Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 8961–8965, December 1987
Biochemistry

High-level expression of sperm whale myoglobin in *Escherichia coli*

(gene synthesis/codon bias/protein engineering)

BARRY A. SPRINGER AND STEPHEN G. SLIGAR

Department of Biochemistry, University of Illinois, Urbana, IL 61801

Communicated by Gregorio Weber, September 4, 1987

ABSTRACT **Sperm whale myoglobin was expressed in *Escherichia coli* from a totally synthetic gene inserted in the expression vector pUC19. The gene was constructed as 23 overlapping oligonucleotides encoding both strands of the DNA. Gene synthesis provides several advantages over traditional eukaryotic gene-cloning techniques, allowing the incorporation of an efficient ribosome binding site, appropriate initiation and termination sequences, restriction enzyme sites for convenient subcloning and future mutagenesis, and frequently used codons for highly expressed *E. coli* genes. The sperm whale myoglobin expressed from the synthetic gene constituted ≈10% of the total soluble protein as holoprotein, indicating that iron-protoporphyrin IX biosynthesis and prosthetic-group incorporation are not limiting in the high-level expression of this heme protein in *E. coli*. We credit the use of frequently used *E. coli* codons for the observed high-level expression. The sperm whale myoglobin produced is stable, easily purified to homogeneity, and indistinguishable from commercially available sperm whale myoglobin by optical and magnetic spectroscopic methods.**

Sperm whale myoglobin (Mb) has been one of the most intensely studied proteins, as attested by the wealth of documented biochemical, biophysical, and spectroscopic data (1). It was the first protein structure determined to high resolution by x-ray crystallographic analyses (2) and, because of its physiological importance and availability, has served as a model system for the study of structure–function relationships in heme proteins. With a desire to use Mb as a model for investigations into the dynamics of protein folding, ligand binding, and conformational transitions, we undertook the complete gene synthesis of sperm whale Mb. We report the successful construction and cloning of this synthetic gene and the high-level expression of authentic sperm whale Mb in *Escherichia coli*. This provides a crucial starting point for site-directed *in vitro* mutagenesis studies designed to probe the structure–function relationships of Mb.

Successful expression of mammalian genes in a bacterial environment has often proved difficult even when promoters and Shine–Dalgarno sequences (ribosome binding sites) from highly expressed *E. coli* genes have been utilized (3–7). Numerous factors other than promoters and Shine–Dalgarno sequences may influence the expression of eukaryotic genes in prokaryotic environments, including mRNA stability, appropriate translational initiation sequences, and the efficiency of the translational machinery as reflected in codon frequencies and tRNA pools. Once expressed, resistance to protease digestion and timely synthesis and assembly of necessary prosthetic groups are paramount to the production of a stable and functional holoprotein. Of related concern is the possible accumulation of insoluble protein aggregates and inclusion bodies (8–11) or the toxic effects of an overexpressed protein. Of considerable interest was the work of Nagai, Thøgersen, and colleagues (12, 13) in which they achieved high-level expression of apoprotein from a human β-globin cDNA clone after fusing the gene to a 5′-terminal region of the bacteriophage λ *cII* gene. The λ cII protein is very stable in *E. coli*; therefore, fusion-protein constructions circumvent protein-degradation problems thought to prevent high-level expression of the globin cDNA clone. The same approach has been used successfully to express human Mb (14).

In an attempt to provide a system for the expression of large quantities of holo-Mb, the total synthesis of a sperm whale Mb gene was undertaken. This construction, consisting of 23 overlapping oligonucleotides, utilized biased *E. coli* codons for highly expressed *E. coli* genes and incorporated an efficient ribosome binding site with optimal spacing 5′ to the initiation codon.* Total gene synthesis also allowed the incorporation of initiation and termination sequences and convenient restriction enzyme sites for subcloning and mutagenesis. This synthesized gene was inserted into the vector pUC19 and resulted in the high-level production of soluble heme-containing sperm whale Mb in *E. coli*. The expressed Mb constitutes ≈10% of the total soluble cell protein and is highly stable. We suggest that the use of preferred codons for highly expressed *E. coli* genes is necessary for the high-level expression of heme-containing globin in *E. coli*. The expressed sperm whale Mb is present in the holoprotein state, indicating that *E. coli* heme biosynthesis and incorporation are sufficient to match the production of a highly expressed protein. This synthetic protein is spectroscopically and functionally indistinguishable from commercially available sperm whale Mb.

