Volume 330, number 1, 90–94 FEBS 12942 September 1993

Cooperative homodimeric hemoglobin from *Scapharca inaequivalvis*

cDNA cloning and expression of the fully functional protein in *E. coli***

Alessandra Gambacurta, Maria Cristina Piro, Franca Ascoli*

Department of Experimental Medicine and Biochemical Sciences, Tor Vergata University, via O Raimondo, 00173 Rome, Italy

Received 6 July 1993; revised version received 22 July 1993

The overexpression of the fully functional, cooperative homodimeric hemoglobin of the bivalve mollusc, *Scapharca inaequivalvis*, has been accomplished in *E. coli* from its cDNA. The latter was isolated by PCR amplification of total RNA and sequenced. The cDNA-derived sequence differed by a single amino acid when compared to that previously obtained from purified protein. Interest in this hemoglobin resides in the unique assemblage of the two identical subunits, with the heme groups facing each other in the inside of the molecule, opposite to that occurring in vertebrate hemoglobins. The results presented here are the basis for future studies of structure/function relationships by site directed mutagenesis.

Globin cDNA; Reverse transcriptase-PCR; Expression in *E. coli*; Recombinant hemoglobin, *Scapharca inaequivalvis*

1. INTRODUCTION

The dimeric hemoglobin (HbI) isolated from the nucleated erythrocytes of the bivalve mollusc, *Scapharca inaequivalvis*, is comprised of two identical chains, each carrying a heme group [1]; as found for other bivalves of the Arcid family [2–5], a tetrameric hemoglobin (HbII), made of two different subunits (A_2B_2 type) is also present in the same organism.

The structural and functional properties of *S. inaequivalvis* HbI have been studied in detail, for the unexpected ability of this homodimeric hemoglobin to bind heme ligands in a cooperative fashion [1]. Studies on the kinetics and equilibrium of oxygen and carbon monoxide binding to HbI [1,6], which behaves in this respect similarly to mammalian hemoglobins, led to the conclusion that there is more than one mechanism able to generate cooperativity in hemoglobins. Indeed, the X-ray crystal structure of HbI at low and high resolution, in the liganded and unliganded state [7–9], has revealed that the two identical subunits have the usual myoglobin folding, apart from an additional pre-A helix, consistent with the amino-terminal extension found in the primary structure of HbI globin [10]. However, the assembly of the subunits is opposite to that occurring in mammalian hemoglobins and involves contacts between the heme-carrying E and F helices of one subunit with the same region of the other subunit. This provides direct heme–heme communication, as the basis for cooperativity in ligand binding, which is accompanied by conformational changes at the subunit interface [9].

Due to such a peculiar subunit interaction in *S. inaequivalvis* HbI, we considered it of interest to explore the possibility of expressing this hemoglobin in vivo, for future site-directed mutagenesis studies at the subunit interface and at the ligand binding site. In this paper we report on the isolation and sequence analysis of HbI cDNA and on the expression in *E. coli* of the fully functional protein.

*Corresponding author Fax: (39) (6) 723-4793

**Sequence data described in this paper have been submitted to the EMBL/GenBank database under the accession number X71386.

2. MATERIALS AND METHODS

2.1. *Strains and vector*

E. coli strains INVαF′ and JM 105 were used in plasmid screening, in DNA sequencing and in the expression experiments. The vector PCRII containing the *lac* promoter and the β-galactosidase gene was used either for cloning the PCR product or for in vivo expression and was purchased from Invitrogen.

2.2. *RNA isolation*

Total RNA was isolated from the erythrocytes of *S. inaequivalvis* by the guanidine isothiocyanate method [11] or by the RNA ZolIB isolation kit from Promega. RNA quality was controlled by electrophoresis of 10 μg on 1% agarose gel containing 2.2 M formaldehyde [12].

2.3. *PCR primers*

A degenerate oligonucleotide, DIM (purchased from Genset) was used as 5′ primer in the PCR experiments. Its sequence was deduced from the N-terminal sequence (amino acid residues at positions 1–6) of *S. inaequivalvis* HbI globin [10]. When a codon third base showed four wobbles, we replaced it with inosine. As downstream primer, CODAT oligonucleotide, also obtained from Genset, was used. The sequences of the two primers are.

Published by Elsevier Science Publishers B.V.

