Fluorescence properties of tryptophan residues in the monomeric *d*-chain of *Glossoscolex paulistus* hemoglobin: an interpretation based on a comparative molecular model

Carolina Bosch Cabral[a,b], Hidetake Imasato[a,*], José Cesar Rosa[b], Hélen Julie Laure[b,e], Carlos Henrique Tomich de Paula da Silva[c], Marcel Tabak[a], Richard Charles Garratt[c], Lewis Joel Greene[b,d]

[a]*Instituto de Química de São Carlos,, Universidade de São Paulo, P.O. Box 780, 13560-970, São Carlos, SP, Brazil*
[b]*Centro de Química de Proteínas, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil*
[c]*Instituto de Física de São Carlos, Universidade de São Paulo, São Paulo, Brazil*
[d]*Departamento de Ginecologia e Obstetrícia e Departamento de Biologia Celular, Molecular e Bioagentes Patogênicos, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil*
[e]*Escola Paulista de Medicina—UNIFESP, São Paulo, Brazil*

Abstract

The primary structure of the 142 residue *Glossoscolex paulistus d*-chain hemoglobin has been determined from Edman degradation data of 11 endo-Glu-C peptides and 11 endo-Lys-C peptides, plus the results of Edman degradation of the intact globin. Tryptophan occupies positions 15, 33 and 129. Homology modeling allowed us to assign the positions of these Trp residues relative to the heme and its environment. The reference coordinates of the indole rings (average coordinates of the $C_{\varepsilon 2}$ and $C_{\delta 2}$ atoms) for W15 and W129 were 16.8 and 18.5 Å, respectively, from the geometric center of the heme, and W33 was located in close proximity to the heme group at a distance which was approximately half of that for W15 and W129. It was possible to identify three rotamers of W33 on the basis of electrostatic and Van der Waals energy criteria. The calculated distances from the center of the heme were 8.3, 8.4 and 9.1 Å for Rot1, Rot2 and Rot3, respectively. Radiationless energy transfer from the excited indole to the heme was calculated on the basis of Förster theory. For W33, the distance was more important than the orientation factor, κ^2, due to its proximity to the heme. However, based on κ^2, Rot2 ($\kappa^2 = 0.945$) was more favorable for the energy transfer than Rot1 ($\kappa^2 = 0.433$) or Rot3 ($\kappa^2 = 0.125$). In contrast, despite its greater distance from the heme, the κ^2 of W129 (2.903) established it as a candidate to be more efficiently quenched by the heme than W15 ($\kappa^2 = 0.191$). Although the Förster approach is powerful for the evaluation of the relative efficiency of quenching, it can only explain pico- and sub-nanosecond lifetimes. With the average lifetime, $\langle \tau \rangle = 3$ ns, measured for the apomonomer as the reference, the lifetimes calculated for each emitter were: W33-1 (1 ps), W33-2 (2 ps), W33-3 (18 ps), W129 (100 ps), and W15 (600 ps). Experimentally, there are four components for oxymonomers at pH 7: two long ones

*Corresponding author. Tel.: +55-16-273-9955; fax: +55-16-273-9982.
E-mail address: hidetake@iqsc.sc.usp.br (H. Imasato).

of 4.6 and 2.1 ns, which contribute approximately 90% of the total fluorescence, one of 300 ps (4%), and the last one of 33 ps (7.4%). It is clear that the equilibrium structure resulting from homology modeling explains the sub-nanosecond fluorescence lifetimes, while the nanosecond range lifetimes require more information about the protein in solution, since there is a significant contribution of lifetimes that resemble the apo molecule. © 2002 Elsevier Science B.V. All rights reserved.

