Mol Gen Genet (1988) 214:158–161

MGG

Short communications

The *Vitreoscilla* hemoglobin gene: Molecular cloning, nucleotide sequence and genetic expression in *Escherichia coli*

Chaitan Khosla and James E. Bailey
Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Summary. *Vitreoscilla* hemoglobin is involved in oxygen metabolism of this bacterium, possibly in an unusual role for a microbe. We have isolated the *Vitreoscilla* hemoglobin structural gene from a pUC19 genomic library using mixed oligodeoxy-nucleotide probes based on the reported amino acid sequence of the protein. The gene is expressed in *Escherichia coli* from its natural promoter as a major cellular protein. The nucleotide sequence, which is in complete agreement with the known amino acid sequence of the protein, suggests the existence of promoter and ribosome binding sites with a high degree of homology to consensus *E. coli* upstream sequences. In the case of at least some amino acids, a codon usage bias can be detected which is different from the biased codon usage pattern in *E. coli*. The downstream sequence exhibits homology with the 3′ end sequences of several plant leghemoglobin genes. *E. coli* cells expressing the gene contain greater than fivefold more heme than controls.

Key words: *Vitreoscilla* – Hemoglobin – Cloning – Nucleotide sequence

Introduction

The filamentous bacterium, *Vitreoscilla*, a member of the Beggiatoa family, is a strict aerobe that is found in oxygen-poor environments such as stagnant ponds and decaying vegetable matter. Growth of this bacterium under hypoxic conditions results in severalfold induction of the synthesis of a homodimeric soluble heme protein (subunit mol. wt. 15775) (Tyree and Webster 1978; Boerman and Webster 1982) which has a remarkable spectral (Webster and Liu 1974), structural (Wakabayashi et al. 1986), and kinetic (Orii and Webster 1986) homology with eucaryotic hemoglobins. The conservation of most features characteristic for eucaryotic hemoglobins as well as its possible role in oxygen utilization in *Vitreoscilla* strongly suggest that this protein is indeed a bacterial hemoglobin.

We report here the molecular cloning and nucleotide sequence of the *Vitreoscilla* hemoglobin gene and some of its flanking regions. Interestingly, functional expression of this gene improves the aerobic growth properties of recombinant *Escherichia coli*, as determined by fed-batch fermentations and respirometer measurements on whole cells (Khosla and Bailey 1988).

Offprint requests to: J.E. Bailey

Materials and methods

Media and growth conditions. E. coli cells were grown at 37° C in L broth (Maniatis et al. 1982). *Vitreoscilla* sp. was grown in a medium containing 1.5% yeast extract, 1.5% peptone, and 0.02% sodium acetate (pH 8.0 with NaOH) at 30° C. Plasmid-containing cells were grown in medium supplemented with 100 mg/l ampicillin. All cells were grown in culture tubes or flasks in a New Brunswick shaker incubator set at 250 rpm.

DNA manipulations. Vitreoscilla genomic DNA was isolated according to the protocol of Silhavy et al. (1984). λgt10 and pUC19 genomic DNA libraries were made according to standard protocols (Huynh et al. 1985; Maniatis et al. 1982). For all other DNA manipulations, routine methods were used (Maniatis et al. 1982), unless otherwise stated below.

DNA sequencing. The *Hin*dIII-*Sph*I fragment from plasmid pRED2 (Fig. 2), which contains the entire structural gene (see Results and discussion), was subcloned in plasmid pUC19 in the same orientation as in pRED2. After digesting this new plasmid (pRED4) at the unique *Hin*dIII and *Mlu*I sites, the ends were 5′ labeled and 3′ labeled. Purified *Hin*dIII-*Mlu*I and *Mlu*I-*Sph*I fragments were then subjected to G-specific (Maxam and Gilbert 1980) and A-specific (Iverson and Dervan 1987) cleavage reactions. Samples were run on 8% and 20% polyacrylamide gels. The nucleotide sequence was read in the directions indicated by dashed arrows in Fig. 2.

Heme assay. Heme extracts were prepared and assayed according to the method of Lamba and Webster (1980). In each case 0.1 g cells (wet weight basis) were used. The heme concentration was estimated on a Shimadzu UV160 spectrophotometer using the extinction coefficient $E_{556-541} = 20.7$ per mM per cm for reduced minus oxidized spectra.