MATERIALS AND METHODS

Gene Design and Construction. Twenty-three oligodeoxynucleotides, typically 45–48 nucleotides in length, were synthesized with an Applied Biosystems (Foster City, CA) 380A DNA synthesizer at the University of Illinois (Urbana–Champaign) Biotechnology Center, using silica-based solid-state chemistry with proton-activated nucleoside phosphoramidites (15, 16). Individual oligonucleotides were purified by reverse-phase HPLC on a Vydac C_4 column (The Separations Group, Hesperia, CA), using a linear gradient of 84% 0.1 M triethylamine/acetic acid (pH 7.0) in 16% acetonitrile to 78% triethylamine/acetic acid (pH 7.0) in 22% acetonitrile over a 15-min period at a flow rate of 1.5 ml/min. Following collection, solvent was removed under vacuum. The remaining material was treated with 300 μl of 80% acetic acid at room temperature for 20 min to hydrolyze the protecting dimethoxytrityl group; an equal volume of ethanol was added and solvent was again removed under vacuum. Samples were resuspended in water, extracted three times with an equal volume of diethyl ether, and desalted by centrifugation

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: Mb, myoglobin.

*The sequence of the synthetic Mb gene is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03566).

through an Isolab's (Akron, OH) "spun column" packed with Sephadex G-50. Samples were precipitated with ethanol, resuspended in water, and frozen at −20°C until used. Four hundred picomoles of each oligonucleotide was phosphorylated using polynucleotide kinase (New England Biolabs). The gene was constructed in three separate fragments: *Pst* I–*Hin*dIII, *Hin*dIII–*Eco*RI, and *Eco*RI–*Kpn* I (Fig. 1). Oligonucleotides for each fragment were annealed by heating to 90°C for 1 min in ligase buffer, followed by slow cooling to room temperature. After ligation by phage T4 DNA ligase (Pharmacia), the three fragments were digested with either *Hin*dIII, *Eco*RI, or *Hin*dIII/*Eco*RI as necessary to generate the correct termini. Each fragment was cloned into pUC18 and used to transform *E. coli* strain TB-1 [*ara*, Δ(*lac–pro*), *strA*, *thi*, φ80d*lacZ*ΔM15, r⁻, m⁺; T. O. Baldwin, Texas A & M, College Station, TX]. Plasmid DNA from each of the three cloned Mb fragments was purified by centrifugation in CsCl equilibrium density gradients and subsequently cut with the appropriate enzymes to release the cloned inserts. The three fragments were combined with *Pst* I/*Kpn* I-digested pUC19 and ligated to generate the intact sperm whale Mb gene (pMb413). DNA manipulations were essentially as described by Maniatis *et al.* (19). Dideoxy sequencing on double-stranded templates was done essentially as described by Hattori and Sakaki (20), using Bethesda Research Laboratories sequencing reagents.

Protein Purification. *E. coli* TB-1 harboring pMb413 was grown at 37°C in LB (10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl per liter) in the presence of ampicillin (200 mg/liter). The dark-brown cells were harvested in late stationary phase and frozen at −70°C until used. Typically, 90 g of cells were thawed and lysed at 4°C for ≈16 hr in a final volume of 360 ml containing 50 mM Tris·HCl, (pH 8.0), 1.0 mM EDTA (pH 8.0), 0.5 mM dithiothreitol, 1.0 mM toluenesulfonyl chloride, 40 units of DNase I per ml, 3 units of RNase A per ml, and 2 mg of lysozyme per ml. The cell debris was removed by centrifugation, resuspended in the above lysis buffer, and sonicated for 30 min on ice. The cell debris was removed again by centrifugation and the supernatants were combined. The blood-red supernatant was stirred at 4°C and slowly brought to 60% saturation with ammonium sulfate by addition of the solid. Once 60% saturation was reached, the solution was allowed to sit at 4°C for 2 hr. The precipitate was collected by centrifugation and solid ammonium sulfate was added to the supernatant as above until 95% saturation was reached. The precipitate was collected by centrifugation and then resuspended in a minimum volume of 20 mM Tris·HCl, pH 8.0/1 mM EDTA. This Mb preparation was applied to a Bio-Gel P-100 (Bio-Rad) gel-filtration column (2.5 × 50 cm) equilibrated in 20 mM Tris·HCl, pH 8.0/1 mM EDTA. The appropriate fractions were pooled, concentrated, and applied to a DEAE-cellulose (Sigma) ion-exchange column (2.5 × 60 cm) equilibrated and resolved in 20 mM Tris·HCl, pH 8.4/1 mM EDTA. This procedure resulted in a homogeneous solution of Mb as monitored by $NaDodSO_4$/PAGE (21).

