Molecular Microbiology (2007) **66**(6), 1382–1395

doi:10.1111/j.1365-2958.2007.05999.x
First published online 19 November 2007

# Mutations that alter the regulation of the *chb* operon of *Escherichia coli* allow utilization of cellobiose

**Aashiq H. Kachroo,[1] Aswani K. Kancherla,[1,2] Nongmaithem S. Singh,[3] Umesh Varshney[3] and Subramony Mahadevan[1]***

[1]*Department of Molecular Reproduction, Development and Genetics,* [2]*Molecular Biophysics Unit, and* [3]*Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India.*

## Summary

**Wild-type strains of *Escherichia coli* are normally unable to metabolize cellobiose. However, cellobiose-positive (Cel$^+$) mutants arise upon prolonged incubation on media containing cellobiose as the sole carbon source. We show that the Cel$^+$ derivatives carry two classes of mutations that act concertedly to alter the regulation of the *chb* operon involved in the utilization of N,N′-diacetylchitobiose. These consist of mutations that abrogate negative regulation by the repressor NagC as well as single base-pair changes in the transcriptional regulator *chbR* that translate into single-amino-acid substitutions. Introduction of *chbR* from two Cel$^+$ mutants resulted in activation of transcription from the *chb* promoter at a higher level in the presence of cellobiose, in reporter strains carrying disruptions of the chromosomal *chbR* and *nagC*. These transformants also showed a Cel$^+$ phenotype on MacConkey cellobiose medium, suggesting that the wild-type permease and phospho-β-glucosidase, upon induction, could recognize, transport and cleave cellobiose respectively. This was confirmed by expressing the wild-type genes encoding the permease and phospho-β-glucosidase under a heterologous promoter. Biochemical characterization of one of the *chbR* mutants, *chbRN238S*, showed that the mutant regulator makes stronger contact with the target DNA sequence within the *chb* promoter and has enhanced recognition of cellobiose 6-phosphate as an inducer compared with the wild-type regulator.**

## Introduction

*Escherichia coli can* respond to stress such as starvation using a variety of strategies. Under conditions of starvation wherein a novel substrate is provided as a sole nutritional source, spontaneous mutants arise that are able to utilize this novel compound. Many genetic systems, upon mutational activation, have been shown to allow *E. coli* to grow on novel substrates (Wright, 2004).

Most wild-type strains of *E. coli* are unable to utilize the cellulose-derived disaccharide cellobiose as a carbon source. Therefore, they cannot grow on minimal cellobiose medium and form white colonies on MacConkey cellobiose indicator medium. However, cellobiose-positive (Cel$^+$) mutants can be obtained after prolonged incubation (~20–30 days) at room temperature (25–30°C) as papillae on MacConkey cellobiose plates and as colonies on minimal cellobiose plates (Kricker and Hall, 1984). The Cel$^+$ phenotype was attributed to mutations in the genetic system originally named the *cel* operon that was believed to be cryptic (Parker and Hall, 1990a). The *cel* operon, subsequently renamed the *chb* operon, is the inducible genetic system involved in the catabolism of N,N′-diacetylchitobiose (Keyhani and Roseman, 1997). The *chb* operon comprises six open reading frames (ORFs) (*chbBCARFG*) with a ~200 bp regulatory region (*chbOP*). The *chbBCA* genes encode the IIB, IIC and IIA domains of the PTS-dependent permease, respectively (Keyhani *et al*., 2000b), *chbR* encodes a dual function activator/repressor (Plumbridge and Pellegrini, 2004), *chbF* encodes a phospho-β-glucosidase assigned to family 4 of the glycosylhydrolase superfamily with wide substrate specificity (Thompson *et al*., 1999) and *chbG* codes for a protein of unknown function. The regulation of the *chb* operon by chitobiose has recently been studied (Plumbridge and Pellegrini, 2004). It has been shown that the three proteins ChbR, CRP and NagC regulate the expression of the *chb* operon. ChbR belongs to the AraC-like dual function activator/repressor family of proteins. In the absence of the substrate N,N′-diacetylchitobiose, ChbR binds to the direct repeats present within *chbOP* and represses transcription along with the negative regulator NagC. In the presence of chitobiose, ChbR, along with CRP, activates transcription from the *chb* promoter (Plumbridge and Pellegrini, 2004).

Activation of the *chb* operon allowing the utilization of cellobiose was reported to occur either via insertion of IS*1*, IS*2* or IS*5* within 72–180 bp upstream of an earlier uncharacterized start site or by base substitutions in *chbR* such that the putative repressor is able to recognize

Accepted 9 October, 2007. *For correspondence. E-mail mahi@mrdg.iisc.ernet.in; Tel. (+91) 80 2293 2607; Fax (+91) 80 2360 0999.

cellobiose, salicin and arbutin as inducers (Parker and Hall, 1990b). The need of prolonged incubation in order to obtain these Cel$^+$ mutants and the complete absence of constitutive Cel$^+$ mutants were, however, unexplained.

The timescale of ~30 days to acquire Cel$^+$ phenotype by *E. coli* either by activating insertion elements or by mutations within the *chbR* locus is long compared with activation by insertion mutations targeting the regulatory region (*bglR*) of the cryptic *bgl* operon of *E. coli*, conferring a salicin-positive (Sal$^+$) phenotype within 24–48 h of incubation on MacConkey salicin plates. This was partly explained when the sequence of the *chb* operon reported earlier (Parker and Hall, 1990a) was compared with the wild-type *E. coli* K-12 genome (Blattner *et al*., 1997). The comparison revealed many differences. These differences are clustered in the permease and in the putative phospho-β-glucosidase. Based on this observation it was hypothesized that additional mutations may be involved in the activation of the *chb* operon allowing utilization of cellobiose (Keyhani and Roseman, 1997).

The results presented in this communication show that acquisition of a Cel$^+$ phenotype involves a minimum of two types of mutations that primarily alter the regulation of the *chb* genes. These two classes of mutations act in a concerted fashion and are necessary and sufficient to confer a Cel$^+$ phenotype; mutations in the structural genes, predicted by earlier work, are not necessary. The biochemical characterization of one of the *chbR* mutants, *chbRN238S*, indicates that the enhanced activation seen in the mutant is due to the combined action of enhanced binding to the target sequences as well as improved recognition of cellobiose as an inducer.

## Results

### *An insertion within* chbOP *or a deletion of* chbR *do not confer a Cel$^+$ phenotype*

To examine whether the *chb* operon can be activated to confer Cel$^+$ phenotype by a single mutational event of insertion within the *chbOP* region as reported earlier (Parker and Hall, 1990b), an artificial insertion of a ~1.1 kb DNA fragment encoding chloramphenicol resistance was made in the *E. coli* strain DY330 carrying a wild-type copy of the *chb* genes. The target for the artificial insertion was the same as that of an activating insertion element seen in a subset of Cel$^+$ mutants (Fig. 1). The strain showed a Cel$^-$ phenotype on MacConkey cellobiose medium (after 24–36 h of incubation at 37°C). Similar results were obtained when the artificial insertion was moved from the strain DY330 (*chbOP*::*cat*) into JF201 by P1 transduction to rule out strain-specific differences. Interestingly, strains carrying the artificial insertion within *chbOP* papillated faster compared with the respective wild-type strains, suggesting a positive role for the insertion in acquiring a Cel$^+$ phenotype.

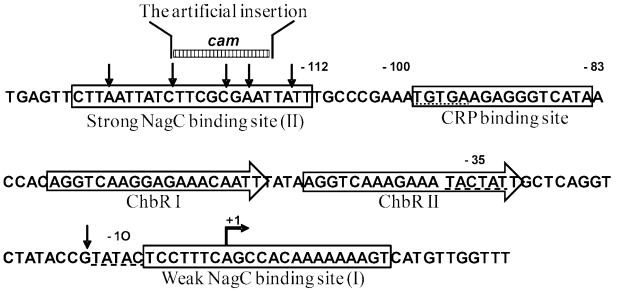


**Fig. 1.** The regulatory region (*chbOP*) of the *chb* operon showing binding sites for NagC, ChbR and CAP proteins (based on Plumbridge and Pellegrini, 2004). A newly characterized transcription start site is also shown. Black arrows within the strong NagCII binding site indicate insertion elements reported in Cel$^+$ mutant strains. The region chosen for the artificial insertion (chloramphenicol resistance marker) is indicated. The dashed underlines indicate the −10 and −35 promoter elements and the core CRP-cAMP binding site.

The possible effect of a null mutation in the putative repressor gene *chbR* in conferring a Cel$^+$ phenotype was also examined. A deletion of the *chbR* locus was constructed as before using a targeted recombination approach in the strain DY330 and subsequently moved to JF201. The *chbR* deletion strains, when plated on MacConkey cellobiose medium, were Cel$^-$ and did not papillate to yield Cel$^+$ mutants even after prolonged incubation of 30–40 days, suggesting an additional regulatory role for *chbR*, consistent with its role in the activation of the operon in the presence of N,N′-diacetylchitobiose (Plumbridge and Pellegrini, 2004). Therefore, single mutations within the *chb* operon are incapable of conferring a Cel$^+$ phenotype, contrary to earlier reports (Parker and Hall, 1990b).

