Published online 12 December 2006 *Nucleic Acids Research, 2007, Vol. 35, No. 1* ***269–278***
doi:10.1093/nar/gkl1023

Transcription factor distribution in *Escherichia coli*: studies with FNR protein

David C. Grainger*, Hirofumi Aiba[1], Douglas Hurd[2], Douglas F. Browning and Stephen J. W. Busby

School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK, [1]Laboratory of Molecular Microbiology, School of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan and [2]Oxford Gene Technology, Begbroke Science Park, Sandy Lane, Yarnton, Oxford OX5 1PF, UK

Received August 25, 2006; Revised November 8, 2006; Accepted November 14, 2006

ABSTRACT

Using chromatin immunoprecipitation (ChIP) and high-density microarrays, we have measured the distribution of the global transcription regulator protein, FNR, across the entire *Escherichia coli* chromosome in exponentially growing cells. Sixty-three binding targets, each located at the 5′ end of a gene, were identified. Some targets are adjacent to poorly transcribed genes where FNR has little impact on transcription. In stationary phase, the distribution of FNR was largely unchanged. Control experiments showed that, like FNR, the distribution of the nucleoid-associated protein, IHF, is little altered when cells enter stationary phase, whilst RNA polymerase undergoes a complete redistribution.

INTRODUCTION

An army of more than 250 transcription factors controls gene expression in *Escherichia coli*. Some of these factors are operon-specific while others, known as global regulators, coordinate the expression of scores of promoters in response to specific environmental cues [reviewed in (1–3)]. The advent of whole-genome DNA sequencing, and associated advances in DNA microarray technology, has enabled investigation of the battery of genes regulated by each of these global factors. The *E.coli* FNR protein (regulator of fumarate and nitrate reduction) is the global transcription factor that manages the distribution of RNA polymerase in response to oxygen starvation. FNR senses oxygen via an N-terminal iron–sulfur cluster. Hence, in anaerobic conditions, FNR is able to bind to specific DNA targets at promoters and modulate transcription. In aerobic conditions, FNR is converted to a form, unable to bind these targets [reviewed in (4,5)]. Bioinformatic analysis has been used to search the *E.coli* genome for DNA sequences that resemble known FNR binding sites (6,7) and DNA microarrays have been used to study differences in the transcriptome that arise when the *fnr* gene is deleted from the genome (8–10). These studies illustrate the complexity of the FNR regulon and predict that, while FNR directly regulates ~100 transcription units, it indirectly affects up to 1000 genes. In this study, we used chromatin immunoprecipitation (ChIP), in conjunction with high-density microarrays (ChIP-chip), to measure the binding of FNR across the *E.coli* chromosome directly, and *in vivo*, for the first time. This allowed us to identify 63 DNA targets for FNR, some of which are adjacent to poorly expressed genes where FNR has minor regulatory effects.

In the second part of the study, we studied the distribution of FNR as growing *E.coli* cells enter stationary phase, and found that it is largely unchanged. Recall that transcription patterns change dramatically when cells cease to grow (11,12) but little is known about the distribution and binding of transcription factors in stationary phase cells. In control experiments, we showed that the binding pattern of IHF, a nucleoid-associated protein, is also unchanged, whereas the distribution of RNA polymerase is radically altered.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

Bacterial strains used in this work are described in Table 1, together with the oligos used to generate different promoter fragments. For ChIP-chip experiments with FNR, MG1655 and JCB1011 cells were grown anaerobically in Luria–Bertani (LB) medium supplemented with 0.4% glucose. Supplementary Figure 1A shows growth of MG1655 and JCB1011 under these conditions, and the time points at which cells were harvested for ChIP-chip experiments. For ChIP-chip experiments with IHF and RNA polymerase, MG1655 cells were grown aerobically in M9 minimal medium to stationary phase (Supplementary Figure 1B). To compare the activity of different promoters:*lacZ* fusions in the presence and absence of FNR, we used *E.coli* JCB387 and the *fnr* derivative JRG1728.

*To whom correspondence should be addressed. Tel: +44 121 414 5435; Fax: +44 121 414 7366; Email: d.grainger@bham.ac.uk

Table 1. Bacterial strains, plasmids and oligonucleotides

Name	Description	Reference
(A) Strains		
MG1655	F- lambda- *ilvG- rfb*-50 *rph*-1	(22)
JCB1011	MG1655 encoding *fnr*3xFLAG	(10)
JCB387	Prototrophic F^-, Δ*nirB-cysG*, *lac*, *chl*$^+$	(23)
JRG1728	*lacX74*, *galK*, *galU*, *rpsL*, Δ(*ara-leu*), Δ(*tyr fnr trg*)	(24)
(B) Plasmids		
pRW50	Low copy *lac* expression vector	(25)
(C) Oligonucleotides		
nohA/B upstream	5′-ggctgc*gaattc*cggcaggcttcaatgacccag-3′	This work
nohA/B downstream	5′-cgcccg*aagctt*cattgttcattccacggccaaaaac-3′	This work
yccF upstream	5′-ggctgc*gaattc*cagacgcaccactttttcgtccgg-3′	This work
yccF downstream	5′-cgcccg*aagctt*cataaaacctcgctttactgtg-3′	This work
helD upstream	5′-ggctgc*gaattc*cgccaacagccagcccagagtg-3′	This work
helD downstream	5′-cgcccg*aagctt*ttccagctcgccccataccag-3′	This work
dbpA upstream	5′-ggctgc*gaattc*acctttctgaccgggttatttg-3′	This work
dbpA downstream	5′-cgcccg*aagctt*cacaatctattctcgtggtcatcg-3′	This work
mrr upstream	5′-cgcccttcag*gaattc*aaaattgg-3′	This work
mrr downstream	5′-cgcccg*aagctt*catagtacatccttgcagaatcag-3′	This work
hsdR upstream	5′-ggctgc*gaattc*cctggctgtcatccagtcctaatg-3′	This work
hsdR downstream	5′-cgcccg*aagctt*attggatttattcatcattgttattaatc-3′	This work

The table lists bacterial strains, plasmids and oligonucleotides used in this work. Segments of DNA amplified in PCR using the oligonucleotides listed in section C are illustrated in supplementary Figure 2.

ChIP

Bacterial cells were treated with formaldehyde, harvested, lysed and their nucleoprotein was extracted as described by Grainger *et al*. (13). Immunoprecipitation was then performed using mouse monoclonal antibodies against the 3× FLAG epitope (Sigma) or the β-subunit of RNA polymerase (Neoclone, Madison, USA) or rabbit polyclonal antibodies against IHF (donated by Steve Goodman). Note that the antibody against IHF is polyclonal and recognizes different IHF heterodimers. Immunoprecipitated DNA samples and total cell nucleoprotein samples were purified and labelled with Cy5 or Cy3 respectively, without amplification, as described by Grainger *et al*. (13).

