

Bacterial expression and spectroscopic characterization of soybean leghaemoglobin *a*

Deborah K. JONES*, Ramin BADII†, Federico I. ROSELL‡ and Emma LLOYD*¹

*Department of Chemistry, University of Leicester, University Road, Leicester, LE1 7RH, England, U.K., †Biological NMR Centre, Department of Biochemistry, University of Leicester, LE1 9HN, and ‡Department of Biochemistry and Molecular Biology, University of British Columbia, 2146 Health Sciences Mall, Vancouver, BC, V6T 1Z3, Canada

A gene encoding leghaemoglobin *a* from soybean has been constructed and the soluble recombinant protein expressed in *E. coli*. The integrity of the recombinant protein has been assessed by a range of spectroscopic techniques. Electrospray mass spectrometry of the protein indicates that the molecular mass of the protein corresponds to the predicted amino acid sequence. Circular dichroism spectra of the ferric derivative and UV-visible spectra of various ferric and ferrous derivatives (pH 6.99, $\mu = 0.10$ M, 25.0 °C) are consistent with published data for the

wild-type protein. For the ferric derivative, UV-visible (298 and 77 K) and EPR (10 K) spectra indicate the existence of a thermal equilibrium between high- and low-spin forms. Titration of the protein (0.10 M NaCl, $\mu = 0.10$ M, 25.0 °C) between pHs 6.68 and 10.35 indicate formation ($pK_a = 8.3 \pm 0.03$) of a 6-coordinate, hydroxide-bound form of the protein at high pH. All of the above data are consistent with the behaviour of the wild-type protein.

INTRODUCTION

The existence of an oxygen-binding protein, similar to haemoglobin, in the root nodules of leguminous plants was first proposed in 1939 [1] and later confirmed independently [2,3]. The term 'leghaemoglobin' was coined in 1945 [4] and since that time these proteins have been the subject of extensive study [5,6]. Of particular interest has been the extremely high oxygen affinity of these proteins, which is now known to derive from a very fast rate constant for ligand binding, together with a moderately slow rate for ligand dissociation [6,7]. Of all the leghaemoglobins, the soybean protein is by far the most well-characterized, with extensive spectroscopic data [5,7,8] and a crystal structure for the nicotinate-bound derivative both available [9,10]. In spite of this, many of the unique functional aspects remain unexplained at the molecular level, and the exact role of protein structure in the control of function is relatively ill-defined.

Recently, cDNA sequences for a number of leghaemoglobins (Lb) have appeared, including *Medicago sativa* (alfalfa) [11–14], *Sesbania rostrata* [15], *Vigna unguiculata* (cowpea) [16], *Lupinus luteus* (lupin) [17,18] and *Glycine max* (soybean) [19,20]. This sequence information has allowed development of expression systems for lupin LbI [21] cowpea LbII [16] and soybean Lba [22,23]. However, these published expression systems have certain limitations: the lupin Lb yields very low amounts of recombinant protein, the cowpea Lb was not subjected to a detailed spectroscopic analysis and the soybean expression system yields insoluble protein in the form of inclusion bodies. In this paper, we report the generation of a synthetic gene for soybean Lba and, for the first time, expression of soluble Lba in *E. coli*. In addition, using a range of spectroscopic techniques not previously applied to a recombinant Lb, we have established that the recombinant protein is an authentic duplicate of wild-type Lba.

EXPERIMENTAL

Assembly of the synthetic gene

The host strain for the initial cloning experiments was *E. coli* NM522 [F' , *lacI*^q, $\Delta(lacZ)M15$, *proAB/supE*, *thi*, $\Delta(lac-proAB)$, $\Delta(hsdMS-mcrB)5$, (r_K^- , m_K^-)]. The construct was expressed in *E. coli* BL21 (DE3) *ompT*, (r_B^- , m_B^-). pET11a and *E. coli* BL21 (DE3) were purchased from Novagen. Oligonucleotides were synthesized on an Applied Biosystems 394 DNA Synthesizer and were purified by precipitation with sodium acetate and ethanol at –70 °C. The gene was assembled in two steps. In the first step, gaps between the six overlapping oligonucleotides coding for the Lba sequence were filled by PCR [20 pmol of each oligonucleotide, 25 mM dNTPs, *Pfu* buffer, *Pfu* DNA polymerase (5 μ l, Stratagene) in 100 μ l total volume]. In the second step, restriction sites were introduced to the 5' and 3' ends of the Lba gene (same conditions as above, but using 5 μ l of the DNA from the first PCR plus two oligonucleotides containing restriction sites for *NdeI* and *BamHI*). All PCR reactions were initiated by a denaturation step of 2 min at 92 °C followed by ten cycles (first PCR) or 15 cycles (second PCR) of denaturation (1 min, 92 °C), annealing [1 min, 30 °C (first PCR) and 68 °C (second PCR)] and extension (1 min, 72 °C). Completion of polymerization was ensured by a further 5 min incubation at 72 °C. The PCR product was purified using a Hybaid Recovery Plasmid Purification Kit (Hybaid, Middlesex).