Spectroscopy. The whole cell UV/visible carbon monoxide difference spectrum was generated by reducing 6 ml of a 50-ml stationary-phase culture of pMb413 in *E. coli* TB-1 with sodium dithionite. The reduced sample was separated into reference and sample cuvettes and a spectral baseline was recorded. The sample cuvette was gassed with oxygen-scrubbed carbon monoxide for 30 sec and a spectrum was recorded with a Cary 219 spectrophotometer. The spectra of purified sperm whale Mb were recorded on a Hewlett–Packard 8450A UV/visible spectrophotometer.

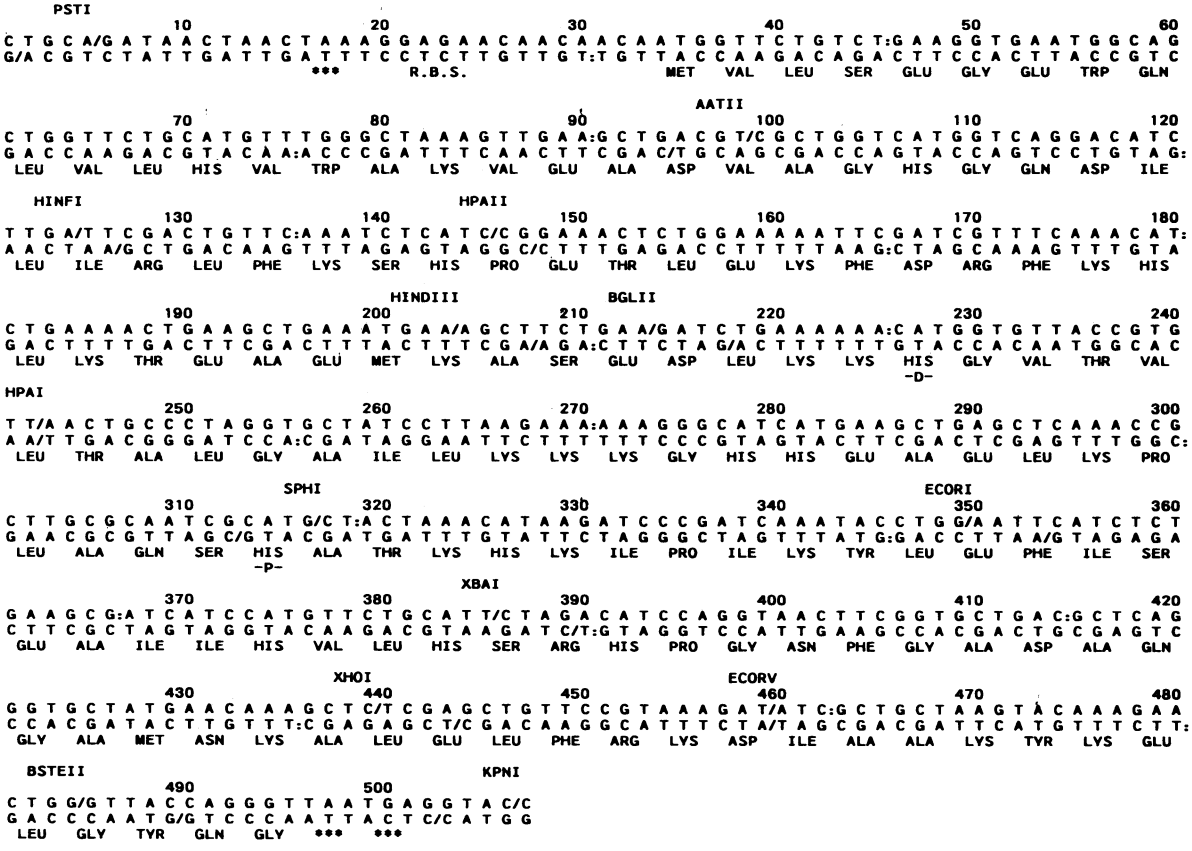

FIG. 1. Nucleotide sequence of the synthesized sperm whale Mb gene as designed from the amino acid sequence, utilizing frequently used codons for highly expressed genes. Individual oligonucleotides are delineated by colons, and asterisks indicate translational stop codons. The ribosome binding site (R.B.S.) is known to allow efficient expression of proteins in *E. coli* (17, 18). The distal and proximal histidines are marked D and P. Restriction enzyme sites designed into the synthetic gene (slash marks) will allow convenient "cassette" mutagenesis.

Electron paramagnetic resonance (EPR) spectroscopy was performed on the synthetic sperm whale Mb (300 μM in 20 mM Tris·HCl, pH 8.0/1 mM EDTA) using a Bruker (Billerica, MA) ER 200 DESR [modulation frequency, 100 kHz; modulation amplitude, 10 G; power 16 decibels (5 mW) at 42 K; microwave frequency, 9.450 GHz]. EPR experiments were carried out at the University of Illinois (Urbana–Champaign) EPR Center.

Amino-Terminal Sequence Analysis. Purified Mb (1 nmol) was subjected to automated Edman degradation in an Applied Biosystems 470A gas/liquid-phase protein sequencer. The phenylthiohydantoin amino acid derivatives were identified by HPLC on an Altex (Berkeley, CA) Ultrasphere ODS-PTH column against standards.

RESULTS AND DISCUSSION