Microarray analysis of immunoprecipitated DNA

Microarrays (Oxford Gene Technology) were designed and produced specifically to analyse DNA obtained from ChIP experiments with *E.coli* MG1655 and its derivatives (13). Labelled DNA obtained from immunoprecipitations was hybridized to the microarray as described previously (13). Arrays were then scanned and probes with low Cy5 and Cy3 values and isolated probes with a high Cy5/Cy3 intensity ratio were removed. Data shown are the average of two independent experiments and are presented as Supplementary Tables 1–3. Replicate datasets had a correlation co-efficient between 0.6 and 0.8.

ChIP-chip data analysis

The average Cy5/Cy3 intensity ratio calculated for each microarray spot was plotted against the corresponding position on the *E.coli* MG1655 chromosome, creating a profile of FNR binding (see Supplementary Table 1). We then searched the profile for 'peaks', formed by two or more consecutive probes, with a Cy5/Cy3 ratio clearly distinguishable from the background signal. A cut-off, corresponding to the lowest Cy5/Cy3 ratio observed for a peak at a known FNR target was set, and all probes that had an intensity ratio greater than this value were selected as FNR targets. When several adjacent probes (i.e. probes forming one peak) passed the cut-off, the target position was defined as the centre of the probe with the highest Cy5/Cy3 ratio.

To identify overlapping peaks for ChIP-chip datasets obtained using nucleoprotein from cells growing in different conditions, we aligned the averaged Cy5/Cy3 signals obtained for each condition and applied an equivalent cut-off to both datasets. We then counted the number of probes that passed the cut-off for both datasets (a leeway of one probe in either direction was allowed). To generate the data in Figure 5C, we selected the top 50, 100, 200, 400, 800, 1600, 3200 and 6400 probes from each dataset and determined the distribution of these probes between coding and non-coding DNA.

Construction and assay of promoter::*lacZ* fusions

Intergenic regions containing putative FNR targets were amplified by PCR from MG1655 genomic DNA using primers listed in Table 1. Fragments were digested with EcoRI and HindIII and cloned into pRW50, a low-copy-number *lac* expression vector, to generate promoter*::lacZ* fusions. β-galactosidase levels in JCB387 and JRG178 cells carrying these recombinants were measured by the Miller method (14). Activities shown are the average of three independent experiments, and error bars show one SD on either side of the mean. Cells were grown anaerobically in LB medium supplemented with 0.4% glucose. Assays were performed in triplicate.

Electrophoretic mobility shift assays (EMSA)

EMSA were carried out as detailed by Browning *et al*. (15). Purified promoter fragments were end-labelled with [γ-^{32}P]ATP and ∼0.5 ng of each fragment was incubated with varying amounts of purified FNR D154A, which allows binding in the presence of oxygen. The reaction buffer

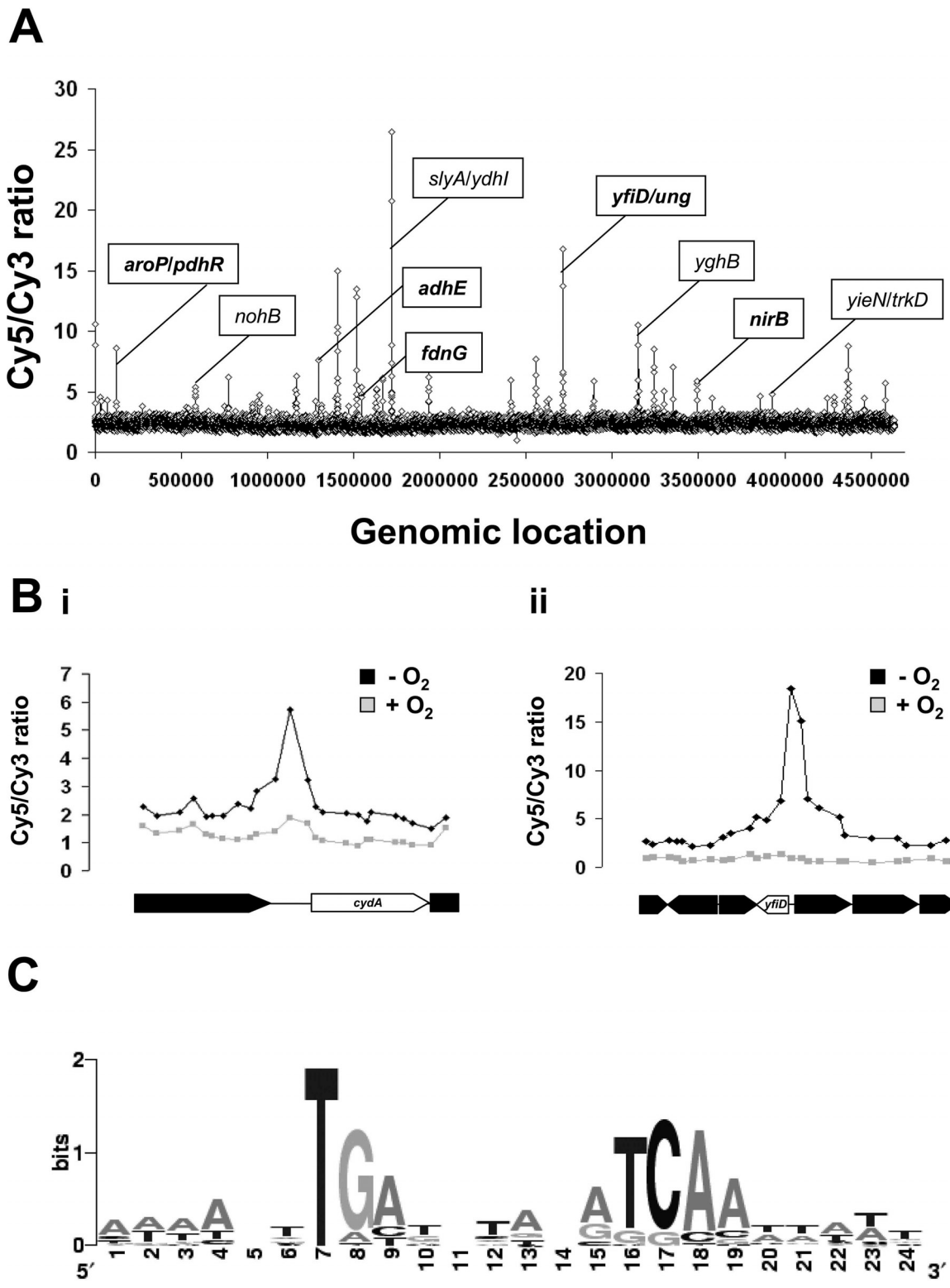

Figure 1. Distribution of FNR binding across the *E.coli* chromosome. (**A**) The figure shows an overview of results from ChIP-chip experiments that measure the profile of FNR binding across the *E.coli* chromosome during exponential growth in anaerobic conditions. Binding signals (*y*-axis) are plotted against their location on the 4.64 Mb *E.coli* chromosome (*x*-axis). The locations of selected signals are labelled in plain typeface (newly identified FNR targets) or in bold face (known FNR targets). A complete list of FNR targets identified is presented in Table 2. (**B**) The figures show expansion of selected regulatory regions, quantifying FNR binding during growth in anaerobic conditions (black) of aerobic conditions (grey). (**C**) The figure shows a DNA sequence motif present at newly identified FNR targets. The DNA sequences from each of the 43 novel FNR targets (see Table 2) were combined and analysed using AlignACE (http://atlas.med.harvard.edu). The motifs identified were then aligned to create a sequence logo (http://weblogo.berkeley.edu). Individual motifs are shown in Table 2.

