Human Hemoglobin Expression in *Escherichia coli*: Importance of Optimal Codon Usage†

Ronald A. Hernan,‡ Hilda L. Hui,§ Mark E. Andracki,∥ Robert W. Noble,§ Stephen G. Sligar,‡ Joseph A. Walder,∥ and Roxanne Y. Walder*,∥

Department of Biochemistry and The Beckman Institute of Advanced Science and Technology, University of Illinois, Urbana—Champaign, Illinois 61801, Departments of Medicine and Biochemistry, School of Medicine, Veterans Administration Medical Center, State University of New York at Buffalo, Buffalo, New York 14215, and Department of Biochemistry, The University of Iowa, Iowa City, Iowa 52242

Received February 7, 1992; Revised Manuscript Received June 12, 1992

ABSTRACT: The overexpression of a nonfusion product of human β-globin in *Escherichia coli* from its cDNA sequence has been accomplished for the first time. Expression of β-globin from its native cDNA required the use of the strong bacteriophage T7 promoter. In this system, β-globin accumulated to approximately 10% of total *E. coli* proteins. α-Globin was not expressed in the T7 system using the native cDNA. For the expression of α-globin, synthetic genes containing optimal *E. coli* codons were constructed. Neither synthetic α- nor β-globin gene alone was expressed from the *lac* or *tac* promoter. Globin expression was achieved when the two synthetic α- and β-globin genes were combined as an operon downstream of the *lac* promoter. The two proteins combined intracellularly with endogenous heme, which was concomitantly overproduced to yield tetrameric hemoglobin as roughly 5–10% of total *E. coli* protein. Cloning the α- and β-globin cDNAs in a construct identical with the *lac* promoter did not yield globin production, establishing the requirement for optimal codon usage. The recombinant β-globin from the T7 expression system was purified and reconstituted in vitro with heme and native α chains. N-Terminal analyses showed that the β-globin produced in the T7 system and the tetrameric hemoglobin produced from the synthetic genes contained an additional β1 methionine residue. Two additional mutants, β1 Val → Met and β1 Val → Ala were produced using the T7 system. Functional and structural properties of the purified hemoglobins will be discussed in the following papers.

Human hemoglobin has been extensively studied using a variety of biochemical and physical techniques. It offers an excellent model system for the analysis of protein–protein interactions, ligand binding, cooperativity, long-range electron transfer, and the thermodynamics of macromolecular assembly (Johnson et al., 1976; Lee & Karplus, 1983; Smith & Ackers, 1985; Speros et al., 1991). The X-ray structures of fully liganded hemoglobin and deoxyhemoglobin and a number of intermediate ligation states of the tetramer have been determined to high resolution, providing the basis for detailed structure–function studies (Baldwin, 1980; Shaanan, 1983; Fermi et al., 1984; Luisi & Shibayana, 1989). Hemoglobin may also find a practical application as a blood substitute. Two principal limitations must be overcome for this purpose: (1) the dissociation of hemoglobin into $\alpha\beta$ dimers, which having a molecular weight of only 32 000 are rapidly cleared from the circulation by the kidneys, and (2) the high oxygen affinity of hemoglobin outside of the red cell due to the absence of 2,3-diphosphoglycerate, which compromises the off-loading of oxygen to the tissues.

Hemoglobin is unique among proteins in that a large number of naturally occurring mutants (>400) have been identified (Lehmann, 1986). This permitted extensive structure–function studies of the molecule which became possible with other proteins only after the advent of site-directed mutagenesis. However, most hemoglobin mutants are extremely rare, often being identified in a single family, and difficult to obtain. This has frequently limited the number of studies with each variant and has often made it impossible to reproduce and extend the results of earlier investigations. Also, many variants of interest have not been identified naturally and the systematic substitution of different amino acids at a single position, which can be accomplished with site-directed mutagenesis, is now possible. These considerations and the possibility of engineering into the molecule properties desirable for its use as a blood substitute have prompted efforts to develop recombinant expression systems for hemoglobin.

Expression of individual human hemoglobin chains in *Escherichia coli* was first accomplished by Nagai and Thøgersen (1984). β-Globin gene expression was achieved by ligating the human cDNA downstream of the 31 amino terminal residues encoded by the bacteriophage λ cII gene. An intervening sequence encoding the recognition site for factor Xa, Ile-Glu-Gly-Arg, was inserted between the two genes to allow for excision of the β chain from the fusion product. The fusion protein formed within the cell precipitated as an inclusion body which could be solubilized under denaturing conditions. After cleavage of the leader sequence with factor Xa, the isolated β chains could then be reconstituted with exogenous heme and native α chains to form the tetramer. A similar system to produce human myoglobin employing the native cDNA sequence was developed soon thereafter (Varaderajan et al., 1985). Expression of α chains in the system described by Nagai and Thøgersen proved more difficult. α-Globin expression required a four-part fusion construct: the sequences for the first 31 amino acid residues of the λ cII gene, the first 20 amino acid residues of β-globin minus the β2 His codon, the factor Xa recognition site, followed by the

† This work was supported by NIH Grants HL 40453 and research funds from the Veterans Administration.

* Address correspondence to this author.
‡ University of Illinois.
§ State University of New York at Buffalo.
∥ The University of Iowa.

0006-2960/92/0431-8619$03.00/0

α-globin gene minus the N-terminal methionine codon (Nagai & Thøgersen, 1987). Although both α- and β-chain hemoglobin mutants can be made utilizing this approach, isolation of the fusion proteins, cleavage of the leader sequences, purification of the subunits, and finally reconstitution of the tetramer is a rather laborious process. As a result, relatively few studies of hemoglobin variants have been reported utilizing this system (Nagai et al., 1987; Olson et al., 1988; Lin et al., 1990; Tame et al., 1991; Imai et al., 1991).

Springer and Sligar (1987) adopted a different stategy for the production of sperm whale myoglobin using a fully synthetic gene in which the codons were optimized for *E. coli*. This also allowed the convenient placement of restriction sites within the gene to simplify the construction of variants by cassette mutagenesis. The frequency of codon usage for many amino acids varies greatly among different species. For highly expressed proteins there is a strong bias in favor of those codons for which the isoaccepting tRNA is in greatest abundance (Ikemura, 1981; deBoer & Kastelein, 1986). Factors such as RNA secondary structure also appear to influence optimal codon selection (Gross et al., 1990). Optimal codon usage, therefore, has been postulated as a means to constitutively regulate gene expression at the posttranscriptional level. In the case of myoglobin the results were dramatic. When the synthetic gene was used, a leader sequence was no longer required for expression. Heme was concomitantly overproduced and the polypeptide chain properly folded to yield the holoprotein within the soluble fraction of the cell (Springer & Sligar, 1987). Presumably, depletion of free heme due to insertion within myoglobin led to upregulation of the heme biosynthetic pathway. This system has greatly simplified the production of myoglobin mutants (Rohlfs et al., 1990; Egeberg et al., 1990). Recently, Hoffman and co-workers (1990) adopted the same approach to produce human hemoglobin in *E. coli*. Synthetic genes for α- and β-globin were constructed using optimal codons for *E. coli*. The genes were then cloned in tandem to form an operon with the *tac* promoter. Both globin chains were overexpressed, and along with overproduction of heme, assembly of the tetramer occurred.

In the present study we have explored a variety of constructs using both native and synthetic genes with several different promoter systems to establish the requirements for hemoglobin expression in *E. coli*. When the strong bacteriophage T7 promoter was used, isolated β chains were found to be expressed in high yields with both the native cDNA sequence and synthetic genes without the requirement for a leader peptide. Use of a strong promoter overrides the need for optimal codon to express the native polypeptide chain. We have produced a variety of β chain variants using this system as well as mixed-metal hybrids. α chains could not be expressed alone even with the use of the T7 promoter. Presumably, without being made as a fusion protein the α chains, which are intrinsically less stable than β chains and myoglobin, are degraded before they are able to precipitate within an inclusion body. Coexpression of β chains proved necessary for the expression of native α chains. This worked only when the synthetic genes were used and led to the formation of the fully assembled tetramer. In the globin system, optimal codon usage not only leads to enhanced expression but appears to be essential to provide a rate of protein synthesis necessary for proper folding and heme insertion.