1. Introduction

The extracellular hemoglobin (Hb) of the *Glossoscolex paulistus* earthworm is a hexagonal bilayer (HBL) oligomeric hemoprotein [1] with a molecular mass of approximately 3.1 MDa [2]. This Hb is similar to several other annelid extracellular hemoglobins, and in particular, to the most extensively studied *Lumbricus terrestris* Hb. Although the first detailed reports on this giant structure were made using Svedberg determination of the molecular mass via the sedimentation coefficient [3], many features of this Hb remain unknown. For instance, the question of the number of heme-bearing and non-heme-bearing peptides has been a matter of controversy for some time. Also, the question of chain heterogeneity has been only recently addressed and experimentally studied using mass spectrometry. Before the publication by Fushitani et al. in 1996 [4], it was believed that the *L. terrestris* Hb contained three *linker* polypeptides. However, these investigators described evidence from two-dimensional gel electrophoresis for the existence of a minor fourth *linker* chain. In addition, Martin et al. [5] made a detailed analysis based on electron-spray ionization mass spectrometry, reaching the conclusion that most of the heme-containing chains are heterogeneous as well as confirming the presence of the fourth linker chain. However, it was demonstrated by reassembly experiments from *abc*-containing fractions and several combinations of *linker* chains that they are structurally interchangeable [6]. The analysis performed by Qiang Xie et al. [7] revealed that chain *d* is actually comprised of three distinguishable peptide chains.

Similarly, although great efforts have been made to characterize the Hb of *G. paulistus*, many aspects of its chemistry still need to be clarified.

The protein can be easily dissociated into a monomer and a disulfide-bonded trimer of globins, all of which contain heme, by gel filtration of the oxidized form at pH 9.0 [8]. Both fractions were demonstrated by SDS-PAGE electrophoresis to be essentially single bands of the monomer, as a 13-kDa globin, and of the trimer.

Homology (or comparative) modeling has proven to be a very useful tool for obtaining approximate structures of proteins based on homologues with known crystallographic coordinates [9,10]. Recently, this approach was used to describe a molecular model for the monomeric *d chain* of the *Lumbricus terrestris* hemoglobin [11], discussing the tryptophan and heme environment and identifying amino acid residues that are expected to be relevant to the stability of the subunit assembly of the giant complex and for cooperativity.

A structure for the HBL hemoglobin from *Lumbricus terrestris* was suggested from early reports based on small-angle X-ray scattering data [12] and stained scanning transmission electron microscopy [13,14]. An improved picture was obtained by scanning transmission cryoelectron microscopy [15–17]. The HBL hemoglobin was described as a 20-nm thick hexagonal bilayer with sides of 30 nm in length. The hexagonal layers were suggested to be rotated by 16° with respect to one another and the entire structure composed of 12 subassemblies localized at the corners of the hexagons. Based on the reassembled HBL from a mixture of *abc* and linker chains, each 1/12 subassembly showed the presence of a hole attributed to the absence of the *d chain* [15]. A three-fold axis was identified at the center of the hole, and this observation was used to infer a $(abcd)_3$ stoichiometry instead of the $(abcd)_4$ suggested by mass spectrometry [18,19]. Several of the observations, suggestions and predictions made with

regard to the quaternary structure of the *abcd* tetrameric unit have been subsequently borne out by the 5.5-Å resolution crystal structure determination of the whole particle [20]. The analysis of this crystallographic data shows that the whole molecule consists of 144 globin and 36 linker chains. The globin subunits in the particle were identified by their general globular geometry and are arranged into 12 sub-structures, each of which is indeed composed of three (*abcd*) tetrameric units. The *linker* chain assembly was roughly described as a globular folded region located between the two hexagonal layers with three emerging rod-like portions oriented towards the center of the HBL, associated with the central cryomicroscopic dense regions, assigned as t_a and t_b [16].

In the present study, the primary structure of the *d-chain* of *G. paulistus* hemoglobin has been determined and used to build a homology model of the protein focusing on data concerning the fluorescence properties of the tryptophan residues in the monomer. Gryczynski and collaborators have evaluated the Förster mechanism of energy transfer with success in several tryptophan- and heme-containing protein systems [21–23]. In these studies, they reported the analysis of fluorescence behavior based on crystallographic data. In the present study, we use molecular modeling as a tool to provide the atomic coordinates necessary to evaluate the tryptophan fluorescence quenching by the heme.

2. Methods and materials

2.1. Protein characterization and sequencing

2.1.1. Amino acid analysis

Amino acid analysis was performed using the phenylthiocarbamyl derivative method after acid hydrolysis in the vapor phase with 6 N HCl containing 1% phenol at 110 °C for 22 h [24]. A mixture containing 2.5 nmol of each amino acid (Standard H, Pierce Chemical Co., Rockford, IL, USA) was derivatized daily and 4% (100 pmol) was injected into the Picotag column (Waters).