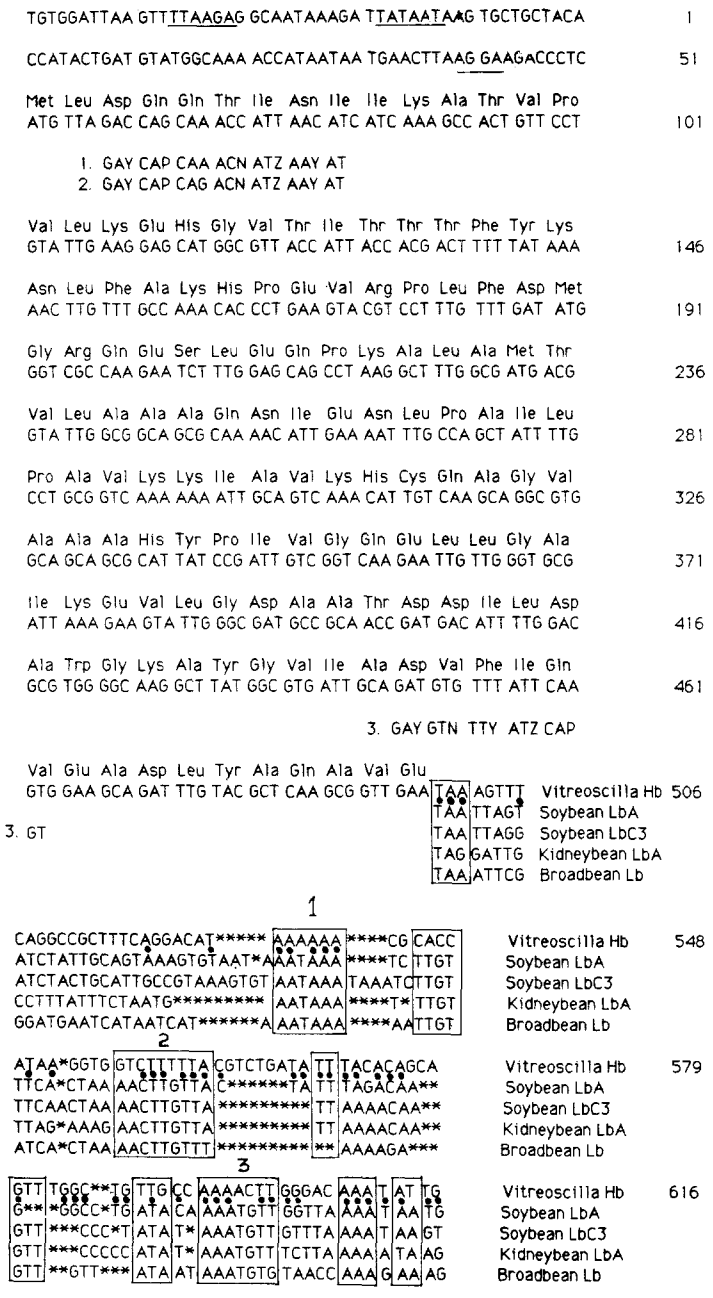

Fig. 1. Nucleotide sequence of the *Vitreoscilla* hemoglobin gene. The nucleotide sequence, a scheme of the sets of mixed oligonucleotides used to isolate the gene, and the amino acid translation of the structural gene are shown. Also shown is an alignment of the 3′ ends of leghemoglobin (Lb) genes from soybean (LbA, LbC_3), broad bean, and kidney bean (LbA). Three sets of mixed oligonucleotide probes (numbered 1, 2, and 3) are indicated by their IUB nomenclature below their respective target sequences. The alignment of the 3′ ends of the Lb genes is essentially the same as that of Kuhse and Puehler (1987) with a few changes to maximize base-for-base similarity. The *Vitreoscilla* sequence was computer-aligned to these sequences to maximize similarity using a standard difference-matrix algorithm. Regions of two or more adjacent bases which are conserved in at least four out of five species are *boxed*. *Dots* are used to indicate base similarity between the *Vitreoscilla* and soybean LbA sequences (see text). The putative ribosome binding site, Pribnow box and the -35 box are *underlined* in the upstream region of the structural gene