Table 2. FNR targets identified by ChIP-chip analysis

Peak centre	Gene	Identified by transcriptome analysis?	Sequence motif identified by AlignACE	Site centre	Distance from nearest transcription start site
(A) Metabolism					
30 774	*carB*	No			
815 981	***ybhK*//*moaA***	**No**			
913 086	***hcp***	**Yes**			
1 003 973	***pyrD***	**No**			
1 297 552	***adhE***	**Yes**			
1 545 174	***yddG*/*fdnG***	**Yes**			
1 934 484	zwf/*yebK*	No			
1 935 348	pykA	No	5′-ATAACTTGAAGCGGGTCAAAGAAG-3′	1 935 377.5	Unknown
2 411 269	*yfbV*/*ackA*	Yes	5′-AAAATTTGCAGTGCATGATGTTAA-3′	2 411 385.5	Unknown
2 619 106	*upp*/*purM*	Yes	5′-TTTCGTTGACTTTAGTCAAAATGA-3′	2 618 979.5	−51.5/−192.5
2 632 233	*guaB*/*xseA*	No	5′-AGAATTTGATCTCGCTCACATGTT-3′	2 632 245.5	−117.5/+28.5
3 242 421	*uxaC*/*exuT*	Yes	5′-TTTTCGTGAGTTAGATCAATAAAC-3′	3 242 481.5	Unknown
3 491 582	***nirB***	**Yes**			
4 285 074	***acs*/*nrfA***	**Yes**			
4 365 790	***aspA*/*fxsA***	**Yes**			
4 380 005	***frdA*/*yjeA***	**Yes**			
4 460 414	***nrdD***	**Yes**			
(B) Unknown function					
528 318	*ylbI*	No			
541 128	*ybbZ*	No			
579 197	*ybcW*	Yes	5′-TGTTGATGATTTATGTCAAATATT-3′	578 914.5	Unknown
1 164 384	*ycfP*	No			
1 311 745	*yciC*/*ompW*	Yes	5′-AAAAATTGATTTAAATCACATTAA-3′	1 311 887.5	Unknown
1 396 615	*ydaA*	No	5′-TATTTTCGATGGTGATGTATTTAC-3′	1 396 756.5	Unknown
1 457 773	*ydaM*/*ydaN*	No			
1 515 200	*ydcX*	No	5′-AAAACTTGATGCACGTCAAAAAAT-3′	1 515 344.5	Unknown
1 627 197	*ydfZ*	Yes	5′-AAGATGTGAGCTTGATCAAAAACA-3′	1 627 144.5	Unknown
1 665 352	*ynfK*	Yes	5′-CCAAATTGAGATAGCGCAAATTTT-3′	1 665 337.5	Unknown
1 717 809	*ydhH*/*slyB*	No	5′-AACAAGTACAGAATGTCAGCTATG-3′	1 718 014.5	Unknown/−215.5
1 777 299	*ydiP*/*ydiQ*	No	5′-CTTTCTTGACGTAAATCAACTCGG-3′	1 777 183.5	Unknown
2 066 717	*yeeH*/*yoeA*	No	5′-AAAACATGATTAAGGTCAAAAATG-3′	2 066 533.5	Unknown
2 415 046	*yfcC*	Yes	5′-CTTAACTGCGCTGCATCAATGAAT-3′	2 415 030.5	Unknown
2 558 454	*yffL*	No	5′-AAACAATGATGTTCGTCAATTTTT-3′	2 558 434.5	Unknown
2 562 348	*yffS*	No	5′-GTTTTTTGTGATCTGCGTCAATAT-3′	2 562 452.5	Unknown
2 714 514	***yfiD*/*ung***	**Yes**			
3 151 520	*yghB*	No	5′-GTCCATGGCTGTTATTCAAGATAT-3′	3 151 585.5	Unknown
3 265 311	*yhaB*	No	5′-AACAATTGATTAACGTCAACTTTT-3′	3 265 180.5	Unknown
3 298 995	*yhbT*/*yhbU*	Yes	5′-TTTAACTGCCTTAAATCAAAAATT-3′	3 299 038.5	Unknown
3 351 934	*yhcC*/*gltB*	Yes	5′-CGGCGATGACCTGGATCAATCGTC-3′	3 351 753.5	Unknown/−399.5
3 578 508	*yhhX*/*yhhY*	No			
3 635 245	*yhiN*/*pitA*	Yes	5′-ATTTTTTGAGTGAAATCCATACAG-3′	3 635 162.5	Unknown
4 248 287	*yjbI*	No	5′-ATAACTTTATTTATATCAGCAATA-3′	4 248 420.5	Unknown
4 368 072	*yjeH*/*groS*	No	5′-GAAATGTGAGGTGAATCAGGGTTT-3′	4 368 101.5	Unknown/−91.5
(C) Transcription factors					
34 153	***yaaV*/*caiF***	**Yes**			
70 210	*araB*/*araC*	No			
121 966	***aroP*/*pdhR***	**Yes**			
1 719 041	*slyA*/*ydhI*	No	5′-ATTTATTAATCTAACGCAATATAT-3′	1 719 013.5	Unknown/−9.5
(D) Membrane proteins					
747 050	*abrB*	No			
770 354	***cydA***	**Yes**			
940 149	***dmsA***	**Yes**			
953 866	***focA***	**Yes**			
1 165 181	***ndh***	**Yes**			
1 277 131	***narX*/*narK***	**Yes**			
2 403 375	***nuoA***	**No**			
2 583 617	*aegA*/*narQ*	No	5′-ATAAACTGTGGCAGATCAAATAAT-3	2 583 671.5	Unknown
3 144 499	*hybO*	No	5′-AATACGTATGTTTGATCAATTTTC-3′	3 144 461.5	−83.5
3 150 195	*exbB*/*metC*	No			
3 928 854	*yieN*/*trkD*	No	5′-GAGCACTGATAATAAGCAATCATT-3′	3 928 956.5	Unknown
4 346 720	***dcuB***	**Yes**			

Table 2. Continued

Peak centre	Gene	Identified by transcriptome analysis?	Sequence motif identified by AlignACE	Site centre	Distance from nearest transcription start site
(E) DNA/RNA manipulation					
579 727	*nohB*	No	5′-TTAAGTTGATGCAGATCAATTAAT-3′	579 735.5	Unknown
1 023 647	*yccF/helD*	No	5′-AGTAATTGATTGAAAGGAATAAGG-3′	1 023 811.5	Unknown
1 407 100	*dbpA*	No	5′-AAAGTTTGAGCGAAGTCAATAAAC-3′	1 407 110.5	Unknown
1 634 803	*nohA/ydfO*	No	5′-TTAAGTTGATGCAGATCAATTAAT-3′	1 634 713.5	−85.5/unknown
4 584 413	*hsdR/mrr*	No	5′-CATCATTGTTATTAATCCATTGCT-3′	4 584 337.5	Unknown