DIM ATG CCI TCI GTI TA(CT) GA(CT) GC
CODAT GCG CTT TTT TTT TTT TTT TT

2.4. *cDNA synthesis and sequence*

1 μg of *S. inaequivalvis* total RNA was used in each experiment. Single-stranded cDNA synthesis and the subsequent PCR was performed with a GeneAmp RNA PCR Kit (Perkin Elmer Cetus). The amount of primers used to amplify cDNA was 550 pmol for DIM and 200 pmol for CODAT; the PCR program for amplification was as follows: denaturation at 95°C for 2 min (1 cycle), denaturation at 95°C for 1 min, annealing at 45°C for 1 min and extension at 72°C for 1 min (23 cycles), extension at 72°C for 2 min (1 cycle). The PCR products were subjected to electrophoresis on agarose gel 1.2%, a ~700 bp band was eluted [12], purified and subcloned in the PCRII vector using the TA Cloning Kit (Invitrogen). The positive clones were isolated and their sequences were determined by the dideoxy-chain termination method [13] with Sequenase 2.0 (US Biochemical).

2.5. *Construction and expression of HbI gene*

The full-length HbI cDNA sequence was removed from the PCRII plasmid as *Eco*RI–*Eco*RI fragment and, after purification [12], cut with *Sfa*NI (present as restriction site at nucleotide +16). After digestion, this *Sfa*NI–*Eco*RI fragment was purified from a 1.8%, TBE agarose gel with JET sorb gel extraction kit (Polymed).

The *Bam*HI–*Sfa*NI synthetic linkers (from Genset):

5′-GATCCATAACTAACTAAAGGAGAACAACAACAATGCCG-3′
GTATTGATTGATTTCCTCTTGTTGTTGTTACGGCAGCC

were ligated to the 5′ end of the *Sfa*NI–*Eco*RI HbI cDNA fragment; the construct was purified from a 2.2% agarose gel and subcloned in the *Bam*HI–*Eco*RI-digested PCRII to generate the expression vector pGAPI. The linker region was a modification of the sequence used by Unger et al. [14] and incorporated the Shine-Delgarno translational regulatory sequence. The complete sequence of this construct was verified by dideoxy-DNA sequencing [12].

2.6. S. inaequivalvis *HbI expression*

pGAPI was transformed into *E. coli* strain JM 105 and the transformed cells were grown in 1 litre batches of broth, containing 100 μg/ml ampicillin, to an absorbance of 0.3 at 550 nm. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM and the cells were incubated for an additional 3–24 h at 37°C. As control, the cloning vector PCRII-HbI cDNA was similarly induced with ITPG and grown at 37°C in parallel experiments. Bacteria were harvested by centrifugation and the pellet resuspended in 50 mM phosphate buffer, 100 mM NaCl, pH 7.8. Cells were lysed for 1–2 h in ice by the addition of lysozyme to 1 mg/ml, followed by sonication for 30 s at medium intensity (three times). Supernatants were collected by centrifugation at 10,000 × *g* for 30 min; aliquots were electrophoresed on a 15% SDS-PAGE and stained with Coomassie brilliant blue.

2.7. *Purification of recombinant* S. inaequivalvis *HbI*

The red supernatant from the lysed cells was extensively dialysed against 1 mM Tris-HCl, pH 7.8, at 4°C and loaded on a CM 52 column equilibrated with the same buffer. A red band was eluted with 100 mM Tris-HCl, pH 7.8. The purity of the hemoglobin fraction was controlled by optical absorption and by 15% SDS-PAGE.

2.8. *Amino-terminal sequence analysis*

The N-terminal sequence was determined using an Applied Biosystems model 476A gas-phase protein sequencer. Protein (20 μg) was applied on a proBlott membrane disk (Applied Biosystems) and the automated sequence analysis was performed for 16 cycles.

2.9. *Difference spectra*

Oxy/deoxy HbI difference spectra were generated by putting the purified hemoglobin solution in both the sample and reference cuvettes; after recording the baseline, the reference cuvette was added of few crystals of sodium dithionite. The spectra were recorded in a Kontron Uvikon 860 spectrophotometer.

2.10 *Ligand binding of the recombinant HbI*

Oxygen binding of the purified hemoglobin was measured with the spectrophotometric method previously described [15] in 100 mM Tris-HCl, pH 7.5, at 25°C. The rate of CO combination, following full or partial photodissociation, were measured at 25°C on the supernatant of the lysed cells using the flash photolysis system previously described [16], in 50 mM phosphate buffer, containing 100 mM NaCl, at pH 7.8.