The PCR product and the pET11a expression vector were digested with *NdeI* and *BamHI* (New England Biolabs) in One-Phor-All Buffer Plus (Pharmacia). The restriction enzymes were heat-inactivated at 70 °C for 20 min followed by a further purification and buffer exchange with the Hybaid Plasmid Recovery Kit. A 1:1 molar ratio ligation and transformation of the DNA samples were carried out using T4 DNA ligase

Abbreviations used: Lb, leghaemoglobin; Lba, soybean leghaemoglobin *a*.

¹ To whom correspondence should be addressed.

(Gibco-BRL) and *E. coli* NM522. DNA sequencing showed that a deletion mutation had occurred at position 116. The deletion was corrected by site-directed mutagenesis (Stratagene Quik-Change kit) and the entire gene resequenced. The Dye DeoxyTM Terminator (Perkin-Elmer) sequencing reactions were carried out on an Applied Biosystems Model 373A DNA Sequencer using Qiagen purified plasmid DNA.

Expression of *Lba* in *E. coli*

Large-scale expression of *Lba* was carried out using frozen glycerol stocks as starting inoculum for an LB plate containing 100 µg/ml ampicillin. Single colonies were then used to inoculate 100 ml of LB media containing 100 µg/ml ampicillin, and the flask was shaken (250 r.p.m., 37 °C) overnight. This culture (50 ml) was then used to seed 750 ml of LB/ampicillin in a 2 l flask, and this culture grown (250 r.p.m., 37 °C) to an optical density at 600 nm of 0.7–0.9, at which time the cells were induced with isopropyl β-D-thiogalactopyranoside (final concentration 1 mM) and grown for a further 24 h (150 rev./min, 28 °C). Cells were harvested at 4 °C by centrifugation at 16000 *g* and resuspended in cold lysis buffer (50 mM Tris/HCl, pH 8.0, 1 mM EDTA). Lysozyme (100 mg) was added, and the cells were gently shaken for 60 min before freezing at –20 °C until required.

Protein isolation and purification

All the following steps were performed at 4 °C. The cells were defrosted, deoxyribonuclease (5 mg, Sigma D4527) and 1 M MgCl₂ (5 ml) added, and the cells shaken until the solution was no longer viscous. The cells were centrifuged (16000 *g*, 45 min) and the red supernatant collected. The pellet was resuspended in lysis buffer (50–100 ml) and the solution recentrifuged (16000 *g*, 45 min). This procedure was repeated until the pellet had no perceptible colour. The supernatant was brought to 45% ammonium sulphate saturation [26.2 g (NH₄)₂SO₄/100 ml], centrifuged (16000 *g* 30 min) and the supernatant brought to 100% ammonium sulphate saturation [by addition of a further 38.8 g (NH₄)₂SO₄/100 ml]. The solution was centrifuged (16000 *g*, 45 min), the pellet resuspended in buffer and dialysed against water (20 litres) for 36 h. Afterwards, the sample was centrifuged and loaded onto a 2.5 × 20 cm FFQ Sepharose column (Pharmacia) which had been equilibrated in 2 mM Tris/HCl pH 9.2. The column was washed with three column volumes of the same buffer and eluted using a linear gradient of 0–0.3 M NaCl in 2 mM Tris/HCl pH 9.2 (350 ml of each). Fractions with an $A_{403}/A_{280} > 2$ were pooled, loaded onto a second FFQ Sepharose column and re-eluted using the same gradient. Fractions with an $A_{403}/A_{280} > 3$ were pooled and loaded onto a 2.5 × 100 cm Sephadex G50 column equilibrated in 20 mM Tris/HCl pH 8.0. Pure fractions eluting from this column showed absorbance ratios $A_{403}/A_{280} > 5.0$ and migrated as a single band on SDS/PAGE. Absorption coefficients were determined using the pyridine-haemochromogen method [24–26]. Samples of the anion-bound adducts were obtained by addition of excess NaN₃ or KCN to oxidized *Lba*. Samples of reduced *Lba* were prepared by addition of a small excess of solid sodium dithionite.

Mass spectrometry

Protein samples ([*Lba*] = 15.22 µM) were prepared by exhaustive exchange into deionized water followed by addition of acetic acid (final concentration 1%) and methanol (final concentration 5%). Spectra were recorded using a customized quadrupole mass spectrometer (sprayer voltage 4900 V, flow rate 5 µl/min).

Mass deconvolution was carried out using a Bio Multiview program package purchased from Perkin-Elmer/Sciex.