The synthetic gene was designed to encode the known amino acid sequence of sperm whale Mb. Synthetic oligodeoxynucleotides, typically 45 or 48 nucleotides long, were designed to encode both strands of the gene (see Fig. 1 and *Materials and Methods*). This synthetic approach to gene cloning circumvents the tedious cDNA cloning procedures normally required for the isolation and expression of mammalian genes in prokaryotic environments. The synthetic approach also permits the strategic positioning of translational initiation and termination sequences and an efficient ribosome binding site with optimal spacing 5′ to the initiation codon. We have utilized the Shine–Dalgarno sequence and spacer region from the *Pseudomonas putida* cytochrome P-450$_{cam}$ gene (17), previously shown to favor high-level expression of the *P. putida* cytochrome P-450$_{cam}$ gene and the rat cytochrome b_5 gene in *E. coli* (17, 18). The construction also included translational stop codons in all three reading frames 5′ to the Shine–Dalgarno sequence to prevent formation of a fusion protein. In addition, the direct synthesis of the entire gene allowed for the incorporation of unique restriction enzyme sites convenient for subcloning and future mutagenesis. The synthetic approach also allowed for the use of biased codons, used for highly expressed *E. coli* genes, which correspond to the larger tRNA pools (22, 23). This permits optimal use of the *E. coli* translational machinery.

The gene for sperm whale Mb was constructed in three fragments (see *Materials and Methods*) to generate the entire coding sequence in the proper orientation to allow transcription from the pUC19 *lac* promoter. Both strands of the entire gene were sequenced from double-stranded templates by the dideoxy procedure. The efficient production of sperm whale Mb in *E. coli* TB-1 is easily visualized by the deep, blood-red color of the bacterial cells and by the whole cell UV/visible carbon monoxide difference spectrum (Fig. 2). The trough at 437 nm and the peak at 423 nm are completely analogous to Mb purified from sperm whale muscle (1). The high-level expression of Mb also is apparent from the NaDodSO$_4$/PAGE analysis of pMb413 soluble extracts (Fig. 3), which showed that synthetic sperm whale Mb migrates as expected with an apparent molecular weight of ≈17,800. A densitometric scan of the gel in Fig. 3 revealed that the Mb band constitutes ≈10% of the total soluble *E. coli* extract protein (data not shown).