The table lists the locations of peaks for FNR binding identified using chip and high-density microarray analysis. All of the peaks identified fell in non-coding DNA or were adjacent to the 5′ end of a gene. Targets are grouped according to the function of the gene(s) adjacent to the FNR target and are listed in chromosomal order within these groups. Entries highlighted with a bold face are experimentally verified FNR targets present in the current version of the Ecocyc database. For newly identified FNR targets (plain typeface) the sequence of the FNR binding site identified by AlignACE is given along with its genomic coordinate and, if known, the distance from the nearest transcription start site. The transcriptome analysis of Constantinidou *et al*. (10) was used for comparison.

contained l0 mM potassium phosphate (pH 7.5), 100 mM potassium glutamate, 1 mM EDTA, 50 mM DTT, 5% glycerol and 25 mg ml^{-1} herring sperm DNA. The final reaction volume was 10 µl. After incubation at 37°C for 20 min, samples were run in 0.25× TBE on a 6% polyacrylamide gel (12 V cm^{-1}) containing 2% glycerol and analysed using a Bio-Rad Molecular Imager FX and Quantity One software.

RESULTS

Isolation of DNA fragments associated with FNR in mid-log phase *E.coli*

Our aim was to use ChIP to measure the distribution of FNR across the chromosome of growing *E.coli* cells. To do this, we exploited strain JCB1011 whose *fnr* gene had been previously modified to encode FNR with a C-terminal 3× FLAG tag (10). Supplementary Figure 3A shows a western blot of total protein from strain JCB1011 and its parent, MG1655, probed with anti-FLAG or anti-FNR antibodies. The results show that intracellular levels of wild-type FNR and the FNR-3× FLAG fusion protein are similar and that the anti-FLAG antibody does not cross-react with other proteins. To check that the activity of FNR was unaffected by the 3× FLAG tag, we compared expression of five FNR-dependent promoters in JCB1011 and MG1655 (Supplementary Figure 3B) and anaerobic growth of the two strains (Supplementary Figure 1A). The results of these tests argue that the function of FNR is unaffected by the tag. Thus, JCB1011 and MG1655 cells were grown anaerobically in LB glucose medium to an OD_{650} of ∼0.4, cultures were treated with formaldehyde, and cellular DNA was extracted and sonicated, yielding DNA fragments of ∼500-1000 bp. After immunoprecipitation with anti-FLAG antibodies, DNA fragments from JCB1011 or control MG1655 cells were purified, labelled with Cy5 and Cy3, respectively, mixed and hybridized to the microarray. After washing and scanning, the Cy5/Cy3 signal intensity ratio was calculated for each probe. In parallel, the experiment was repeated using aerobically grown cells. Complete datasets are shown in Supplementary Table 1. Figure 1A gives an overview of the profile for FNR binding and some examples are shown in Figure 1B. Peaks for FNR binding are discrete and easily distinguishable from the background signal.

Identification and sequence analysis of FNR targets

To determine the location of peaks for FNR binding in an unbiased manner, a Cy5/Cy3 cut-off was applied to the 'anaerobic' dataset. A total of 204 probes passed this cut-off, corresponding to 63 separate peak locations, all of which were in non-coding DNA or close to the 5′ end of a gene (Table 2). Of the 63 peaks identified, 20 correspond to locations listed as FNR targets in the current version of the Ecocyc database [www.ecocyc.org, (16)] and 9 of the remaining 43 peaks locate to FNR targets predicted by Constantinidou *et al*. (10). To identify FNR binding sites at the 43 loci not currently listed by Ecocyc, we selected and then combined 500 bp DNA sequences corresponding to the centre of each peak. We then used AlignACE to search for sequence motifs present in these DNA sequences. A sequence logo representing the motif that we identified is shown in Figure 1C. The motif clearly matches the known FNR consensus binding motif of $TTGAT(n)_4ATCAA$. Using this approach, we were able to identify FNR binding sites at 33 of the 43 putative FNR targets identified by our ChIP-chip analysis. The base sequences and locations of the binding sites that we identified are shown in Table 2. Recall that the activity of FNR is triggered by lack of oxygen (4,5). The control dataset from cells grown aerobically showed that binding at the different targets was suppressed (e.g. see Figures 1B and 2A).

Association of FNR with some previously uncharacterized DNA targets

Five of the targets for FNR binding are adjacent to genes encoding proteins involved in the manipulation of RNA or DNA and, in each case, a likely FNR binding site was identified (see Table 2 section E; Figure 2A). DNA fragments covering each of these targets were amplified and end-radiolabelled, or cloned into the *lac* expression vector pRW50 to create promoter::*lacZ* fusions. To detect the binding of FNR to the predicted target promoters *in vitro*, the radiolabelled DNA fragments were used in EMSA assays with purified FNR protein (Figure 2B). In all cases, addition of purified FNR retarded the migration of the purified DNA fragments. To investigate the effect of FNR on transcription from each target promoter, each of the pRW50 encoded

