Acta Biochimica Polonica

Vol. 44 No. 1/1997

153–158

QUARTERLY

Short communication

Characterization of the *Escherichia coli* gene encoding a new member of the short-chain dehydrogenase/reductase (SDR) family*⊡

Agnieszka Sirko✉, Anna Węgleńska, Monika Hryniewicz and Danuta M. Hulanicka

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5a, 02-106 Warsaw, Poland

Received: 21 January, 1997

Keywords: short-chain dehydrogenase/reductase, SDR, nucleotide sequence, *E. coli*

The nucleotide sequence of a chromosomal DNA fragment located upstream from the *cysPTWAM* operon of *Escherichia coli* was established. Sequence analysis indicates the presence of an open reading frame which has been designated *ucpA* (upstream *cysP*). The potential protein product exhibits strong sequence homology to the members of a large protein family, short-chain dehydrogenases/reductases. Involvement of Crp, FruR and IHF in the regulation of *ucpA* transcription *in vivo* was demonstrated.

Short-chain dehydrogenases/reductases (SDR) constitute a large protein family consisting of about 60 members sharing 15–30% of overall amino-acid sequence homology [1, 2]. Many of the first known SDRs were prokaryotic but recent additions include both animal and plant enzymes [2]. Several of the family members have been detected from nucleotide sequence just as open reading frames whose predicted products can be ascribed oxidoreductase functions by homology analysis.

Alignment of these proteins revealed clustered similarities indicating regions of special functional and structural importance [2]. Strongly conserved N-terminal halves, where a common GlyXXXGlyXGly pattern occurs, are responsible for coenzyme binding [3]. Several recent results suggest that the specificity for NAD(H) or NADP(H) can be determined by the presence of Asp or Arg residues, respectively, in the C-terminal end of βαβ-fold of the nucleotide binding domains of dehydrogenases [4–6]. The C-terminal halves of SDRs presumably have individual functions. In spite of variability in this region some similarities were noticed, e.g. TyrXXX-Lys pattern located close to the center of the proteins chains is strictly conserved [1, 2]. Site-specific mutagenesis of *Drosophila melanogaster* alcohol dehydrogenase confirmed the importance of that pattern and a

*This work was supported in part by a grant to M.H. from the State Committee for Scientific Research, project No. 6P20302206.

⊡The sequence data reported in this paper has been submitted to the EMBL/GenBank Data Libraries under the accession number X99908.

✉To whom correspondence should be addressed. Tel. (48-22) 659 70 72 ext. 5727 Fax: (48) 39 121 623. E-mail: asirko@ibbrain.ibb.waw.pl

catalytic mechanism based on these residues has been proposed [7]. Conservation of the critical residues suggests that the short-chain dehydrogenases/reductases are functionally related, while remaining overall residue similarities also suggest their structural relationship [2].

In this paper we describe cloning and sequencing (Fig. 1) of the chromosomal *Escherichia coli* region located on 52 min, upstream of *cysPTWAM* operon [8] and a partial overlapping of this sequence (from position 764) with the sequence M32101 is observed. We did not identify in EMBL/Genbank Sequence Libraries any sequences overlapping 5′ end of the sequence reported in this paper. The sequence analysis using the Genetic Computer Group software package [9] revealed the presence of an ORF of 789 bp (Fig. 1) designated by us *ucpA* (upstream *cysP*). The predicted amino-acid sequence consists of 263 residues with calculated molecular mass of 27.9 kDa. It shows a significant sequence homology with all members of the SDR family [1, 2]. The presence of the Asp residue at the C-end of a βαβ-fold of the nucleotide binding domain (Asp-37) suggests that the protein functions with NAD(H) as a coenzyme [5]. A molecular phylogenetic tree of most known members of SDR family (including a hypothetical product of *ucpA* gene) was constructed (Fig. 2) with the program Clustal W [10]. The tree shows that the UcpA is phylogenetically closest to the following SDR proteins: FixR from *Bradyrhizobium japonicum* (FIXR BRAJA), *N*-acylman-