Synthetic sperm whale Mb protein was purified to homogeneity for spectroscopic analysis. The metmyoglobin Soret band is observed at 409 nm, and the deoxy-ferrous Soret band is at 434 nm, with a broad visible band at 556 nm. The MbO$_2$ has a characteristic Soret band at 418 nm, with α and β bands at 580 nm and 542 nm, respectively, and the MbCO exhibits a Soret band at 423 nm, with α and β bands at 578 nm and 541 nm, respectively (Fig. 4). These spectra are identical to those of commercially available sperm whale Mb (1).

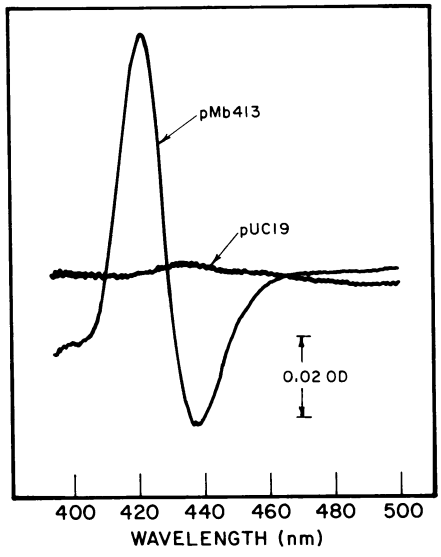

FIG. 2. Whole-cell UV/visible carbon monoxide difference spectra of the *E. coli* TB-1 strain harboring pMb413 or pUC19. Absorbance scale (OD) is as indicated in the figure. The difference spectrum represents the Fe(II)CO − Fe(II) heme protein states.

Sequential Edman degradation of the amino terminus of the purified synthetic sperm whale Mb protein revealed that the initiator methionine is not processed. Sequencing of purified protein (1 nmol) yielded Met (500 pmol), Val (311 pmol), Leu (415 pmol), Ser (75 pmol), Glu (250 pmol), Gly (216 pmol), Glu (210 pmol), Trp (114 pmol), and Gln (160 pmol), which, except for Met, are identical to the first eight residues of sperm whale Mb, indicating that a *lac* fusion protein is not produced from this construction. Total amino acid analysis confirmed the expected amino acid composition as predicted from the coding sequence (data not shown). The purified protein has a Soret band at 409 nm, indicating that it is isolated as metmyoglobin. This protein is high-spin with an EPR *g* value of 5.87 as expected for native metmyoglobin. A small fraction of low-spin protein (<1%) was observed with *g* values of 1.84, 2.16, and 2.59, indicative of OH^--bound Mb. These data indicate that the synthetic sperm whale Mb is identical to native Mb.

It is apparent that rapid protein degradation does not limit

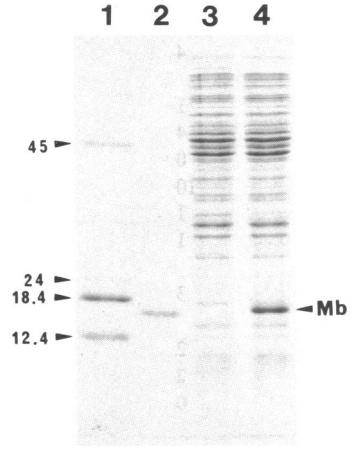

FIG. 3. NaDodSO$_4$/PAGE analysis (Coomassie blue R-250 stain). Lane 1: 1 μg of each molecular weight standard (horse heart cytochrome *c*, M_r 12,400; β-lactoglobulin, M_r 18,400; soybean trypsin inhibitor, M_r 24,000; ovalbumin, M_r 45,000). Lane 2: 1 μg of sperm whale Mb (Sigma). Lane 3: *E. coli* TB-1 pUC19 cell extract. Lane 4: *E. coli* TB-1 pMb413 cell extract.