Electronic spectroscopy

Electronic spectra were collected using a Cary 219 spectrophotometer interfaced to a microcomputer (Olis, Bogart, GA) and fitted with a circulating thermostatted water bath (±0.1 °C). Wavelength maxima are accurate to ±0.5 nm. Low-temperature (77 K) spectra were recorded with a SLM-Aminco Model DW-2C dual wavelength spectrophotometer (2 mm path length).

pH titration

Titration (0.10 M NaCl) were carried out by addition (~2 µl) of 0.10 M NaOH. The pH was recorded before and after measurement of each spectrum. A nonlinear least squares fitting program (Scientist v.4.02, MicroMath) was used to fit the data to a single-proton process, eqn. (1),

$$Z = \frac{A + B \times 10^{(pH - pK_a)}}{1 + 10^{(pH - pK_a)}} \quad (1)$$

where *Z* is the absorbance and *A* and *B* are the absorbancies of the acidic and basic forms respectively.

EPR spectroscopy

EPR spectra were collected using a Bruker Model ESP 300E spectrometer equipped with an Oxford Instruments ESR900 continuous flow cryostat, an Oxford Instruments ITC4 temperature controller, and an HP5352B frequency counter. Spectra (9.45 GHz) were collected using 1.00 mW microwave power with a modulation frequency and amplitude of 100 kHz and 0.46 mT respectively. Protein samples (200 µl, 2 mM) were prepared in 50 mM sodium phosphate, pH 7.0, and then diluted 1:1 with 100% glycerol.

Circular dichroism spectroscopy

CD spectra (sodium phosphate, pH 6.99, µ = 0.10 M) were measured with a JASCO J720 spectropolarimeter that was equipped with a computer-interfaced Neslab RTE111 water bath and a water-jacketed cylindrical quartz cuvette (0.1 or 1 mm pathlength). Reported spectra were the average of 5 scans. Thermal denaturations were monitored at 222 nm by applying a thermal gradient of 50 °C/h. Each melting curve was smoothed, and the first derivative was calculated using the JASCO software utilities to determine the inflection point (*T_m*) of the curve.

Magnetic circular dichroism spectroscopy

MCD spectra were measured with a JASCO J720 spectropolarimeter fitted with an electromagnet (Alpha Magnetics, Hayward, CA) operating at approx. 1 T. Protein samples were prepared in D₂O (50 mM sodium phosphate, pD 7.0 or 10.83) and data collected using a quartz cuvette (1 cm pathlength) at 25 °C.

RESULTS

Gene synthesis and protein expression

The DNA sequence of the *Lba* gene was derived from the published DNA sequence [20] and from the UWGCG Molecular Biology Package [26], which was used to incorporate codons

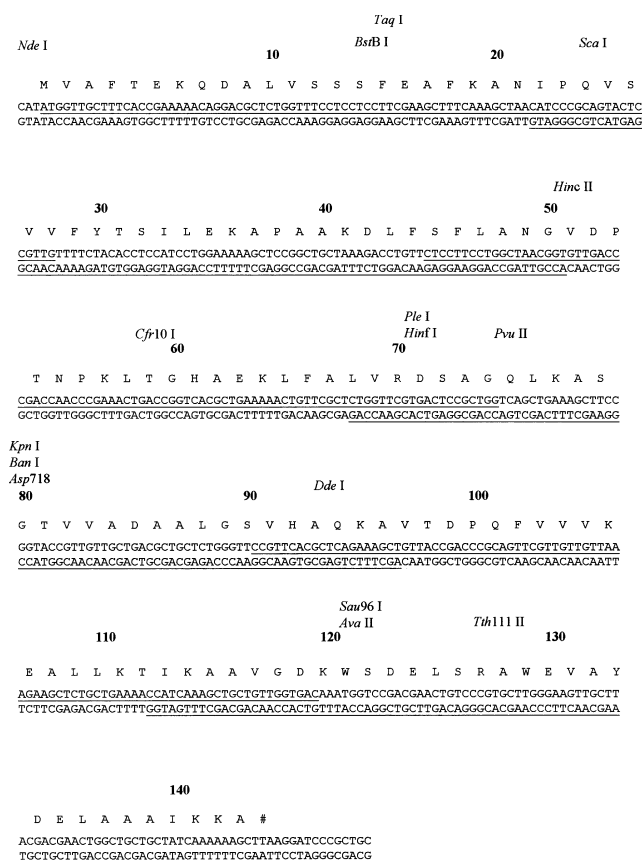


Figure 1 Nucleotide sequence of the synthetic gene encoding for soybean Lba, together with the corresponding amino acid sequence

The gene was designed on, and uses the same numbering system as, the cDNA sequence of Hyldig-Nielsen et al. [20], and utilizes codons found in highly expressed *E. coli*. The individual oligonucleotides used for synthesis of the gene are underlined. Unique restriction sites are indicated and stop codons are marked with a #.

which are present in highly expressed genes of *E. coli*. Based on the newly derived DNA sequence, six partially overlapping oligonucleotides were designed, ranging in size from 85 to 90 bases and covering both strands of the gene. Overlapping oligonucleotides were designed to allow the second strand of the DNA to be synthesized enzymatically during the PCR reaction (Figure 1).

