Aplysia limacina myoglobin cDNA cloning: an alternative mechanism of oxygen stabilization as studied by active-site mutagenesis

Francesca CUTRUZZOLÀ, Carlo TRAVAGLINI ALLOCATELLI, Andrea BRANCACCIO* and Maurizio BRUNORI†

Istituto Pasteur–Fondazione Cenci Bolognetti and Dipartimento di Scienze Biochimiche 'A. Rossi Fanelli', Università di Roma 'La Sapienza', Piazzale A. Moro, 5-00185 Roma, Italia

The isolation and cloning of the cDNA coding for myoglobin (Mb) from the mollusc *Aplysia limacina* is reported here. Five amino acid differences from the previously published protein sequence have been found in positions 22, 26, 27, 77 and 80 by back translating the cDNA; some of these may be relevant for overall structure stabilization in this Mb. High-level expression of the holoprotein in *Escherichia coli* has been achieved in the presence of the haem precursor δ-aminolevulinic acid, underlying the importance of tuning haem and apoprotein biosynthesis to achieve high-level expression of haemproteins in bacteria. The recombinant protein is identical to the protein purified from the mollusc buccal muscle. Native *A. limacina* Mb has an oxygen dissociation rate constant of 70 s^{-1} [as compared with the value of 15 s^{-1} for sperm whale Mb, which displays His(E7) and Thr(E10)] (amino acid positions are referred to within the eight helices A–H of the globin fold). In order to understand the mechanism of oxygen stabilization in *A. limacina* Mb, we have prepared and investigated three active-site mutants: two single mutants in which Val(E7) and Arg(E10) have been replaced by His and Thr, respectively, and a double mutant carrrying both mutations. When Arg(E10) is substituted with Thr, the oxygen dissociation rate constant is increased from 70 s^{-1} to more than 700 s^{-1}, in complete agreement with the previously proposed role of the former residue in ligand stabilization. In the His(E7)-containing single and double mutants, both displaying high oxygen dissociation rates, the stabilization of bound oxygen by the distal His is insufficient to slow down the ligand dissociation rate constant to the value of sperm whale Mb. These results essentially prove the hypothesis that in *A. limacina* Mb a mechanism of oxygen stabilization involving Arg(E10), and thus different from that mediated by His(E7), has evolved.

INTRODUCTION

Myoglobin (Mb) has been employed in past and present days as a model system to clarify relevant biochemical processes such as ligand binding and protein folding. Among others, the Mb from the buccal muscle of the gastropod mollusc *Aplysia limacina* has been widely studied. The protein contains 146 amino acids [1,2] and has a typical globin fold, in spite of a low (20 %) primary sequence homology with the 'prototype' Mb from sperm whale. It also unfolds in a fully reversible manner [3], raising the possibility of studying in detail reversible folding in a conjugated protein similar to but different from sperm whale Mb.

Among other features, the most interesting one is the presence of only one histidyl residue, located at the proximal haem iron co-ordination position; a valyl residue, at topological position E7 (amino acid positions are referred to within the eight helices A–H of the globin fold), replaces the evolutionarily conserved distal histidine, which, in most vertebrate Mbs, has been shown to play a major role in the control of ligand binding [4,5]. Stabilization and modulation of haem reactivity is strictly related to the hydrogen-bonding capability and the polarity of the distal pocket. As far as oxygen is concerned, the formation of a hydrogen bond in the oxygenated derivative is critical to its intrinsic stability [6] and for the prevention of autoxidation [7]; in deoxyMb the polarity of the pocket is involved in the stabilization of a water molecule, which is displaced upon oxygen binding. Hydrogen bonds, steric hindrance, water stabilization and electrostatic interactions are also involved to a different extent in the reactivity of ferric Mb (metMb) with anionic ligands [8].

Crystallographic and ^{1}H-NMR studies on native *A. limacina* Mb ferric derivatives [9,10], and extensive equilibrium, kinetic and NMR investigations on sperm whale Mb site-directed mutants [11,12], have suggested that an arginyl residue at topological position E10 might be crucial for an alternative mechanism of ligand stabilization in this Mb. However, the fine control of the protein moiety over binding of oxygen, the physiologically relevant ligand, could not be fully clarified in these studies [13].

In order to study the active-site reactivity and, as a subsequent step, the folding of this Mb, we have cloned *A. limacina* Mb cDNA and we have achieved its expression in *Escherichia coli*. The proposed role of Arg(E10) has been investigated by site-directed mutagenesis, introducing at this position a Thr residue. Moreover, the effect of a distal His at the topological position E7, shown to be functionally crucial in vertebrate Mbs, has been tested with two mutants where Val(E7) has been replaced by His, both in the presence and in the absence of Arg(E10). The functional properties of these three key active-site mutants, in particular with regard to oxygen reactivity, have been studied by spectroscopy and kinetics.

EXPERIMENTAL

cDNA isolation and cloning

A. limacina molluscs were obtained from the Stazione Zoologica in Naples (Italy); the radular muscles were extracted from animals and immediately chilled at −70 °C. Total RNA was

Abbreviations used: Mb, myoglobin; metMb, ferric myoglobin; ALA, δ-aminolevulinic acid.

* Present address: Department of Biophysical Chemistry, Biozentrum der Universitat Basel, CH-4056 Basel, Switzerland.

† To whom correspondence should be addressed.

Aplysia limacina cDNA sequence has been deposited in the EBI (European Bioinformatics Institute) databank with the accession number X79304.