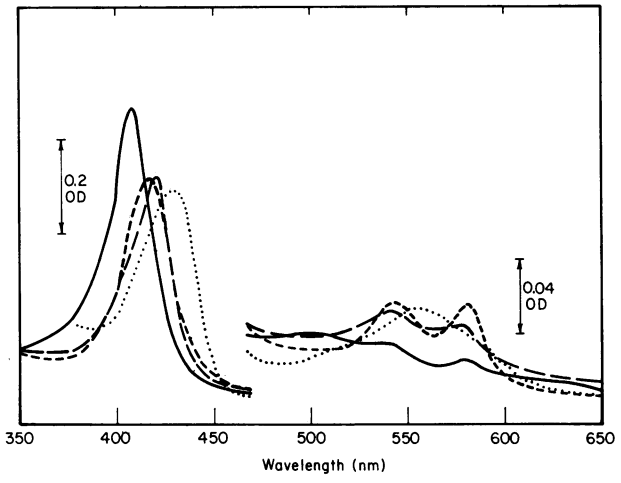

FIG. 4. Optical spectra of the synthesized sperm whale Mb gene product expressed in *E. coli*. Spectra were generated using ≈4 nmol of purified protein. —, Oxidized metmyoglobin; ····, dithionite-reduced Mb; - - -, MbO_2; – – –, MbCO.

the expression of synthetic Mb as was suspected for the low-level expression of the human globin cDNA clones (12, 14). In addition, our results and those of Beck von Bodman *et al.* (18) show that iron-protoporphyrin IX biosynthesis and prosthetic group incorporation into apoprotein are adequate for the high-level expression of functional *b*-type cytochromes and heme proteins in *E. coli*. An RNA secondary structure prediction was undertaken according to Zuker and Stiegler (24), in an attempt to determine whether mRNA secondary structure is responsible for the low-level expression of the human globin cDNA clones as compared to synthetic sperm whale Mb. Complete mRNAs transcribed from human Mb, human β-globin, and synthetic sperm whale Mb are capable of considerable secondary structure (data not shown). Although large folding free energy differences exist (e.g., 202 kcal/mol for human Mb vs. 128 kcal/mol for synthetic Mb), there are no obvious structural differences between the mRNAs that might account for the low-level human globin cDNA expression. The relevance of complete mRNA secondary structure in prokaryotes is questionable, since transcription and translation are tightly coupled.

In lieu of the above information, we credit the high-level expression of heme-containing globin on the use of preferred codons for highly expressed *E. coli* genes, which reflects the correspondingly larger tRNA pools (22, 23). A comparison of codon usage between human Mb (25), human β-globin (26),

Table 1. Codon usage for human Mb, human β-globin, and synthetic sperm whale Mb

Amino acid	Codon	Human Mb	Human β-globin	Synthetic Mb	Amino acid	Codon	Human Mb	Human β-globin	Synthetic Mb
Gly	GGG	7	1	1	Ser	UCU	1	1	5
	GGU	4	4	10		UCC	1	2	0
	GGC	4	8	0		UCA	1	0	0
						UCG	2	0	1
Pro	CCC	3	0	0		AGC	2	0	0
	CCA	2	2	1		AGU	0	2	0
	CCU	0	5	0					
	CCG	0	0	3	Asp	GAU	1	5	3
						GAC	7	2	3
Lys	AAA	0	3	16					
	AAG	20	8	2	Val	GUU	1	3	6
						GUC	2	2	1
Ala	GCU	2	6	14		GUG	4	13	1
	GCC	7	9	1					
	GCA	3	0	0	Glu	GAA	3	2	13
	GCG	0	0	2		GAG	12	6	2
Phe	UUU	3	5	0	Arg	CGG	1	0	0
	UUC	4	3	6		AGG	1	3	0
						CGA	0	0	1
Leu	CUU	1	0	2		CGU	0	0	2
	CUC	4	3	2		AGA	0	0	1
	CUA	0	0	1					
	CUG	10	15	11	Asn	AAU	0	1	0
	UUA	1	0	1		AAC	3	5	2
	UUG	1	0	1					
					Ile	AUU	1	0	1
Met	AUG	3	1	2		AUC	7	0	8
Thr	ACU	2	3	4	Gln	CAA	0	0	1
	ACC	2	3	1		CAG	6	3	4
	ACA	0	1	0					
					His	CAU	6	2	12
Trp	UGG	2	2	2		CAC	3	7	0
Cys	UGU	0	2	0	Tyr	UAU	0	2	0
	UGC	1	0	0		UAC	2	1	3

Codon usage comparison of human Mb (25), human β-globin (26), and synthesized sperm whale Mb (Fig. 1) genes. Only those amino acids and corresponding codons that are present in the three genes are presented. The underlined codons are the most frequently used codons for highly expressed *E. coli* genes (23). Values indicate the number of times each codon is used in a gene.