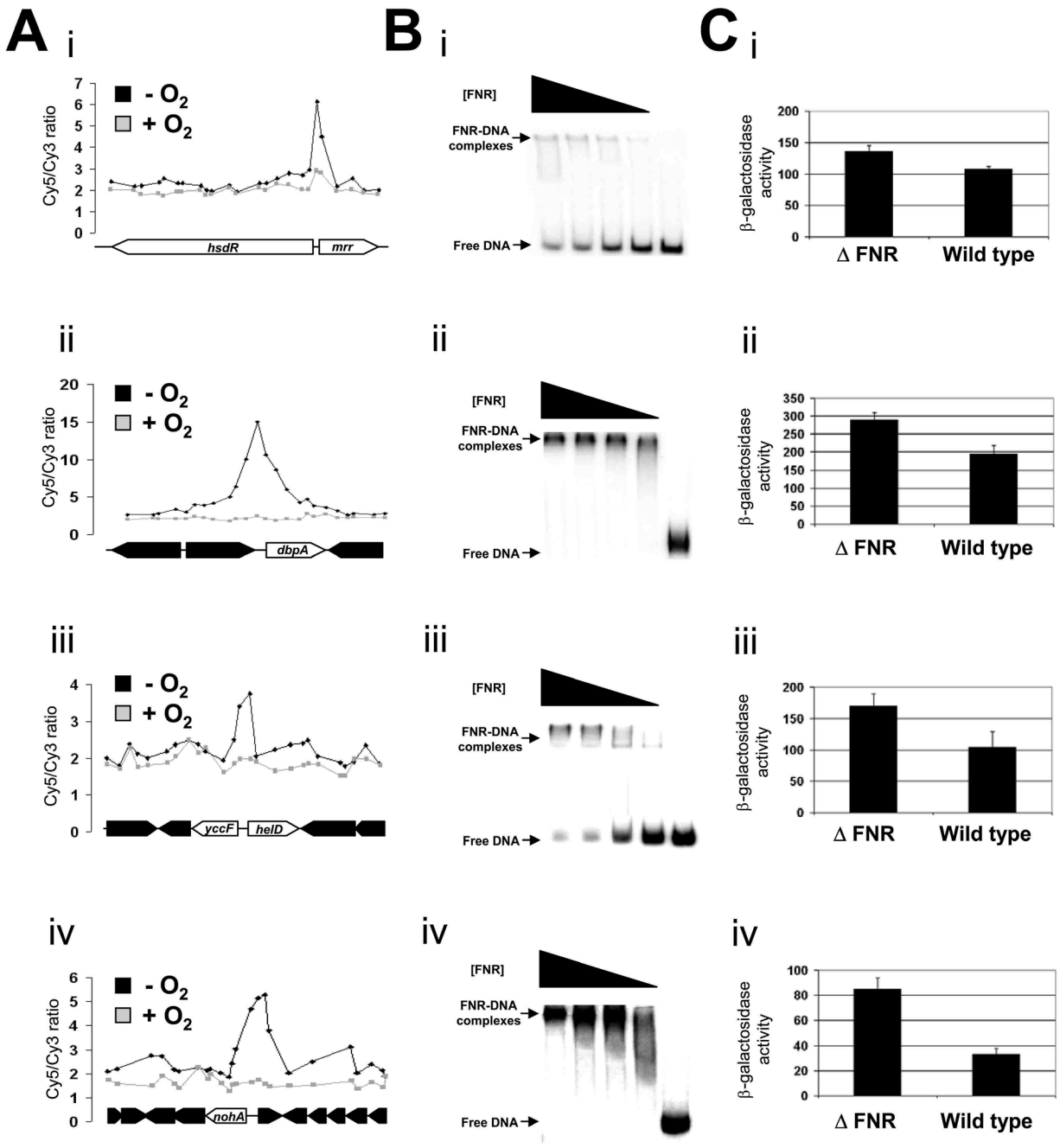

Figure 2. Association of FNR with a novel group of DNA targets. (**A**) The *in vivo* DNA binding profile of FNR across intergenic regions upstream of the *hsdR* and *mrr* (i), *dbpA* (ii), *yccF* and *helD* (iii) and *nohA* (iv) genes. Binding is illustrated by Cy5/Cy3 signals obtained from ChIP-chip experiments and these signals are plotted against the corresponding features of the *E.coli* chromosome. (**B**) Binding of purified FNR D154A protein to DNA fragments corresponding to intergenic regions upstream of the *hsdR* and *mrr* (i), *dbpA* (ii), *yccF* and *helD* (iii), and *nohA*/*nohB* (iv) genes.EMSA with end-labelled EcoRI–HindIII DNA fragments are illustrated. DNA fragments were incubated with 0, 0.2, 0.4, 1.0 or 2.0 μM FNR as indicated. Note that, because the *nohA* and *nohB* intergenic regions have a similar sequence, the amplified DNA fragment represents both locations. (**C**) Activity of promoter::*lacZ* fusions in wild-type (JCB387) and Δ*fnr* (JRG1728) *E.coli* cells. For intergenic regions between divergent genes we show the activity of the *mrr* (i) and *helD* (iii) promoters. The activities of the upstream promoters were 147 and 235 for *hsdR* (i) and 210 and 175 for *yccF* (iii) in strains JRG1728 and JCB387, respectively. We measured 45 and 30 U of β-galactosidase activity for strains JRG1728 and JCB387 carrying pRW50 with no promoter insert.

promoter::*lacZ* fusions was transformed into *E.coli* strain JCB387 or the *fnr* derivative JRG1728. Expression of *lacZ* in each strain was measured. Our data (Figure 2C) show that deletion of *fnr* had marginal affects on transcription and that the activity of each cloned promoter was low in our conditions.

Distribution of FNR in stationary phase

The analysis of FNR binding was repeated using stationary phase cultures of JCB1011 and MG1655. The dataset for the FNR experiment is presented in Supplementary Table 1, alongside the data from growing cells, and an overview is shown in Figure 3A. The results show that the profile of