extracted by the guanidinum isothiocyanate method [14]. Single-strand cDNA synthesis and the subsequent PCR was performed with a GeneAmp RNA PCR Kit (Perkin Elmer Cetus). After the reverse transcription reaction [with 1 μg of total RNA and 300 pmol of dT *n*-mer (where *n* is the number of Ts)], the first PCR amplification was performed using as 3′ primer the dT *n*-mer and as 5′ primer a degenerate oligonucleotide (Genset) whose sequence [5′-ATG(C/T)TITCIGCIGCIGA(A/G)-GCIGA-3′] was back translated from the N-terminal amino acid sequence (residues 2–8) of *A. limacina* Mb [1]. When a codon third base presented four wobbles we positioned an inosine there. A second round of PCR amplification was performed on 1/20 of the above amplification mixture using 400 pmol of the same N-terminal primer and of another degenerate primer (Genset) whose sequence [5′-TTA(T/C)TTICCIGCIGC(T/C)TTIA-(A/G)IGC-3′] was deduced from the C-terminal residues 140–146 on the basis of *A. californica* genes' codon usage. The first amplification was performed for 25 cycles as follows: 95 °C for 1 min, 45 °C for 1 min and 72 °C for 2 min. The second amplification was the same except that the annealing temperature was raised to 48 °C. Final concentrations of reagents were in both cases 0.2 mM dNTP, 1 mM $MgCl_2$ and 2.5 units of *Taq* polymerase (Promega) in a total volume of 100 μl. With this approach, a 440 bp product was amplified; the purified fragment was cloned in the pCRII vector using the TA Cloning Kit (Invitrogen) and transformed in *E. coli* INVαF′ cells. Five recombinant clones were isolated and sequenced by the dideoxy chain termination method with Sequenase 2.0 (U.S. Biochemicals). One construct was then subjected to an additional PCR with 200 pmol of the same C-terminal primer as above and a new N-terminal primer (5′-TAACTAACTAAAGGAGAACAACAACAATGT-CGCTGTCGGCTGCTGAGGCGGATCTCG- 3′), in the presence of 1.3 mM $MgCl_2$ (annealing temperature of 44 °C). The PCR amplification product was subcloned in the pCRII vector and transformed in *E. coli* INVαF′ cells. The complete sequence of a construct, named pApMb, correctly positioned for expression, was verified by dideoxy-DNA sequencing with Sequenase 2.0 reagents (U.S. Biochemicals).

Mb expression and purification

Mb expression from pApMb was analysed both in the original *E. coli* INVαF′ and in TB-1 strains. The transformed cells were grown in LB broth containing 100 μg/ml ampicillin and 10 mM δ-aminolevulinic acid (ALA) (Sigma). The overexpression of haem-containing proteins was analysed in whole cells by reduced/CO differential absorption spectra as described elsewhere [15], and in total cell extracts by SDS/PAGE [16] and Western blot with anti-(*A. limacina* Mb) antibodies raised in rabbit. To assess the optimal ALA concentration for protein expression, pApMb-containing TB-1 cells were grown as described above, in the presence of 0.1, 0.5, 1, 5 or 10 mM ALA. As a preliminary step to site-directed mutagenesis, protein expression was also tested after subcloning of the Mb gene in pUC19 (pUC-ApMb), with results identical to those obtained with pApMb.

E. coli fermentations were carried out for 18 h in 50 l of LB broth containing 100 μg/ml ampicillin and 0.1 mM ALA. For protein purification, pUC-ApMb- or pApMb-containing TB-1 cells were harvested by centrifugation, resuspended in 300% (v/w) of lysis buffer [50 mM Tris/HCl (pH 8.0)/1 mM EDTA/0.5 mM dithiothreitol/1 mM PMSF/40 units/ml deoxyribonuclease/3 units/ml ribonuclease A] and left stirring overnight at 4 °C after the addition of lysozyme (2 mg/ml). The lysed material was then extensively sonicated. The protein was initially purified from the red-cell extract by ammonium sulphate precipitations (55 and 95%). Two ion-exchange chromatography steps were then performed: the first on DEAE-Sepharose Fast-Flow (Pharmacia) in 15 mM sodium phosphate, pH 6.5, to which the Mb does not bind, and the second on CM-Sepharose Fast Flow in 20 mM sodium acetate, pH 4.5, with a pH step elution at pH 5.6. Mb purity was determined by SDS/PAGE as described previously [16].

Mb characterization

Spectroscopic studies were carried out on a computer-controlled (OLIS) double-beam Cary 14 (Varian) or Jasco V550 UV/Vis spectrophotometer. Ferric and ferrous Mb derivatives were obtained as described elsewhere [17]. Absorption spectra were recorded at 20 °C in the following buffers: 0.1 M sodium phosphate, pH 6.2; 0.1 M sodium phosphate, pH 7.0; and 0.1 M 3-(cyclohexylamino)-1-propanesulphonic acid, pH 10.8. Absorption spectra for the 'acid–alkaline' transition were measured as a function of pH at 20 °C in 0.1 M sodium sulphate by addition of concentrated NaOH. Autoxidation rate constants were measured at 37 °C in 0.1 M sodium phosphate buffer, pH 7.0, with 1 mM EDTA; oxygen equilibria were measured at 20 °C in 0.1 M sodium phosphate, pH 7.0, by the tonometric method [17]. Oxygen dissociation rate constants were measured at 20 °C in 0.1 M sodium phosphate, pH 7.0, on an Applied Photophysics DX.17MV stopped-flow apparatus. To overcome problems related to the instability of the oxygen derivative, different experimental designs were applied: (1) direct measure of O_2 dissociation by mixing MbO_2 versus a sodium dithionite ($Na_2S_2O_4$) solution in degassed buffer; (2) a pulse experiment in which deoxyMb in the presence of sodium dithionite was mixed against buffers containing different O_2 concentrations; and (3) a pulse–double mixing experiment in which deoxyMb was first mixed with O_2-containing buffer, but in which observation was started only after a second mixing step against sodium dithionite solutions in degassed buffer.

Mutagenesis

Site-directed mutants were produced starting from the *A. limacina* Mb cDNA subcloned in pUC19 using the U.S.E. Mutagenesis Kit (Pharmacia). Protein expression, purification and characterization were performed with the same methods outlined above for the wild-type protein.

RESULTS AND DISCUSSION

cDNA cloning and expression

The cDNA coding for *A. limacina* Mb was isolated from mollusc radular muscle total RNA. The presence of 28 S rRNA was not evident in gels, recalling previous findings in other molluscs and invertebrates [18]; this observation suggested the presence of a 'hidden break' in the larger ribosomal RNA leading, during sample preparation, to smaller species co-migrating with the 18 S rRNA.