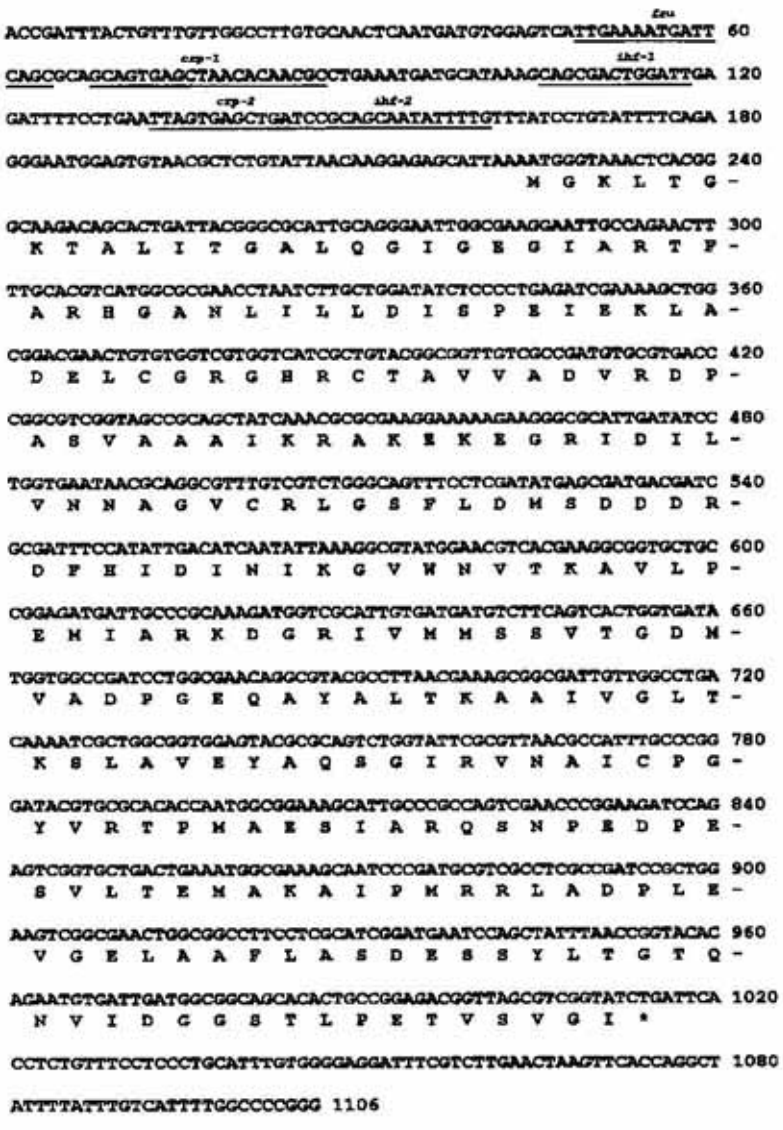

Figure 1. Nucleotide sequence and deduced amino-acid sequence of *ucpA*.

The potential binding sites for Crp (*crp-1* and *crp-2*), FruR (*fru*) and IHF (*ihf-1* and *ihf-2*) are underlined. The following consensus sequences has been used for localization of the potential binding sites: ANTGTGAN$_6$TCACANT for Crp [12], RSTGAAWCSNTHHW for FruR [13] and WATCAAN$_4$TTR for IHF [14]. The letter codes for nucleotide ambiguity are following: N, any nucleotide; R, A or G; S, C or G; W, A or T; H, A or C or T.

nosamine 1-dehydrogenase from *Flavobacterium* sp. (DHMA FLAS1), sporulation-specific protein SPX19 from *Saccharomyces cerevisiae* (SP19 YEAST), *cis*-1,2-dihydroxy-3,4-cyclohexadiene-1-carboxylate dehydrogenases, involved in aromatic hydrocarbons catabolism from two bacterial species: *Pseudomonas putida* (XYLL PSEPU) and from *Acinetobacter calcoaceticus* (BEND ACICA) and two hypothetical proteins of unknown function: Yjgi from *E. coli* (YJGI ECOLI) and Ynt3 from *Anabaena* sp. (YNT3 ANASP).

The analysis of the nucleotide sequence preceding *ucpA* revealed no typical –10 and –35 regions of σ^{70}-dependent promoters. It is known that sequences of positively controlled promoters deviate significantly from the consensus: $TTGACAN_{17}TATAAT$ [11]. In fact, the potential binding sites (Fig. 1) for three global, pleiotropic bacterial regulatory proteins: Crp [12], FruR [13] and IHF [14] have