For both α- and β-chain genes of human hemoglobin and myoglobin, the amino terminal sequence is Met-Val. In mammalian cells the methionine is cleaved cotranslationally, leaving valine as the amino terminal residue of each of the mature proteins. In *E. coli* with Val as the second amino acid methionine is not cleaved, although we will show that the formyl group is removed from the Met residue. In the case of myoglobin, the additional methionine residue has no observable effects on the functional properties of the protein (Olson et al., 1988). For hemoglobin, however, the amino terminal residues of both the α and β chains play important roles in modulating the oxygen affinity of the molecule by different allosteric effectors. The amino terminal valine residue of the α chains contributes to the Bohr effect through the binding of chloride ions, while the amino terminal valine of the β chains, one of the positively charged residues within the β cleft, contributes to the binding of 2,3-diphosphoglycerate (DPG) (Arnone, 1972; Chiancone et al., 1972; O'Donnell et al., 1974; Nigen et al., 1980). Only by the construction and cleavage of a fusion protein can the native amino terminal sequence be obtained. To utilize systems in *E. coli* in which the soluble holoprotein is produced, it is necessary to modify that sequence. For this reason we have constructed three amino terminal variants of the β chains, which after processing in *E. coli* have the N-terminal sequences β1 + Met, β1 Val → Met, and β1 Val → Ala. In the following papers the functional and structural properties of these three mutants were studied in detail to determine which would provide the best baseline hemoglobin for further mutagenesis experiments.

EXPERIMENTAL PROCEDURES

Strains and Vectors. The T7 expression system, *E. coli* strains BL21(DE3) and BL21(DE3)pLysE, and the vector pET8c were kindly supplied by F. W. Studier. *E. coli* strain TB-1 was used in routine plasmid screening, in DNA sequencing, and in the synthetic gene assembly. M13mp19 and recombinant derivatives were transformed into JM103. The vectors pKK233-2 and pKK223-3, containing the *tac* promoter system, were purchased from Pharmacia. Competent JM109 cells were from Stratagene (La Jolla, CA).

Cloning of the Native α- and β-Globin cDNAs in pUC19, pKK233-2, and the T7 Expression System. Plasmids B43 and B6-6, containing full-length cDNA sequences for α and β human hemoglobin genes, respectively, were generously supplied by Dr. Bernard Forget. The cDNAs were cloned into the *Pst*I site of pKT218 (Michelson et al., 1983) by GC tailing. The native α- and β-globin cDNAs were both removed from B43 and B6-6 as *Nco*I–*Pst*I fragments. The *Pst*I–*Nco*I synthetic linkers

```
   5' GATAACTAACTAAAGGAGAACAACAAC
TGCACTATTGATTGATTTCCTCTTGTTGTTGGTAC  5'
```

were ligated to the 5′ ends of the two genes and subcloned into *Pst*I-digested pUC19 to generate plasmids pHA-4 and pHB-7 containing the cDNAs of the α- and β-globin genes in the correct orientation for expression from the *lac* promoter. The upstream linker region was a modification of the sequence used by Springer and Sligar (1987) in the cloning of sperm whale myoglobin in *E. coli* and incorporated the Shine–Dalgarno translational regulatory sequence.

The α- and β-globin cDNAs were ligated into the *tac* promoter vector, pKK233-2, as *Nco*I–*Pst*I fragments isolated from pHA-4 and pHB-7, resulting in pKK-A4 and pKK-B8, respectively.

For the cloning in the T7 expression system, the β-globin cDNA was excised from pHB-7 by digestion with *Sph*I, followed by removal of the single-stranded DNA ends with mung bean nuclease, and the DNA was further digested with *Nco*I. The fragment was purified by agarose gel electro-

phoresis and inserted between the *Nco*I and *Eco*RV sites of pET8c to give pETβ102. A similar *Nco*I-blunt-ended *Sph*I fragment containing the α-globin gene was isolated from pHA-4 and cloned into pET8c to give pETα107. In both pETβ102 and pETα107 the T7 transcription termination sequence, TΦ, was removed.

Construction of the Synthetic α-Globin Gene. The synthetic α-globin gene was constructed in three sections: *Xba*I–*Bam*HI, *Bam*HI–*Hin*dIII, and *Hin*dIII–*Pst*I (Figure 3a). A 400-pmol aliquot of each oligonucleotide used in assembling the gene was phosphorylated with 10 units of T4 polynucleotide kinase in 50 μL containing 50 mM Tris–HCl, pH 7.6/10 mM $MgCl_2$/5 mM DTT/100 μM EDTA for 60 min at 37 °C. Each segment was assembled by first annealing 100 pmol of the corresponding oligonucleotides (see Figure 3a) in 50 μL by heating at 95 °C for 5 min and slowly cooling the solution to 35 °C over 2 h. The annealed fragments, in 10–50-fold molar excess, were ligated with 0.25 pmol of pUC19, which had been digested with the appropriate restriction enzymes. The reaction was carried out in a final volume of 10 μL containing 50 mM Tris–HCl, pH 7.6/10 mM $MgCl_2$/1 mM rATP/1 mM DTT/5% (w/v) PEG 8000 and 1 unit of T4 DNA ligase. The ligated plasmids were transformed into *E. coli* TB-1. The complete α-globin gene was assembled by excising the cloned inserts with the appropriate restriction enzymes and purifying the inserts from agarose gels. The three inserts, 5 pmol each, were combined and ligated in a reaction volume of 10 μL as described above, with 0.5 pmol of pUC18, which had been digested with *Xba*I and *Pst*I, to yield the complete α-globin gene placed under control of the *lac* promoter (pHS471). Both strands of the gene were sequenced using the dideoxy method with universal primers.

The synthetic α-globin gene was cloned into the *tac* vector, pKK223-18, a derivative of pKK223-3 containing the polylinker sequence from pUC18 inserted between the *Eco*RI and *Hin*dIII restriction sites. The globin gene from pHS471 was excised with *Xba*I and *Pst*I and then ligated into *Xba*I- and *Pst*I-digested pKK223-18 to generate pKKHS471-18.

Construction of the Synthetic β-Globin Gene. The synthetic β-globin gene was similarly constructed in three segments: *Pst*I–*Bgl*II, *Bgl*II–*Bam*HI, *Bam*HI–*Hin*dIII (Figure 3b). Each of the three sections was assembled as described earlier for the synthetic α-globin gene and ligated into pUC137, a derivative of pUC19 with a modified polylinker sequence. The complete β-globin gene was assembled by ligating the three sections into pUC18, digested with *Pst*I and *Hin*dIII, giving the plasmid pWHS486. This places the gene in the proper orientation to be transcribed from the *lac* promoter. The sequence of both strands of the gene in the final construct was verified.

Construction of the β chain in the *tac* vector was analogous to cloning the α-chain gene. The β-globin gene was removed from pWHS486 by digestion with *Pst*I and *Hin*dIII and ligated into pKK233-18 between the *Pst*I and *Hin*dIII restriction sites to generate pKKWHS486-18.

Cloning of the Synthetic β-Globin Gene in the T7 System. The synthetic β-globin gene excised from pWHS486 was inserted in pET8c by ligation as a *Nco*I–*Hin*dIII fragment. As in the cloning of the α- and β-globin cDNAs in pET8c, the TΦ sequence was deleted.

Construction and Expression of the Hemoglobin Operon Using Native cDNA and Synthetic Genes. The two native cDNAs were cloned sequentially in the *tac* promoter vector by inserting the *Pst*I fragment from pHB-7 containing the β-globin gene 3′ of the α-globin gene in pKK-A4, yielding pKK-AB17. Similarly, the *Pst*I fragment from pHA-4 containing the α-globin gene was ligated into the *Pst*I site of pKK-B8, 3′ of the β-globin gene, giving pKK-BA6. In this and in all the other α- and β-globin operons constructed, the orientation and complete sequence of the genes were verified by dideoxy-DNA sequencing.