The nucleotide sequence of the cDNA is shown in Figure 1 together with the deduced sequence of 146 amino acids; in the same Figure the upstream regulatory sequence added to achieve bacterial expression is also reported. The protein sequence shown here is different from that of 145 amino acids originally reported by Tentori et al. [1] and from that of 146 amino acids subsequently corrected by Bolognesi et al. [2]; altogether five positions were found to be different from the latter. Three substitutions (Asn → Asp in positions 22 and 80 and Asp → Asn in position 77) are identical to those found in the sequence of *A. juliana* and *A.*

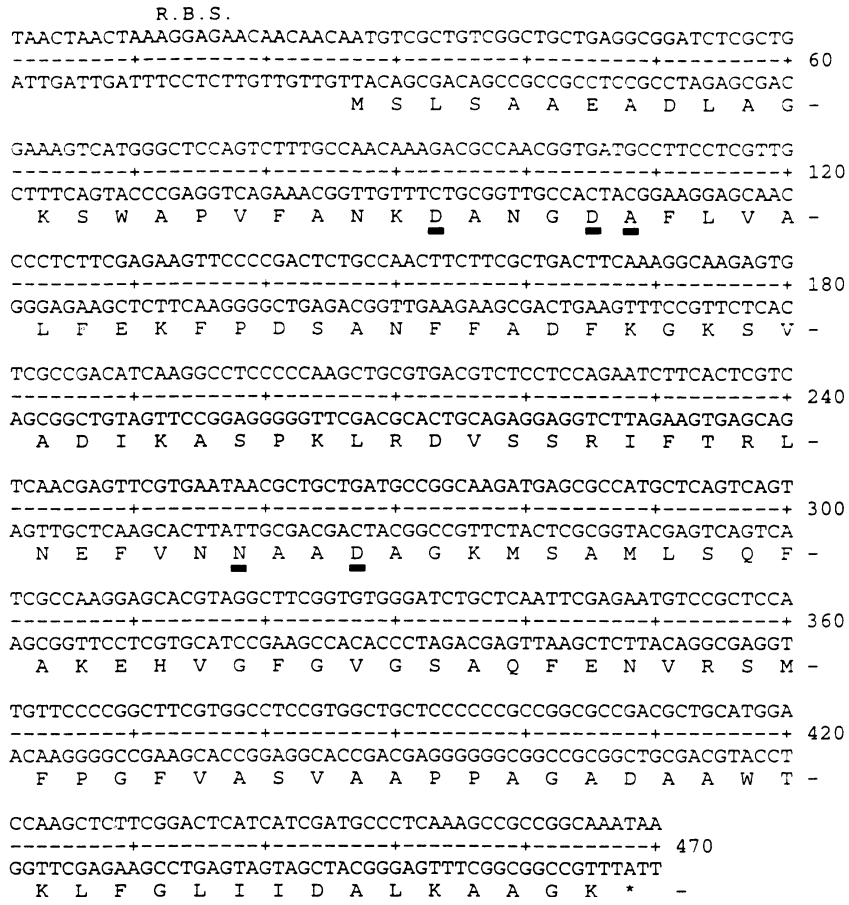

Figure 1 Nucleotide sequence of the cDNA coding for *A. limacina* Mb

Nucleotides 1–32 represent the sequence added to obtain protein expression in *E. coli;* this sequence contains (i) three stop codons in the three reading frames; (ii) an efficient ribosome binding site (R.B.S.) with optimal spacing to an initiator Met codon [15]; and (iii) the first N-terminal amino acid, Ser, which was omitted in the original N-terminal primer because of redundancy problems. Underlined amino acid residues correspond to the positions found to be different between the cDNA and the previously published amino acid sequence (see text).

kurodai Mbs [19]. Positions 26 and 27, which we assign as Asp and Ala on the basis of the cDNA sequence, are inverted with respect to the sequence published by Tentori et al. [1]; all the clones sequenced show these substitutions. Indeed, redetermination of the amino acid sequence of the purified tissue protein (results not shown) and a careful inspection of the native *A. limacina* Mb electron-density map in the 26–27 region (M. Bolognesi, personal communication), fully confirm the cDNA-deduced amino acid sequence. The sequence Asp–Ala at positions 26 and 27 is peculiar to *A. limacina* Mb, since Ala is present at position 26 in *A. juliana* and *A. kurodai* Mbs, and Asn-27 is found in *A. kurodai* Mb. Moreover, analysis of the frequencies of these residues in 226 different globin chains [20] has shown that position 26 is usually a small non-charged amino acid, whereas position 27 is preferentially occupied by a negatively charged residue (Glu more frequently than Asp). Interestingly, residues 26 and 27 are near the conserved Gly-25, which is the point of closest contact between helices B and E, and may possibly have some implications for the stability of this Mb.

In order to obtain bacterial expression of *A. limacina* Mb, a regulatory sequence has been introduced by PCR at the 5′ end of the cDNA (see Figure 1). The expression construct was placed downstream of a *lac* promoter in the vector pCRII, yielding the recombinant plasmid pApMb; high-level expression in two *E. coli* strains (INVαF′ and TB-1) transformed with pApMb was analysed in the presence of the protoporphyrin IX precursor, ALA (G. Colotti and W. Royer, personal communication). As shown in Figure 2(a), the expression of the holoprotein, monitored in whole cells by reduced/CO differential absorption spectra, is increased in the presence of 10 mM ALA. Since the untransformed *E. coli* strain TB-1 also displays a small increase in the spectroscopic signal upon ALA addition (results not shown), probably due to an aspecific induction of haem-containing proteins, the effect of ALA on *A. limacina* globin production was also monitored by Western blotting of cell extracts with antibodies directed against *A. limacina* Mb. As shown in Figures 2(b) and 2(c), the result obtained with the latter method, which specifically detects accumulation of the apoprotein, parallels that obtained from spectroscopy of whole cells.

An increase in Mb expression similar to that obtained with 10 mM ALA was observed with lower ALA concentrations (down to 0.1 mM), whereas higher concentrations led to the production of a major red contaminant, which may represent uncomplexed protoporphirin IX [21].