3. RESULTS

3.1. *Cloning and sequencing of* S. inaequivalvis *HbI cDNA*

Gel electrophoresis of total RNA extracted from *S. inaequivalvis* erythrocytes (Fig. 1) shows the presence of only 18S ribosomal RNA, in agreement with previous findings on other molluscs and invertebrates [17]. In fact, the apparent absence of 28S rRNA in *S. inaequivalvis* can be explained as being due to a 'break' in its primary structure, which would convert the 28S into an 18S component, during the sample preparation.

The RNA preparation was used for the PCR experiments. After cloning of the 700 bp band (Fig. 2, lane b), three positive clones were sequenced. The sequence of the 678 bp cDNA (pAC3) is reported in Fig. 3, with the deduced protein sequence of 146 residues, identical to that of *S. inaequivalvis* homodimeric HbI globin, except for the substitution of an Asp residue for an Asn residue at position 56 [10]. The same Asp-56 residue was found in *A. trapezia* HbI and HbII A- and B-globins [18–20]. Two other clones were 98% and 95% identical, most likely reflecting either individual polymorphism or allelic forms. The corresponding protein sequences show only a single further amino acid substitution, i.e. His for Asn at position 44, due to a single nucleotide substitution at position 133. The same His-44 is present in *A. trapezia* HbI [18]. In all cases the mRNA contained the

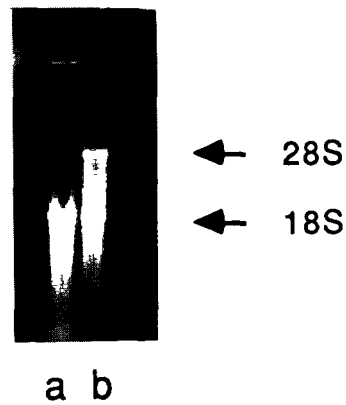

Fig. 1. 1% Agarose/2.2 M formaldehyde gel electrophoresis of 10 μg of total RNA from *S. inaequivalvis* (lane a) and bovine lung total RNA as control (lane b). The positions of 28S and 18S are indicated on the right

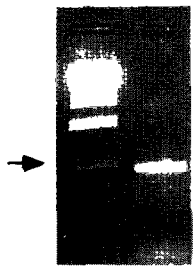

Fig. 2. 1.2% Agarose gel electrophoresis of the PCR products from *S inaequivalvis* total RNA. Lane a = marker lambda DNA–*Bst*EII digest (2 μg) Lane b = HbI globin PCR product.

polyadenylation consensus sequence typical of eukaryotes.

3.2. S. inaequivalvis *HbI expression*

In vivo expression was performed using the vector pGAPI as described in section 2. Upstream of the initiator methionine codon a sequence containing a Shine-Delgarno ribosome binding site and the spacer region from *P. putida* cytochrome $P_{450_{cam}}$ gene was added to maximize translational efficiency [14]. Translational stop codons were introduced in all three reading frames upstream of the ribosome binding site to inhibit the formation of a LacZ fusion protein [21].

Expression experiments were conducted at 37°C after IPTG induction, usually for 16 h. HbI synthesis was immediately detected from the red color of the pelleted cells. The yield was about 10% of the total *E. coli* proteins. The level of HbI obtained from the supernatant of the lysed cells after 3 h induction is shown in Fig. 4, lane c. Incorporation of endogenous heme was confirmed by the absorption spectrum in the Soret and visible region (data not shown). The absorption spectrum of the purified recombinant HbI, with the Soret band at 416 nm and the two bands at 542 and 577 nm (Fig. 5), is indicative of a fully oxygenated protein. Addition of CO gas generates the typical absorption spectrum of a CO-liganded hemoglobin (not shown). In the inset of the Fig. 5, the oxy/deoxy difference spectrum is shown, with peaks at 413 and at 434 nm, in agreement with the spectral properties of the native hemoglobin [1]. Gas-phase sequence analysis of the N-terminus of the purified recombinant protein indicated the correct amino acid sequence, with only 4% retention of the N-terminal methionine with respect to proline in the first cycle.