### *Isolation and characterization of Cel$^+$ mutants from different* E. coli *strains*

In an attempt to understand the mechanism of activation of the *chb* operon, large-scale isolation of ~120 Cel$^+$ mutants from different strains of *E. coli* was carried out (Table 1). Mutants were obtained as papillae on MacConkey cellobiose plates [MCP] or colonies on M9 minimal cellobiose plates [M9]. Typically mutants appeared as papillae in ~24 days on MCP and as colonies on M9 plates in ~14 days. The extent of cellobiose utilization by different Cel$^+$ mutants was tested by plating on MacConkey cellobiose plates and spotting dilutions of different Cel$^+$ mutants on M9 minimal cellobiose medium. The Cel$^+$ phenotypes of 42 mutants tested are indicated in Table 3. Variation in the cellobiose utilization phenotype of different Cel$^+$ mutants suggested the presence of different types of activating mutations. Contrary to earlier reports (Kricker and Hall, 1987), none of the cellobiose-utilizing mutants tested could utilize salicin and arbutin.

 

**Table 1.** Bacterial strains used in this study.

| Strain | Genotype | Source |
|---|---|---|
| JF201 | F⁻ Δ*lac*X74 Δ(*bgl-pho*) 201 *ara thi gyrA* | Reynolds *et al*. (1986) |
| DH5α | *F′/endA1 hsdR*17 ($r_k^-m_k^+$) *supE44 thi-1 recA1gyrA* (Nal$_r$)*relA1*(Δ*lacZYA argF*)*U169 deo* F*dlac*Δ(*lacZ*)M15 | Woodcock *et al*. (1989) |
| MG1655 | (F⁻ wt *bgl*° *lam*⁻*rph-1*) Kan$^r$ | *E. coli* Genetic Stock Center |
| JM-chb21 | JM101 [Δ*lac-pro, thi-1, sup*E, F′(*traD36, proAB*⁺, *lacI*$^q$, *lacZ*DM15)] (*chbR::cat*) *pchbOPB′–lacZ* | Plumbridge and Pellegrini (2004) |
| JM-chb22 | JM101 [Δ*lac-pro, thi-1, supE*, F′(*traD36, proAB*⁺, *lacI*$^q$, *lacZ*DM15)] (*chbR::cat*; *nagC::tet*) *pchbOPB′ lacZ* | Plumbridge and Pellegrini (2004) |
| JM-chb3 | JM101 [Δ*lac-pro, thi-1, supE*, F′(*traD36, proAB*⁺, *lacI*$^q$, *lacZ*DM15)] (*nagC::tet*) *pchbOPB′ lacZ* | Plumbridge and Pellegrini (2004) |
| JF496 | *nagB2 asn50::*Tn*5* | *E. coli* Genetic Stock Center |
| DY330 | W3110 Δ*lacU169 gal490* λ*cI857* Δ(*cro-bioA*) | Yu *et al*. (2000) |
| DOPCam | DY330 (*chbOP*::*cat*) | This work |
| AHK5 | DY330 (*chbOP*::*cat*) Cel⁺ | This work |
| DRCam | DY330 (*chbR*::*cat*) | This work |
| DRKan | DY330 (*chbR*::*kan*) | This work |
| JOPCam | JF201 (*chbOP*::*cat*) | This work |
| JRCam | JF201 (*chbR*::*cat*) | This work |
| MG-MCP-01-32 | Cel⁺ mutants of MG1655 | This work |
| MG-M9-01-30 | Cel⁺ mutants of MG1655 | This work |
| J-MCP-01-28 | Cel⁺ mutants of JF201 | This work |
| J-M9-01-25 | Cel⁺ mutants of JF201 | This work |
| AHK3 PI | Cel⁺ natural isolate | This work |
| MS201 | Cel⁺ natural isolate | This work |
| NC 2.1 | Cel⁺ natural isolate | This work |
| NC 7.1 | Cel⁺ natural isolate | This work |
| Rosetta (DE3) | T7 polymerase-inducible expression strain | Novagen |

### *Mutations linked to the* chb *locus are involved in the Cel⁺ phenotype*

The cellobiose utilization phenotype of the Cel⁺ mutants, at least in part, was shown to be linked to the *chb* operon using P1 transduction. P1 lysate prepared using the Cel⁻ donor strain JF201 (*chbR*::*cat*), in which the chloramphenicol resistance gene is 100% linked to the *chb* operon, was used to transduce the Cel⁺ mutants. All chloramphenicol-resistant transductants showed loss of the Cel⁺ phenotype, confirming linkage of at least one of the associated mutations to the *chb* operon.

### *Many Cel⁺ strains carry mutations in* nagC

The NagC repressor that regulates the *nag* operon involved in *N*-acetylglucosamine metabolism was also shown to regulate the *chb* operon (Plumbridge and Pellegrini, 2004). The insertions seen in many Cel⁺ mutants disrupt the strong NagC binding site within the *chb* regulatory region. However, only a small number of Cel⁺ mutants obtained carried insertion elements within the regulatory region (*chbOP*) that disrupted the strong NagC binding site. If repression by NagC has to be eliminated for acquiring a Cel⁺ phenotype, one possibility is that these strains have mutations in the *nagC* locus itself. To test this possibility, the wild-type *nagC* locus was transduced into the Cel⁺ mutants using the donor strain JF496, which carries a Tn*5* insertion within the *asn* locus that has 85% linkage to wild-type *nagC*. The transductants showed a Cel⁻ phenotype at a frequency of ~85%, indicating that the Cel⁺ strains had lesions within *nagC* that are associated with the Cel⁺ phenotype, which is lost upon introduction of the wild-type *nagC* locus.

The MG1655 strain used for isolation of Cel⁺ mutants was inherently kanamycin resistant, pre-empting transduction experiments using the *kan* marker. Alternatively the wild-type *nagC* locus was introduced on a plasmid by transformation. All ampicillin-resistant transformants were Cel⁻, indicating that the chromosomal *nagC* carries recessive mutations that are partly responsible for the Cel⁺ phenotype. These results confirm that mutations within the *chb* operon and loss-of-function mutations within the *nagC* locus play a critical role in conferring a Cel⁺ phenotype.

### *Characterization of the mutations linked to the* chb *operon in Cel⁺ mutants*

The inability of *chbR* deletion strains to yield Cel⁺ mutants even after prolonged incubation (~30–40 days) suggested an essential role for ChbR in conferring a Cel⁺ phenotype. To examine if the Cel⁺ mutants harbour mutations within *chbR*, the *chbR* locus (~900 bp) from different Cel⁺ mutants and the Cel⁻ parent strains was amplified and sequenced. The nucleotide sequence of the *chbR* locus from the Cel⁻ parent strains showed a 100% match with

 

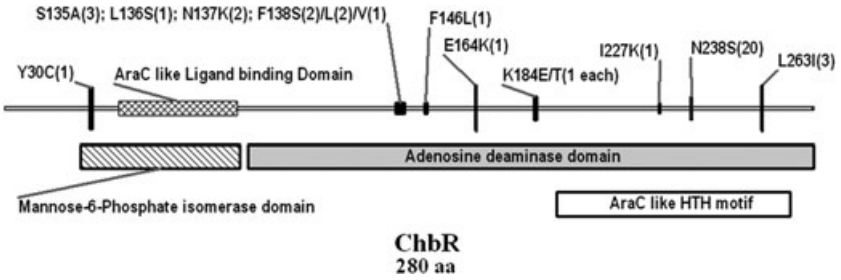


**Fig. 2.** Different amino acid substitutions in the ChbR protein from various Cel$^+$ mutants due to point mutations at the *chbR* locus are shown. The numbers in parenthesis are the number of strains showing the same type of amino acid substitution. Conserved domain database analysis of ChbR showed that the N-terminal has similarity to the mannose 6-phosphate isomerase domain associated with carbohydrate transport and metabolism. The C-terminal half of the protein showed partial similarity to the adenosine deaminase domain. The C-terminal end of ChbR also carries an AraC-like helix–turn–helix (HTH) motif.

the *E. coli* K12 sequence available in the NCBI database. The sequences of the *chbR* gene obtained from 42 Cel$^+$ mutants analysed showed a single-base mutation within the coding region of *chbR*, resulting in a single-amino-acid substitution at the protein level (Table 3). As indicated in Fig. 2, the mutations are scattered across the different domains of the ChbR protein. The role of these single-amino-acid substitutions in conferring a cellobiose utilization phenotype was analysed both at the phenotypic level and with respect to their effect on transcription from the *chb* promoter.

### *The effect of point mutations within the* chbR *locus on* chb *promoter activity*

To investigate the role of mutations within the *chbR* and *nagC* loci in acquiring a Cel$^+$ phenotype, single-copy chromosomal reporter strains that contain a *chbOPB′–lacZ* transcriptional/translational fusion at the *attB* site (Plumbridge and Pellegrini, 2004) were used. Primarily three strains were used in these studies. One of them, JM-chb22, had disruptions of the chromosomal *chbR* and *nagC* loci and the other, JM-chb21, had a *chbR* disruption, but an intact *nagC* locus. The third strain, JM-chb3, had a disruption of *nagC*, but an intact *chbR* locus. Two *chbR* mutants that showed a strong Cel$^+$ phenotype were chosen for the studies. The mutant *chbR* alleles were introduced on plasmids and their effect on transcription *in trans* was observed. The plasmid pChbRN238S carries the *chbRN238S* allele which is a predominant mutation seen in several Cel$^+$ mutants (Table 2). Another *chbR* clone, pChbRY30C, carries the *chbRY30C* allele amplified from the Cel$^+$ natural isolate AHK3PI. The wild-type and mutant *chbR* clones were introduced into different reporter strains by transformation and β-galactosidase assays were performed in the presence or absence of 10 mM cellobiose as inducer.

