 300, 29-40.
12. Jordan, P. A., Thomson, A. J., Ralph, E. T., Guest, J. R. & Green, J. (1997). FNR is a direct oxygen sensor having a biphasic response curve. *FEBS Lett* 416, 349-52.
13. Khoroshilova, N., Popescu, C., Munck, E., Beinert, H. & Kiley, P. J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O<sub>2</sub>: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci U S A* 94, 6087-92.
14. Fisher, S. H., Brandenburg, J. L. & Wray, L. V., Jr. (2002). Mutations in *Bacillus subtilis* glutamine synthetase that block its interaction with transcription factor TnrA. *Mol Microbiol* 45, 627-35.
15. Wray, L. V., Jr., Zalieckas, J. M. & Fisher, S. H. (2001). *Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell* 107, 427-35.
16. Brinkman, A. B., Ettema, T. J., de Vos, W. M. & van der Oost, J. (2003). The Lrp family of transcriptional regulators. *Mol Microbiol* 48, 287-94.
17. Newman, E. B. & Lin, R. (1995). Leucine-responsive regulatory protein: a global regulator of gene expression in *E. coli*. *Annu Rev Microbiol* 49, 747-75.

18. Willins, D. A., Ryan, C. W., Platko, J. V. & Calvo, J. M. (1991). Characterization of Lrp, and Escherichia coli regulatory protein that mediates a global response to leucine. *J Biol Chem* 266, 10768-74.
19. Levnikov, V. M., Blagova, E., Joseph, P., Sonenshein, A. L. & Wilkinson, A. J. (2006). The structure of CodY, a GTP- and isoleucine-responsive regulator of stationary phase and virulence in gram-positive bacteria. *J Biol Chem* 281, 11366-73.
20. Molle, V., Nakaura, Y., Shivers, R. P., Yamaguchi, H., Losick, R., Fujita, Y. & Sonenshein, A. L. (2003). Additional targets of the Bacillus subtilis global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. *J Bacteriol* 185, 1911-22.
21. Sonenshein, A. L. (2005). CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. *Curr Opin Microbiol* 8, 203-7.
22. Ratnayake-Lecamwasam, M., Serror, P., Wong, K. W. & Sonenshein, A. L. (2001). Bacillus subtilis CodY represses early-stationary-phase genes by sensing GTP levels. *Genes Dev* 15, 1093-103.
23. Shivers, R. P. & Sonenshein, A. L. (2004). Activation of the Bacillus subtilis global regulator CodY by direct interaction with branched-chain amino acids. *Mol Microbiol* 53, 599-611.
24. Alexeeva, S., Hellingwerf, K. J. & Teixeira de Mattos, M. J. (2003). Requirement of ArcA for redox regulation in Escherichia coli under microaerobic but not anaerobic or aerobic conditions. *J Bacteriol* 185, 204-9.
25. Colloms, S. D., Alen, C. & Sherratt, D. J. (1998). The ArcA/ArcB two-component regulatory system of Escherichia coli is essential for Xer site-specific recombination at psi. *Mol Microbiol* 28, 521-30.
26. Castilla-Llorente, V., Munoz-Espin, D., Villar, L., Salas, M. & Meijer, W. J. (2006). Spo0A, the key transcriptional regulator for entrance into sporulation, is an inhibitor of DNA replication. *Embo J* 25, 3890-9.
27. Fujita, M., Gonzalez-Pastor, J. E. & Losick, R. (2005). High- and low-threshold genes in the Spo0A regulon of Bacillus subtilis. *J Bacteriol* 187, 1357-68.
28. Georgellis, D., Lynch, A. S. & Lin, E. C. (1997). In vitro phosphorylation study of the arc two-component signal transduction system of Escherichia coli. *J Bacteriol* 179, 5429-35.
29. Molle, V., Fujita, M., Jensen, S. T., Eichenberger, P., Gonzalez-Pastor, J. E., Liu, J. S. & Losick, R. (2003). The Spo0A regulon of Bacillus subtilis. *Mol Microbiol* 50, 1683-701.
30. Yamashita, S., Kawamura, F., Yoshikawa, H., Takahashi, H., Kobayashi, Y. & Saito, H. (1989). Dissection of the expression signals of the spoA gene of Bacillus subtilis: glucose represses sporulation-specific expression. *J Gen Microbiol* 135, 1335-45.
31. Cheng, Y. S., Yang, W. Z., Johnson, R. C. & Yuan, H. S. (2000). Structural analysis of the transcriptional activation on Fis: crystal structures of six Fis mutants with different activation properties. *J Mol Biol* 302, 1139-51.
32. Choe, L. H., Chen, W. & Lee, K. H. (1999). Proteome analysis of factor for inversion stimulation (Fis) overproduction in Escherichia coli. *Electrophoresis* 20, 798-805.
33. O'Reilly, M. & Devine, K. M. (1997). Expression of AbrB, a transition state regulator from Bacillus subtilis, is growth phase dependent in a manner resembling

- that of Fis, the nucleoid binding protein from *Escherichia coli*. *J Bacteriol* 179, 522-9.
34. Travers, A., Schneider, R. & Muskhelishvili, G. (2001). DNA supercoiling and transcription in *Escherichia coli*: The FIS connection. *Biochimie* 83, 213-7.
  35. Bobay, B. G., Mueller, G. A., Thompson, R. J., Murzin, A. G., Venters, R. A., Strauch, M. A. & Cavanagh, J. (2006). NMR structure of AbhN and comparison with AbrBN: FIRST insights into the DNA binding promiscuity and specificity of AbrB-like transition state regulator proteins. *J Biol Chem* 281, 21399-409.
  36. Strauch, M. A. (1995). AbrB modulates expression and catabolite repression of a *Bacillus subtilis* ribose transport operon. *J Bacteriol* 177, 6727-31.
  37. Friedman, D. I. (1988). Integration host factor: a protein for all reasons. *Cell* 55, 545-54.
  38. Landy, A. (1989). Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem* 58, 913-49.
  39. Hamoen, L. W., Van Werkhoven, A. F., Bijlsma, J. J., Dubnau, D. & Venema, G. (1998). The competence transcription factor of *Bacillus subtilis* recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes Dev* 12, 1539-50.
  40. Susanna, K. A., den Hengst, C. D., Hamoen, L. W. & Kuipers, O. P. (2006). Expression of transcription activator ComK of *Bacillus subtilis* in the heterologous host *Lactococcus lactis* leads to a genome-wide repression pattern: a case study of horizontal gene transfer. *Appl Environ Microbiol* 72, 404-11.
  41. Shimizu, M., Miyake, M., Kanke, F., Matsumoto, U. & Shindo, H. (1995). Characterization of the binding of HU and IHF, homologous histone-like proteins of *Escherichia coli*, to curved and uncurved DNA. *Biochim Biophys Acta* 1264, 330-6.
  42. Oshima, T., Ishikawa, S., Kurokawa, K., Aiba, H. & Ogasawara, N. (2006). *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res* 13, 141-53.
  43. Zuber, F., Kotlarz, D., Rimsky, S. & Buc, H. (1994). Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol Microbiol* 12, 231-40.
  44. Owen-Hughes, T. A., Pavitt, G. D., Santos, D. S., Sidebotham, J. M., Hulton, C. S., Hinton, J. C. & Higgins, C. F. (1992). The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* 71, 255-65.
  45. Contreras-Moreira, B., Branger, P. A. & Collado-Vides, J. (2007). TFmodeller: comparative modelling of protein-DNA complexes. *Bioinformatics*.
  46. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res* 14, 1188-90.