The two native cDNAs were also cloned in tandem into pUC19. *Pst*I fragments, containing the upstream translational regulatory sequences and the entire genes, were removed from pHA-4 and pHB-7 and ligated in the presence of T4 DNA ligase. The dimer-sized products were purified from a 1% agarose gel and ligated to *Pst*I-digested pUC19. The orientation and the presence of a single copy of each gene were first determined by restriction enzyme digestion. Two plasmids with the genes under control of the *lac* promoter (pAB-5 and pAB-12) were isolated, containing α 5′ of β and β 5′ of α, respectively.

The derivatives of pKK233-2 containing the globin genes were transformed into *E. coli* strains JM109 and TB-1 and grown for expression. The transformed strains were grown in 1 L of broth, containing 200 μg/mL ampicillin, to an absorbance of 0.3 at A_{600nm} and divided into two equal volumes. Isopropyl β-D-thiogalactopyranoside (IPTG) was added to one sample to a final concentration of 1 mM, and both samples were then incubated for an additional 2–24 h at 37 °C. Derivatives of pUC19 carrying the cDNAs were grown at 37 °C for 16 h in 1 L of broth containing 200 μg/mL ampicillin and harvested by centrifugation; aliquots were electrophoresed on 15% SDS–polyacrylamide gels. The total cell lysate from equal numbers of cells was electrophoresed on 15% SDS–polyacrylamide gels and stained with Coomassie Brilliant Blue.

The complete operon, composed of the synthetic α- and β-globin genes under control of the *lac* promoter, was constructed by ligating the *Xba*I–*Pst*I fragment from pHS471 containing the synthetic α-globin gene, and the *Pst*I–*Hin*dIII fragment from pWHS486 containing the synthetic β-globin gene, into pUC18 digested with *Xba*I and *Hin*dIII to generate pHS957. A hemoglobin operon composed of the synthetic genes was also constructed in the *tac* vector, pKK223-18, in a procedure analogous to the cloning into the *lac* vector, yielding pKKHS951.

Synthesis of β-Globin in the T7 Expression System. The expression strains BL21(DE3) and BL21(DE3)pLysE were transformed with pETβ102. The transformed strains were grown in 1 L of broth, containing 100 μg/mL ampicillin and 4 mg/mL glucose, to midlog phase, induced with 0.4 mM IPTG, and then incubated for an additional 2–3 h at 37 °C. The cells were collected by centrifugation and lysed in 80 mM Tris–HCl, pH 6.8/0.1 M DTT/2% SDS/10% glycerol, and an aliquot was analyzed by SDS–polyacrylamide gel electrophoresis.

Isolation of β-Globin from the T7 Expression System. E. coli BL21(DE3) transformed with pETβ102 was grown in 100-L batches in the University of Iowa Fermentation Facility using the same media and induction protocol described above. Approximately 500–1000 g of cell paste was obtained from each 100-L growth. The method used for the isolation of β-globin grown in *E. coli* was based upon that of Nagai et al. (1985) with simplifications introduced. Frozen cells were thawed and resuspended in 0.8 times (w/v) 50 mM Tris–HCl, pH 8.0, 25% sucrose, and 1 mM EDTA. Cells were lysed at room temperature for 30 min by the addition of lysozyme to 1 mg/mL in the presence of 2.5 μg/mL DNase/1 mM $MnCl_2$/10 mM $MgCl_2$. An equal volume of 0.2 M NaCl/1% DOC/1.6% NP-40 (v/v)/20 mM Tris–HCl, pH 7.5/2

mM EDTA was added and the mixture centrifuged at 17600*g* for 10 min. The pellet containing β-globin as an insoluble inclusion body was washed four times in 0.5% Triton X-100/1 mM EDTA until a tight pellet was formed. The pellet was solubilized in 8 M urea/25 mM Tris acetate, pH 5.0/1 mM DTT/0.1 mM EDTA. The opaque mixture was centrifuged at 17600*g* for 10 min, and the clear supernatant was loaded onto a CM Sepharose CL-6B column (4 × 10 cm) equilibrated with the 8 M urea starting buffer. The β-globin was eluted with a 900-mL linear gradient of starting buffer containing 0–0.4 M NaCl. The fractions containing the β-globin were pooled, concentrated using an Amicon PM10 membrane to 5 mg/mL, adjusted to pH 8 with solid Tris base, and used in the reconstitution of tetrameric hemoglobin.

Assembly of Tetrameric Hemoglobin Using E. coli β-Globin from the T7 System. α-SH chains were isolated as previously described (Bucci & Fronticelli, 1965; Geraci et al., 1969). The synthetic hemoglobin was assembled by adding β-globin (0.3 mg/mL), α-SH, and cyanohemin in 1.2 molar excess in CO-saturated deionized water (Nagai et al., 1985). After 16 h at 4 °C, the mixture was concentrated 6-fold using an Amicon PM 30 membrane, reduced anaerobically with dithionite, and loaded onto a Sephadex G-25 column (4 × 60 cm) equilibrated with 10 mM Tris–HCl, pH 8.6. The hemoglobin was eluted with the same buffer. The hemoglobin fraction was then loaded onto a DE52 cellulose column equilibrated in 10 mM Tris–HCl, pH 8.6, and eluted with a linear gradient (1300 mL) of 15 mM Tris–HCl, pH 8.6, to 80 mM Tris–HCl, pH 8.0. Alternatively, the material from the Sephadex column was purified by HPLC. The reduced and concentrated hemoglobin-containing peak was loaded onto a SP5PW TSK gel column (Toso-Haas, 2 × 15 cm), previously equilibrated with 25 mM sodium phosphate, pH 6.0, and eluted at a flow rate of 5 mL/min with a 65-minute linear gradient from 25 mM sodium phosphate, pH 6.0, to 25 mM Tris–HCl, pH 8.6.

CO Difference Spectrum. A whole-cell CO difference spectrum of *E. coli* TB-1 harboring the plasmid pHS957 was generated by adding a few crystals of dithionite to 3 mL of stationary-phase cells, and a baseline was recorded. The cuvette was removed and CO bubbled through the solution for 30 s. The spectrum was recorded on a Beckman DU60 spectrophotometer.

N-Terminal Mutants of β-Globin. N-Terminal mutants of β-globin were constructed using the oligonucleotide in vitro mutagenesis system from Amersham. The *Xba*I–*Hin*dIII fragment, containing the upstream T7 regulatory sequences and β-globin gene from pETβ102, was inserted into M13mp19 to give M13XHβ. The β1 Val → Met construct was produced using the mutagenic oligonucleotide 5′-AGGAGATAACCATGCACCTGACTCCTGAG and the β1 Val → Ala mutation with 5′-ATACCAAATGGCGCACCTGACT. The DNA sequence of the mutant β-globin genes was verified, and the *Xba*I–*Hin*dIII fragment was purified from M13 and ligated into *Xba*I- and *Hin*dIII-digested pET8c. Expression of the mutant globin was detected following transformation into BL21(DE3) and BL21(DE3)pLysE by analysis of whole-cell protein extracts on 15% SDS–polyacrylamide gels as previously described.