This result clearly suggests that an increased rate of haem biosynthesis leads to a stabilization of the apoprotein in the bacterial cell, thus allowing a higher yield in holoprotein. This evidence, in accordance with recent findings [21,22], underlines

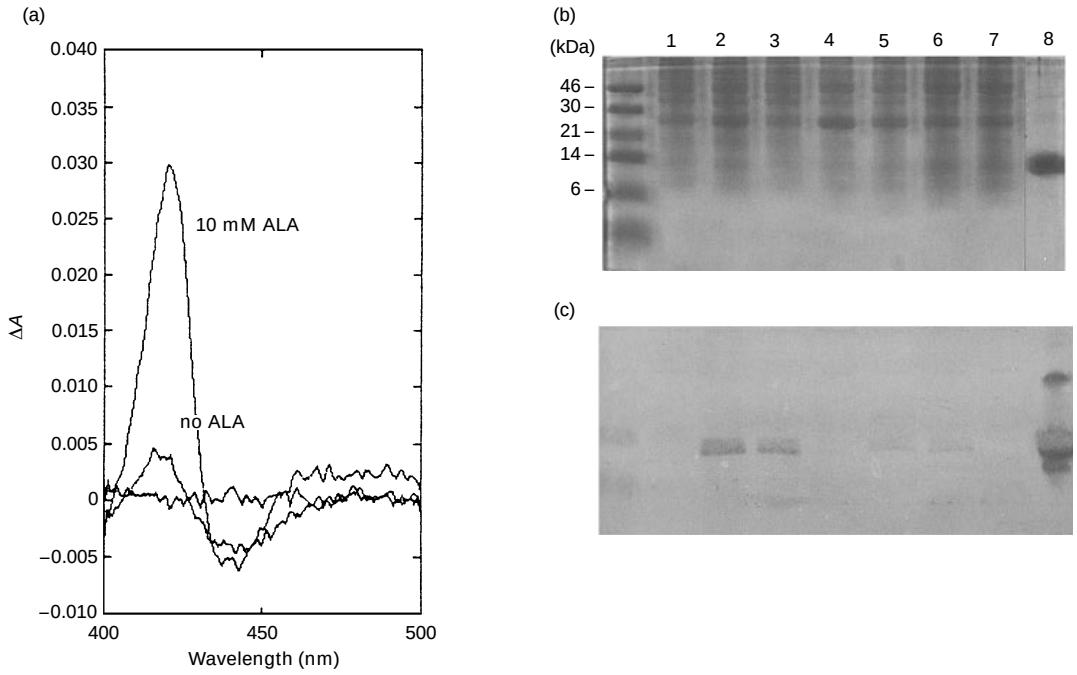

Figure 2 Analysis of *A. limacina* Mb expression in *E. coli*

Protein expression was analysed in *E. coli* strains INVαF′ and TB-1 in the presence and absence of 10 mM ALA. (**a**) Result of the CO/reduced-difference spectra of whole cells for sample pApMb in TB-1 in the absence and in the presence of 10 mM ALA; the same result was obtained for the sample pApMb in INVαF′ (results not shown). (**b** and **c**) SDS/PAGE of total protein extracts from (1) TB-1 + ALA; (2) pApMb in TB-1 + ALA; (3) pApMb in INVαF′ + ALA; (4) plasmid pApMb-rev (containing the cDNA in the opposite orientation) in TB-1 + ALA; (5) pApMb in INVαF′ — ALA; (6) pApMb in TB-1 — ALA; (7) TB-1 — ALA; and (8) purified *A. limacina* Mb from tissue. Proteins were stained with Coomassie Brilliant Blue (**b**) and analysed by Western blot with anti-(*A. limacina* Mb) antibodies (**c**). Molecular masses of markers (Amersham) are in kDa.

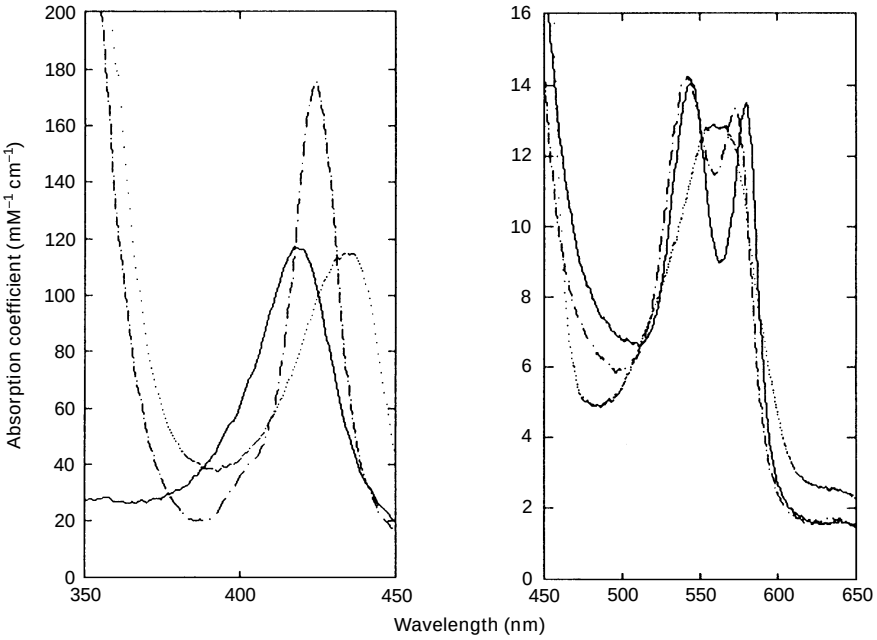

Figure 3 Absorption spectra of ferrous derivatives of wild-type recombinant *A. limacina* Mb

Oxy- (continuous line), deoxy- (dotted line) and CO (dashed line)-derivatives are shown. All spectra were recorded at 20 °C in 0.1 M sodium phosphate, pH 7.0.