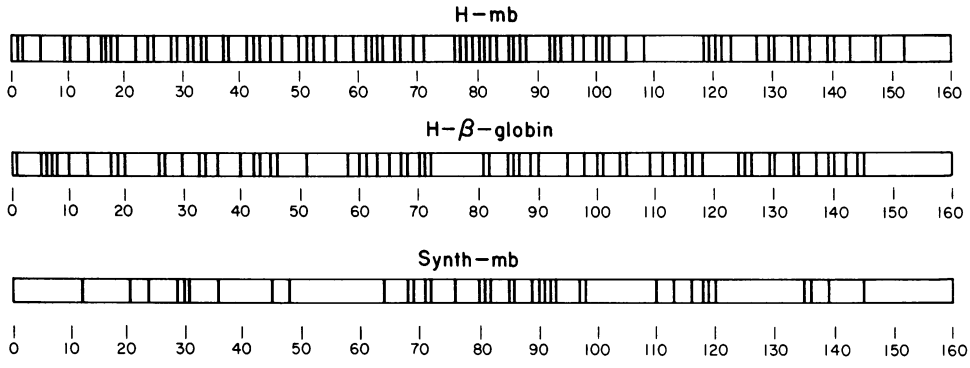

FIG. 5. The distribution of "poor" codons in the genes for human Mb, human β-globin, and synthetic sperm whale Mb. Codons were considered poor as described by deBoer and Kastelein (23). Position of each poor codon is indicated by a vertical bar. Numbers indicate the nucleotide position within the coding region. Several regions of poor codons are present in the human globin genes that are not present in the synthetic sperm whale Mb gene.

and synthetic sperm whale Mb indicates that numerous infrequently used *E. coli* codons are present in the human genes (Table 1). In particular, major differences in codon usage are observed for the frequently used amino acids lysine, alanine, valine, and glutamic acid and for the less frequently used amino acids proline and phenylalanine. The only disparity in this theme is the use of histidine codons. The infrequently used histidine codon (CAU) was chosen based on the work of Ikemura (27) and Grantham *et al.* (28) in which the most frequently used histidine codon for highly expressed *E. coli* genes was not evident. Although several infrequently used codons are present in the synthetic sperm whale Mb gene, these codons comprise <25% of the entire gene and are widely dispersed throughout the gene, whereas "poor" codons comprise >50% of the human Mb and >45% of the human β-globin genes. In addition, there are clusters of poor codons present in both human globin genes (Fig. 5). Although it would be interesting to speculate that ribosomes might stall or release at these regions of infrequently used codons while waiting for a specific charged tRNA, the work of deBoer and Kastelein (22) refutes this possibility. They mutated 25 consecutive codons at the 5′ end of the *Saccharomyces cerevisiae* gene coding for the highly expressed protein 3-phosphoglycerate kinase (PGK) to unusual codons and no decrease in PGK expression in yeast was observed, and only when 30% of the entire gene at the 5′ end was changed did they note decreased expression of PGK. deBoer and Kastelein suggested that the overall number of unusual codons, rather than their position in the gene, is important for high-level expression. The presence of poor *E. coli* codons exceeding 45% of the entire gene for the human Mb and the human β-globin genes might explain their inability to express in high levels.