RESULTS

Hemoglobin Expression in the T7 System. The complete cDNA sequence for human β-globin was ligated between the *Nco*I and *Eco*RV sites of the T7 expression vector, pET8c, giving the plasmid pETβ102. The cloning was facilitated by the presence of a *Nco*I site at the initiator methionine codon of the cDNA. The vector pET8c also contains a unique *Nco*I site, which is located downstream of the T7 Φ10 promoter and S10 translational regulatory sequences, yielding a construct for the expression of β-globin as a nonfusion protein, pETβ102. Expression of β-globin was detected after IPTG induction of *E. coli* BL21(DE3) or BL21(DE3)pLysE transformed with pETβ102 grown at 37 °C. Following a 2–3-h induction with IPTG, β-globin was synthesized to approximately 10% of total *E. coli* proteins (Figure 1). The globin precipitated within the cell in the form of an inclusion body. Since the TΦ sequence of the vector pET8c was removed in pETβ102, transcription continued around the plasmid and consequently β-lactamase was also overproduced (Figure 1). The level of β-globin measured by analyzing an aliquot of the culture on SDS–polyacrylamide gels was greater in BL21(DE3) than in BL21(DE3)pLysE. A similar level of inducible expression was observed when the synthetic β-globin gene, utilizing optimal codons for *E. coli*, was inserted into pET8c and transformed into BL21(DE3). No increase in the expression level or solubility of the protein was observed when the expression strains were grown at lower temperatures. β-Globin synthesis was approximately the same when the cells were grown at 30 °C and decreased noticeably at 25 °C. α-Globin was not expressed in the T7 system. The native cDNA was cloned into pET8c and pETα107 and transformed into the same expression strains. Following induction with IPTG and analyses on SDS–polyacrylamide gels, no globin synthesis was detected by staining with Coomassie Brilliant Blue.

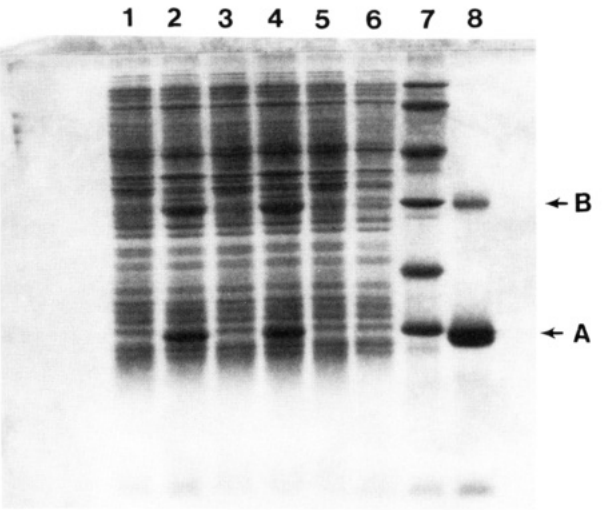

FIGURE 1: SDS–polyacrylamide gel analysis of recombinant β-globin expression using the T7 system: (lane 1) native β-globin cDNA without addition of IPTG to the culture; (lane 2) native β-globin cDNA with IPTG; (lane 3) synthetic β-globin gene without IPTG; (lane 4) synthetic β-globin gene with IPTG; (lane 5) pET8c without IPTG; (lane 6) pET8c with IPTG; (lane 7) molecular weight markers; (lane 8) human hemoglobin. The native cDNA and synthetic β-globin genes were cloned into the T7 expression vector pET8c as described in the text. Whole-cell lysates from the expression strain *E. coli* BL21(DE3) carrying the various plasmids were separated on a 15% SDS–polyacrylamide gel and visualized with Coomassie Brilliant Blue. Arrows A and B designate the positions of the β-globin and β-lactamase bands, respectively.

Isolation of β-Globin from the T7 Expression System and Assembly into Hemoglobin. The procedure for the isolation of β-globin expressed in the T7 system was greatly simplified from that described by Nagai and Thøgersen (1984) because in the present case the product of expression is not a fusion protein. This obviated the need to isolate or purchase factor Xa. Cleavage of the fusion protein, which typically takes 2–3 days to reach completion, was rendered unnecessary. Finally,

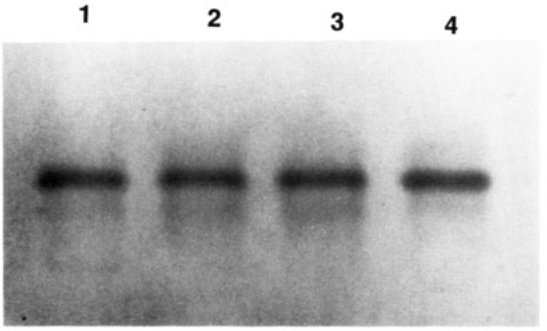

FIGURE 2: Nondenaturing polyacrylamide gel analysis (Laemmli, 1970) of assembled hemoglobins using β chains purified from the T7 system. The running gel was 7.5% acrylamide/0.002% bisacrylamide/0.375 M Tris, pH 8.8. The stacking gel was 4% acrylamide/0.001% bisacrylamide/0.125 M Tris, pH 6.8. The electrode buffer was 0.025 M Tris/0.24 M glycine, pH 8.3, and the electrophoresis was run at 5 °C and 200 V for 75 min. Gels were stained with Coomassie Brilliant Blue: (lane 1) native human hemoglobin; (lane 2) $\alpha_2\beta$(1 + Met)$_2$; (lane 3) $\alpha_2\beta$(1 Val → Met)$_2$; (lane 4) $\alpha_2\beta$(1 Val → Ala)$_2$.

fractionation by gel filtration after chromatography over CM Sepharose was eliminated since it did not yield a significant purification of the globin protein. The globin produced by the procedure described here is not pure. Faint satellite bands can be observed on SDS–polyacrylamide gels. However, these impurities do not appear to affect the assembly of the β-globins into intact hemoglobins, and for this reason, further purification of the individual β-globin chains was deemed unnecessary.

Recently, Fronticelli et al. (1991) reported a fusion protein system to produce β-globin very similar to that developed by Nagai and Thøgersen, also requiring cleavage of the initial protein by factor Xa. They indicated the need to use assembly times of 1 week or greater, but in the present study the yield of intact hemoglobin did not increase after an incubation of 16 h (overnight). The usual yield of the assembled hemoglobins after purification was roughly 50% of that predicted from the starting amount of β-globin. This translates into a yield with this expression system of roughly 1.5 mg of hemoglobin from each gram of wet cell paste. By starting with 300 g of cells, it has become possible to routinely prepare nearly 0.5 g of a hemoglobin variant.

The results of nondenaturing polyacrylamide gel electrophoresis of three variants, $\alpha_2\beta$(1 + Met)$_2$, $\alpha_2\beta$(1 Val → Met)$_2$, and $\alpha_2\beta$(1 Val → Ala)$_2$ along with native HbA are shown in Figure 2. Each of the mutants is present as a single band. The CO, oxy, and deoxy spectra of each of these variants are identical to those of native HbA (data not shown). Many other variants have also been constructed using this procedure. Besides those presented in Figure 2, these have included a synthetic Hb-Rothschild which is $\alpha_2\beta$(1 + Met, 37 Trp → Arg)$_2$. The latter hemoglobin exists almost entirely as dimers when saturated with ligand at the concentration used for assembly. Because the yield of this material was as great as that for other variants, it appears that assembly is dependent only upon the formation of $\alpha^1\beta^1$ dimers and is insensitive to disruption of the $\alpha^1\beta^2$ interface.

N-Terminal Analysis of Recombinant Globins. Ten cycles of amino terminal sequencing analysis of the hemoglobin assembled with β-globin obtained from the initial cDNA sequence showed the following residues: Val/Met, Leu/Val, Ser/His, Pro/Leu, Ala/Thr, Asp/Pro, Lys/Glu, Thr/Glu, Asn/Lys, and Val/Ser. The sequence corresponds to that of native α chain and β chain containing an extra methionine at the N-terminus. This result unambiguously indicates that there is no fusion protein associated with the expression of β-globin. The formyl group has been removed by *E. coli* from the N-terminal methionine residue. The absence of any secondary sequences indicates that further processing of the amino terminus does not occur. Two additional N-terminal mutants were constructed: β1 Val → Met and β1 Val → Ala (see Figure 2). Ten cycles of N-terminal sequencing of the β1 Val → Met mutant revealed the sequence of Met/Val, His/Leu, Leu/Ser, Thr/Pro, Pro/Ala, Glu/Asp, Glu/Lys, Lys/Thr, Ser/Asn, and Ala/Val. The N-terminal methionine is not processed from the β chain. Since valine at position 2 was deleted, this leads to the conservative amino acid substitution in the mature protein of β1 Val → Met. Ten cycles of N-terminal sequencing of the β1 Val → Ala mutant produced the sequence Ala/Val, His/Leu, Leu/Ser, Thr/Pro, Pro/Ala, Glu/Asp, Glu/Lys, Lys/Thr, Ser/Asn, and Ala/Val. In this mutant, the N-terminal methionine is removed from the β chain. Detailed functional studies and the crystal structures of these two mutants and the parent hemoglobin (β1 + Met) will be described in the following papers.