Initial attempts to grow the cells at 37 °C and 250 r.p.m. resulted in the formation of insoluble inclusion bodies. Subsequently, the growth conditions were altered and the cells allowed to grow more slowly (28 °C, 150 r.p.m.). This resulted in formation of substantial (20 mg pure Lba per litre of culture) amounts of haem-containing protein.

Mass spectrometry

Mass spectrometry of Lba revealed two components. The major peak, with a mass of 15235.4 ± 2.74 , corresponds to that expected for the apo-protein with the N-terminal methionine cleaved ($MW_{\text{calc}} = 15233.1$). The minor peak, with mass 15364.6 ± 2.76 , corresponds to the mass of the apo-protein with the N-terminal methionine intact ($MW_{\text{calc}} = 15364.1$).

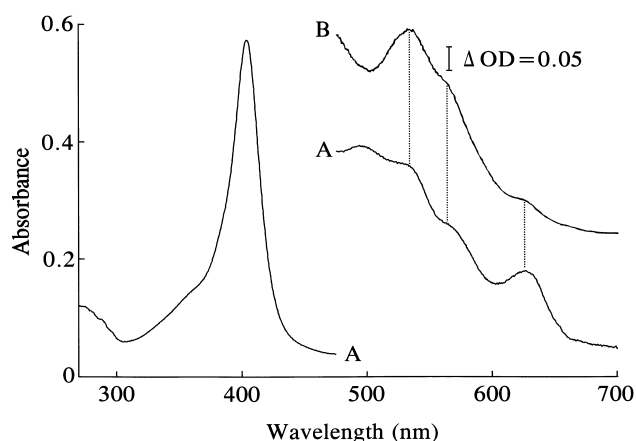


Figure 2 UV-visible spectra of ferric Lba at (A) 298 K and (B) 77 K

The absorbance scale for the 298 K spectrum in the visible region (475–700 nm) has been multiplied by ten. The absorbance scale for the 77 K spectrum is indicated separately. Conditions: sodium phosphate pH 6.99, $\mu = 0.10$ M.

Table 1 Wavelength maxima (nm) and, in parentheses, absorption coefficients ($\text{mM}^{-1} \cdot \text{cm}^{-1}$) for various ferric and ferrous derivatives of Lba (phosphate pH 6.99, $\mu = 0.10$ M, 25.0 °C)

Derivative	Recombinant Lba
Fe(III)	403 (157), 495 (10.7), 530 (9.91), 560 (7.13), 626 (4.92)*
Fe(III)-CN ⁻	415 (117), 539 (11.9)†
Fe(III)-N ₃ ⁻	413.5 (126), 549 (12.4), 572 (9.36)‡
Fe(III)-OH ⁻	408.5 (117), 539 (11.3), 570 (10.1)§
Fe(II)	427 (111), 554.5 (13.3)
Fe(II)-CO	415 (199), 537 (13.6), 561 (12.7) ¶

* Literature values: 403, 495, 530, 560, 627 nm [27,28].
 † Literature values: 415–417, 538–540 nm [5].
 ‡ Literature values (for kidney bean Lba): 414.5, 539, 573 [42].
 § Literature values: 412, 572, 541 nm [5,43].
 || Literature values: 427, 555–556 nm [5].
 ¶ Literature values: 417, 538, 562–563 nm [5,44].

Electronic absorption spectra

The spectrum of the ferric derivative of recombinant Lba (Figure 2A; pH 6.99, $\mu = 0.10$ M, 25.0 °C), reveals a mixture of species. Transitions characteristic of both high-spin ($\lambda_{\text{max}} = 403, 495, 626$ nm) and low-spin ($\lambda_{\text{max}} = 530, 560$ nm) species are apparent, as has been previously reported [27,28] for the wild-type protein ($\lambda_{\text{max}} = 403, 495, 530, 560, 627$ nm). At 77 K, the low spin component ($\lambda_{\text{max}} = 408, 530, 561$ nm) dominates (Figure 2), with only small amounts of high-spin species remaining ($\lambda_{\text{max}} \sim 620$ nm). Similar behaviour has been previously reported for the wild-type protein [6,28]. The spectrum of the oxy derivative compared well with that of the wild-type protein, with wavelength maxima at 540 and 573 nm compared with 540 and 574 for the wild-type [44]. Absorption maxima and corresponding absorption coefficients for the various ferric and ferrous derivatives of Lba (Table 1), are also similar to those for the wild-type protein.