In the strain JM-chb22 (*nagC*::*tet; chbR*::*cat*), presence of both *chbR* mutants resulted in a high basal level of expression. More importantly, the transformant carrying ChbRN238S showed an approximately threefold induction over the basal level in the presence of 10 mM cellobiose whereas no induction was seen in the presence of wild-type ChbR (Fig. 3). The overall activated level in the

**Table 2.** List of plasmids used in this study.

| Plasmid | Genotype and description | Source |
|---|---|---|
| pBR322 | Tet$^r$, Amp$^r$ | Boliver *et al.* (1977) |
| pChbR WT | *chbR*$^+$ in pBR322 | This work |
| pChbRN238S | *chbR N238S* in pBR322 | This work |
| pChbRY30C | *chbR Y30C* in pBR322 | This work |
| pChbRL136S | *chbR L136S* in pBR322 | This work |
| pRA197T | *chbR A197T* in pBR322 | This work |
| pRNC2.1 | *chbR A197T;Q172H* in pBR322 | This work |
| pNWT | *nagC*$^+$ in pBR322 | This work |
| pACDH | *plac* followed by MCS, pACYC origin of replication | Singh *et al.* (2005) |
| placB-F | *plac-chbB*$^+$*C*$^+$*A*$^+$ *chbR::kan chbF*$^+$ in pACDH | This work |
| pOP-A | *chbOP*$^+$*B*$^+$*C*$^+$*A*$^+$ in pBR322 | This work |
| pChbF | *chbF*$^+$ in pBR322 | This work |
| pChbNC2.1 | *chbOPBCARF* (amplified from NC2.1 Cel$^+$) in pBR322 | This work |
| pDRIVE | pUC origin TA cloning vector (Qiagen) *plac–lacZ′* | Qiagen |
| pJES307 | T7 polymerase expression vector for expressing specific native proteins | Tabor and Richardson (1985) |
| pJES (ChbR) | pJES301 *chbR*$^+$ (NdeI–BamHI) | Plumbridge and Pellegrini (2004) |
| pJES (N238S) | pJES301 *chbR N238S* (NdeI–BamHI) | This work |

 

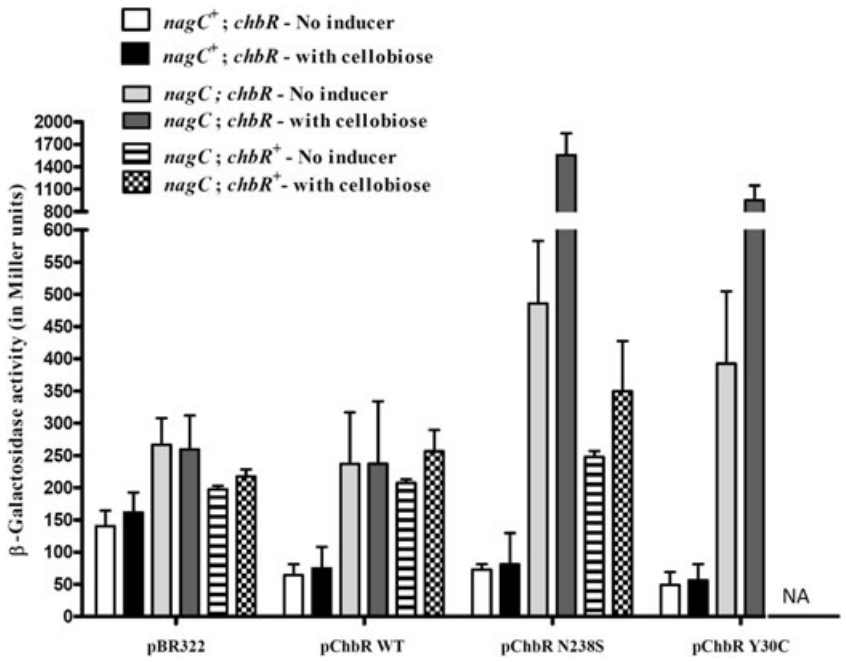


**Fig. 3.** β-Galactosidase assays using reporter strains carrying the wild-type and mutant *chbR* clones, in M9 medium containing 0.4% glycerol and 0.4% casamino acids, with or without 10 mM cellobiose. The reporter strains JM-chb 21 (*chbR*::*cat*), JM-chb22 (*nagC*::*tet; chbR*::*cat*) and JM-chb3 (*nag*::*tet*) carrying a single-copy *chbOPB′–lacZ* transcriptional/translational fusion, were used. The results are the mean of three to six independent measurements.

presence of the mutant was approximately sevenfold higher as compared with that in the presence of the wild-type ChbR. The presence of ChbRY30C resulted in an approximately twofold induction by cellobiose over the basal level. This induction by cellobiose was specific as no induction was seen in the presence of similar concentrations of other substrates such as *N*-acetylglucosamine or salicin. A modest repression was seen in the presence of glucose (data not shown).

All Cel$^+$ strains harbouring mutations at the *chbR* locus also contained loss-of-function mutations at the *nagC* locus. This suggested that the phenotypic effect of the mutations within *chbR* can be elicited only in the absence of negative regulation by NagC. The β-galactosidase assays performed in JM-chb21 strain containing the wild-type copy of *nagC* indicated that neither of the *chbR* mutants could activate transcription in the presence of cellobiose (Fig. 3). The repression by NagC was epistatic over the transcriptional activation by mutant ChbR in the presence of cellobiose. Transformants carrying the wild-type and mutant *chbR* clones in fact showed an approximately twofold reduction in promoter activity in the presence of the wild-type *nagC* (Fig. 3). This reduction could be attributed to the role of ChbR as a repressor in the presence of NagC.

In order to see if the mutant ChbR proteins could activate transcription in the presence of a genomic copy of wild-type *chbR*, β-galactosidase assays were carried out in the strain JM-chb3 (*nagC*::*tet; chbR*$^+$). ChbRN238S could not activate transcription in the presence of 10 mM cellobiose (Fig. 3), indicating that the wild-type *chbR* is dominant over *chbR N238S*.

*The role of mutations within the* chbR *and* nagC *loci in conferring a Cel$^+$ phenotype*

In an attempt to observe the effect of *chbR* mutations on the cellobiose utilization phenotype, transformants of JM-chb22 (*nagC*::*tet; chbR*::*cat*) carrying the wild-type and mutant *chbR* clones described above were streaked on MacConkey cellobiose plates. Transformants carrying the mutant *chbR* clones conferred a Cel$^+$ phenotype to the strain whereas those carrying the wild-type *chbR* clones remained Cel$^-$. Interestingly the ability of cellobiose utilization could be correlated with the transcriptional activation seen in the presence of cellobiose. The higher the induction in the presence of cellobiose, as seen in reporter assays, the faster was the Cel$^+$ phenotype observed. The mutant ChbRN238S conferred a Cel$^+$ phenotype within 12 h of incubation at 37°C as compared with the mutant ChbRY30C that conferred a Cel$^+$ phenotype after 24 h of incubation at 37°C. However, mutant *chbR* clones after transformation into a strain JM-chb21 carrying the wild-type *nagC* locus could not confer cellobiose utilization phenotype (Fig. S1).

These results suggested that the wild-type permease (ChbBCA) and phospho-beta-glucosidase (ChbF), upon induction could recognize, transport and cleave cellobiose respectively. This was confirmed by cloning the wild-type genes encoding the permease and phospho-β-glucosidase under a heterologous promoter P*lac*. Transformants of the Cel$^-$ strain DH5α carrying the low-copy plasmid expressing the permease and the phospho-β-glucosidase could utilize cellobiose efficiently (Fig. S1). Transformants expressing either the *chb* per-

 

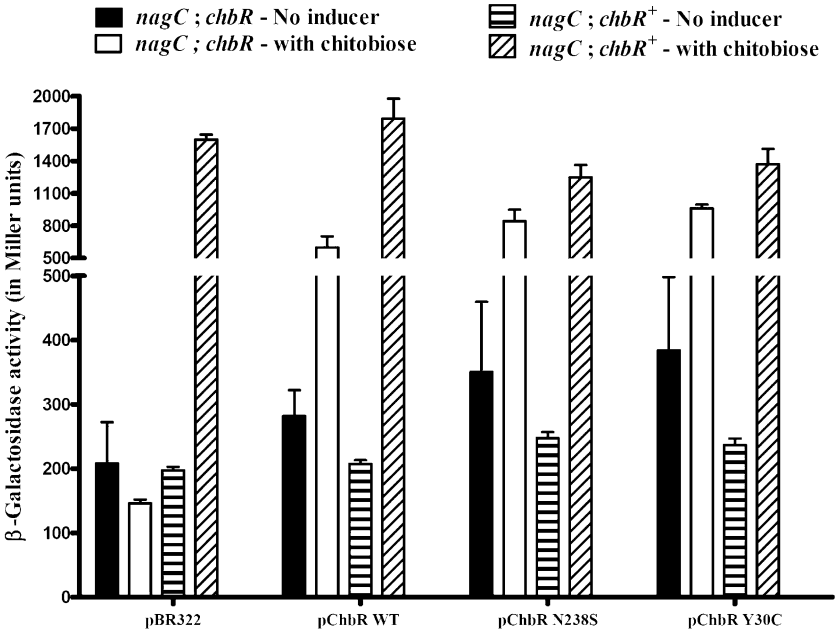


**Fig. 4.** β-Galactosidase assays performed using reporter strains carrying the wild-type and mutant *chbR* clones, grown in M9 medium containing 0.4% glycerol and 0.4% casamino acids, with or without 0.1% N,N′-diacetylchitobiose (2.4 mM). The reporter strains JM-chb22 (*nagC*::*tet; chbR*::*cat*) and JM-chb3 (*nagC*::*tet*) carrying a single-copy *chbOPB′–lacZ* transcriptional/translational fusion were used. The results are the mean of four independent measurements.