Synthetic α- and β-Globin Genes. The synthetic genes were designed to encode the 141 amino acids for α-globin and the 146 amino acids of β-globin plus the additional methionine residue necessary for expression in *E. coli* (Figure 3A,B). The two genes contained amino acid codons found in highly expressed *E. coli* proteins (deBoer & Kastelein, 1986). Unique restriction enzyme sites were inserted roughly at 50 base pair intervals throughout the genes for subcloning and cassette mutagenesis. The synthetic genes were each first cloned into the high copy number plasmid pUC18 (see Experimental Procedures), which contains the *lac* promoter. Upstream of the initiator methionine codon, a sequence containing a Shine–Dalgarno ribosome binding site and the spacer region from the *Pseudomonas putida* cytochrome P-450$_{cam}$ gene was added to maximize translational efficiency (Unger et al., 1986). The same 5′ regulatory sequence has been used in the high-level expression of rat liver cytochrome b_5 (Beck von Bodman et al., 1986) and sperm whale myoglobin (Springer & Sligar, 1987). Translational stop codons were introduced in all three reading frames upstream of the ribosome binding site to inhibit the formation of a *lac* Z fusion protein or an α–β fusion protein in pHS957.

Individually α- and β-globin genes in pUC18 did not express any protein when transformed in *E. coli* TB-1 or JM109. When driven by the *lac* promoter, neither the native cDNA sequences nor the synthetic genes in which optimal codons were employed expressd globin chains to a level detectable by Coomassie Brilliant Blue staining on SDS–polyacrylamide gels. Globin synthesis was also undetectable using more sensitive Western blotting techniques. Ligation of the individual genes into the *tac* expression vectors pKK233-2 or pKK223-3 (Experimental Procedures) also did not allow the polypeptide chains to be synthesized from the native or synthetic genes.

Coexpression of α- and β-Globin Genes. A number of hemoglobin operons were constructed with both the native cDNAs and the two synthetic globin genes (see Table I). Both the *lac* and *tac* promoter systems were tested. No expression of α- or β-globin was detectable by SDS gel electrophoresis with either promoter using the cDNAs regardless of the order in which the two genes were placed within the operon. In sharp contrast, both α and β chains were synthesized with concomitant overproduction of heme and assembly of the holoprotein when the synthetic genes utilizing optimal *E. coli*

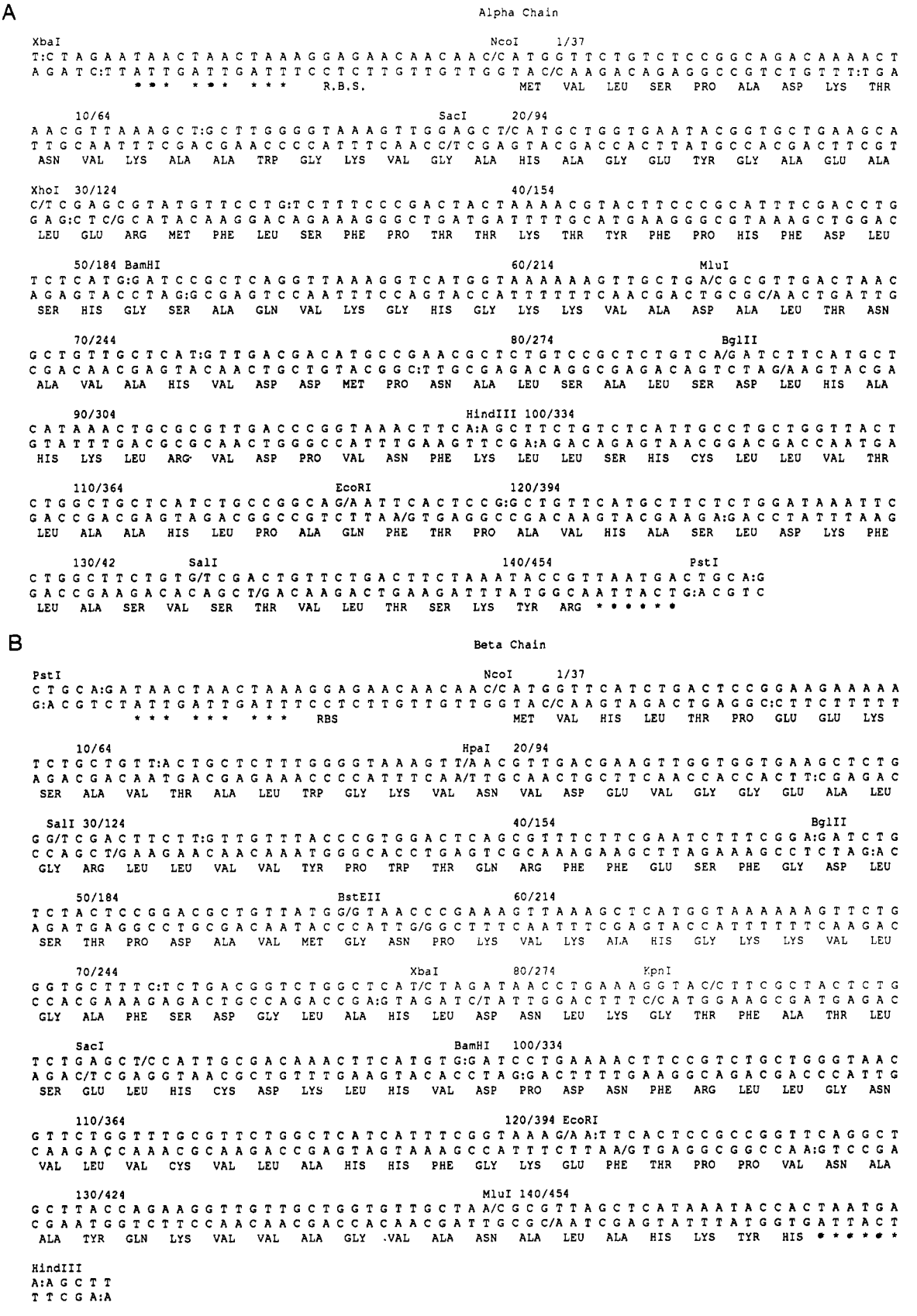

FIGURE 3: (A) Nucleotide and protein sequence of the synthetic human α-globin gene, using biased *E. coli* codons. (B) Nucleotide and protein sequence of the synthetic human β-globin gene, using biased *E. coli* codons. The Shine–Dalgarno ribosome binding site (RBS) and the translational stop codons (***) are indicated. The individual oligonucleotides, ligated together to construct the genes, are marked by colons, and the restriction enzyme sites, which allow for convenient cassette mutagenesis are illustrated with a slash.

codons were employed. Expression of hemoglobin was immediately suspected because the cell pellets were pigmented with either a reddish or dark brown color. Production of both α- and β-globin chains was confirmed by SDS–polyacrylamide gel electrophoresis (not shown). Incorporation of endogenous heme was indicated by a difference spectrum between reduced and reduced minus CO-treated cell suspensions (Figure 4). The difference spectrum shows a maximum in

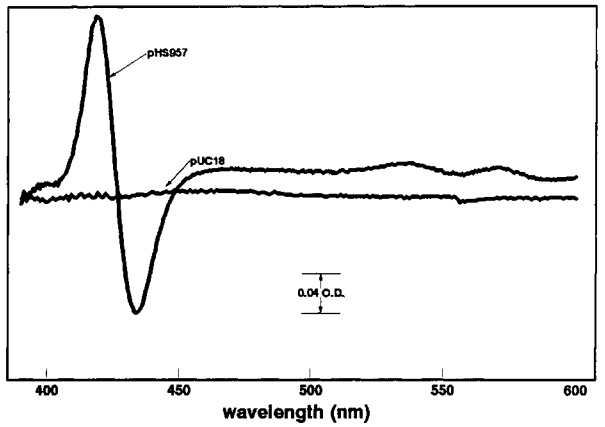

FIGURE 4: Whole-cell CO – deoxy difference spectra of *E. coli* TB-1 cells transformed with the plasmid pHS957 expression human hemoglobin and *E. coli* containing pUC18. Absorbance scale (OD) is indicated.