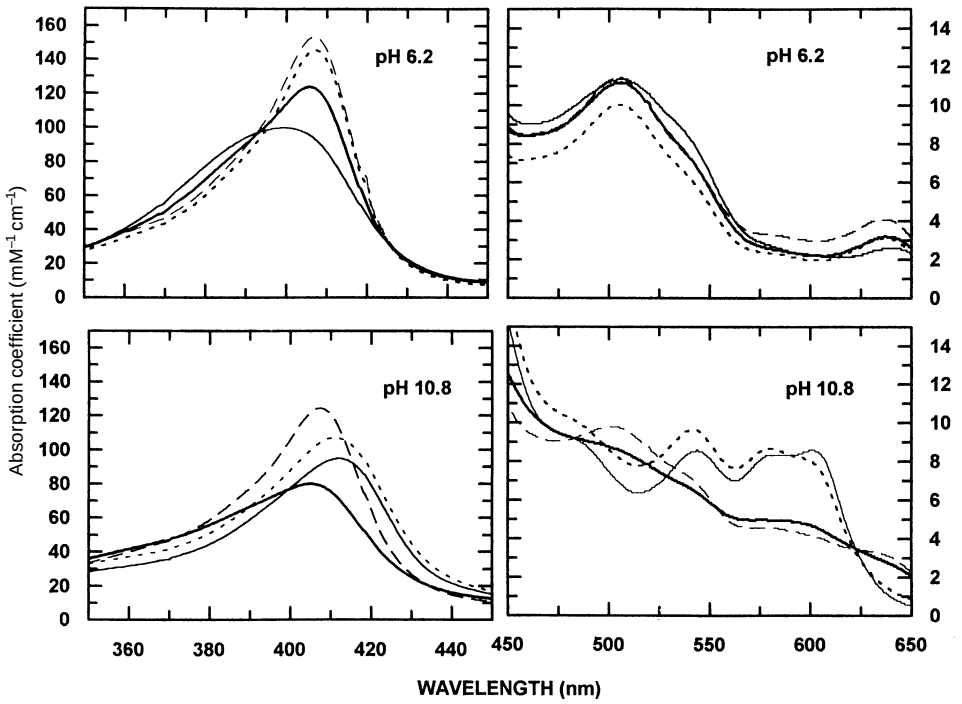

Figure 4 Absorption spectra of the acid and alkaline ferric derivatives of wild-type and mutant *A. limacina* Mb

Spectra were recorded in 0.1 M sodium phosphate, pH 6.2, and 0.1 M 3-(cyclohexylamino)-1-propanesulphonic acid, pH 10.8. Spectra of wild-type Mb (continuous thin line) as well as of R(E10)T (continuous bold line), V(E7)H (dotted line) and V(E7)H/R(E10)T (dashed line) mutants are shown.

the importance of a balanced haem and globin synthesis for the optimal expression of recombinant haemproteins in *E. coli*. To our knowledge, this is the first successful attempt of direct expression of a Mb cDNA in *E. coli*: in other cases, fusion proteins (human and porcine Mbs) [23,24] or synthetic genes with optimized codon usage (sperm whale and horse Mbs) [15,25] have had to be employed.

Wild-type Mb characterization

The recombinant *A. limacina* Mb is indistinguishable from the native protein, as shown by biochemical, spectroscopic and functional data. A novel purification procedure has been developed for the recombinant *A. limacina* Mb, as outlined in the Experimental section. Interestingly, the recombinant protein is correctly matured in *E. coli*: amino acid sequence analysis of the N-terminus of the recombinant Mb indicates that the first residue is a Ser and that the N-terminal Met is $> 90\,\%$ processed, in agreement with the work of Hirel et al. [26] in which the catalytic efficiency of methionyl aminopeptidase has been correlated with the side-chain length of the penultimate amino acid in the substrate sequence. It's to be noted that whereas the N-terminal Ser is acetylated in the Mb extracted from the *A. limacina* buccal muscle [1,2], this is not the case for the recombinant protein expressed in *E. coli* (where post-translational modifications cannot take place), as is evident from the availability of the N-terminus for amino acid sequence analysis by Edman degradation.

The optical spectra of the recombinant *A. limacina* Mb are in agreement with the spectra reported for the native protein [27]; both ferrous (Figure 3) and ferric (Figure 4, continous thin line) derivatives have been determined. The unusual peak ratio observed for the oxygen-bound form in the visible region ($\epsilon_{542} > \epsilon_{582}$ nm) is a distinctive feature of this invertebrate Mb, since this derivative displays inverted peak heights ($\epsilon_{542} < \epsilon_{582}$ nm) in other Mbs (e.g. those from sperm whale or horse).

To analyse further the properties of the recombinant *A. limacina* Mb, we have measured its oxygen affinity in comparison with the native protein: the values, shown in Table 1 also in comparison with those for sperm whale Mb, are almost identical to those previously determined for the native protein [17,27]. The haem iron autoxidation rate, measured at pH 7.0 and 37 °C, showed similar values for both the recombinant and the native Mb (1.9 and 2.2 h^{-1}, Table 1). This result confirms the observation that in *A. limacina* Mb, which lacks the distal His, the autoxidation rate is 40 times higher than that of sperm whale Mb, although both proteins have similar oxygen affinity (Table 1). The analysis of the autoxidation mechanism in mammalian Mbs, carried out by Brantley et al. [7], clearly showed that the highest autoxidation rate is expected for Mbs in which bound O_2 is exposed to solvent, but not hydrogen bonded to the side chain of the residue in position E7: in *A. limacina* Mb the autoxidation rate might therefore reflect a higher overall solvent exposure of the distal pocket, partially compensated by favourable interaction with a specific distal residue side chain such as that of Arg(E10). Although no structural data is available for the ferrous derivatives of this Mb, at least three water molecules have been located by crystallography in the pocket of different derivatives of the ferric form [9], suggesting that in the oxygenated derivative the distal pocket may be also open to solvent molecules.

Production and characterization of Mb mutants

In order to elucidate the molecular bases of oxygen stabilization in *A. limacina* Mb, and starting from previous studies [9–12] that

Table 1 Functional parameters for oxygen and hydroxyl binding to different haemproteins

Values of k_{on} were calculated from equilibrium and dissociation rate constants. Values of $k_{autox.}$ were obtained in 0.1 M sodium phosphate, pH 7.0, at 37 °C in air.

	O_2				OH^-
	k_{on}	k_{off}	K	$k_{autox.}$	
	($10^{-7} \times M^{-1}s^{-1}$)	(s^{-1})	($10^{-6} \times M^{-1}$)	(h^{-1})	pK
Sperm whale native*	1.7	15	1.1	0.05	8.9
A. limacina native	7	70 ± 5	1.0 ± 0.1	2.2	7.5†
A. limacina wild type	10	80 ± 5	1.3 ± 0.1	1.9	7.4 ± 0.1
R(E10)T	–	> 700	–	20 ± 3	10.2 ± 0.1
V(E7)H	–	260 ± 40	–	5.9 ± 1	9.4 ± 0.1
V(E7)H/R(E10)T	–	200 ± 40	–	26.2 4.3	10.9 ± 0.1

* From [4].
† From [30].

pointed to a possible role for an arginyl residue in topological position E10, three site-directed mutants were prepared: Arg(E10) → Thr [R(E10)T], Val(E7) → His [V(E7)H] and Val(E7) → His/Arg(E10) → Thr [V(E7)H/R(E10)T].