mease (*chbBCA*) or the phospho-β-glucosidase (*chbF*) alone could not confer Cel$^+$ phenotype. These results are consistent with the report that ChbF has a wide substrate specificity that includes phospho-cellobiose and phospho-salicin (Thompson *et al*., 1999). The results are also consistent with the observation that cellobiose can inhibit chitobiose transport by *chbBCA* (Keyhani *et al*., 2000a).

### *The effect of mutations in* chbR *on the recognition of chitobiose as an inducer*

To investigate whether Cel$^+$ mutant strains retained the ability to respond to chitobiose as an inducer, β-galactosidase assays were carried out in the reporter strain (JM-chb22) transformed with the wild-type and the mutant *chbR* clones described above. These assays showed that both mutants could activate transcription approximately twofold in the presence of 2.4 mM N,N′-diacetylchitobiose over and above the high basal level (Fig. 4), indicating that they retain the ability to recognize chitobiose. When the assays were performed in a reporter strain JM-chb3 carrying the wild-type *chbR* at the chromosomal locus and a disruption of *nagC*, the activity in the presence of chitobiose showed an induction of approximately seven- to eightfold (Fig. 4), indicating that the wild-type *chbR* at the genomic locus activates transcription better compared with *chbR* present on the plasmid. This is likely to be related to the low level of transcription from the endogenous plasmid promoter and the lack of positive autoregulation of *chbR* expression in the clones.

### *Variation in the amino acid sequence of ChbR across different species and genera and the role of the* chb *operon in Cel$^+$ natural isolates of* E. coli

To examine whether Cel$^+$ natural isolates of *E. coli* also carry mutations within the *chb* locus similar to those seen in *E. coli* K12, several Cel$^+$ strains were obtained from different sources and were further characterized. In addition to cellobiose utilization, some of the natural isolates could also utilize salicin and arbutin although they carried deletions of the *bgl* genes.

Sequence comparison at the *chbR* locus in all the Cel$^+$ natural isolates of *E. coli* showed changes at the DNA level resulting in the amino acid substitutions indicated in Table 3. The single-amino-acid change A197T seen in two of the natural isolates was also seen in the pathogenic strain O157::H7 (Sakai) and some strains of *Shigella sonnei*. Another Cel$^+$ natural isolate NC2.1 showed an additional change (Q172H) (Table 3). The *chbR* locus from *S. sonnei*, *E. coli* O157:H7 (Sakai) and the Cel$^+$ natural isolates (NC2.1, NC7.1) was amplified by PCR and cloned in pBR322. The *chbR* clones, when transformed into JM-chb22, could not confer a Cel$^+$ phenotype. However, the *chbR* clones obtained from the Cel$^+$ natural isolate MS201 (L136S) and AHK3 PI (Y30C) could confer a Cel$^+$ phenotype to JM-chb22.

The entire *chb* operon, along with the *chb* promoter, was amplified by PCR from the Cel$^+$ natural isolate NC2.1 and cloned in pBR322. The clone, when introduced into the strain JM-chb22 carrying disruptions of the chromosomal *nagC* and *chbR* loci, could not confer a Cel$^+$ phenotype, suggesting the presence of additional genetic

 

**Table 3.** Single-base-pair mutations in *chbR* resulting in single-amino-acid substitutions from different Cel$^+$ mutants and Cel$^+$ natural isolates.

| Cel$^+$ strain (*) | GenBank accession No. | Amino acid change in ChbR | Presence of insertion within *chbOP* | Cel$^+$ phenotype[a] on MacConkey cellobiose plates (after 24–48 h) |
|---|---|---|---|---|
| MG-MCP-02 | EF470556 | K184T | – | + |
| MG-MCP-07 | EF470554 | I227L | – | ++ |
| MG-MCP-12 | EF470558 | N137K | – | + |
| MG-M9-19, 21, 28 (3) | EF470561 | S135A | – | +± |
| J-MCP-01; J-M9-02-09, 11–14, 17, 18, 20, 21, 24 and 27; AHK5 (20) | EF470559 | N238S | – | ++ |
| J-MCP-03 and J-M9-16 (2) | EF470551 | F138S | – | ++ |
| J-MCP-17 | EF470555 | K184E | – | ++ |
| J-M9-01 | EF470558 | N137K | – | ++ |
| J-M9-10 | EF470552 | F138V | – | + |
| J-M9-19 and 22 (2) | EF470550 | F138L | – | ++ |
| J-M9-23 | EF470553 | F146L | – | ++ |
| MS201 | EF470557 | L136S | – | ++ |
| AHK3 PI | EF470562 | Y30C | – | ++ |
| NC2.1 | EF470560 | Q172H; A197T | – | ++ |
| NC7.1 | EF470549 | A197T | – | ++ |
| MG-M9-25 | EF577379 | E164K | + | + |
| MG-M9-26, 29 and MG-MCP-09 (3) | EF577380 | L263I | + | ++ |

**a.** The phenotypes are indicated as (+) light pink, (+±) pink, and (++) red.
*The numbers in the parentheses indicate the number of Cel$^+$ strains carrying identical *chbR* mutations.

systems for cellobiose utilization in this strain. Thus, many Cel$^+$ natural isolates of *E. coli* may have activated genetic systems other than the *chb* system such as the *asc* operon (Parker and Hall, 1988) for cellobiose utilization.

### *Biochemical studies using purified ChbR and ChbRN238S*

In an attempt to investigate the mechanism of transcriptional activation mediated by mutations in *chbR*, detailed biochemical studies using purified ChbR and ChbRN238S were carried out. Transformants carrying the *chbR* and *chbRN238S* recombinant plasmids were induced with IPTG. As in the earlier study (Plumbridge and Pellegrini, 2004) the wild-type protein could be purified by the use of the three columns, heparin sepharose, mono-S and hydroxyl-apatite (see *Experimental procedures*). The same protocol, however, did not allow purification of ChbRN238S as the mutant failed to bind heparin sepharose, suggesting possible differences in the biochemical properties of the mutant. ChbRN238S was purified using a mono-S column (Fig. S2). The estimated molecular mass of both the wild-type and the mutant ChbRN238S was ~32 KDa on 12% SDS-PAGE, as reported earlier (Plumbridge and Pellegrini, 2004).

### *Differential binding of ChbR and ChbRN238S to* chbOP

The intrinsic property of the ChbR to bind to the specific site within the *chb* regulatory region, *chbOP*, was exploited to find possible differences between wild-type ChbR and ChbRN238S, using the electrophoretic mobility shift assay (EMSA). The 188 bp *chbOP* regulatory region used as the target sequence in the assays is shown in (Fig. 5A).

The EMSAs were carried out using two concentrations (5 nM and 10 nM) of the wild-type and the ChbRN238S proteins (Fig. 5B). Wild-type ChbR formed two types of DNA–protein complexes with the specific target DNA: one that remained in the wells and could not be competed out by excess of cold specific DNA (unlabelled *chbOP*) and one that entered the gel and could be competed out by excess of cold specific DNA (Fig. 5B, lanes 1–4 and 5–8). On the other hand, ChbRN238S formed only a single DNA–protein complex that could be competed out by excess cold specific DNA (Fig. 5B, lanes 9–12 and 13–16). ChbRN238S was more efficient in forming the discrete DNA–protein complex compared with ChbR, as the complex can be detected even at the lower protein concentration (Fig. 5B, lanes 1–4 versus lanes 9–12). In addition, a lower concentration of cold specific DNA could compete away the complex in the case of ChbR, indicating that the affinity of ChbRN238S for the target DNA is higher. Both ChbR and ChbRN238S were unable to bind a 120 bp DNA fragment generated by NruI digestion of the 188 bp template and a 60 bp synthetic oligonucleotide dimer, carrying the direct repeats known to be the binding site for ChbR (data not shown).