Table I: Hemoglobin Expression with Native and Synthetic Genes Utilizing Different Promoter Systems[a]

	native				synthetic		
promoter	α	β	αβ	βα	α	β	αβ
T7	–	+				+	
tac	–	–	–	–	–	–	±
lac	–	–	–	–	–	–	+

[a] Key: +, expression; –, no expression; ±, reduced expression. The order of the genes within the coexpression constructs is designated either αβ or βα.

the Soret region at 419 nm with α and β absorption bands at 570 and 540 nm, respectively, and a minimum at 434 nm, which is consistent with native hemoglobin (Antonini & Brunori, 1971). The pyridine hemochromogen assay indicated that the heme was protoporphyrin IX (data not shown). Spectra generated in the absence of dithionite indicated that the majority of the hemoglobin produced within the cells was in the deoxy form. From the whole-cell CO difference spectra, it appeared that the production of hemoglobin was approximately 7-fold greater from the *lac* expression vector than with the *tac* vector.

Amino acid analysis of the purified recombinant human hemoglobin showed the correct amino acid composition. Ten cycles of Edman degradation showed the following amino terminal residues: Met, Val, Leu/His, Ser/Leu, Pro/Thr, Ala/Pro, Asp/Glu, Lys/Glu, Thr/Lys, and Asn/Ser. The sequence corresponds to that of native human hemoglobin with an additional methionine on both the α and β chains. The absence of any secondary sequence implies that there is no processing of the amino terminal methionine of either the α or β chains. The formyl group is removed from the two chains and there is no *lac* Z fusion protein or an α–β fusion protein.

DISCUSSION

β-Globin Expression. In this paper we report the first high-level expression of recombinant human β-globin using the native cDNA sequence without synthesis of a fusion protein. This greatly simplified isolation of the β chains and reconstitution of the tetramer. We are now able to produce β-chain mutants with this system very readily, typically within 1 week starting from the *E. coli* cell paste, in amounts of up to 0.5 g.

Expression of β-globin from the cDNA sequence required use of the very strong bacteriophage T7 promoter. Insertion of the β-globin gene under control of the *E. coli lac* or *tac* promoter did not give rise to any detectable level of β-globin synthesis (Table I). The T7 promoter is among the strongest promoters in *E. coli* and generally results in rates of transcription 5–10-fold greater than that obtainable with the *lac* or *tac* promoter (Golomb & Chamberlin, 1974). Other factors may also have contributed to the greater level of expression of β-globin with the T7 system. The expression strains BL21(DE3) and BL21(DE3)pLysE are deficient in the lon protease and also lack the ompT outer membrane protease that can degrade proteins during purification (Grodberg & Dunn, 1988). Production of β-globin in both BL21-(DE3) and BL21(DE3)pLysE indicates that the protein is not toxic to *E. coli* and can be expressed constitutively without requiring the T7 lysozyme gene to regulate expression.

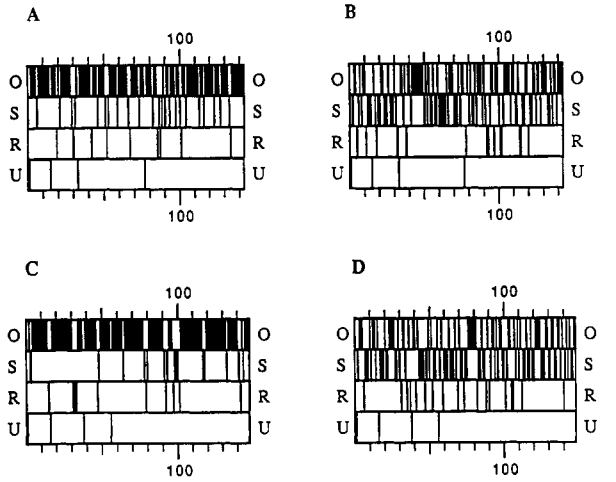

FIGURE 5: Codon distribution of the synethetic α (A), native α (B), synthetic β (C), and native β (D) genes. Codons are labeled as optimal (O), suboptimal (S), rare (R), and unique (U) (Sharp & Li, 1987). Each codon is represented by a bar. Numbers indicate the codon position within the reading frame.

RNA secondary structure has been proposed as an inhibitor of gene expression. To determine whether this is the case for the constructs using the *lac* and *tac* promoters, an mRNA secondary structure prediction was undertaken (Zucker & Steigler, 1981; Frier et al., 1986). The secondary structure predictions failed to show any significant structure around either the Shine–Dalgarno ribosome binding site or the initiator methionine for either the native cDNA sequence or the synthetic gene.

The β-globin synthesized in the T7 system precipitates into inclusion bodies presumably as a partially denatured or as an improperly folded intermediate. This may be due to hydrophobic interactions between the partially folded polypeptide chains (Goldberg et al., 1983). The presence of heme in β chains is known to stabilize the properly folded state, and when it is removed, changes in the secondary structure occur (Yip et al., 1972). In the presence of heme, β chains are able to assemble into a relatively stable homotetramer. Since the T7 system makes very large amounts of protein in a short period of time, the rate of heme synthesis may be too slow for proper folding to occur, leading to the formation of inclusion bodies rather than the soluble holoprotein β_4.

It is apparent that optimal codon usage is not the rate-limiting step in the expression of recombinant β-globin. Both the native cDNA sequence and the synthetic sequence express to the same level when placed under control of the T7 promoter. The frequency of codon distribution for both the native cDNA sequence and the synthetic sequence is shown in Figure 5. Although there is a much higher frequency of optimal codon usage in the synthetic gene, there is no increase in the level

of expression of the protein. In highly expressed genes in *E. coli*, there is a strong bias toward the use of optimal codons (Ikemura, 1981). However, the rate of transcription of the gene generally exerts a far more important degree of control on the level of protein synthesis. This has often made it possible to express heterologous proteins with a very poor codon distribution in *E. coli*, provided a strong promoter is used (Walder et al., 1983; Ernst & Kawashima, 1988).

α-Globin Expression. The absence of expression of α-globin alone, using either the native cDNA sequence or the synthetic gene with a variety of constructs, presents several possible interpretations. We have used both strong promoters and high copy number plasmids in hope of producing significant quantities of mRNA for translation. Secondary structure predictions (Zucker & Stiegler, 1981; Frier et al., 1986) of both the native and synthetic genes were conducted to establish whether mRNA secondary structure could be the origin of low expression. The synthetic gene had a smaller free energy of folding than the native gene, −134.6 versus −201.1 kcal/mol. Some secondary structure around both the Shine–Dalgarno ribosome binding site and the initiator AUG codon of both genes was predicted, suggesting that the initiation of translation might be somewhat inhibited. However, it has been reported that ribosomes are capable of reading through stem loop structures once the 30S subunit is bound to the 5′ end of the message. Moreover, the reliability of secondary structure predictions is uncertain, particularly in prokaryotic systems where transcription and translation are tightly coupled. A short half-life of the mRNA might contribute to the lack of expression of α-globin; however, one might expect this to be at least partially overcome by use of high copy number plasmids and strong promoters such as T7. Rapid proteolytic degradation of the improperly folded protein before it is able to precipitate as an inclusion body is the most likely explanation for the lack of expression of apo α-globin. This is consistent with the fact that the apoprotein can be made as a fusion construct incorporating a portion of the β-globin gene (Nagai & Thøgersen, 1987). It is well-known that α-globin is intrinsically less stable than β-globin or myoglobin. However, production of the completely assembled tetramer using the synthetic operon indicates that complete α chains can be synthesized in *E. coli*. Concomitant expression of β-globin may be necessary to provide a template for proper folding of the α chains and insertion of the heme group, protecting the polypeptide chain against proteolytic attack.