Results and discussion

Cloning the Vitreoscilla *hemoglobin gene*

Three sets of mixed oligonucleotide probes were synthesized which had a predicted homology to one 5′ and one 3′ domain in the hemoglobin gene (Fig. 1). Southern blot hybridization to *Eco*RI-digested genomic DNA established that the structural gene was present in a single fragment of around 7.24 kb and also enabled the discrimination between probe sets 1 and 2 (Fig. 1). Probe set 1 was identified as the correct set. Further, the minimum selective wash temperatures for probe sets 1 and 3 were also determined to be 50° and 46° C respectively. We were unable to isolate the gene from a λgt10-*Eco*RI library, probably due to the sharp reduction in packaging efficiency of this vector for inserts bigger than 7.6 kb (Huynh et al. 1985). Hence, Southern blot hybridizations to genomic DNA digested with *Bam*HI, *Hin*dIII, *Eco*RI and various combinations of these three enzymes were performed. The smallest gene-containing fragment was a *Hin*dIII fragment of 2.2 kb. A pUC19-*Hin*dIII library of *Vitreoscilla* DNA was test-plated on rich media containing ampicillin (100 mg/l), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (40 mg/l) and isopropyl-1-thio-β-D-galactopyranoside (IPTG) (1 mM). More than 70% of the colonies were probable recombinants, as estimated by visual inspection. About 40000 colonies were then screened. Three positive clones were identified and restreaked for single colonies. The plasmids from a small number of clones from each group were examined by *Hin*dIII digestion. One of these, pRED1, carried three *Hin*dIII inserts including one of 2.2 kb. This 2.2 kb fragment was purified and cloned into pUC19 in both orientations (pRED2 and pRED3). Digestion of pRED2 with various enzymes and Southern blot analysis confirmed that the entire hemoglobin structural gene in present on a *Hin*dIII-*Sph*I fragment (approx. 1.1 kb). The restriction map of plasmid pRED2 is shown in Fig. 2.

Interestingly, the *Vitreoscilla* genome had a statistically small number of restriction sites for *Bam*HI. There were only nine observable distinct bands smaller than 5.5 kb and four such bands smaller than 4 kb. A similar observation has also been made in the case of the genome of *Methanococcus voltae* which has an unusually small number of sites for the 4 bp recognizing enzyme, *Sau*3AI, which is an isoschizomer of *Bam*HI (Jarrell et al. 1987).

Transformation of Vitreoscilla

Attempts to transform *Vitreoscilla* with plasmid pRED2 using both the $CaCl_2$ (Shilhavy et al. 1984) and the Hanahan (1983) protocols were unsuccessful. However, neither of the competent cell making and transformation processes had resulted in a noticeable decrease in the viability of the cells. Since the existence of a natural plasmid in this *Vitreoscilla* strain has recently been reported (Dikshit and Webster 1987), this difficulty may have resulted from plasmid incompatibility.

Genetic expression of the hemoglobin gene in E. coli

E. coli cells containing plasmids pRED1, pRED2, pRED3 and pUC9 as well as *Vitreoscilla* cells were grown to stationary phase and cell extracts were examined on an SDS-polyacrylamide gel for the existence of the hemoglobin polypeptide. As can be seen in Fig. 3 (lanes 2–5), hemoglobin is

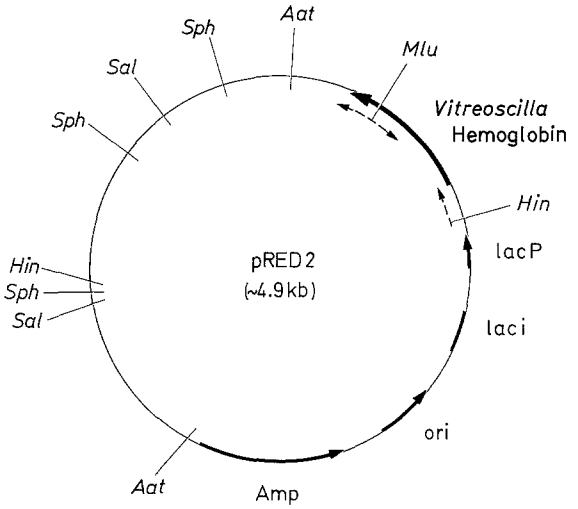

Fig. 2. Restriction map of plasmid pRED2. This plasmid is a derivative of pUC19. The position and orientation of the *Vitreoscilla* hemoglobin structural gene is as shown. The symbols Aat, Hin, Mlu, Sal, and Sph indicate approximately the recognition sites for the restriction enzymes *Aat*II, *Hin*dIII, *Mlu*I *Sal*I, and *Sph*I, respectively. The *Hin*dIII insert does not contain any recognition sites for the enzymes *Bam*HI, *Eco*RI, *Hae*II, *Nde*I, *Pst*I, *Pvu*I, *Sma*I, *Sst*I, or *Xba*I. The *dashed arrows* near the *Hin*dIII and *Mlu*I sites represent the directions of sequencing of labeled fragments (see Materials and methods)