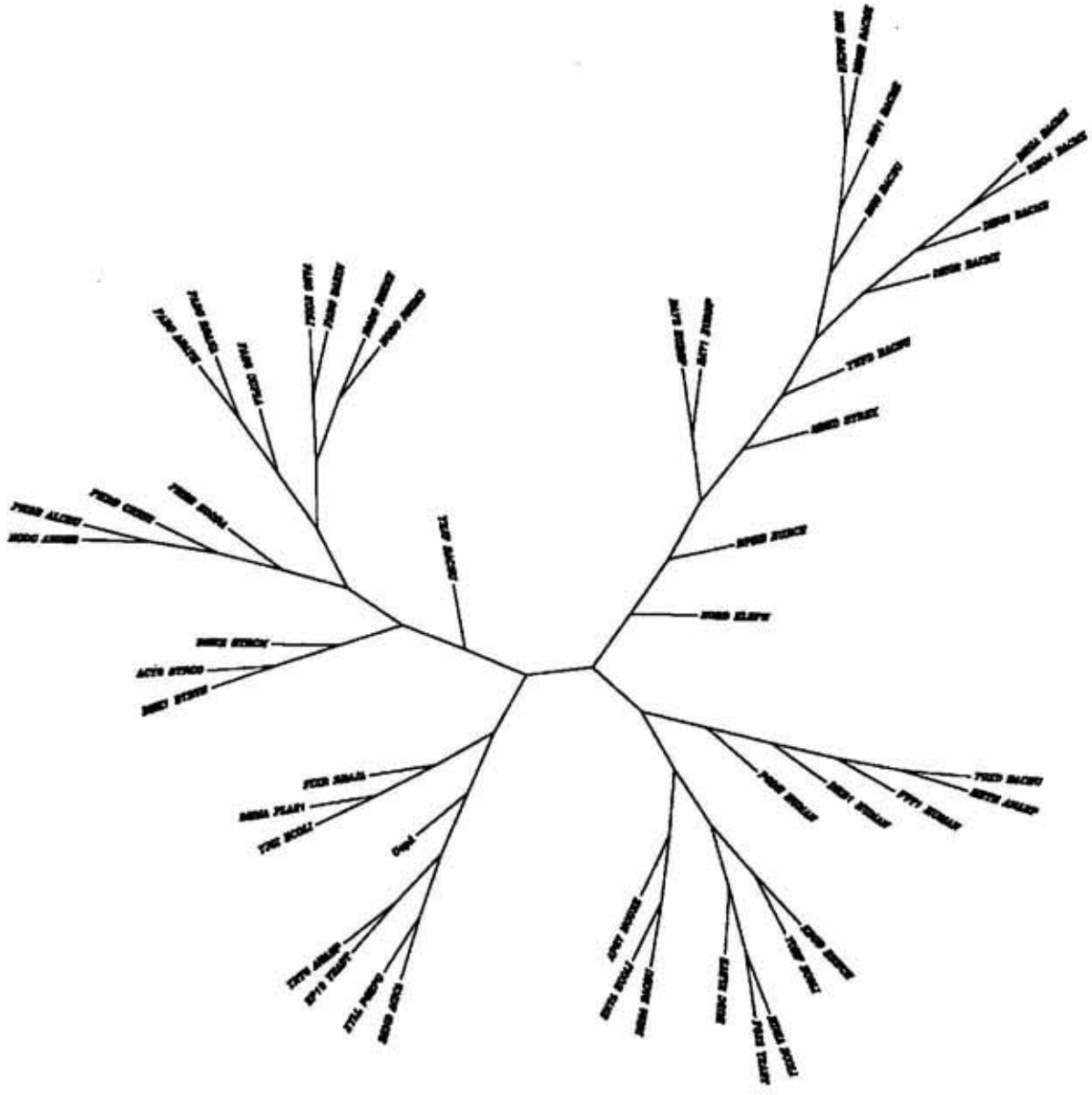

Figure 2. Phylogenetic tree of the proteins demonstrating the best scores of similarity to UcpA.

The database (SWISS-PROT) searching was performed with FastA program within the GCG sequence analysis software package [9] and the tree was constructed with the program Clustal W [10].

been identified. The effects of the mutations in genes encoding these regulators on expression of the transcriptional *ucpA-lacZ* fusion constructed by us were studied and are shown in Fig. 3. It appears that both Crp and FruR are required for full activation of *ucpA* transcription. We observed about 8- and 2-fold lower activities of β-galactosidase in the *crp* or *fruR* mutants, respectively, in comparison to the wild-type strain (Fig. 3). A slight FruR effect is in agreement with the data indicating that the genes exhibiting minimal FruR dependence are usually strongly dependent on other transcriptional regulator(s) [13]. In our case Crp would be the "major regulator" of *ucpA* transcription.

In *ihfA* and *ihfB* mutants devoided of α and β subunit of IHF, respectively, the expression of *ucpA-lacZ* was almost 2-fold higher in comparison to the wild-type strain; it suggests that IHF negatively affects the expression of *ucpA*.