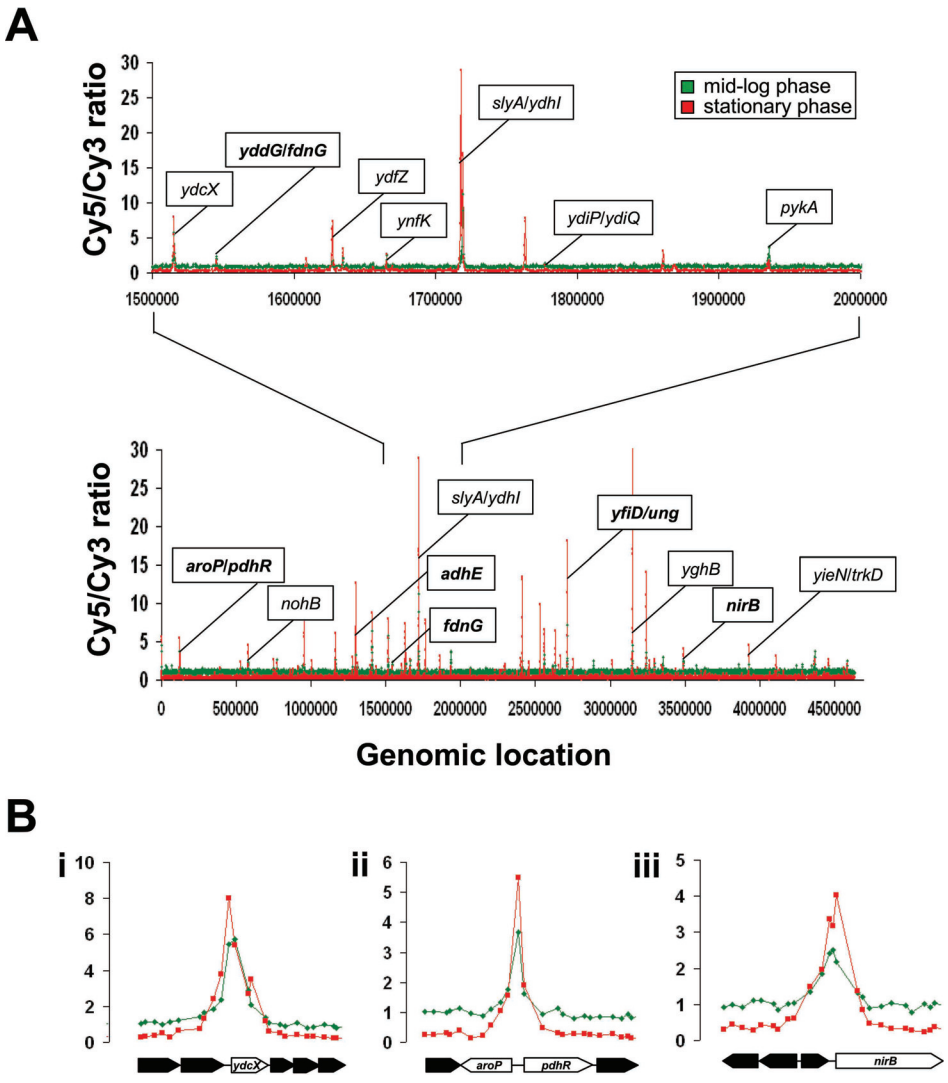

Figure 3. Comparison of chromosome-wide FNR distribution in growing and stationary phase *E.coli*. (**A**) The figure shows an overview of results from ChIP-chip experiments that measure the profile of FNR binding across the *E.coli* chromosome during rapid growth (green) or during stationary phase (red). Binding signals (*y*-axis) are plotted against their location on the 4.64 Mb *E.coli* chromosome (*x*-axis). The locations of selected signals are labelled in plain typeface (newly identified FNR targets) or in bold face (known FNR targets). The lower section shows data for the entire *E.coli* genome and the upper section shows an expanded 0.5 Mb window. (**B**) Expansion of selected regulatory regions.

FNR binding in stationary phase is similar to the profile in growing cells (e.g. are shown in Figure 3B).

As controls for this experiment, similar analyses were performed with the nucleoid-associated protein, IHF and with RNA polymerase, which is known to be redistributed in stationary phase *E.coli*. Note that previously we had used ChIP-chip to study IHF and RNA polymerase in exponentially growing cells (17). The datasets for IHF and RNA polymerase binding are shown in Supplementary Tables 2 and 3, respectively, alongside data generated using growing *E.coli* cells (17). The results show that the profile of IHF binding in stationary phase is similar to the profile in growing cells while the profile of RNA polymerase is radically altered.

Similarities between the datasets for FNR, IHF and RNA polymerase binding, in growing and stationary phase cells, were quantified by calculating correlation coefficients (Figure 4A) and by comparing the position of probes passing the cut-off for each dataset (Figure 4B). For FNR and IHF, the correlation between ChIP-chip datasets profiling binding in growing and stationary phase cells is high [Figure 4A (i) and (ii)]. Much less correlation is observed when ChIP-chip datasets for RNA polymerase binding in stationary phase and mid-log phase cells are compared [Figure 4A (iii)]. Consistent with this, for FNR and IHF, many of the same probes pass the cut-off for both datasets. In contrast, few probes pass the cut-off for both the mid-log and stationary phase RNA polymerase experiments (Figure 4B).

DISCUSSION

In this work, we have applied ChIP-chip technology to produce the first chromosome-wide direct analysis of DNA binding *in vivo* by the global *E.coli* transcription regulator, FNR. The advantage of this approach to studying the FNR

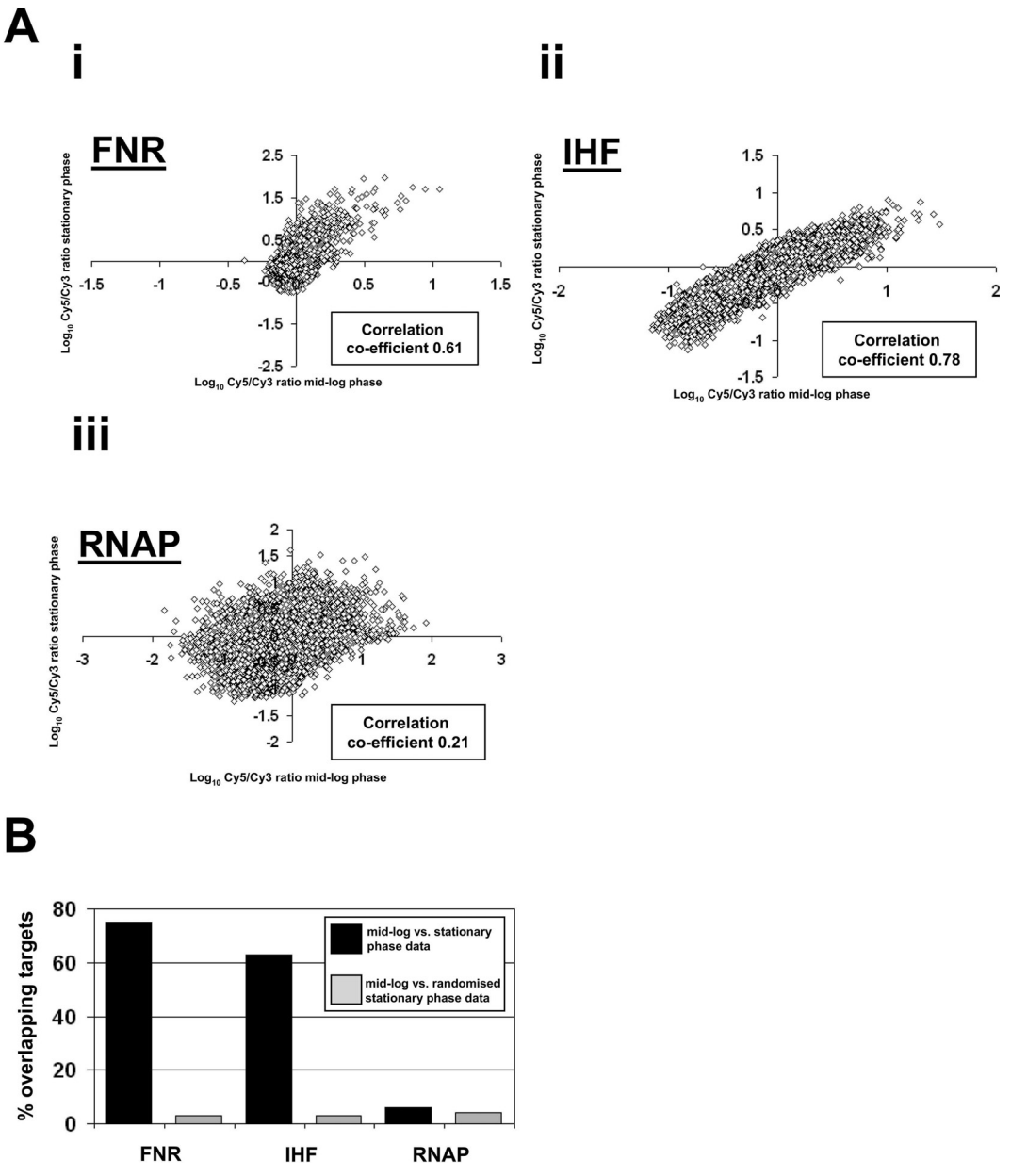

Figure 4. Comparison of chromosome-wide FNR, IHF and RNA polymerase distribution in growing and stationary phase *E.coli*. (**A**) Correlation between ChIP-chip datasets profiling the chromosome-wide distribution of FNR (i) IHF (ii) and RNA polymerase (iii) in growing and stationary phase cells. The $\log_{10}$ value of the mid-log phase (*x*-axis) binding signal measured at each probe by ChIP-chip analysis is plotted against the corresponding $\log_{10}$ ratio for stationary phase (*y*-axis). (**B**) Overlap of DNA targets occupied by FNR, IHF and RNA polymerase during rapid growth and stationary phase. The ChIP-chip datasets obtained for FNR, IHF and RNA polymerase binding during rapid growth and stationary phase were aligned and an equivalent cut-off was applied to each dataset. The number of probes passing the cut-off for both stationary phase and mid-log phase datasets was then determined (black bar). As a control, this analysis was repeated after the genomic position of probes in the stationary phase dataset had been randomized (grey bar).