### *DNase I footprinting assays with wild-type ChbR and ChbRN238S*

To further ascertain the difference between ChbR and ChbRN238S in their abilities to bind the specific site within the *chbOP* regulatory region, DNase I footprinting studies were carried out. These assays indicated that the ability of ChbRN238S to bind a direct repeat within *chbOP* was better than the wild-type counterpart (Fig. 6, compare lanes 1 and 2; 3 and 4). Furthermore, ChbRN238S pro-

 

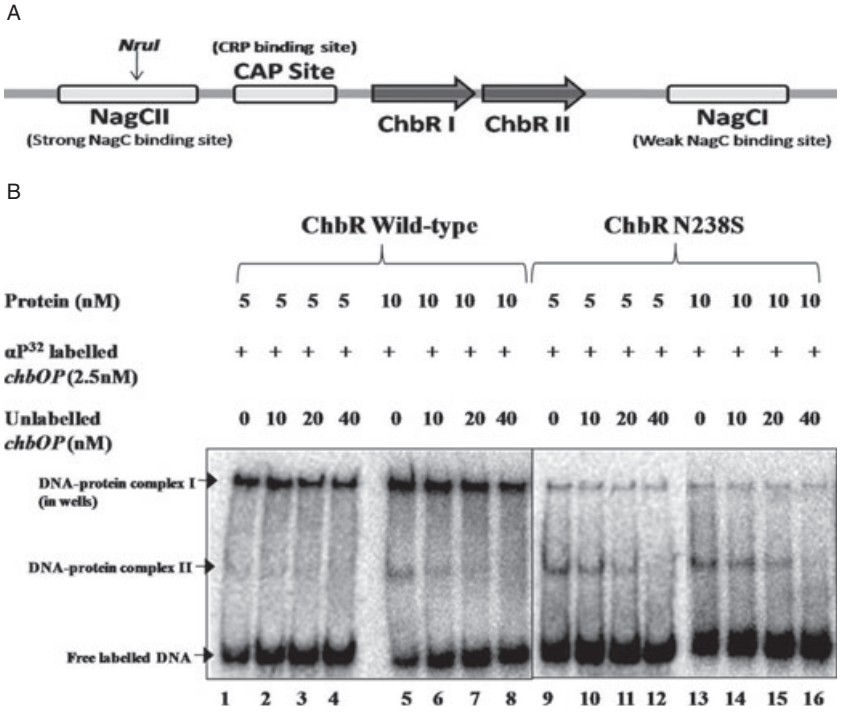


**Fig. 5.** A. The 188 bp regulatory region used for electrophoretic mobility shift assays. The NruI-digested DNA fragment (restriction site shown) was also used for the assays.
B. EMSA carried out with ChbR and ChbRN238S using body-labelled 188 bp *chbOP* (2.5 nM). The assay was carried out at two different concentrations of the protein (5 and 10 nM). Unlabelled *chbOP* was used at three different concentrations (10, 20 and 40 nM) to compete out the labelled DNA.

tected an additional 5–10 bp region downstream to the direct repeats compared with the wild-type ChbR (Fig. 6, lane 1 versus 2; 3 versus 4). The effect of longer incubation with DNase I on the protection of the specific binding site within *chbOP* was minimal, suggesting tight DNA binding. These results corroborate well with the DNA binding studies performed with the two proteins wherein the formation of a discrete DNA–protein complex was better in the case of ChbRN238S compared with ChbR (Fig. 5). The footprinting assays carried out with the wild-type and the mutant ChbRN238S, in the presence and absence of cellobiose 6-phosphate, did not show any alteration in the protection of the specific binding site within *chbOP* (Fig. 6, compare lanes 1–4 and 5–8). The DNase I footprinting assay may not be sensitive enough to pick up changes in the conformation of the mutant protein. Circular Dichroism studies were carried out to detect possible conformational change induced by the effector.

### *Conformational studies with ChbR and ChbRN238S*

Circular Dichroism spectroscopy was carried out in the far UV range and the results are presented in Fig. 7. The wild-type ChbR protein was predominantly α-helical with the two troughs corresponding to two known wavelengths (209 nm and 222 nm). ChbRN238S appeared to be much more structured compared with the wild-type protein as indicated by the higher negative values. The ability of a single-amino-acid change resulting in such changes in the conformation of the protein was noticeable. Upon incubating the proteins with the DNA corresponding to the *chb* regulatory region, there was a distinct conformational change in both the wild-type and the mutant proteins. No such change was induced by DNA fragments that did not show binding in EMSAs (data not shown). Upon the addition of cellobiose 6-phosphate (100 μM) to the DNA-bound wild-type ChbR and ChbRN238S, a distinct change in the spectrum was observed only in the case of the mutant protein (Fig. 7). This effect by cellobiose 6-phosphate was specific as there was no change in the spectrum when glucose 6-phosphate was used (data not shown). Negligible difference was observed in the spectra of the wild-type ChbR bound to DNA in the presence of cellobiose 6-phosphate or glucose 6-phosphate. These results are also consistent with the β-galactosidase assays carried out with the wild-type and mutant *chbR* that showed induction by cellobiose.

The denaturation kinetics of the wild-type and the mutant ChbR over a temperature range of 10–90°C at 222 nm indicated that the overall stabilities of the two proteins are similar (data not shown).

## Discussion

The experiments outlined in this study were carried out to determine the genetic mechanism of activation of the *chb* operon that enables the transport and catabolism of cellobiose. Analysis of cellobiose-utilizing derivatives obtained from a large-scale isolation of $Cel^+$ mutants,

 

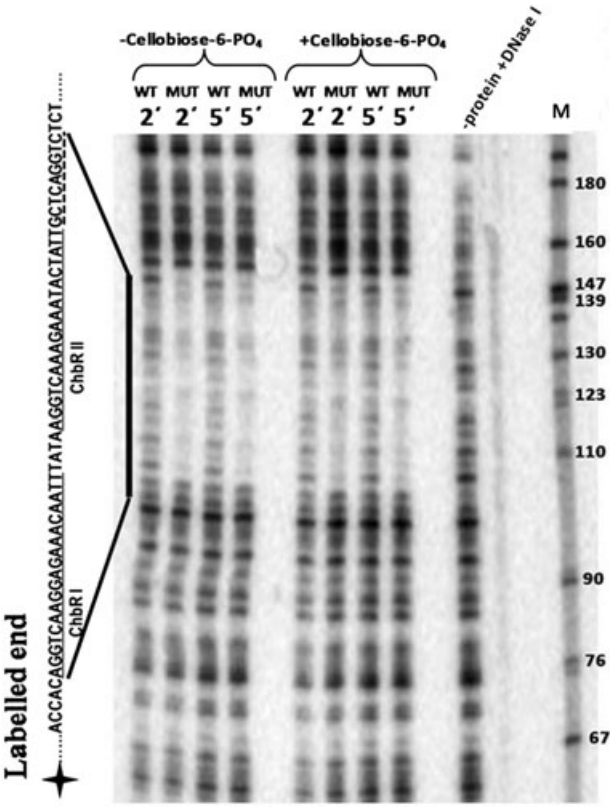


**Fig. 6.** DNase I footprinting assay carried out with ChbR and ChbRN238S (30 nM). The end-labelled *chb* regulatory region of ~340 bp (10 nM) was allowed to form the complex with the proteins at room temperature for 15 min. The complex was incubated with DNase I (5 ng $\mu l^{-1}$) for 2 and 5 min at 37°C before terminating the reaction. The protected regions of the 340 bp *chb* regulatory region along with the labelled end are shown. The dotted line below the sequence indicates the extra 5–10 bp protection by ChbRN238S.

carried out using different strains of *E. coli*, revealed the presence of two classes of mutations that act concertedly. These were mutations that resulted in the loss of NagC repression and gain-of-function mutations in the *chbR* locus. These two classes of mutations were necessary and sufficient to confer a Cel$^+$ phenotype. Contrary to earlier speculations, additional mutations in the structural genes were not necessary.

Is there any sequential order in which the two mutations occur? The chances of random mutations being loss-of-function mutations are usually higher and therefore accumulation of *nagC* mutations could be the first step in this process. So is the case of insertional disruption of the NagC binding site within *chbOP*. This may give a low-level growth advantage to the mutants, leading to the accumulation of gain-of-function mutation within the *chbR* locus. The frequency of mutations in a transcriptionally activated system is higher compared with loci that are silent; a process that is termed transcription-associated mutations or TAM (Beletskii and Bhagwat, 1996; Klapacz and Bhagwat, 2005). The partial derepression of *chb* transcription in *nagC* mutants could lead to a higher mutation rate at the *chbR* locus, leading to the selection of the second mutation in *chbR*. The sequential activation of the *chb* genes to confer a Cel$^+$ phenotype is similar to the activation of the *bgl* operon that carries an insertion within the *bglF* locus by two classes of mutations: activation of the silent promoter and excision of the insertion within *bglF* (Hall, 1988). This observation of a relatively high frequency of the activating mutations was among the initial results that suggested the existence of the phenomenon of adaptive mutations in bacteria (Foster, 1993).