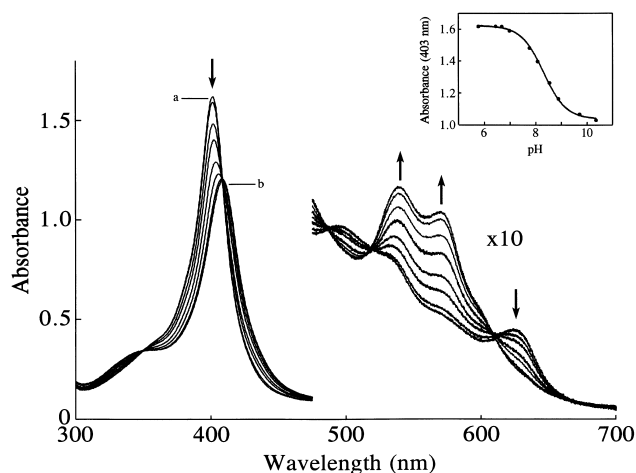


Figure 3 Spectrophotometric pH titration of ferric Lba

(a) pH 6.68, (b) pH 10.35 and intermediate pHs 6.98, 7.76, 8.07, 8.55, 8.89, 9.71. Arrows indicate the direction of change in absorbance with pH. Inset: plot of the variation in absorbance at 403 nm with pH (additionally including pHs 5.77 and 6.44). The solid line represents a nonlinear least squares fit to eqn. (1). Conditions: $[Lba] = 10.32 \mu\text{M}$, 0.10 M NaCl, $\mu = 0.10 \text{ M}$, 25.0 °C.

The titration behaviour of the recombinant Lba (Figure 3), is similar to that previously reported [29,30]. Hence, as the pH is raised, titration of the distal water molecule and subsequent formation of a hydroxide-bound derivative is observed [see eqn. (2)].



A fit of these data to eqn. (1) yielded a value for this $\text{p}K_a$ of 8.3 ± 0.03 [$\mu = 0.10 \text{ M}$ (NaCl), 25.0 °C], compared to the reported value [8] for the wild-type protein of 8.34 ($\mu = 0.10 \text{ M}$,

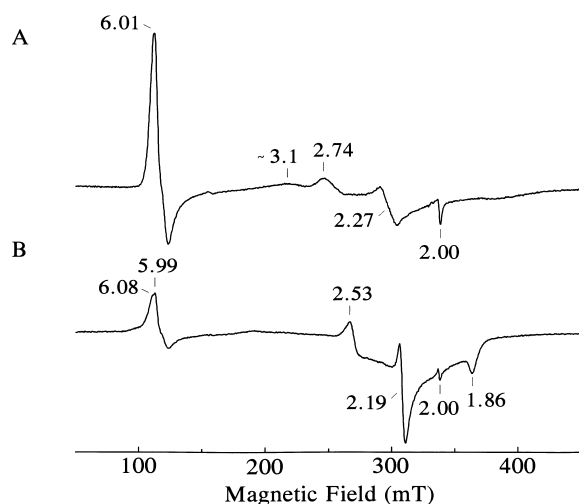


Figure 4 Low-temperature X-band EPR spectra and g values for ferric Lba

(A) At pH 6.99 and 4 K and (B) at pH 9.92 and 10 K ($[Lba] \sim 2 \text{ mM}$, 50 mM sodium phosphate pH 7.0, 50% glycerol). Conditions: microwave frequency 9.45 GHz, microwave power 1.00 mW, modulation frequency 100 kHz, modulation amplitude 0.46 mT.

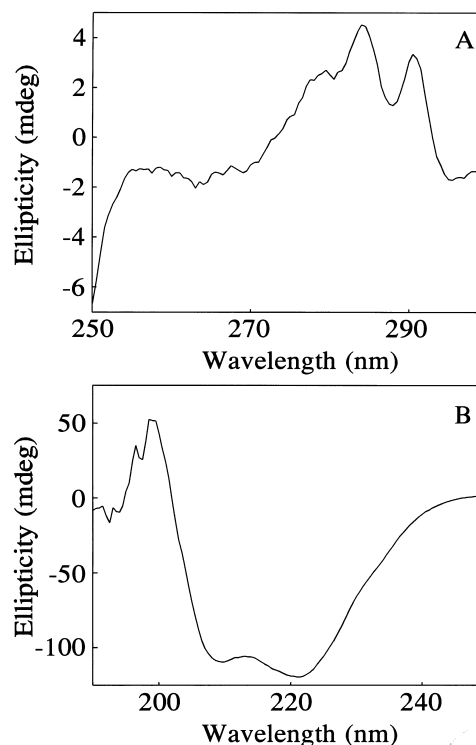


Figure 5 CD spectra of ferric Lba

(A) In the near-UV region; (B) in the far-UV region. Conditions: $[Lba] = 30.2 \mu\text{M}$, sodium phosphate, pH 6.99, $\mu = 0.10 \text{ M}$, 25.0 °C.

20 °C). Wavelength maxima for the hydroxide-bound derivative of Lba are given in Table 1.