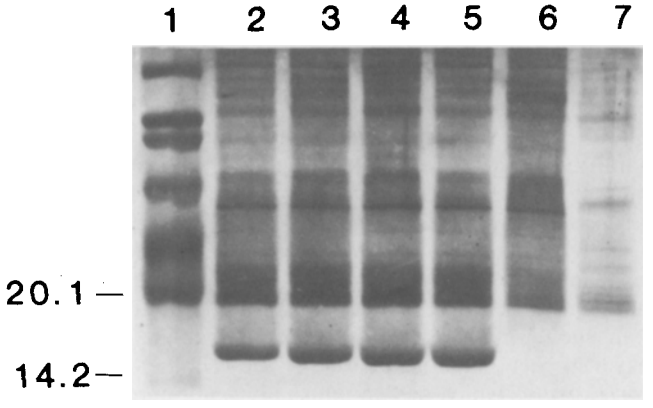

Fig. 3. Expression of hemoglobin in *Escherichia coli*. The total cellular protein content of cells containing various plasmid constructs was resolved on a 12.5% SDS-polyacrylamide gel. Lane 1, molecular weight standards. Lanes 2–4 and 6, *E. coli* cells containing plasmids pRED1, pRED2, pRED3 and pUC9 respectively. Lane 5, *E. coli* cells containing plasmid pRED2 (which has the hemoglobin gene cloned in the same orientation as the *lac* promoter, see Fig. 2) grown in the presence of 1 mM IPTG. Lane 7, the protein content of *Vitreoscilla*. Gels were run according to the standard protocol of Laemmli (1970). Protein was visualized by the silver staining method of Merril et al. (1983)

expressed as a major cellular protein in all the recombinant cells. Since both pRED2 and pRED3 express about equal amounts of this polypeptide and since the level of expression from pRED2 is about the same with and without IPTG, the gene is probably expressed from its natural promoter in *E. coli*. This conclusion is reinforced by an examination of the upstream sequences of the gene (see below). Such high level expression of the gene product in *E. coli* contrasts with the low level of expression obtained under better oxygenated growth conditions (Khosla and Bailey 1988) and suggests that the oxygen-dependent regulatory mechanism of this gene is functional in *E. coli*. We are presently studying this regulatory system.

Effect of the cloned gene on heme metabolism

Recombinant cells containing the hemoglobin gene were all observed to be red in color compared with JM101 or JM101/pUC9. Spectrophotometric assays for the heme content of these cells showed elevated heme concentrations in cells containing the *Vitreoscilla* hemoglobin gene. The heme content of JM101/pRED2 was 21 nmol/g cells compared with 3.8 nmol/g cells for JM101/pUC9. Correspondingly, the hemoglobin activity of JM101/pRED2 was fivefold higher than JM101/pUC9 (Khosla and Bailey 1988).

Nucleotide sequence of the structural hemoglobin gene and flanking regions

The complete nucleotide sequence of the structural hemoglobin gene along with its amino acid translation is shown in Fig. 1. The deduced amino acid sequence is in perfect agreement with the sequence reported by Wakabayashi et al. (1986). The region downstream of the structural gene has been computer-aligned with the untranslated 3′ ends of four leghemoglobin (Lb) genes which are known to have conserved sequence motifs (Kuhse and Puehler 1987).

Several interesting observations can be made from the nucleotide sequence shown in Fig. 1. Firstly, a significant homology can be observed between the 3′ end sequence of this gene and other plant globin genes. As shown in the figure, regions 1, 2 and 3 represent highly conserved sequence motifs in several plant leghemoglobin genes (Kuhse and Puehler 1987) which are also fairly conserved in *Vitreoscilla*. This suggests common evolutionary ancestry for this gene and plant leghemoglobin genes. It should be noted, however, that no functional significance has been assigned to this region in the case of the plant genes (Stougaard et al. 1987). Earlier, Wakabayashi et al. (1986) indicated that amongst several animal and plant globin amino acid sequences, the *Vitreoscilla* hemoglobin shows the maximum sequence homology (24%) with lupin leghemoglobin. Within the 3′ region indicated in Fig. 1, the *Vitreoscilla* hemoglobin shows the greatest homology (45%) with the soybean leghemoglobin A gene (indicated by dots in Fig. 1). Since the corresponding figure for the soybean LbA gene and the kidney bean LbA gene is 52%, this figure is indicative of significant homology. It should be noted that a putative procaryotic terminator, TTTTTA (Glass 1982), exists within region 2 (Fig. 1) with a possible stem-loop region. ACCAtaaggTGGT, preceding it (corresponding to another conserved region in the plant sequences). The *Vitreoscilla* and the soybean LbA sequences share a greater than 60% homology downstream of this putative terminator.

On examining the codon usage pattern within this structural gene, a distinct bias toward the usage of specific codons for some of the more abundant amino acids in the protein can be observed. For example, out of 14 leucine codons, 13 are UUG and 1 is UUA. Similarly, all phenylalanine codons are UUU and 5 out of 7 proline codons are CCU. In all 3 cases the preferred triplet for more abundant proteins in *E. coli* is different (Gouy and Gautier 1982).