The role of mutations in *chbR* was analysed using the two alleles that resulted in a strong Cel$^+$ phenotype: *chbR*N238S, the most abundant mutation obtained and the other, *chbR*Y30C. Clones carrying the mutant *chbR* could activate transcription in the presence of cellobiose in a reporter strain carrying disruptions of both the chromosomal *nagC* and *chbR* loci. This activation of transcription was lost when the assays were performed in a strain carrying the chromosomal *nagC* and a disruption of *chbR*,

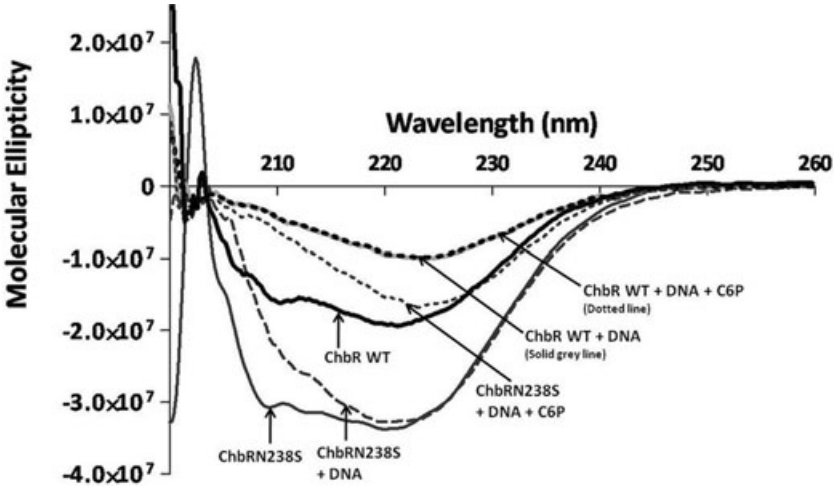


**Fig. 7.** Molecular ellipticity values of Circular Dichroism spectra (200–300 nm) of the wild-type ChbR and the mutant ChbRN238S at 1 μM concentration plotted on the *y*-axis. The *chbOP* (188 bp) DNA was added at 1 μM concentration. Cellobiose 6-phosphate [C6P] (100 μM) was used as an effector molecule.

 

underscoring the necessity of the loss of NagC repression in acquiring a Cel$^+$ phenotype.

ChbRN238S could not activate transcription in a reporter strain carrying the wild-type *chbR* at the chromosomal locus and a disruption of *nagC* in the presence of cellobiose indicating the dominant nature of wild-type *chbR* over *chbR*N238S. This result suggests that ChbR may function as a dimer/multimer. Interestingly, the mutant ChbR proteins retained the ability to induce transcription in the presence of chitobiose.

ChbR belongs to the family of AraC-XylS like transcription regulators (Gallegos *et al*., 1997; Tobes and Ramos, 2002). Broadly two characteristics are associated with these transcription factors – their ability to bind two different sites within the promoter region regulated by the presence or absence of the effector molecule and the induction of a conformational change upon binding the effector molecule. EMSA and DNase I protection assays confirmed that ChbRN238S has acquired the ability to form a more efficient discrete DNA–protein complex. ChbRN238S could additionally protect an extra stretch of ~10 bp DNA adjacent to the direct repeat. These results are consistent with the enhanced basal transcriptional activity seen in the case of ChbRN238S in reporter gene assays. Ellipticity measurements with the ChbR proteins indicated that, in the presence of cellobiose 6-phosphate and DNA, there is a conformational change in the case of ChbRN238S, possibly towards an activation state, which was absent in the case of the wild-type protein. This is also reflected in the reporter assays where a distinct induction over the high basal level could be observed in the presence of cellobiose. The enhanced transcriptional activity of ChbRN238S leading to a Cel$^+$ phenotype is therefore due to a higher basal level of expression that can be correlated to tighter binding to DNA, which is enhanced by the effector cellobiose 6-phosphate.

The mutant *chbR* alleles, when introduced into a strain lacking the chromosomal *nagC* and *chbR*, conferred on it a Cel$^+$ phenotype. This experiment indirectly suggested that the wild-type permease (*chbBCA*) and the phosphor-β-glucosidase (*chbF*) could transport and hydrolyse cellobiose. Direct evidence was provided by cloning the genes encoding the wild-type permease and β-glucosidase under a heterologous promoter. Therefore the acquisition of the Cel$^+$ phenotype is the result of alteration in the regulation of the *chb* operon. This observation is consistent with recent bioinformatic studies (Babu *et al*., 2006; Lozada-Chávez *et al*., 2006) which suggest that mutations in the transcription regulatory networks (TRNs) are primarily responsible for the variation in bacterial phenotypes. The results also underscore the flexibility of structural genes in recognizing substrates with similar chemical structures.

Preliminary studies were carried out with the Cel$^+$ natural isolates of *E. coli*. These studies showed that although a similar pathway of acquisition of a Cel$^+$ phenotype exists in some strains, additional genetic systems in cellobiose metabolism are also involved in other isolates.

The mechanism of activation of the *chb* operon to confer the ability of cellobiose utilization is presented as a model (Fig. 8). Wild-type *E. coli* is capable of generating an inducing signal for derepression of the *chb* operon when N,N′-diacetylchitobiose is used as an inducer. The *N*-acetyl glucosamine 6-phosphate that is generated after phosphorylation and subsequent hydrolysis of N,N′-diacetylchitobiose acts as an inducing signal for NagC, thereby relieving NagC repression. However, cellobiose is not capable of generating that inducing signal. Wild type *E. coli* can abrogate NagC binding to the strong NagCII binding site within *chbOP* either by transposition of IS elements within *chbOP* or by mutations in *nagC*. The derepression of the *chb* operon due to the loss of NagC regulation is still not sufficient to confer a Cel$^+$ phenotype. To change the specificity of the operon for cellobiose, gain-of-function mutations at the *chbR* locus are necessary.

The *chbR* locus emerges as a focal point for the evolution of cellobiose utilization in *E. coli*. The ChbR of *E. coli* K12, when compared across different natural isolates of *E. coli* and related genera, showed differences. The preliminary studies with *chbR* isolated from *S. sonnei*, *E. coli* O157:H7 and two Cel$^+$ natural isolates suggested the existence of polymorphism at the *chbR* locus. The *chbR* locus, when compared across unrelated genera, shows many differences. Wild-type strains of some of these bacteria (*Citrobacter*, *Yersinia pestis*, *Klebsiella*) are known to be Cel$^+$ and the *chb* operon is conserved in these bacteria. Some of the changes seen at the amino acid level in ChbR in *Yersinia* and *Citrobacter* were similar to the changes observed in ChbR isolated from various Cel$^+$ mutants of *E. coli* (Fig. S3). These observations suggest that the *chb* system has undergone divergent evolution over a period of time wherein one group of organisms has evolved to utilize primarily chitobiose and another group to utilize cellobiose. However, mutations in the regulatory genes might allow them to switch from one to the other depending on the environment. Bacteria can use this as an efficient strategy for rapid evolution under selective pressure.

## Experimental procedures

### *Construction of strains carrying disruption of* chbOP *and deletion of* chbR

The strain carrying an artificial insertion of the chloramphenicol acetyl transferase gene within the *chb* regulatory region (*chbOP*::*cat*) was constructed by targeted homologous recombination in DY330 using the plasmid pKD3 (Yu *et al*.,

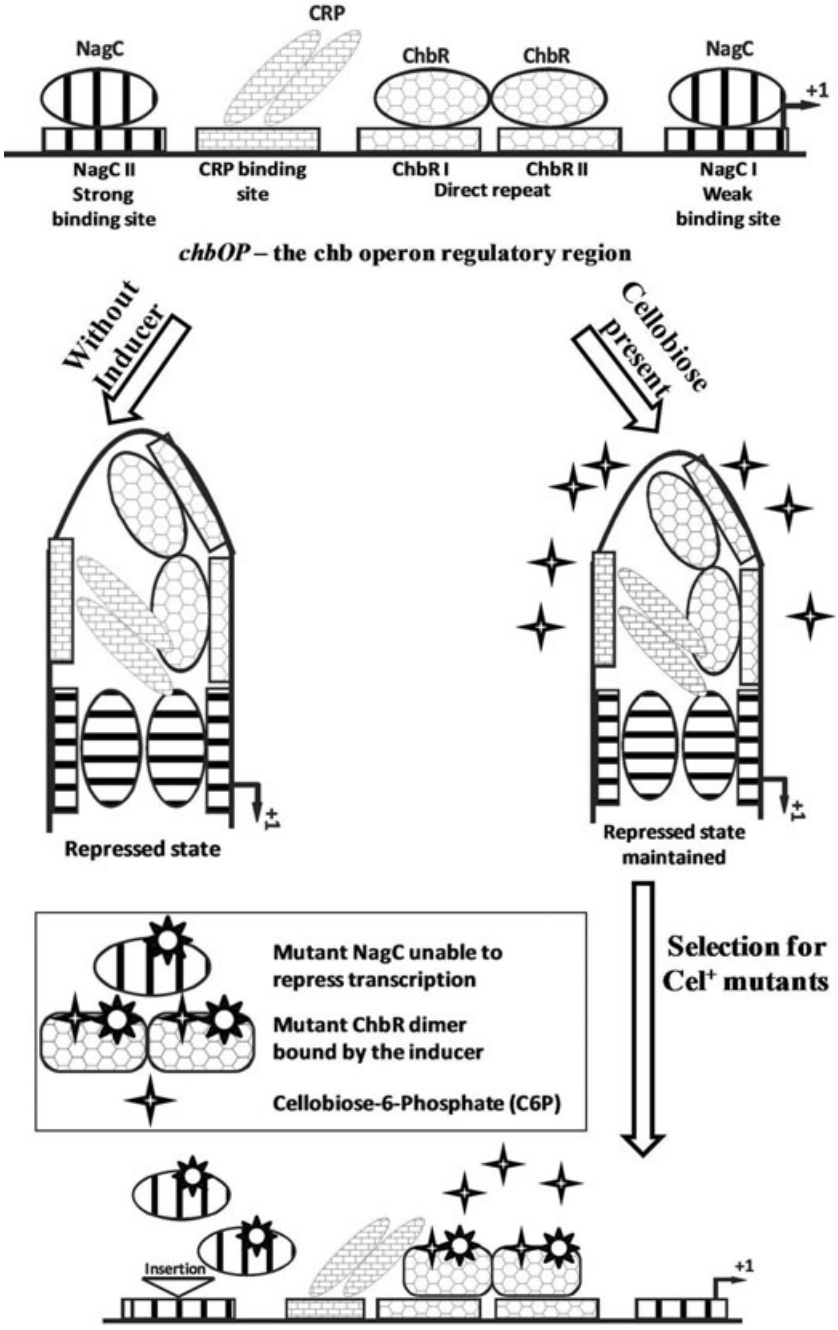


**Fig. 8.** A model for the mutational activation of the *chb* operon of *E. coli* allowing utilization of cellobiose. See text for details.