Oxygen reactivity of these mutants was assessed by determining the oxygen dissociation and autoxidation rates (Table 1). The removal of the Arg side chain in the single mutant R(E10)T dramatically increased the oxygen dissociation rate constant, indicating that favourable hydrogen-bonding capability associated with the long and flexible guanidinium side chain of ArgE(10) is lost; the autoxidation rate constant in this mutant was also increased 10-fold. The introduction of a His residue at position E7 perturbs oxygen reactivity with respect to wild type, since, both in the presence [V(E7)H] and absence [V(E7)H/R(E10)T] of Arg at position E10, fast oxygen dissociation rate constants were measured. Nonetheless, the observed dissociation rate constant for the double mutant ($k = 200\ s^{-1}$) was lower than that of the single mutant R(E10)T ($k > 700\ s^{-1}$), supporting the idea that a limited stabilization may be exerted by the imidazole side chain in the pocket. Moreover, analysis of the autoxidation rate for the double mutant showed a biphasic behaviour with two rate constants (Table 1) whose values were similar to those measured for the individual mutations, suggesting that His(E7) may adopt multiple conformations in the absence of the Arg(E10) side chain. Ligand-dependent movements of His(E7) have been demonstrated clearly for other haemproteins, such as erythrocruorin from the insect *Chironomus thummi thummi* and, to a lesser extent, sperm whale Mb [28,29].

The stability of the oxygenated derivative is measured by the value of the autoxidation rate, and therefore only in the case of wild-type and V(E7)H oxygen dissociation can the kinetics be measured starting from the MbO_2 complex. The time course of oxygen dissociation determined in the presence of dithionite is often biphasic in the mutants, and the complex time course was found to depend on experimental conditions. For the wild-type Mb a single phase is observed ($80\ s^{-1}$) under all experimental conditions, which can unequivocally be assigned to oxygen dissociation on the basis of the comparison between the kinetic and static difference spectra (Figure 5a, filled circles and line). For V(E7)H, analysis of the data obtained when oxygen dissociation kinetics were measured starting from MbO_2 showed a fast phase ($260\ s^{-1}$), which can be identified as oxygen dissociation on the basis of the comparison between the kinetic and static difference spectra (Figure 5a, open circles), and a slower one ($25–30\ s^{-1}$), which can be interpreted (on the same basis) as the reduction of the ferric derivative by excess sodium dithionite (see Figure 5b, open circles). The double mixing experimental design (see the Experimental section) allows the following reactions to be performed in a sequential mode:

$$\overset{\text{I}}{Mb^{2+} + O_2 \leftrightarrow MbO_2} + Na_2S_2O_4 \overset{\text{II}}{\rightarrow} Mb^{2+} + \text{products}$$

Scheme 1

Observation starts only after the second mixing (II), and the interval between the first and second mixing (delay time) can be varied from a few milliseconds to minutes. According to Scheme 1, the fraction of MbO_2 achieved after the first mixing with oxygen will depend on oxygen concentration for haemproteins with low oxygen affinity (i.e. all the mutants). This being the case, the recorded signal (i.e. the observed absorbance change at 435 nm, which monitors the oxy- to deoxy-Mb transition) should increase with increasing oxygen concentration up to saturation. The results obtained for V(E7)H yield a biphasic time course: the fast process, assigned to oxygen dissociation from MbO_2 (Figure 5a, squares), displays the expected amplitude dependence upon oxygen concentration, and, at high oxygen concentration (i.e. 650 μM after mixing), this is the only process observed and corresponds to maximal amplitude. The slowest process observed ($3–7\ s^{-1}$), which is much smaller in amplitude (by 4-fold), is harder to interpret, since its kinetic difference spectrum (Figure 5b, squares) does not correspond to statically identified derivatives; nonetheless its amplitude decreases at the highest oxygen concentrations and at long delay times. In the case of R(E10)T and V(E7)H/R(E10)T, oxygen dissociation can be measured only with the double mixing experimental design, due to high autoxidation rates: a behaviour similar to that of V(E7)H was observed with a slower phase (which is again smaller in amplitude) that tends to vanish at the highest oxygen concentration. Our analysis indicates that for all three mutants the faster process (Figure 5a and Table 1) indeed corresponds to the oxygen dissociation from MbO_2, while the slowest process we believe indicates, among other hypotheses, the reaction of a side product of dithionite occurring in that fraction of molecules which (due to low oxygen affinity) is not saturated with oxygen at lower concentrations.

The oxygen dissociation rate obtained for the R(E10)T mutant clearly demonstrates that in *A. limacina* Mb oxygen stabilization is mediated by Arg(E10). In the case of V(E7)H, steric hindrance

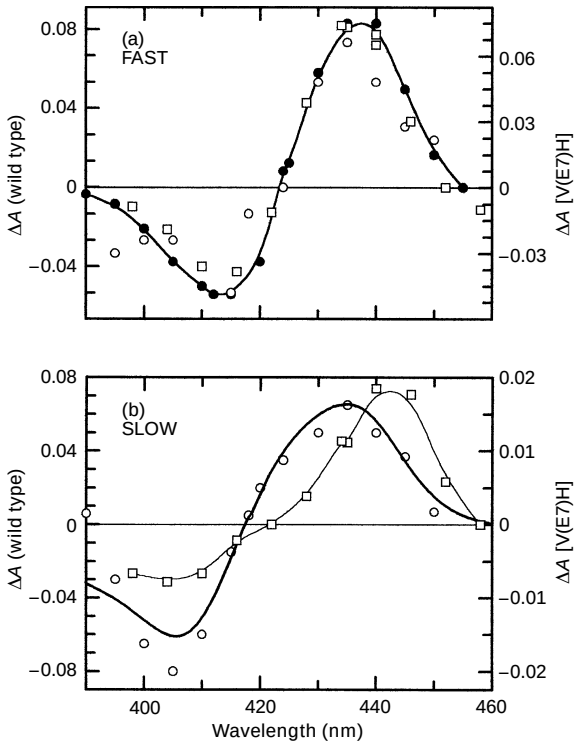

Figure 5 Kinetic difference spectra of wild-type and V(E7)H *A. limacina* Mbs

(**a**) Normalized absorbance changes of the fast process in the oxygen dissociation kinetics for wild type (●), for V(E7)H starting from MbO_2 (○) and for V(E7)H in the double mixing pulse experiment (□). The continous line represents the static difference spectrum of oxy-deoxy-Mb. (**b**) Normalized absorbance changes of the slow process observed in the oxygen dissociation experiments for V(E7)H starting with MbO_2 (○) or in the double mixing pulse experiment (□). The thick line represents the static difference spectrum of deoxy-ferric Mb; the thin line is a polynomial interpolation of the (□) data.