3.3. *Functional properties of recombinant* S. inaequivalvis *HbI*

Oxygen binding properties of recombinant HbI are almost identical to those of the natural protein [1,22]: at pH 7.5 and 25°C, the p_{50} value and the Hill coefficient, n, are 10.5 torr and 1.6, respectively. In Fig. 6 the kinetics of CO combination following flash photolysis, is shown at three different extents of photodissociation, at pH 7.8 and 25°C. At full photodissociation, the initial rate constant (l'_1) is $\sim 1.6 \times 10^5\ M^{-1} \cdot s^{-1}$, while at 5% photodissociation it reaches the limiting value (l'_2) of $\sim 3 \times 10^5\ M^{-1} \cdot s^{-1}$. This roughly two-fold increase in the CO combination velocity constant as the reaction proceeds is in full agreement with the results obtained with the native protein and is indicative of a cooperative CO binding [6].

4. DISCUSSION

In this paper we report the first expression of a functional dimeric hemoglobin in *E. coli* from its native cDNA. The protein sequence derived from the HbI cDNA sequence is the same as that of the purified protein [10], except for residue 56, which is aspartic acid, on the basis of the codon found. This difference stresses once again the importance of comparing protein sequences deduced from cDNA sequences with those ob-

```
    *  *  *
ATGCCGTCGGTGTACGATGCAGCAGCTCAGCTTACCGCTGACGTAAAGAAGGATCTGAGAGACAGCTGGAAAGTCATTGGTAGTGACAAA  90
 M  P  S  V  Y  D  A  A  A  Q  L  T  A  D  V  K  K  D  L  R  D  S  W  K  V  I  G  S  D  K   30

AAAGGAAACGGTGTTGCCTTGATGACAACTCTGTTCGCCGATAATCAAGAAACTATTGGATACTTCAAACGTTTGGGAGACGTCAGCCAG 180
 K  G  N  G  V  A  L  M  T  T  L  F  A  D  N  Q  E  T  I  G  Y  F  K  R  L  G  D  V  S  Q   60
                                          H
GGCATGGCAAACGACAAACTGAGAGGACATTCCATCACTCTTATGTACGCATTGCAGAACTTCATTGATCAGCTCGACAACCCCGATGAC 270
 G  M  A  N  D  K  L  R  G  H  S  I  T  L  M  Y  A  L  Q  N  F  I  D  Q  L  D  N  P  D  D   90

TTGGTCTGTGTTGTAGAAAAATTTGCCGTCAACCACATCACCAGAAAAATCAGTGCAGCTGAATTTGGAAAAATCAATGGCCCAAATAAA 360
 L  V  C  V  V  E  K  F  A  V  N  H  I  T  R  K  I  S  A  A  E  F  G  K  I  N  G  P  N  K  120

AAGGTTTTGGCAAGCAAGAATTTCGGTGATAAATACGCCAACGCATGGGCAAAACTTGTAGCTGTTGTCCAGGCTGCTTTATAAAGTGGG 450
 K  V  L  A  S  K  N  F  G  D  K  Y  A  N  A  W  A  K  L  V  A  V  V  Q  A  A  L  *        147

ATTTCCGAAGACAACTTTCCCACTACAACCAGATGAATATAAACTTTACATCAACTTTACCATGTAAGGAGTCTTTAACATGGAGACAAT 540
CATCAGGGTAGCGTGATTTTGACATTAGTCTATAATCACGGGTGTTGATTTTATGTAAATTTTTGGTTGAATTTTTATCGTTATATTTCA 630
AATATCTTAATAAATTCTATAAGCATAAAAAAAAAAAAAAAAAA                                                 678
```

Fig 3. Nucleotide sequence of *S maequivalvis* HbI cDNA and deduced protein sequence. The three bases indicated by an asterisk (*) could be different in HbI DNA, due to the presence of inosine residues at positions 6, 9, 12 of the PCR primer. The underlined residue (Asp) at position 56 is at variance with the biochemical data [10] (see text). In other clones, the underlined nucleotide 133 is C, and the corresponding residue is His for Asn, as indicated The polyadenylation site is underlined.

Fig. 4. 15% SDS-PAGE analysis of recombinant *S inaequivalvis* HbI expression using the *E coli* JM105 system. *S. inaequivalvis* erythrocytes (lane a). IPTG induced cells transformed with PCRII HbI cDNA (lane b) or with pGAPI (lane c). Molecular weight markers, Sigma VII-L (lane d). Recombinant purified HbI (lane e) In lane a, the larger, upper band corresponds to the tetrameric HbII A- and B-globins, while the lower band corresponds to HbI globin [1]

tained biochemically. According to the crystal structure, residue 56 is in the middle of the so-called CD loop (D helix is absent in this hemoglobin) participating with its carboxyl group in the formation of hydrogen bonds characteristic of a 3_{10} helix [8]; the possible role of the Asp-56 carboxyl group remains to be elucidated.