EPR spectroscopy

The low-temperature (10 K) spectrum of Lba at pH 6.99 (Figure 4A), showed the presence of a high-spin component ($g = 6.01$ and 2.00), similar to that previously reported [31] for the wild-type protein ($g = 5.99$ and 2.00). Low-spin features are also present in the spectrum, notably at $g = 2.74$ and 2.27. This species has been detected previously ($g = 2.72$ and 2.26 [31]; $g = 2.76$ and 2.31 [32]; $g = 2.69$ and 2.24 [28]) and has been assigned to the formation of a bis-imidazole species. Some high-spin species still persists even at 10 K, consistent with previous work [31] and the UV-visible data at 77 K presented above. A weak feature at $g \sim 3.1$ is also apparent, which has been detected previously [28,31] and probably derives from a second low-spin form. As the pH is raised (Figure 4B), formation of a hydroxide-bound derivative is observed ($g = 2.53$, 2.19 and 1.86), consistent with the electronic absorption data (above) and with previous findings ($g = 2.51$, 2.19 and 1.86 [31]; $g = 2.63$, 2.23 and 1.84 [32]; $g = 2.54$, 2.19 and 1.84 [28]; $g = 2.54$, 2.24 and 1.84 [33]).

Circular dichroism spectroscopy

The CD spectra of ferric Lba in the near- and far-UV regions are shown in Figures 5A and 5B respectively. In the near-UV region (250–300 nm), maxima (at 284 and 290.5 nm) and minima (at 287 and 295.5 nm) are observed, consistent with published data for the wild-type protein (maxima at 283 and 290 nm; minima 286 and 294 nm) [34]. In the far-UV region, minima at 209.5 and

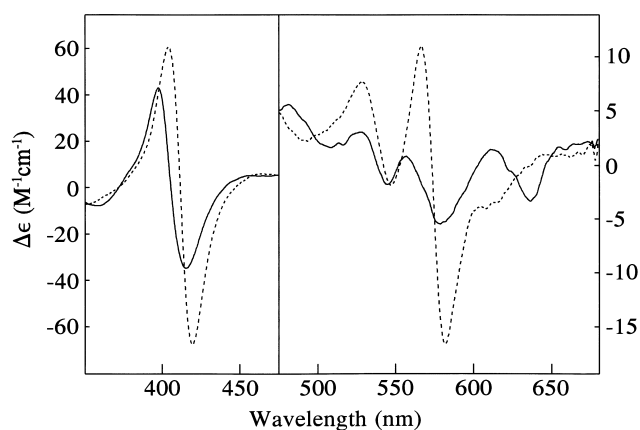


Figure 6 Room-temperature MCD spectra of recombinant Lba

(—), at pD 7.0; (---), at pD 10.83. Conditions: [Lba] = 112 μ M, 50 mM sodium phosphate, pD 7.0 or 10.83, 25 $^{\circ}$ C, 1 T.

221 nm are observed, consistent with previous work on the wild-type protein (minima at 208 and 220 nm) [34].

Thermal denaturation of Lba (monitored at 222 nm) and the melting temperature, obtained from the first derivative of the melting curve [35], was found to be 67 ± 0.5 $^{\circ}$ C (not shown).

Magnetic circular dichroism spectroscopy

The MCD spectra of recombinant Lba at pDs 7.0 and 10.8 are presented in Figure 6. In the visible region, the spectrum shows similar features to those previously reported for the wild-type protein [32], with positive peaks at 397, 527, 556 and 613 nm, and negative troughs at 415, 545, 579 and 636 nm. At alkaline pH, formation of a low-spin species is observed (positive peaks at 404, 429 and 567 nm; negative troughs at 419, 547 and 581 nm), consistent with the formation of a hydroxide-bound derivative at high pH.

DISCUSSION

The present report describes the synthesis of a gene encoding soybean Lba, expression of the soluble protein in *E. coli* and complete spectroscopic characterization of the protein. The synthetic gene was based on the sequence for the cDNA [20], and not on the published amino acid sequence [36], which, although subsequently corrected [37], does not match that for the cloned DNA. We have, for the first time, obtained large yields of soluble, haem-containing protein, and this has allowed us to examine in detail the spectroscopic properties of the molecule and compare them to those published for the wild-type protein. The pET series of vectors [38] have been extensively used for high-level expression of recombinant proteins in *E. coli* [21,39,40] and the incorporation of *E. coli* codon bias in the synthetic gene together with the powerful T7 RNA polymerase promoter of pET11a greatly improves the yield of recombinant protein. Interestingly, soluble holo-protein was only obtained at lower temperatures, and alteration of growth conditions to slow down expression of protein seemed to prevent the formation of insoluble inclusion bodies.

The integrity of the recombinant apo-protein was authenticated using MS. This indicated that the mass of the protein matched the predicted amino acid sequence, and that no post-

translation modifications had occurred. Interestingly, cleavage of the N-terminal methionine was observed. Similar behaviour has been reported for myoglobin [40].