of the imidazole side chain in the distal pocket of *A. limacina* Mb may be responsible for the very small effect obtained with His at position E7. Previous data on the crystal structure of the imidazole derivative of ferric *A. limacina* Mb [9] have shown that this bulky ligand induces a structural perturbation of the distal pocket, which is particularly evident for the CD region where deviations up to 2.6 Å were observed. This is in agreement with our results on the modelling of the Val → His mutation in the crystal structure of *A. limacina* Mb, where sterically unfavourable interactions with Phe(CD1) were clearly observed (F. Cutruzzolà, C. Travaglini Allocatelli and M. Brunori unpublished work). It is interesting to notice that in the structure of the imidazole derivative [9] the electron density for Arg(E10) suggests a poorly ordered location of its side chain, in accordance with the observation that burial of this functionally relevant side chain depends upon ligand size; these structural data explain why Arg(E10) cannot exert its stabilizing role in the presence of His in topological position E7 [as in the mutant V(E7)H].

The optical spectra of the mutant proteins were found to be identical to the wild-type spectra, with the exception of the ferric derivative (metMb): the absorption maximum in the Soret region measured at pH 6.2 (Figure 4) is at 406 nm for the single mutant R(E10)T and at 407 nm for the V(E7)H and V(E7)H/R(E10)T mutants, with a red shift of 6 or 7 nm with respect to the wild-type protein ($\lambda_{max.} = 400$ nm at the same pH). Native *A. limacina* metMb is five-co-ordinate at pH values below the acid–alkaline transition point [30,31], at variance with sperm whale Mb and other Mbs with either His or Gln at position E7; in the latter Mbs a water molecule is co-ordinated to the haem iron, leading to an hexa-co-ordinate high-spin complex. Water co-ordination is associated with absorption maxima at 407–410 nm in the Soret, while in the absence of water (five-co-ordinate complex) the maximum is blue shifted to 395–400 nm [30]. Therefore we conclude that these *A. limacina* mutants have a co-ordination similar to that of ferric sperm whale Mb. The shorter and polar Thr side chain at position E10 in R(E10)T in place of the long and flexible Arg may increase solvent accessibility to the distal pocket, thus leading to water co-ordination. Similarly, the distal His introduced in V(E7)H and V(E7)H/R(E10)T may increase pocket polarity and favour water co-ordination. To further characterize the properties of the ferric derivative of the wild-type recombinant Mb and of the site-directed mutants, we have also investigated the acid–alkaline transition, i.e. the pH dependence of the optical spectrum, which relates to OH^- binding to the Fe^{3+} haem iron. As shown in Table 1, the observed pK values are shifted upward by two (single mutants) or three (double mutant) pH units with respect to that of wild-type Mb.

In metMb, affinity for OH^- is generally related both to the presence of a co-ordinated water molecule at neutral pH and to the stabilization of the $Mb(Fe^{3+}):OH^-$ complex by hydrogen bonding from distal pocket residues [His(E7) in mammalian Mbs]. In the five-co-ordinate *A. limacina* Mb the pK of the acid–alkaline transition is 7.5 [30] (7.4 in this work), arguing for a high affinity for OH^- of this Mb, whereas in the hexa-co-ordinated sperm whale Mb a value of 8.9 is observed, consistent with the fact that H-bonded water displacement must occur in the latter Mb to yield the OH^- complex. This mechanistic interpretation has been confirmed previously with experiments on selected sperm whale site-directed mutants [11]; as an example, the sperm whale single mutant His(E7) → Val (in which the hydrogen-bonding capability has been removed by replacement of the distal histidine) is a five-co-ordinate, low-affinity Mb (pK $\gg 9.5$) [11]. Along the same lines, introduction of an Arg at position E10 in the His(E7) → Val sperm whale Mb mutant (which restores hydrogen bonding to OH^-, although it is not capable of stabilizing a water molecule) yielded an Mb [His(E7) → Val/Thr(E10) → Arg] that is five-co-ordinate but displays a higher affinity for OH^- (p$K = 8.1$) [11].

As discussed above, the spectrum of ferric *A. limacina* Mb mutants suggests water co-ordination: the observed higher values for the acid–alkaline pK values might therefore be explained by suggesting that (i) water displacement has to occur in these mutants before binding of OH^- (as in the other hexa-co-ordinated Mbs); and (ii) no stabilizing contribution can be offered by distal side residues that are either absent [R(E10)T] or sterically hindered [V(E7)H and V(E7)H/R(E10)T]. Interestingly, the presence of the positively charged arginine in position E10 has some relevance to OH^- binding, since in the single mutant V(E7)H a slightly lower pK than the double mutant V(E7)H/R(E10)T is observed: a similar effect has been noticed for some site-directed mutants of sperm whale Mb where the presence of a positive charge at the entrance of the distal pocket has been shown to facilitate binding of negatively charged ligands such as azide (N_3^-) [8].

Conclusions

The availability of a heterologous expression system to produce wild-type and specifically mutated *A. limacina* Mb is relevant to the understanding of monomeric haemproteins reactivity in a protein framework different to the mammalian one, which in the

last years has been widely dissected [4]. Indeed, the oxygen dissociation rate obtained for the R(E10)T mutant clearly demonstrates that in *A. limacina* Mb oxygen stabilization is mediated by Arg(E10), thus providing the first direct evidence that this mechanism (proposed mainly on the basis of crystallographic data, see [9]) has been selected by evolution to modulate reactivity of the physiological ligand in this invertebrate Mb lacking the distal histidine. The small effect of introducing an His at position E7 and the failure to reproduce the slow oxygen dissociation typical of sperm whale Mb may tentatively be understood in terms of steric hindrance of the imidazole side chain in the distal pocket of *A. limacina* Mb.