Amino Terminal Processing. The rules for processing the amino terminal formyl methionine in *E. coli* and the initial methionine in eukaryotic proteins are similar. Generally the methionine residue is cleaved if the second amino acid is small and neutral. There are, however, some differences between the two systems. In eukaryotes, if the second residue is valine, as is the case for α- and β-globin and myoglobin, the methionine residue is cleaved and valine becomes the first amino acid residue of the mature protein. This is not true in *E. coli*. In this case, the presence of valine as the second amino acid does not provide a signal for cleavage of the N-terminal methionine. Both the α and β chains of the complete tetramer produced using the synthetic operon, and the β chains produced using the T7 expression system, contain an additional methionine residue at the amino terminus. In both cases the N-terminal formyl group was cleaved. This is indicated by the fact that the proteins could be sequenced and the expected yield of amino acids was obtained. In the β1 Val → Met mutant the formyl group is also removed, but with histidine as the second amino acid of the β chains, the N-terminal methionine is not cleaved. This leads to the conservative amino acid substitution of β1 Val → Met. A second N-terminal mutant of the β chains was constructed, β1 Val → Ala. As expected, the N-terminal methionine was cleaved leaving alanine as the amino terminus. Cytochrome b_5, which has alanine as the second amino acid residue, also shows complete processing of the N-terminal methionine (Beck von Rodman et al., 1986).

Expression of Tetrameric Hemoglobin. The presence of both α- and β-globins is necessary for the expression of hemoglobin, and neither of the chains alone is capable of forming a holoprotein in *E. coli*. We used a high copy number plasmid and the strong *lac* promoter to drive transcription of a polycistronic message containing both the synthetic α- and β-globin genes. Under constitutive conditions we are able to produce hemoglobin with the complete complement of heme. This is consistent with the idea that a steady-state level of transcription is required for the incorporation of heme in overexpressed heme-containing proteins. The decrease in production of hemoglobin with the *tac* vector is probably due to a lower copy number of the plasmid. Inherently, the *tac* promoter is actually somewhat stronger than the *lac* promoter (Russel et al., 1985). Concomitant with the overexpression of the two globin chains there must be a large increase in the level of heme biosynthesis to produce the holoprotein. As large amounts of protoporphyrin IX are diverted to the production of hemoglobin, heme synthesis must be markedly elevated, suggesting that the biosynthetic pathway is subject to negative feedback regulation. From this and earlier work on cytochrome b_5 and myoglobin, it is clear that *E. coli* is capable of generating appropriate amounts of heme for the overexpression of heme-containing proteins (Beck von Bodman et al., 1986; Springer & Sligar, 1987).

Since the synthetic operon is able to express hemoglobin and the native cDNA sequences *in identical constructs* are not, there must be some feature within the native operons which inhibits expression of tetrameric hemoglobin. Secondary structure predictions of the mRNA (Zucker & Steigler, 1981; Frier et al., 1986) were performed to determine if there is any secondary structure around the Shine–Dalgarno ribosome binding site or the initiator methionine that might indicate a possible explanation for the absence of expression of the native operon. Both the native cDNA operon and the synthetic operon showed considerable secondary structure, with free energies of folding of −310.8 kcal/mol for the native genes and −276.4 kcal/mol for the synthetic gene for full-length transcripts. Although both transcripts show considerable structure, there is no apparent structural difference around the ribosome binding site or the initiator methionine that might explain the absence of expression of the native operons which was observed with both orientations of the α- and β-globin genes.

We attribute the high-level expression of hemoglobin with the synthetic genes to the biased usage of optimal codons (Ikemura, 1981). The actual codon usage for both the native cDNA and the synthetic genes is shown in Table II. The optimal codons chosen were those determined by deBoer and Kastelein (1986), and the biased codon usage is mainly reflected in the amino acids valine, alanine, threonine, proline, and lysine, while the codon CAU was used for histidine since no obvious codon bias is apparent (Ikemura, 1981; Grantham et al., 1980, Sharp et al., 1988). The codon usage was 83% optimal for the synthetic β-globin gene and 85% optimal for the synthetic α-globin gene. By comparison, the codon usage was only 46% optimal for both the native α- and

Table II: Codon Usage of Native and Synthetic α-Globin and β-Globin[a]

amino acid	codon	native α	synth α	native β	synth β	amino acid	codon	native α	synth α	native β	synth β
Ala	*GCT*	3	17	5	14	Pro	CCT	2		5	1
	GCC	11		8			CCC	3			
	GCA			3	1		CCA			2	
	GCG	7	1	1	1		*CCG*	2	7		6
Cys	TGT			2		Gln	CAA	1			
	TGC	1	1		2		*CAG*		1	3	3
Asp	GAT		2	5	3	Arg	CGT	1	2		2
	GAC	8	6	2	4		*CGC*		1		
Glu	*GAA*		3	2	7		CGA				1
	GAG	4	1	6	1		CGG	1			
Phe	TTT			5			AGG	1		3	
	TTC	7	7	3	8	Ser	*TCT*	3	7	1	5
Gly	*GGT*	2	5	4	12		*TCC*	4	2	2	
	GGC	5		8			TCA		1		
	GGA		2		1		TCG		1		
	GGG			1			AGT			2	
His	CAT		10	2	8		*AGC*	4			
	CAC	10		7	1	Thr	*ACT*		8	3	6
Lys	*AAA*	1	10	3	10		*ACC*	9		3	1
	AAG	10	1	8	1		ACA			1	
Leu	TTG		1				ACG		1		
	CTT	1	2		4	Val	*GTT*	1	11	3	17
	CTC	2	1	3	1		GTC	3		2	
	CTA	1			1		*GTA*		1		
	CTG	14	14	15	11		GTG	9	1	13	1
Met	ATG	3	3	2	2	Trp	TGG	1	1	1	1
Asn	AAT			1		Tyr	TAT	1		2	
	AAC	4	4	5	6		*TAC*	2	3	1	3

[a] Comparison of codon usage for the native and synthetic α- and β-globin genes. The number of times each codon is used within each of the genes is indicated. Italicized codons are those used for highly expressed genes (de Boer & Kastelein, 1986).

β-globin genes. The codon distribution for the synthetic genes and the native cDNAs is shown in Figure 5. It is apparent that both synthetic genes contain <20% suboptimal or rare codons and that they are widely dispersed. The systematic removal of rare codons in the coding region of tetanus toxin fragment C gene (Makoff et al., 1989) showed an increase in the expression level as more rare codons were replaced by optimal codons; however, there was also some increase in the level of mRNA reported. Others (Williams et al., 1988) have replaced the coding region of the interleukin-2 gene with a synthetic gene that increases the preferred codons from 43% to 85% and have demonstrated an increase of 16-fold in the expression level using the synthetic gene. In addition, Northern blot analysis of the mRNA content of both the native and synthetic constructs indicated no change in the message half-lives. However, the replacement of the first 25 codons of the highly expressed 3-phosphoglycerate kinase gene in yeast to rare codons did not appear to reduce the expression level (deBoar & Kastelein, 1986). Only when 30% of the entire gene was replaced with rare codons was expression decreased. As noted above, the rate of transcription is generally the major determinant of the level of gene expression. The *absolute requirement* for a high level of optimal codon usage for the expression of hemoglobin, and apparently other heme-containing proteins, is the exception rather than the rule. Sorenson et al. (1989) have shown that rare codons may cause pausing of the ribosome during translation. The use of optimal codons may ensure a proper rate of protein synthesis which allows for correct folding and insertion of the heme group within the polypeptide chains and permits assembly of the soluble holoprotein, which is stable to proteolytic digestion within *E. coli*.