Our attempts to identify the metabolic pathway in which the potential UcpA protein is involved by phenotypic characterization of a deletion-insertion *ucpA* mutant using API-32E tests (Biomerieux) were unsuccessful. The exact function of this gene remains to be characterized further.

In summary, we present here the nucleotide sequence of an ORF encoding a new member of a short-chain dehydrogenase/reductase family. Transcription of the *ucpA* gene is positively controlled by Crp and FruR and negatively affected by IHF.

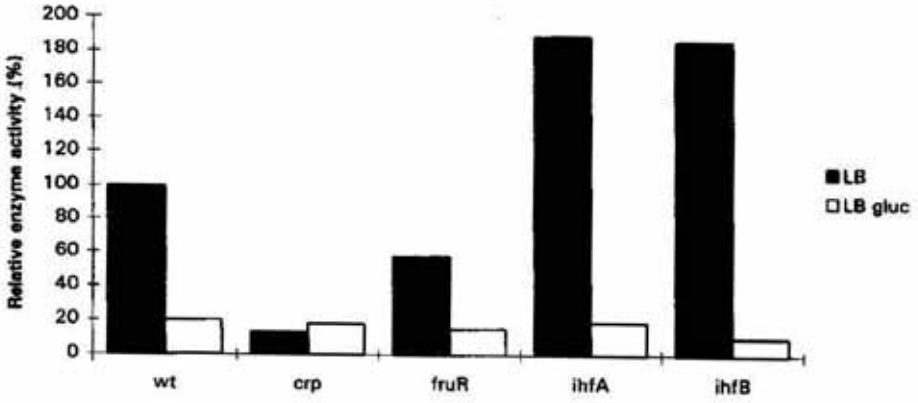

Figure 3. The effects of *crp*, *fruR*, *ihfA* and *ihfB* mutations on the expression of *ucpA-lacZ* fusion.

The fusion has been constructed *in vitro* and then introduced into bacteriophage lambda attachment site (*attL*) of *E. coli* chromosome using the method of Simons *et al.* [15]. Mutations *crp*, *fruR*, *ihfA* and *ihfB* were introduced into the strain containing *ucpA-lacZ* fusion by P1 transduction. Bacteria were grown in LB or LB supplemented with 0.5% glucose (LB gluc). Assays of β-galactosidase were performed in logarithmic cultures as described by Miller [16] and are shown in relation to the enzyme activity of the LB-grown *ucpA-lacZ* fusion strain in wild-type background (designated as 100%).

We thank Drs S. Adhya, M. Saier Jr. and M. Freundlich for bacterial strains. We are grateful to Dr M. Gniadkowski for his help with API tests.

REFERENCES

1. Persson, B., Krook, M. & Jörnvall, H. (1991) *Eur. J. Biochem.* **200**, 537–543.
2. Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J. & Ghosh, D. (1995) *Biochemistry* **34**, 6003–6013.
3. Wierenga, R.K., De Maeyer, M.C.H. & Hol, W.G.J. (1985) *Biochemistry* **24**, 1346–1357.
4. Bocanegra, J.A., Scrutton, R. & Perham, R.N. (1993) *Biochemistry* **32**, 2737–2740.
5. Scrutton, N.S., Berry, A. & Perham, R.N. (1990) *Nature* **343**, 38–43.
6. Chen, Z., Lee, W.R. & Chang, S.H. (1991) *Eur. J. Biochem.* **202**, 263–267.

7. Chen, Z., Jiang, J.C., Lin, Z.-G., Lee, W.R., Baker, M.E. & Chang, S.H. (1993) *Biochemistry* **32**, 3342–3346.

8. Hryniewicz, M., Sirko, A., Palucha, A., Böck, A. & Hulanicka, D. (1990) *J. Bacteriol.* **172**, 3358–3366.

9. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.

10. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) *Comput. Appl. Biosci.* **10**, 19–29.

11. Raibaud, O. & Schwartz, M. (1984) *Annu. Rev. Genet.* **18**, 173–206.

12. Kolb, A., Busby, S., Buc, H., Garges, S. & Adhya, S. (1993) *Annu. Rev. Biochem.* **62**, 749–795.

13. Ramseier, T.M., Bledig, S., Michotey, V., Feghali, R. & Saier, M.H., Jr. (1995) *Mol. Microbiol.* **16**, 1157–1169.

14. Goosen, N. & van de Putte, P. (1995) *Mol. Microbiol.* **16**, 1–7.

15. Simons, R.W., Houman, F. & Kleckner, N. (1987) *Gene* **53**, 85–96.

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