The present study agrees with and extends the work of Smerdon et al. [13], in which mimicry of *Aplysia* Mb reactivity starting from a mammalian framework has been attempted. The low sequence homology between the molluscan and mammalian Mbs, which mirrors different evolutionary pathways, is responsible for subtle differences in side-chain stereochemistry of key residues, yielding proteins for which the molecular basis of reactivity is different. Transplantation of specific functions into a structurally related molecule, whenever possible, requires a more extensive reorganization that, also in view of the extremely low free-energy changes involved, may be achieved only with a finite, but greater than two, number of substitutions [32].

Finally the heterologous high-level expression of *A. limacina* Mb offers the interesting possibility of investigating the protein-folding problem in an Mb with the typical globin fold, in spite of the remarkably different primary structure from the prototype sperm whale Mb, to assess the contribution of the amino acid sequence to the stability of the folded state and to the choice of folding pathways in monomeric proteins.

We wish to thank Dr. A. Gambacurta and Professor F. Ascoli for useful suggestions in cDNA cloning; Dr. E. Schininà and Professor D. Barra for protein sequencing; and Professor P. Ascenzi and Dr. R. A. Staniforth for critical reading of the manuscript. Special thanks to Mr. B. Volpi and Professor M. Tomasi (Istituto Superiore di Sanità, Roma, Italy) for their competent collaboration and help in the optimization of fermentations of *E. coli* strains. Financial support from MURST (40% Liveproteins) is gratefully acknowledged. C.T.A. is the recipient of a fellowship from the Fondazione A. Buzzati Traverso.

REFERENCES

1 Tentori, L., Vivaldi, G., Carta, S., Marinucci, M., Massa, A., Antonini, E. and Brunori, M. (1973) Int. J. Pep. Prot. Res. **5**, 187–200
2 Bolognesi, M., Onesti, S., Gatti, G., Coda, A., Ascenzi, P. and Brunori, M. (1989) J. Mol. Biol. **205**, 529–544
3 Brunori, M., Giacometti, G. M., Antonini, E. and Wyman, J. (1972) J. Mol. Biol. **63**, 139–152
4 Springer, B. A., Sligar, S. G., Olson, J. S. and Phillips, G. N., Jr. (1994) Chem. Rev. **94**, 699–714
5 Perutz, M. F. (1989) Trends Biochem. Sci. **14**, 42–44
6 Phillips, S. E. V. and Shoenborn, B. P. (1981) Nature (London) **292**, 81–82
7 Brantley, R. E., Jr., Smerdon, S. J., Wilkinson, A. J., Singleton, E. W. and Olson, J. S. (1993) J. Biol. Chem. **268**, 6995–7010
8 Brancaccio, A., Cutruzzolà, F., Travaglini Allocatelli, C., Brunori, M., Smerdon, S. J., Wilkinson, A. J., Dou, Y., Keenan, D., Ikeda-Saito, M., Brantley, R. E., Jr. and Olson, J. S. (1994) J. Biol. Chem. **269**, 13843–13853
9 Conti, E., Moser, C., Rizzi, M., Mattevi, A., Lionetti, C., Coda, A., Ascenzi, P., Brunori, M. and Bolognesi, M. (1993) J. Mol. Biol. **233**, 498–508
10 Qin, J., La Mar, G. N., Ascoli, F., Bolognesi, M. and Brunori, M. (1992) J. Mol. Biol. **224**, 891–897
11 Travaglini Allocatelli, C., Cutruzzolà, F., Brancaccio, A., Brunori, M., Qin, J. and La Mar, G. N. (1993) Biochemistry **32**, 6041–6049
12 Qin, J., La Mar, G. N., Cutruzzolà, F., Travaglini Allocatelli, C., Brancaccio, A. and Brunori, M. (1993) Biophys. J. **65**, 2178–2190
13 Smerdon, S. J., Krzywda, S., Brzozowski, A. M., Davies, G. J., Wilkinson, A. J., Brancaccio, A., Cutruzzolà, F., Travaglini Allocatelli, C., Brunori, M., Li, T., Brantley, R. E., Jr., Carver, T. E., Eich, R. F., Singleton, E. and Olson, J. S. (1995) Biochemistry **34**, 8715–8725
14 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
15 Springer, B. A. and Sligar, S. G. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 8961–8965
16 Shagger, H. and von Jagow, G. (1987) Anal. Biochem. **166**, 368–379
17 Antonini, E. and Brunori, M. (1971) Hemoglobin and Myoglobin in their Reactions with Ligands, Elsevier, Amsterdam
18 Gambacurta, A., Piro, M. C. and Ascoli, F. (1993) FEBS Lett. **330**, 90–94
19 Takagi, T., Iida, S., Matsuoka, A. and Shikama, K. (1984) J. Mol. Biol. **180**, 1179–1184
20 Bashford, D., Chothia, C. and Lesk, A. M. (1987) J. Mol. Biol. **196**, 199–216
21 Nakahigashi, K., Nishimura, K., Miyamoto, K. and Inokuchi, H. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 10520–10524
22 Komar, A. A., Kommer, A., Krasheninnokov, I. A. and Spirin, A. S. (1993) FEBS Lett. **326**, 261–263
23 Varadarajan, R., Szabo, A. and Boxer, S. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5681–5684
24 Dodson, G., Hubbard, E., Oldfield, T. J., Smerdon, S. J. and Wilkinson, A. J. (1988) Protein Eng. **2**, 233–237
25 Guillemette, J. G., Matsushima-Hibiya, Y., Atkinson, T. and Smith, M. (1991) Protein Eng. **4**, 585–592
26 Hirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G. and Blanquet, S. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 8247–8251
27 Rossi Fanelli, A. and Antonini, E. (1957) Biokhimija **22**, 336–344
28 Steigemann, W. and Weber, E. (1979) J. Mol. Biol. **127**, 309–338
29 Johnson, K. A., Olson, J. S. and Phillips, G. N., Jr. (1989) J. Mol. Biol. **207**, 459–463
30 Giacometti, G. M., Ascenzi, P., Brunori, M., Rigatti, G., Giacometti, G. and Bolognesi, M. (1981) J. Mol. Biol. **151**, 315–319
31 Rousseau, D. L., Ching, Y., Brunori, M. and Giacometti, G. M. (1989) J. Biol. Chem. **264**, 7878–7881
32 Komiyama, N. H., Miyazaki, G., Tame, J. and Nagai, K. (1995) Nature (London) **373**, 244–246

Received 3 July 1995/8 September 1995; accepted 5 October 1995