2000). The oligonucleotide primers were chosen such that the disruption mimicked a natural insertion seen in Cel$^+$ mutants. A PCR fragment carrying the *cat* gene flanked by *chbOP* sequences was introduced into DY330 after induction of the λ *red* recombination system and recombinants were selected on LB chloramphenicol medium. The strains DY330 (*chbR*::*cat*) was generated using a similar strategy. A PCR fragment covering the *cat* gene was amplified using pKD3 as a template and 56 nucleotides long primers. The 5′ 36 nucleotides of the forward and reverse primers corresponded to the N- and C-terminals of the *chbR* gene respectively. A similar strain, DY330 (*chbR*::*kan*), was constructed using the plasmid pKD4. The strain JF201 (*chbOP*::*cat*) was made by transducing the (*chbOP*::*cat*) allele into JF201, an *E. coli* strain deleted for the chromosomal *bgl* operon. Similarly, JF201 (*chbR*::*cat*) was constructed by transducing the *chbR*::*cat* allele from the parent DY330 (*chbR*::*cat*) into JF201.

### *Construction of plasmids*

The plasmid carrying wild-type *nagC* was constructed by PCR amplifying the *nagC* locus from JF201 using the primers AHK7 (5′-GCGAATTCATGACACCAGGCGGACAAGC-3′) and AHK8 (5′-GCGGATCCTTAATTTTCCAGCAAATGC-3′).

The ~1.1 kb PCR fragment was initially cloned in pDRIVE (Qiagen) and subcloned in pBR322.

The plasmid placB-F carrying the wild-type *chb* permease (*chbB*CA) and the phosphor-β-glucosidase (*chbF*) without the *chb* regulatory region was constructed by PCR amplifying a ~5.1 kb DNA fragment from DY330 (*chbR*::*kan*; Cel$^-$) using the forward primer AHK9 (5′-CGAGCTCGATGGAAAA GAAACACATT-3′) and the reverse primer AHK11 (5′-GGAATTCCAGCCTCGGTTAATGTGC-3′). The PCR fragment was digested using SacI and EcoRI restriction enzymes and cloned at these sites within pACDH, a low-copy vector derived from pACYC (Singh *et al.*, 2005). The individual clones containing the permease (*chbF*) and the phospho-β-glucosidase (*chbF*) were constructed by PCR amplification of the genes using appropriate sets of primers flanking the ORFs. The complete *chb* operon from the Cel$^+$ natural isolate NC2.1 was cloned by PCR amplifying a 4.5 kb DNA fragment containing *chbBCARF* using the primer set ASK1 (5′-GCGAATTCAACAACGGAAACC GGCC-3′) and AHK5 (5′-CGGATCCTTAATCGCCGGATGCAAGG-3′) using genomic DNA isolated from NC2.1. The EcoRI- and BamHI-digested PCR product was cloned at these sites in pBR322.

The plasmids used in the overexpression of ChbR and ChbRN238S were made as follows. The 850 bp *chbR* locus from wild-type and the mutant strains was amplified by PCR using Pfu DNA polymerase (MBI, Fermentas) and the forward primer AHK10 carrying an NdeI site (5′-CCATATG ATGCAGCCAGTGATTAACGC-3′) and the reverse primer AHK6 carrying a BamHI site (5′-CGGGATCCATATG TGAATTGTCAGGT-3′). The NdeI- and BamHI-digested DNA fragments were cloned into an expression vector pJES307 (Tabor and Richardson, 1985; Plumbridge and Pellegrini, 2004). The clones were confirmed by sequencing.

### *Method for large-scale isolation and characterization of Cel$^+$ mutants*

The *E. coli* strain JF201 was streaked on MacConkey cellobiose plates and was spread on M9 cellobiose plates (without casamino acids). Cellobiose used (Sigma Aldrich) was free of glucose. The plates were incubated for the first 24 h at 37°C and subsequently incubated at room temperature (~25°C). Cel$^+$ mutants appeared as papillae on existing colonies on the MacConkey cellobiose plates and as distinctly growing colonies on M9 cellobiose plates. The Cel$^+$ mutants started to appear after 10 days on M9 cellobiose plates and after 20 days on MacConkey cellobiose plates. A total of 53 single papillae or colonies were picked and streaked on fresh MacConkey cellobiose plates. The papillae picked up from MacConkey cellobiose plates are named JF-MCP-01 to JF-MCP-28 and the Cel$^+$ mutants picked up from M9 cellobiose plates are named J-M9-01 to J-M9-25. A total of 62 Cel$^+$ mutants were similarly isolated from the strain MG1655. The Cel$^+$ mutants picked up from MacConkey cellobiose plates were named MG-MCP-01 to MG-MCP-32 and the Cel$^+$ mutants picked up from M9 minimal cellobiose plates were named MG-M9-01 to MG-M9-30.

### *DNA sequencing and analysis*

DNA sequencing was carried out at Macrogen, Korea and the in-house facility at the Indian Institute of Science, Bangalore, India. Both strands were sequenced in all cases. Analysis of the DNA and protein sequences was performed using the Clone Manager-5 software, SciEd.; Vector NT1/Align software, Invitrogen; CLUSTALW (http://www.ebi.ac.uk/clustalw/ http://www.ch.embnet.org/software/ClustalW.html); and NCBI BLAST (http://www.ncbi.nlm.nih.gov/BLAST/).

### *β-Galactosidase assay*

Assays for β-galactosidase activity were carried out as described by Miller (1972). Cells were grown in LB broth or in M9 minimal medium supplemented with 0.4% glycerol and 0.4% casamino acids. Cellobiose (10 mM) or 0.1% (2.4 mM) N,N′-diacetylchitobiose (Seikagaku Corporation, Japan) were used as inducer when required. Average values of Miller units of activity were computed based on at least three independent measurements in each case.

### *Overexpression and purification of wild-type ChbR and ChbRN238S*

Wild-type ChbR was purified as per published protocol (Plumbridge and Pellegrini, 2004). The plasmid pJES ChbR was introduced into the *E. coli* strain Rosetta (DE3) by transformation, selecting for resistance to ampicillin and chloramphenicol. Cultures (1000 ml) in 2× LB broth containing ampicillin (100 μg ml$^{-1}$) were grown at 37°C to 0.7 $OD_{600}$. T7 RNA polymerase synthesis was induced by the addition of 0.5 mM IPTG. The cultures were harvested after 3 h, pelleted at 5000 r.p.m. for 5 min. The pellet was re-suspended in 25 ml of buffer B (20 mM MES, pH 6.5, 1 mM EDTA and 10% glycerol) containing 200 mM NaCl and a protease inhibitor cocktail (Sigma Aldrich) at 4°C. All subsequent steps were carried out at 4°C. The mixture was sonicated and centrifuged at 14 000 r.p.m. for 30 min. The supernatant was passed through a 0.45 μm filter (Sartorius) and the filtrate was applied on to a 1 ml heparin sepharose column (USB) pre-equilibrated with buffer B (with 200 mM NaCl). ChbR eluted at 600 mM NaCl. The subsequent step of purification on 5 ml of mono-S was similar to the published protocol (Plumbridge and Pellegrini, 2004). However, at the step of hydroxyl-apatite column purification, a modification was made at the step (iii) of washing carried out post the loading of the sample. Four steps of equilibration and washing were used: (i) protein was loaded on to the column and washed in buffer B with 400 mM NaCl, (ii) washed with buffer B containing 3 M NaCl, (iii) washed with buffer B containing 200 mM potassium phosphate, adjusted to pH 6.5, and (iv) eluted with a gradient 200–800 mM phosphate in buffer B. The solution containing pure wild-type ChbR protein was dialysed against buffer B containing 0.3 M NaCl and 50% glycerol and stored at –20°C.