It has been hypothesized that translational pausing may be necessary for the proper folding of some proteins (29), which suggests that translational pausing due to infrequently used codons in the human globin genes may disrupt the normal folding pattern of these proteins when translated in *E. coli*, thus exposing sites for protease digestion. The addition of a fusion protein that is highly stable in *E. coli* might protect these exposed sites. Although the protein fusion system employed for efficient expression of the human globin genes seems to resist protease digestion, the fusion protein and/or the infrequently used codons may disturb proper folding and subsequent heme incorporation, thereby resulting in the expression of apoprotein. The presence of frequently used *E. coli* codons in the synthetic sperm whale Mb gene permits proper folding and heme insertion, thus generating holoprotein. It appears that the use of frequently used *E. coli* codons is the most important parameter in the high-level expression of stable, heme-containing globin in *E. coli*.

We thank Mr. Saw Kyin and the Biotechnology Center at the University of Illinois (Urbana–Champaign) for oligonucleotide synthesis and amino acid sequencing and Ms. Susan Martinis for generating the EPR data. This work was supported in part by the National Institutes of Health Grants GM33775 and GM31756.

1. Antonini, E. & Brunori, M. (1971) *Hemoglobin and Myoglobin in Their Reactions with Ligands* (North-Holland, Amsterdam).
2. Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davis, D. R., Phillips, D. C. & Shore, V. C. (1960) *Nature (London)* **185**, 422–427.
3. Remaut, E., Tsao, H. & Friers, W. (1983) *Gene* **22**, 103–113.
4. Queen, C. J. (1983) *J. Mol. Appl. Genet.* **2**, 1–10.
5. Shatzman, A., Ho, Y. S. & Rosenberg, M. (1983) in *Experimental Manipulation of Gene Expression*, ed. Inouye, M. (Academic, New York), pp. 1–14.
6. Shine, J. & Dalgarno, L. (1975) *Nature (London)* **254**, 34–38.
7. Steitz, J. A. (1979) in *Biological Regulation and Development 1: Gene Expression*, ed. Goldberger, R. F. (Plenum, New York), pp. 349–399.
8. Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D., Grubman, M., McKercher, P., Morgan, D., Robertson, B. & Bachrach, H. (1981) *Science* **214**, 1125–1129.
9. Simons, G., Remaut, E., Allet, B., Devos, R. & Fiers, W. (1984) *Gene* **28**, 55–64.
10. Schoner, R., Ellis, L. & Schoner, B. (1985) *Bio/Technology* **3**, 151–154.
11. Masui, Y., Mizuno, T. & Inouye, M. (1984) *Bio/Technology* **3**, 81–85.
12. Nagai, K. & Thøgersen, H. C. (1984) *Nature (London)* **309**, 810–812.
13. Nagai, K., Perutz, M. F. & Poyart, C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7252–7255.
14. Varaderajan, R., Szabo, A. & Boxer, S. G. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5681–5684.
15. Matteucci, M. A. & Caruthers, M. H. (1980) *Tetrahedron Lett.* **21**, 719–722.
16. Beaucage, S. L. & Caruthers, M. H. (1981) *Tetrahedron Lett.* **22**, 1859–1862.
17. Unger, B. P., Gunsalus, I. C. & Sligar, S. G. (1986) *J. Biol. Chem.* **261**, 1158–1163.
18. Beck von Bodman, S., Schuler, M. A., Jollie, D. R. & Sligar, S. G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9443–9447.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
20. Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152**, 232–238.
21. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
22. deBoer, H. A. & Kastelein, R. A. (1986) in *Maximizing Gene Expression*, eds. Reznikoff, W. & Gold, L. (Butterworth, Stoneham, MA), pp. 225–285.
23. Bulmer, M. (1987) *Nature (London)* **325**, 728–730.
24. Zucker, M. & Stiegler, P. (1981) *Nucleic Acids. Res.* **9**, 133–148.
25. Akabashi, E. (1985) *Gene* **33**, 2241–2249.
26. Lawn, R. M., Efstratiadis, A., O'Connel, C. & Maniatis, T. (1980) *Cell* **21**, 647–651.
27. Ikemura, T. (1981) *J. Mol. Biol.* **151**, 389–409.
28. Grantham, R., Gautier, C. & Gouy, M. (1980) *Nucleic Acids Res.* **8**, 1893–1912.
29. Purvis, I. M., Bettany, A. J. E., Santiago, T. C., Coggins, J. R., Duncan, K., Eason, R. & Brown, A. J. P. (1987) *J. Mol. Biol.* **193**, 413–417.