regulon is that it avoids complications due to genes that are indirectly controlled by FNR or genes that are regulated by multiple transcription factors. Moreover, FNR binding at sites adjacent to poorly transcribed genes, or genes where FNR has little impact on transcription, can be detected and the effects of environmental conditions can be studied. We identified 63 locations at which FNR binds to the *E.coli* chromosome, including a group of five targets adjacent to genes encoding proteins that manipulate DNA and RNA. None of these five targets were identified as FNR regulated by previous transcriptome analyses (8,9,10), none are listed as FNR targets by the Ecocyc database (16), and, to our knowledge, FNR-dependent regulation of such proteins has not been documented. At 10 of the 63 targets of FNR binding, we were unable to identify a match to the canonical FNR binding sequence. These may be locations at which FNR binds cooperatively with another factor. We note that the profile of FNR binding presented here consists of discrete peaks (Figure 1A). In contrast, the published binding profile for the related transcription activator, CRP, is far more complex, due to the existence of ~10 000 low affinity binding sites for CRP scattered throughout the genome (6,13). Consistent with

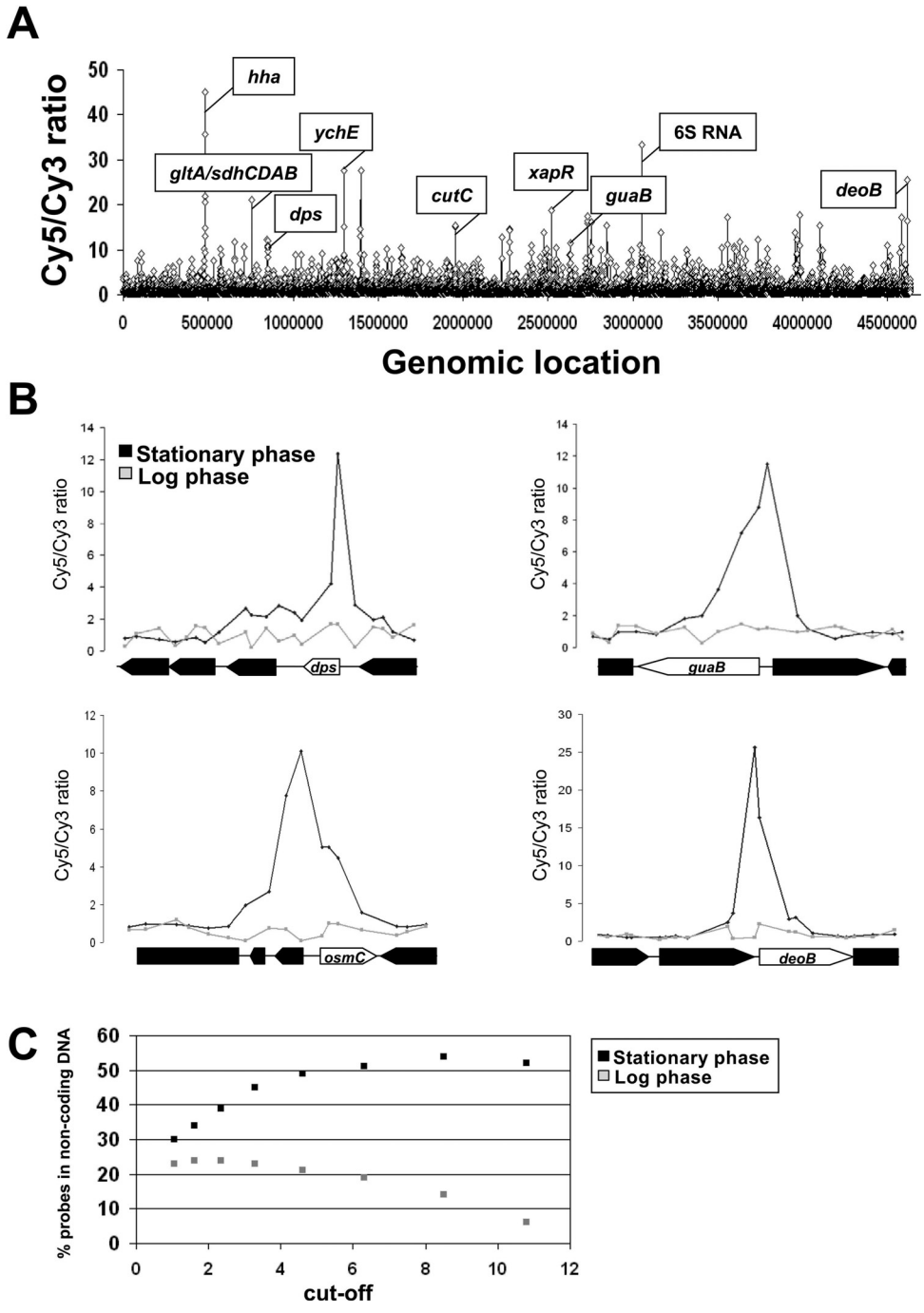

Figure 5. Increased association of RNA polymerase with non-coding DNA during stationary phase. (**A**) The figure shows an overview of results from ChIP-chip experiments that measure the profile of RNA polymerase binding across the *E.coli* chromosome during stationary phase. Binding signals (*y*-axis) are plotted against their location on the 4.64 Mb *E.coli* chromosome (*x*-axis). The locations of selected signals are labelled. (**B**) Skewed distribution of RNA polymerase across transcribed regions during stationary phase. The figure illustrates of selected regions highlighted in (A). Data for RNA polymerase binding during stationary phase are shown in black and RNA polymerase binding during mid-log phase is shown in grey. (**C**) Increased association of RNA polymerase with non-coding DNA during stationary phase. The ChIP-chip datasets for RNA polymerase binding during rapid growth (grey) and stationary phase (black) were aligned and a range of Cy5/Cy3 cut-offs were applied to select the upper 50, 100, 200, 400, 800, 1600, 3200 and 6400 probes for each dataset. We then determined the distribution of probes passing the cut-offs between coding and non-coding DNA.

this, Robison *et al*. (6) predicted only ∼500 low affinity targets for FNR in the *E.coli* chromosome. Although our experiments identified 63 discrete targets for FNR binding, these targets include only 20 out of the 65 validated targets listed in the Ecocyc database (16). This highlights an important limitation of the ChIP-chip methodology. Its inability to detect all FNR-DNA interactions is likely to be due to inefficient crosslinking at some locations, or epitope masking (18).