Expression of the hemoglobin polypeptide as a major cellular protein in *E. coli* from its own promoter suggests

an homology between the expression control sequences of *E. coli* and *Vitreoscilla*. In Fig. 1, putative Pribnow and -35 boxes as well as ribosome-binding sites have been underlined. These show near-perfect homology to consensus *E. coli* sequences (Glass 1982). Comparison of lanes 2–5 and lane 7 in Fig. 3 indicates that the amount of hemoglobin in *E. coli* is much greater than that in *Vitreoscilla* grown under similar conditions. In our experience, growth of *Vitreoscilla* to early stationary phase in culture tubes placed in an air-circulated shaker incubator is usually not sufficient to induce a detectable reddish tint to the cells; however such a tint is observed in cases where the *Vitreoscilla* cells are grown under more oxygen-limited conditions. Whether this difference in expression in *E. coli* and *Vitreoscilla* is a gene dosage effect or due to a difference in regulation of gene expression in these two strains is as yet unclear. Introducing this gene into *Vitreoscilla* on a plasmid would be useful in this context.

Since the protein is active in *E. coli* and confers growth enhancement to *E. coli* in oxygen-poor environments (Khosla and Bailey 1988), molecular cloning and sequencing of this gene will facilitate a better understanding of its exact in vivo metabolic role. Work in this direction is currently in progress in our laboratory.

Acknowledgements. This work was supported by the Energy Conversion and Utilization Technology (ECUT) program of the United States Department of Energy. The authors are grateful to Dr. Webster for providing the *Vitreoscilla* sp. strain.

References

Boerman S, Webster DA (1982) Control of heme content in *Vitreoscilla* by oxygen. J Gen Appl Microbiol 28:35–42

Dikshit KL, Webster DA (1987) Molecular cloning and restriction mapping of *Vitreoscilla* plasmid pVT1. Abstracts Annual meeting of American Society for Microbiology, Atlanta, Georgia, p 147

Glass RE (1982) Gene function – *E. coli* and its heritable elements. University of California Press, Berkeley, Los Angeles, pp 51–109

Gouy M, Gautier C (1982) Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res 10:7055–7074

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Huynh TV, Young RA, Davis RW (1985) Construction and screening cDNA libraries in λgt10 and λgt11. In: Glover D (ed) DNA cloning: A practical approach. IRL Press, Oxford, pp 49–78

Iverson BL, Dervan PB (1987) Adenine specific DNA chemical sequencing reaction. Nucleic Acids Res 15:7823–7830

Jarrell KF, Julseth C, Pearson B, Kuzio J (1987) Paucity of the *Sau*3AI recognition sequence (GATC) in the genome of *Methanococcus voltae*. Mol Gen Genet 208:191–194

Khosla C, Bailey JE (1988) Heterologous expression of a bacterial haemoglobin improves the growth properties of recombinant *Escherichia coli*. Nature 331:633–635

Kuhse J, Puehler A (1987) Conserved sequence motifs in the untranslated 3′ end of leghemoglobin transcripts isolated from broadbean nodules. Plant Sci 49:137–143

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Lamba P, Webster DA (1980) Effect of growth conditions on yield and heme content of *Vitreoscilla*. J Bacteriol 142:169–173

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning – a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65:499–560

Merril CR, Goldman D, Sedman SA, Ebert MH (1983) Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. Science 211:1437–1438

Orii Y, Webster DA (1986) Photodissociation of oxygenated cytochrome o(s) (*Vitreoscilla*) and kinetic studies of reassociation. J Biol Chem 261:3544–3547

Silhavy TJ, Berman ML, Enquist LW (1984) Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Stougaard J, Sandal NN, Gron A, Kuehle A, Marcker KA (1987) 5′ analysis of the soyabean leghemoglobin lbc_3 gene: regulatory elements required for promoter activity and organ specificity. EMBO J 6:3565–3569

Tyree B, Webster DA (1978) The binding of cyanide and carbon monoxide to cytochrome o purified from *Vitreoscilla*. J Biol Chem 253:6988–6991

Wakabayashi S, Matsubara H, Webster DA (1986) Primary sequence of a dimeric bacterial hemoglobin from *Vitreoscilla*. Nature 322:481–483

Webster DA, Liu CY (1974) Reduced nicotinamide adenine dinucleotide cytochrome o reductase associated with cytochrome o purified from *Vitreoscilla*. J Biol Chem 249:4257–4260

Communicated by H. Hennecke

Received February 29, 1988