In summary we have shown that optimal codon usage is necessary for the expression of intact hemoglobin within *E. coli*, and we have developed a new expression system for the convenient construction of β-chain mutants not requiring the synthesis of a fusion protein. In both cases the amino terminal methionine of the globin chains is retained. Therefore, we have constructed two additional β-chain mutants in which the extra methionine residue is deleted and which, after amino terminal processing in *E. coli*, results in the relatively conservative amino acid substitutions of β1 Val to Met and β1 Val to Ala. The functional properties and structures of these three variants are described in the following papers to determine which most closely mimics native hemoglobin and provides the best baseline for further studies.

ACKNOWLEDGMENT

We thank Dr. Saw Kym of the University of Illinois Genetic Engineering Facility for synthesis of oligonucleotides, amino acid analyses, and protein sequencing. All of the expression strains described in this paper, as well as the clones within the library of hemoglobin mutants that we have constructed, are available. Address requests to the corresponding author.

REFERENCES

Antonini, E. & Brunori, M. (1971) *Hemoglobin and Myoglobin in Their Reaction with Ligands*, North-Holland, Amsterdam.

Arnone, A. (1972) *Nature 237*, 146–149.

Baldwin, J. M. (1980) *J. Mol. Biol. 136*, 103–128.

Beck von Bodman, S., Schuler, M. A., Jollie, D. R., & Sligar, S. G. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9443–9447.

Bucci, E., & Fronticelli, C. (1965) *J. Biol. Chem. 240*, 551–552.

Chiancone, E., Norme, J. E., Forsen, S., Antonini, E., & Wyman, J. (1972) *J. Mol. Biol. 70*, 675–688.

deBoer, H. A., & Kastelein, R. A. (1986) in *Maximizing Gene Expression* (Reznikoff, W., & Gold, L. Eds.) pp 225–227, Butterworth, Stoneham, MA.

Egeberg, K. D., Springer, B. A., Sligar, S. G., Carver, T. E., Rohlfs, R. J., & Olson, J. S. (1990) *J. Biol. Chem. 265*, 11788–11795.

Ernst, J. F., & Kawashima, E. (1988) *J. Biotechnol. 7*, 1–10.

Fermi, G., Perutz, M. F., Shaanan, B., & Fourme, B. (1984) *J. Mol. Biol. 175*, 159–174.
Frier, S. M., Klerzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.
Fronticelli, C., O'Donnell, J. K., & Brinigar, W. S. (1991) *J. Protein Chem. 10*, 495–501.
Geraci, G., Parkhurst, L. J., & Gibson, Q. H. (1969) *J. Biol. Chem. 244*, 4664–4667.
Goldberg, D. P., Smith, D. H., & King, J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7060–7064.
Golomb, M., & Chamberlin, M. (1974) *J. Biol. Chem. 249*, 2858–2862.
Grantham, R., Gautier, C., & Gouy, M. (1980) *Nucleic Acids Res. 8*, 1893–1912.
Grodberg, J., & Dunn, J. J. (1988) *J. Bacteriol. 170*, 1245–1253.
Gross, G., Mielke, C., Hollatz, I., Blocker, H., & Frank, R. (1990) *J. Biol. Chem. 265*, 17627–17636.
Hoffman, S. J., Looker, D. L., Roehrich, J. M., Cozart, P. E., Durfee, S. L., Tedesco, J. L., & Stetler, G. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8521–8525.
Ikemura, T. (1981) *J. Mol. Biol. 151*, 389–409.
Imai, K., Fushitani, K., Miyazaki, G., Ishimori, K., Kitagawa, T., Wada, Y., Morimoto, H., Morishima, I., Shih, D. T., & Tame, J. (1991) *J. Mol. Biol. 218*, 769–778.
Johnson, M. L., Halvorson, H. R., & Ackers, G. K. (1976) *Biochemistry 15*, 5363–5371.
Laemmli, U. K. (1970) *Nature 227*, 680.
Lee, A. W., & Martin, K. (1983) *Proc. Nat. Acad. Sci. U.S.A. 80*, 7055–7059.
Lehmann, H. (1986) in *Hemoglobin: Molecular, Genetic and Clinical Aspects* (Bunn, E. H., & Forget, B. G., Eds.) pp 381–421, W. B. Saunders Co., Philadelphia, PA.
Lin, S.-H., Yu, N.-T., Tame, J., Shih, D., Renaud, J.-P., Pagnier, J., & Nagai, K. (1990) *Biochemistry, 29*, 5562–5566.
Luisi, B., & Shibayana, N. (1989) *J. Mol. Biol. 200*, 723–736.
Michelson, A. M., Markhem, A. F., & Orkin, S. H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 472–476.
Nagai, K., & Thøgersen, H. C. (1984) *Nature 309*, 810–812.
Nagai, K., & Thøgersen, H. C. (1987) *Methods Enzymol. 153*, 461–481.
Nagai, K., Luisi, B., Shih, D., Miyazaki, G., Imai, K., Poyart, C., De Young, A., Kwiatkowsky, L., Noble, R. W., Lin, S.-H., & Yu, N.-T. (1987) *Nature 329*, 858–860.
Nigen, A. M., Manning, J. M., & Alben, J. O. (1980) *J. Biol. Chem. 255*, 5525–5529.
O'Donnell, S., Mandoro, R., Shuster, T. M., & Arnone, A. (1979) *J. Biol. Chem. 254*, 12204–12208.
Olson, J. S., Mathews, A. J., Rohlfs, R. J., Springer, B. A., Egeberg, K. D., Sligar, S. G., Tame, J., Renaud, J.-P., & Nagai, K. (1988), *Nature 336*, 265–266.
Rohlfs, R. J., Mathews, A. H., Carver, T. E., Olson, J. S., Springer, B. A., Egeberg, K. D., & Sligar, S. G. (1990) *J. Biol. Chem. 265*, 3168–3176.
Russel, D. R., Miller, D. P., & Bennet, G. N. (1985) *Biochemistry 24*, 1410–1417.
Shaanan, B. (1983) *J. Mol. Biol. 171*, 31–59.
Sharp, P. M., & Li, W. H. (1987) *Nucleic Acids Res. 15*, 1218–1295.
Sharp, P. M., Cowe, E., Higgins, D. G., Sheids, D. C., Wolfe, K. J., & Wright, F. (1988) *Nucleic Acids Res. 16*, 8207–8211.
Smith, F. R., & Ackers, G. K. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5347–5351.
Sorenson, M. A., Kurland, C. G., & Pederson, S. (1989) *J. Mol. Biol. 207*, 365–377.
Speros, P. C., LiCata, V. J., Yonetani, T., & Ackers, G. K. (1991) *Biochemistry 30*, 7254–7262.
Springer, B. A., & Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8961–8965.
Tame, J., Shih, D. T., Pagnier, J., Fermi, G., & Nagai, K. (1991) *J. Mol. Biol. 218*, 761–767.
Unger, B. P., Gunsalus, I. C., & Sligar, S. G. (1986) *J. Biol. Chem. 261*, 1158–1163.
Varaderajan, R., Szabo, A., & Boxer, S. G. (1985) *Proc. Natl. Acad. Sci. U.S.A 82*, 5681–5684.
Walder, R. Y., Walder, J. A., & Donelson, J. E. (1984) *J. Biol. Chem. 259*, 8015–8026.
Williams, D. P., Regeir, D., Akiyoshi, D., Genbauffe, F., & Murphy, J. R. (1988) *Nucleic Acids Res. 16*, 10453–10467.
Yip, Y. K., Waks, M., & Beychok, S. (1972) *J. Biol. Chem. 247*, 7237–7244.
Zucker, M., & Steigler, P. (1981) *Nucleic Acids Res. 9*, 1333.