ChbRN238S overexpression was performed by a procedure similar to that used in the case of wild-type ChbR. However, ChbRN238S could not be purified via heparin sepharose column although the purification was repeated many times with different salt concentrations (minimum of 150 mM NaCl below which the protein precipitates). The mutant protein was then loaded on to a mono-S cation exchange column and a linear gradient of 200 mM to 1 M NaCl was applied. The protein eluted at 450 mM NaCl concentration. The extractions containing partially purified

 

protein (~80%) were dialysed against buffer C (50 mM Tris pH 6.5 containing 200 mM NaCl) and applied onto a mono-Q anion exchange column (BIO-RAD). The protein eluted in the flow-through. The purified protein fractions (~90%) were dialysed independently with buffer B containing 50% glycerol and 300 mM NaCl and stored at –20°C for further use.

Different dilutions of wild-type ChbR and ChbRN238S were run on 12% SDS-PAGE for checking the purity and estimating the amount of the proteins. The concentrations were also checked using a Nanodrop spectrophotometer.

### *Labelling of DNA probes*

For generating internally labelled probe, PCR reactions were carried out in the presence of [$\alpha$-$^{32}$P]-dATP and the PCR fragments were purified from 2% agarose gel using Qiagen and Amersham gel purification kits. The eluate was further purified by passing through a Sephadex G50 column. The probe, after ethanol precipitation and wash, was re-suspended in appropriate volumes of sterile water. The purity of the DNA fragments was tested by electrophoresis on 12–15% polyacrylamide (19:1) or on 6% polyacrylamide gels containing 8 M Urea. Oligonucleotides were end-labelled using T4 polynucleotide kinase (MBI Fermentas) and [$\gamma$-$^{32}$P]-ATP. The labelling was performed as per the protocol provided by the manufacturer. The labelled DNA was purified by passing through Sephadex G50 mini column.

### *Electrophoretic mobility shift assay (EMSA)*

The template used in EMSA was a body-labelled 188 bp DNA fragment carrying *chbOP*, PCR amplified using the primers ASK1 (5′-GCGAATTCAACAACGGAAACCGGCC-3′) and AHK14 (5′-CGGATCCGGGCTGAAAGGAGTATACG-3′). In addition, a ~120 bp *chbOP* fragment generated by digestion of 188 bp fragment with NruI and a 60 bp synthetic oligonucleotide dimer containing the direct repeat known to be a specific binding site for ChbR were also used. The DNA concentration in binding reactions was 2.5 nM and protein concentrations were 5 and 10 nM in a final volume of 10–15 µl. The assays were performed in 25 mM HEPES buffer containing 100 mM K glutamate (pH 8.0). The mixture was incubated at 4°C for 15–30 min. Competition with unlabelled specific DNA was carried out by incubating the mixture with different concentrations of cold DNA (0–40 nM) for another 15 min. The samples were directly loaded onto a 6% native PAGE (1× TBE buffer system). The gels were run at 4°C and bands were visualized using a Bio-image Analyser (Fuji film, Japan). The intensity of the bands was measured using Fuji and Kodak ID imaging software.

### *DNase I footprinting assay*

The protocol for the assay was the same as that described by Plumbridge and Pellegrini (2004). The template used in the assays was a ~350 bp DNA fragment obtained by PCR using the end-labelled forward primer ASK1 (5′-GCGAATTCAACAACGGAAACCGGCC-3′) and the reverse primer ASK2 (5′-CGGGATCCTGATACCAGTAAAGAGG-3′). The proteins ChbR and ChbRN238S were used at a concentration of 30 nM. The DNase I concentration was 5 ng µl$^{-1}$ and the reactions were incubated at 37°C for 2 and 5 min. Assays were carried out in the presence or absence of 100 µM cellobiose 6-phosphate. The reactions were stopped by the addition of phenol. After ethanol precipitation of the aqueous layer overnight at –20°C, the samples were analysed on 6% denaturing polyacrylamide gels and quantified as above.

### *Circular Dichroism*

Circular Dichroism spectroscopy was carried out using a JASCO J-810 CD spectrometer. Measurements were made using 400 µl volumes of proteins (1 µM) in quartz cuvettes of 1 mm path length. All the assays were carried out in buffer B (containing 200 mM NaCl). The 188 bp DNA fragment carrying the *chbOP* region at 1 µM concentration was used in the assays. The DNA–protein complex was allowed to incubate for 15–30 min on ice after which the CD spectroscopy was carried out in the wavelength range of 200–300 nm at 20°C. To one set of the DNA–protein solutions, purified cellobiose 6-phosphate at a final concentration of 100 µM was added and incubated on ice for another 15 min before the spectroscopic measurement. The CD values were transformed into molecular ellipticity values using software from JASCO.

## Acknowledgements

We thank J. Plumbridge for many valuable suggestions and for providing single-copy reporter strains and J. Thompson for the gift of purified cellobiose 6-phosphate. We also thank D.N. Rao, U.K. Madhusoodanan and the Department of Biochemistry, Indian Institute of Science, for help with CD spectroscopy. We are grateful to the reviewers of the earlier version of the manuscript who suggested several improvements. This work was funded by a grant form the Council of Scientific and Industrial Research (CSIR), Government of India.

## References

Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277:** 1453–1474.

Babu, M.M., Teichmann, S.A., and Aravind, L. (2006) Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J Mol Biol* **358:** 614–633.

Beletskii, A., and Bhagwat, A.S. (1996) Transcription-induced mutations: increase in C to T mutations in the non-transcribed strand during transcription in *Escherichia coli*. *Proc Natl Acad Sci USA* **93:** 13919–13924.

Boliver, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., and Boyer, H.W. (1977) Construction and characterization of new cloning vehicles II. A multipurpose cloning system. *Gene* **2:** 95–113.

Foster, P. (1993) Adaptive mutation: the use of adversity. *Annu Rev Microbiol* **47:** 467–504.

Gallegos, M., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J. (1997) AraC/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* **61:** 393–410.

Hall, B.G. (1988) Adaptive evolution that requires multiple spontaneous mutations. I. Mutations involving an insertion sequence. *Genetics* **120:** 887–897.

Keyhani, N.O., and Roseman, S. (1997) Wild-type *Escherichia coli* grows on the chitin disaccharide, *N,N′*-diacetylchitobiose, by expressing the *cel* operon. *Proc Natl Acad Sci* **94:** 14367–14371.

Keyhani, N.O., Wang, L., Lee, Y.C., and Roseman, S. (2000a) The chitin disaccharide N,N'-diacetylchitobiose is catabolized by *Escherichia coli* and is transported/phosphorylated by the phosphoenolpyruvate: glycose phosphotransferase system. *J Biol Chem* **275:** 33084–33090.

Keyhani, N.O., Bacia, K., and Roseman, S. (2000b) The transport/phosphorylation of N,N'-diactylchitobiose in *Escherichia coli*. *J Biol Chem* **275:** 33102–33109.

Klapacz, J., and Bhagwat, A.S. (2005) Transcription promotes guanine to thymine mutations in the non-transcribed strand of an *Escherichia coli* gene. *DNA Repair* **4:** 806–813.

Kricker, M., and Hall, B.G. (1984) Directed evolution of cellobiose utilization in *Escherichia coli*. *Mol Biol Evol* **1:** 171–182.

Kricker, M., and Hall, B.G. (1987) Biochemical genetics of the cryptic gene system for cellobiose utilization in *Escherichia coli* K-12. *Genetics* **115:** 419–429.

Lozada-Chávez, I., Janga, S.C., and Collado-Vides, J. (2006) Bacterial regulatory networks are extremely flexible in evolution. *Nucleic Acids Res* **34:** 3434–3445.

Miller, J. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Parker, L.L., and Hall, B.G. (1988) A fourth *Escherichia coli* gene system with the potential to evolve β-glucoside utilization. *Genetics* **119:** 485–490.

Parker, L.L., and Hall, B.G. (1990a) Characterization and nucleotide sequence of the cryptic *cel* operon of *Escherichia coli* K12. *Genetics* **124:** 455–471.

Parker, L.L., and Hall, B.G. (1990b) Mechanisms of activation of the cryptic *cel* operon of *Escherichia coli* K12. *Genetics* **124:** 473–482.

Plumbridge, J., and Pellegrini, O. (2004) Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP. *Mol Microbiol* **52:** 437–449.

Reynolds, A.E., Mahadevan, S., Le Grice, S.F.J., and Wright, A. (1986) Enhancement of bacterial gene expression by insertion elements or by mutations in a CAP-cAMP binding site. *J Mol Biol* **191:** 85–95.

Singh, N.S., Das, G., Seshadri, A., Sangeetha, R., and Varshney, U. (2005) Evidence for a role of initiation factor 3 in recycling of ribosomal complexes stalled on mRNAs in *Escherichia coli*. *Nucleic Acid Res* **33:** 5591–5601.

Tabor, S., and Richardson, C.C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* **82:** 1074–1078.

Thompson, J., Ruvinov, S.B., Freedberg, D.I., and Hall, B.G. (1999) Cellobiose-6-phosphate hydrolase (CelF) of *Escherichia coli*: characterization and assignment to the unusual family 4 of glycosylhydrolases. *J Bacteriol* **181:** 7339–7345.

Tobes, R., and Ramos, J.L. (2002) AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucleic Acid Res* **30:** 318–321.

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., *et al.* (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acid Res* **17:** 2469–2478.

Wright, B.E. (2004) Stress-directed adaptive mutations and evolution. *Mol Microbiol* **52:** 643–650.

Yu, D., Elis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* **97:** 5978–5983.

## Supplementary material

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2958.2007.05999x
(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

 