The authenticity of the holo-protein was examined in a number of ways, and all of the spectroscopic data presented in this paper are consistent with previous findings for the wild-type protein. CD spectra (Figure 5) are consistent with spectra for wild-type Lba [34], indicating that there are no structural differences between the recombinant and wild-type proteins. The melting temperature of the protein, measured as 67 ± 0.5 $^{\circ}$ C, is similar to that previously reported (~ 75 $^{\circ}$ C) for wild-type myoglobin [41]. The electronic and MCD spectra of the recombinant Lba are also identical to the wild-type protein. Furthermore, the protein behaves normally in the sense that it readily binds exogenous ligands at the haem iron in both the ferric and ferrous oxidation states (Table 1). For the ferric derivative, the existence of a mixed-spin thermal equilibrium at ambient temperatures is readily apparent. Hence, the UV-visible (Figure 2) and EPR (Figure 3) spectra of ferric Lba show features characteristic of high-spin and low-spin components. The high-spin component is in thermal equilibrium with the low-spin form and preferential formation of the low-spin derivative at low temperature is observed (Figures 2B and 3A). The high-spin form is known to derive from a 6-coordinate, met-aquo derivative, similar to that observed in mammalian oxygen carriers [24]. The low-spin derivative is also 6-coordinate, but contains a histidine residue, thought to be the distal histidine [6], as the sixth ligand. This spin equilibrium is further complicated by the existence of a pH-dependent equilibrium of the met-aquo form. Hence, on raising the pH, a second low-spin derivative is observed (Figures 2 and 3B), this time deriving from titration, with $pK_a = 8.3$, of the distal water molecule to hydroxide. MCD spectroscopy, which is a sensitive diagnostic for axial ligands in haem proteins, is also consistent with formation of a hydroxide-bound species.

Conclusions

We have successfully synthesized a gene for soybean Lba, and have obtained large yields of soluble, haem-containing protein from *E. coli*. Using a multitude of spectroscopic techniques, we have established that the recombinant protein is identical to the wild-type. With the fundamental properties of the recombinant protein now defined, generation of site-specific variants of Lba will allow the functional role of a number of key amino acids to be examined.

We gratefully acknowledge the contributions of Professors Grant Mauk, Phillip Bragg (University of British Columbia) and Gordon Roberts (University of Leicester) for their generous support of this work, Lars Konermann for assistance with mass spectrometry and Kuldip Singh for technical assistance. This work was supported by grants from the Royal Society, the Nuffield Foundation and the University of Leicester (to E.L.), the MRC in Canada (Grant MT-14021 to A. G. Mauk) and the Wellcome Trust (to G. C. K. Roberts). E.L. gratefully acknowledges the Association of Commonwealth Universities for a Development Fellowship, which allowed collaboration between the University of British Columbia and the University of Leicester. D.K.J. acknowledges the University of Leicester, for a PhD studentship. Oligonucleotide synthesis and automated sequencing runs were performed by the Protein and Nucleic Acid Chemistry Laboratory of the University of Leicester.

REFERENCES

- 1 Kubo, H. (1939) *Acta Phytochim.* **11**, 195
- 2 Keilin, D. and Wang, Y. L. (1945) *Nature (London)* **155**, 227
- 3 Virtanen, A. I. (1945) *Nature (London)* **155**, 747
- 4 Virtanen, A. I., Jorma, J. and Laine, T. (1945) *Suom. Kemistil.* **B18**, 49–50
- 5 Dilworth, M. J. (1980) *Methods Enzymol.* **69**, 812–823
- 6 Appleby, C. A., Blumberg, W. E., Bradbury, J. H., Fushman, W. H., Peisach, J., Wittenberg, B. A., Wittenberg, J. B. and Wright, P. E. (1982) in *Hemoglobin and Oxygen Binding* (Ho, C., ed.), pp. 435–441, Elsevier, North Holland