Although major changes in gene expression and nucleoid structure occur when *E.coli* cells enter stationary phase

(11,12,19), it is not known if changes in the distribution of global DNA binding proteins occur. Our observation that the DNA binding profile for FNR is largely unaltered in stationary phase cells shows that the compaction of the stationary phase chromosome does not occlude FNR binding sites. This surprising result prompted us to examine the nucleoid-associated protein IHF and we came to the same conclusion; the distribution of the protein is similar in growing and stationary phase cells. In sharp contrast, as expected, the binding profile of RNA polymerase was completely altered. In growing cells, most RNA polymerase is associated with ~90 transcription units encoding factors required for protein synthesis, motility and ATP production (13). In stationary phase, this RNA polymerase is liberated and is distributed more equitably between the different genes (Figure 5A). Interestingly, for most transcription units, RNA polymerase binding is skewed toward the 5′ end (e.g. are shown in Figure 5B) and thus, in stationary phase, the proportion of RNA polymerase bound to non-coding parts of the genome is increased (Figure 5C). Consistent with this, Lee and Gralla (20) showed that some σ^S-dependent promoters have the ability to trap RNA polymerase. We note that, in their study of RNA polymerase distribution in stationary phase *S.cerevisiae*, Radonjic and colleagues identified a similar phenomenon (21). This suggests that trapping of RNA polymerase at promoters may be an evolutionarily conserved mechanism to regulate transcription in response to growth rate.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

The authors thank Steve Goodman and Jeff Green for providing antisera and Jeff Cole, Bob Landick, David Lee, Tim Overton, Bernhard Palsson and Joseph Wade for helpful discussions. This work was supported by a Wellcome Trust programme grant, a BBSRC Japan Partnering Award, a grant-in-aid of scientific research from the Japanese Ministry of Education, Culture, Sports, Science and Technology, and the Daiwa Anglo-Japanese Foundation. Funding to pay the Open Access publication charges for this article was provided by our Wellcome Trust programme grant.

Conflict of interest statement. None declared.

REFERENCES

1. Browning,D.F. and Busby,S.J. (2004) The regulation of bacterial transcription initiation. *Nature Rev. Microbiol*., **2**, 57–65.
2. Martinez-Antonio,A. and Collado-Vides,J. (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol*., **6**, 482–489.
3. Babu,M.M. and Teichmann,S.A. (2003) Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucleic Acids Res*., **31**, 1234–1244.
4. Browning,D., Lee,D., Green,J. and Busby,S. (2002) Secrets of bacterial transcription initiation taught by the *Escherichia coli* FNR protein. In Hodgson,D. and Thomas,C. (eds), *Signals, Switches, Regulons and Cascades: Control of Bacterial Gene Expression*. Cambridge University Press, Cambridge, UK, pp. 127–142.
5. Green,J. and Paget,M. (2004) Bacterial redox sensors. *Nature Rev. Microbiol*., **2**, 954–966.
6. Robison,K., McGuire,A.M. and Church,G.M. (1998) A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete *Escherichia coli* K-12 genome. *J. Mol. Biol*., **284**, 241–54.
7. Tan,K., Moreno-Hagelsieb,G., Collado-Vides,J. and Stormo,G. (2000) A comparative genomics approach to prediction of new members of regulons. *Genome Res*., **11**, 566–584.
8. Salmon,K., Hung,S.P., Mekjian,K., Baldi,P., Hatfield,G.W. and Gunsalus,R.P. (2003) Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. *J. Biol. Chem*., **278**, 29837–29855.
9. Kang,Y., Weber,K.D., Qiu,Y., Kiley,P.J. and Blattner,F.R. (2005) Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function. *J. Bacteriol*., **187**, 1135–1160.
10. Constantinidou,C., Hobman,J.L., Griffiths,L., Patel,M.D., Penn,C.W., Cole,J.A. and Overton,T.W. (2006) A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth. *J. Biol. Chem*., **281**, 4802–4815.
11. Selinger,D.W., Cheung,K.J., Mei,R., Johansson,E.M., Richmond,C.S., Blattner,F.R., Lockhart,D.J. and Church,G.M. (2000) RNA expression analysis using a 30 bp resolution Escherichia coli genome array. *Nat. Biotechnol*., **18**, 1262–1268.
12. Weber,H., Polen,T., Heuveling,J., Wendisch,V.F. and Hengge,R. (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol*., **187**, 1591–1603.
13. Grainger,D.C., Hurd,D., Harrison,M., Holdstock,J. and Busby,S.J. (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli chromosome*. *Proc. Natl Acad. Sci.USA*, **102**, 17693–17698.
14. Miller,J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. Browning,D.F., Grainger,D.C., Beatty,C.M., Wolfe,A.J., Cole,J.A. and Busby,S.J. (2005) Integration of three signals at the *Escherichia coli* nrf promoter: a role for Fis protein in catabolite repression. *Mol. Microbiol*., **57**, 496–510.
16. Keseler,I.M., Collado-Vides,J., Gama-Castro,S., Ingraham,J., Paley,S., Paulsen,I.T., Peralta-Gil,M. and Karp,P.D. (2005) EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Res*., **33**, 334–337.
17. Grainger,D.C., Hurd,D., Goldberg,M. and Busby,S.J. (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res*., **34**, 4642–4652.
18. Buck,M.J. and Lieb,J.D. (2004) ChIP-chip: considerations for the design, analysis and application of genome-wide chromatin immunoprecipitation experiments. *Genomics*, **83**, 349–360.
19. Kim,J., Yoshimura,S.H., Hizume,K., Ohniwa,R.L., Ishihama,A. and Takeyasu,K. (2004) Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. *Nucleic Acids Res*., **32**, 1982–1992.
20. Lee,S.J. and Gralla,J.D. (2004) Osmo-regulation of bacterial transcription via poised RNA polymerase. *Mol. Cell*, **14**, 153–162.
21. Radonjic,M., Andrau,J.C., Lijnzaad,P., Kemmeren,P., Kockelkorn,T.T., van Leenen,D., van Berkum,N.L. and Holstege,F.C. (2005) Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon S. *cerevisiae* stationary phase exit. *Mol. Cell*, **18**, 171–183.
22. Blattner,F.R., Plunkett,G., III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F. *et al*. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1474.
23. Page,L., Griffiths,L. and Cole,J.A. (1990) Different physiological roles for two independent pathways for nitrite reduction to ammonia by enteric bacteria. *Arch. Microbiol*., **154**, 349–354.
24. Spiro,S. and Guest,J.R. (1988) Inactivation of the FNR protein of *Escherichia coli* by targeted mutagenesis in the N-terminal region. *Mol. Microbiol*., **2**, 701–707.
25. Lodge,J., Fear,J., Busby,S., Gunasekaran,P. and Kamini,N.R. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol. Lett*., **74**, 271–276.