Insertion of the HbI cDNA (duly modified for maximal expression efficiency) under the control of the lac promoter into the same PCRII plasmid used for cDNA cloning, gave rise to a significant level of HbI synthesis. Recombinant HbI was isolated from the supernatant of the cell lysate in high yield, up to 10 mg per 10 g of *E. coli* cell paste. Like human hemoglobin and sperm whale myoglobin expressed in *E. coli* [21,23], HbI globin is capable of incorporating heme provided by the host cell, and generating a native-like dimer. The intrinsic nature of HbI, which is made of two identical subunits may have favoured the formation of the oxygenated hemoglobin, without iron oxidation or protein denaturation.

Contrary to that found previously in the *E. coli* expression of other globins [21,24,25], the N-terminal methionine residue of *S. inaequivalvis* recombinant HbI is almost completely processed (~96%), leaving proline as the amino-terminus.

Recombinant HbI is fully functional. The p_{50} value measured at 25°C compares very well with that reported previously (9.8 torr) in 0.1 M phosphate under similar conditions [22]. More importantly, oxygen binding is cooperative as for the native protein [1], with an Hill coefficient very close to that of the native HbI ($n \approx 1.5$). Cooperativity in ligand binding is clearly evident also in the CO combination reaction, with rate constants very similar to those of an authentic sample of the protein [6].

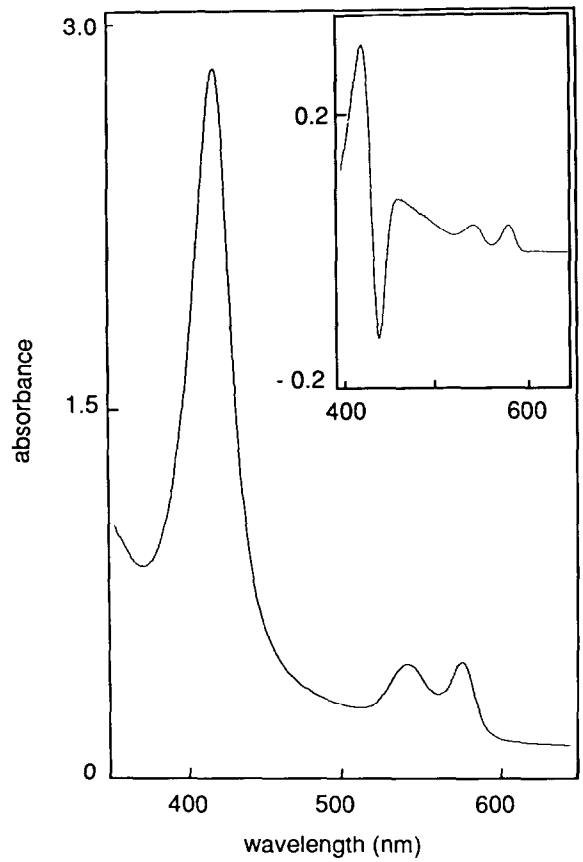

Fig. 5. Spectral properties of recombinant *S inaequivalvis* HbI. The spectrum in the Soret and visible region of the purified oxygenated protein is in 100 mM Tris-HCl, pH 7.5. The oxy/deoxy difference spectrum is shown in the inset.

These data are indicative of a properly assembled and functional molecule. Natural extension of the work presented here will be the expression of HbI mutants with specific amino acid substitutions at the heme pocket and at the subunit interface.

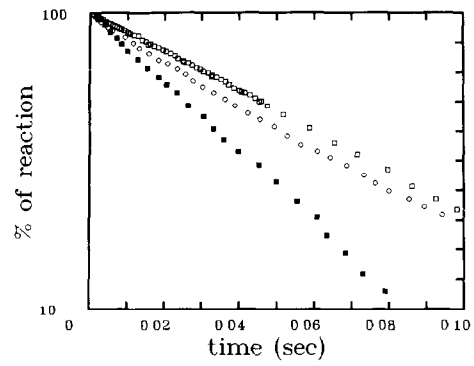

Fig 6. Kinetics of CO combination with recombinant *S inaequivalvis* HbI following flash photolysis of HbICO in the lysed IPTG induced *E coli* JM105 cells transformed with pGAPI. The % of the combination reaction (in a logarithmic scale) is plotted as a function of time after different degrees of photodissociation: 100% (□), 20% (○), 5% (■). Protein concentration, 2 μM; CO concentration, 91 μM; buffer, 0.1 M Tris-HCl, pH 7.8 at 25°C. Observation wavelength: 436 nm.