- 7 Appleby, C. A. (1974) in *The Biology of Nitrogen Fixation* (Quispel, A., ed.), pp. 521–554, North-Holland, Amsterdam
- 8 Appleby, C. A. and Dilworth, M. J. (1979) in *A Treatise on Dinitrogen Fixation* (Burns, R. C., ed.), pp. 691–765, Wiley, New York
- 9 Ollis, D. L., Appleby, C. A., Colman, P. M., Cutten, A. E., Guss, J. M., Ventakappa, M. P. and Freeman, H. C. (1986) *Aust. J. Chem.* **36**, 451–468
- 10 Ellis, P. J., Appleby, C. A., Guss, J. M., Hunter, W. N., Ollis, D. L. and Freeman, H. C. (1997) *Acta Crystallogr.* **D53**, 302–310
- 11 Kiss, G. B., Végh, Z. and Vincze, É. (1987) *Nucleic Acids Res.* **15**, 3620
- 12 Davidowitz, E. J., Dow, A. and Lang-Unnasch, N. (1989) *Nucleic Acids Res.* **17**, 3307
- 13 Davidowitz, E. J., Creissen, G., Vincze, É., Kiss, G. B. and Lang-Unnasch, N. (1991) *Plant Mol. Biol.* **16**, 161–165
- 14 Löbler, M. and Hirsch, A. M. (1992) *Plant Mol. Biol.* **20**, 733–736
- 15 Welters, P., Metz, B. A., Schell, J. and de Bruijn, F. J. (1989) *Nucleic Acids Res.* **17**, 1253
- 16 Arrendondo-Peter, R., Moran, J. F., Sarath, G., Luan, P. and Klucas, R. V. (1997) *Plant Physiol.* **114**, 493–500
- 17 Konieczny, A. (1987) *Nucleic Acids Res.* **15**, 6742
- 18 Konieczny, A., Jensen, E. O., Marcker, K. A. and Legocki, A. B. (1987) *Mol. Biol. Rep.* **12**, 61–66
- 19 Wiborg, O., Hyldig-Nielsen, J. J., Jensen, E. O., Paludan, K. and Marcker, K. A. (1982) *Nucleic Acids Res.* **10**, 3487–3495
- 20 Hyldig-Nielsen, J. J., Jensen, E. O., Paludan, K., Wiborg, O., Garrett, R., Jørgensen, P. and Marcker, K. A. (1982) *Nucleic Acids Res.* **10**, 689–701
- 21 Sikorski, M. M., Topunov, A. F., Strozycycki, P. M., Vorgias, C. E., Wilson, K. S. and Legocki, A. B. (1995) *Plant Sci.* **108**, 109–117
- 22 Prytulla, S., Dyson, H. J. and Wright, P. E. (1996) *FEBS Lett.* **399**, 283–289
- 23 Hargrove, M. S., Barry, J. K., Brucker, E. A., Berry, M. B., Phillips, G. N., Olson, J. S., Arrendondo-Peter, R., Dean, J. M., Klucas, R. V. and Sarath, G. (1997) *J. Mol. Biol.* **266**, 1032–1042
- 24 Antonini, M. and Brunori, E. (1971) *Hemoglobin and Myoglobin and their Reactions with Ligands* (Neuberger, A. and Tatum, E., eds.), p. 10, North-Holland, Amsterdam
- 25 de Duve, C. (1948) *Acta Chem. Scand.* **2**, 264–289
- 26 Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
- 27 Appleby, C. A., Nicola, N. A., Hurrell, J. G. R. and Leach, S. J. (1975) *Biochemistry* **14**, 4444–4450
- 28 Appleby, C. A., Blumberg, W. E., Peisach, J., Wittenberg, B. A. and Wittenberg, J. B. (1976) *J. Biol. Chem.* **251**, 6090–6096
- 29 Pertillä, U. (1981) *Acta Chem. Scand.* **B35**, 435–440
- 30 Ellfolk, N. (1961) *Acta Chem. Scand.* **B31**, 975
- 31 Maskall, C. S., Gibson, J. F. and Dart, P. J. (1977) *Biochem. J.* **167**, 435–445
- 32 Sievers, G., Gadsby, P. M. A., Peterson, J. and Thomson, A. J. (1983) *Biochim. Biophys. Acta* **742**, 637–647
- 33 Gadsby, P. M. A. and Thomson, A. J. (1990) *J. Am. Chem. Soc.* **112**, 5003–5011
- 34 Ellfolk, N. and Sievers, G. (1975) *Biochim. Biophys. Acta* **213**, 213–227
- 35 Gwynneth, M. E., Huber, H. E., DeFreo-Jones, D., Vuocolo, G., Goodhart, P. J., Maigetter, G. S., Oliff, A. and Heimbrook, D. C. (1992) *J. Biol. Chem.* **267**, 7971–7974
- 36 Ellfolk, N. and Sievers, G. (1971) *Acta Chem. Scand.* **25**, 3532–3548
- 37 Ellfolk, N. and Sievers, G. (1974) *Acta Chem. Scand.* **28**, 1245–1246
- 38 Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
- 39 Badii, R., Basran, J., Casarotto, M. G. and Roberts, G. C. K. (1995) *Protein Express. and Purif.* **6**, 237–243
- 40 Lloyd, E., Burk, D. L., Ferrer, J. C., Maurus, R., Doran, J., Carey, P., Brayer, G. D. and Mauk, A. G. (1996) *Biochemistry* **35**, 11901–11912
- 41 Hildebrand, D. P., Burke, D. L., Maurus, R., Ferrer, J. C., Brayer, J. D. and Mauk, A. G. (1995) *Biochemistry* **34**, 1997–2005
- 42 Lehtovaara, P. (1977) *Acta Chem. Scand.* **B31**, 21–27
- 43 Sternberg, H. and Virtanen, A. I. (1952) *Acta Chem. Scand.* **6**, 1342
- 44 Fushman, W. H. and Appleby, C. A. (1979) *Biochemistry* **18**, 1309–1321