Acknowledgements This work was performed by M.C.P. in partial fulfilment of her thesis in Biotechnological Applications. The authors are deeply indebted to Prof. M. Brunori for measuring CO combination kinetics by flash photolysis on the recombinant HbI, to Dr. E Clementi for oxygen affinity measurements and to Dr. E. Schinina for the N-terminal sequence analysis. They also acknowledge grants from the Italian Ministry of University and Scientific and Technological Research (MURST, 40% and 60%)

REFERENCES

[1] Chiancone, E., Vecchini, P., Verzili, D., Ascoli, F. and Antonini, E (1981) J. Mol. Biol. 152, 577–592.
[2] Ohnoki, S., Mitomi, Y., Hata, R. and Satake, K. (1973) J. Biochem. (Tokyo) 73, 717–725.
[3] Furuta, H , Ohe, M. and Kajita, A. (1977) J. Biochem. (Tokyo) 82, 1723–1730.
[4] Djangmah, J S., Gabbott, P.A. and Wood, E.J (1978) Comp. Biochem Physiol. 60B, 245–250
[5] Como, P.F. and Thompson, E.O.P. (1980) Aust. J. Biol. Sci. 33, 643–652.
[6] Antonini, E., Ascoli, F., Brunori, M., Chiancone, E., Verzili, D., Morris, R.J. and Gibson, Q. (1984) J. Biol Chem. 259, 6730–6738.
[7] Royer Jr., W.E , Love, W.E. and Fenderson, F F (1985) Nature 316, 277- 280.
[8] Royer Jr., W.E , Hendrickson, W.A. and Chiancone, E. (1989) J. Biol Chem. 264, 21052–21061.
[9] Royer Jr., W.E., Hendrickson, W.A. and Chiancone, E. (1990) Science 249, 518- 521.
[10] Petruzzelli, R , Goffredo, B.M., Barra, D , Bossa, F., Verzili, D., Ascoli, F. and Chiancone, E. (1985) FEBS Lett. 184, 328–332.
[11] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[12] Maniatis, T., Fritsch, E.F and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[13] Sanger, F., Nicklen, S. and Coulson, A R. (1977) Proc. Natl Acad. Sci. USA 74,5463–5467.
[14] Unger, B.P., Gunsalus, I.C. and Sligar, S.G. (1986) J. Biol. Chem. 261, 1158–1163.
[15] Rossi Fanelli, A. and Antonini, E. (1958) Arch. Biochem Biophys. 77, 478–492.
[16] Brunori, M. and Giacometti, G.M. (1981) Methods Enzymol 76, 582–595.
[17] Ishikawa, H. (1977) Comp. Biochem. Physiol. 58B, 1–7
[18] Mann. R.G., Fisher, W.K., Gilbert, A.T. and Thompson, E.O P. (1986) Aust. J. Biol. Sci 39, 109–115.
[19] Como, P F. and Thompson, E.O P (1980) Aust. J Biol Sci. 33, 653–664.
[20] Gilbert, A.T. and Thompson, E.O P. (1985) Aust J. Biol Sci. 38, 221–236.
[21] Hernan, R.A., Hui, H L., Andracki, M.E., Noble, R W , Sligar, S.G., Walder, J.A. and Walder, R.Y (1992) Biochemistry 31, 8619–8628
[22] Ikeda-Saito, M., Yonetani, T., Chiancone, E , Ascoli, F., Verzili, D. and Antonini, E. (1983) J. Mol. Biol. 170, 1009–1018.
[23] Springer, B A. and Sligar, S.G (1987) Proc. Natl. Acad. Sci. USA 84, 8961–8965.
[24] Hoffman, S J , Looker, D.L , Roehrich, J.M., Cozart, P.E., Durfee, S.L., Tedesco, J.L. and Stetler, G.L. (1990) Proc Natl. Acad. Sci USA 87, 8521–8525.
[25] Guillemette, J.G., Matsushima-Hibiya, Y., Atkinson, T and Smith, M. (1991) Protein Engng 4, 585- 592.