2.1.2. Protein sequencing

Sequences were determined by automatic Edman degradation using a Procise device model 491 (Perkin-Elmer—Applied Biosystem Division, Foster City, CA, USA) and either gas-phase or pulse liquid chemistry with on-line identification of phenylthiohydantoin derivatives. The standard contained 10 pmol of each PTH-amino acid derivative and 10 pmol of lactalbumin was used to determine the efficiency of the instrument.

2.1.3. Reverse phase chromatography

Protein and peptide chromatography was carried out using a Pharmacia HPLC system employing two pumps (model 2248) and a dual wavelength spectromonitor (UV–Vis VWM2141) controlled by an LCC2252 gradient controller (Pharmacia-LKB, Uppsala, Sweden). In all cases, solvent A was 0.1% TFA and solvent B was acetonitrile/ water/TFA (80:20:0.85).

2.2. Sample preparation

2.2.1. Monomer fraction

Hemolymph was extracted into 0.1 mM sodium citrate anticoagulant and centrifuged at $2800 \times g$ at 5 °C in a clinical centrifuge to eliminate solids. After overnight dialysis at 5 °C against the 0.1 M Tris buffer, pH 7.0, containing 0.1 mM EDTA, the supernatant solution was centrifuged at $137\,000 \times g$ at 5 °C for 4 h, and the pellet of hemoglobin was collected. The pellet was suspended with a minimum amount of 0.1 M Tris buffer, pH 7.0, containing 0.1 mM EDTA. Using the same buffer as the solvent, the sample was submitted to gel filtration on Sephadex G-200 (0.9×100 cm) at room temperature. The effluent containing polymeric hemoglobin (28–42 ml) was oxidized with a five-fold excess of potassium ferricyanide and resubmitted to gel filtration on Sephadex G-200 in the same buffer at pH 9.0. The hemoglobin monomer was eluted from 72 to 84 ml [15,19].

2.2.2. Apomonomer preparation for fluorescence studies

The apomonomer was prepared from the oxy-monomer obtained by Sephadex G-200 gel filtration at pH 9.0 as described above. The heme was

extracted with methylethylketone after adjusting the pH of the protein solution to 2.3 [25]. The organic solvent was removed by dialysis against 0.1 M Tris–HCl buffer, pH 7.0.

2.2.3. Apomonomer preparation for sequence determination

The heme group was removed by RP-HPLC in a hand-packed C_4 (Bakerbond) column (4.6×60 mm). Monomer purification was carried out using RP-HPLC on a Vydac C_4 (250×4.6 mm, 5 μm) column equilibrated with 40% solvent B, and eluted with a gradient increase of solvent B (40–60% B for 35 min, 60–100% B for 20 min).

2.3. Sequence determination: peptide mapping

The complete primary structure was obtained by N-terminal sequencing of the intact globin plus the Edman sequencing of peptides obtained by two specific enzymatic cleavages: endoproteinase Lys-C (*Achromobacter* protease I, Wako Company, Japan) and endoproteinase Glu-C (*Staphylococcus aureus* V8, Boehringer Mannheim Biochimica, 1091 E17.4.1056506). An Applied Biosystems Procise sequencer was used.

2.3.1. Cysteine modification

The apoprotein monomer (40 nmol) was reacted with 1.16 μmol of DTT in 0.25 mM Tris–HCl buffer, pH 8.5, containing 0.1% sodium EDTA at 50 °C under N_2 in the dark for 4 h and then reacted with 80 μl vinylpyridine for an additional 4 h, under nitrogen at room temperature in the dark.

2.3.2. Glu-C endoproteinase digestion

A 160-μg aliquot (10 nmol) of the globin was hydrolyzed with endoproteinase Glu-C at a 1:10 (w/w) enzyme/protein ratio in 25 mM ammonium bicarbonate, pH 7.7, for 18 h at 25 °C. The reaction was stopped with 5% TFA and the peptide mixture was separated by HPLC on a Vydac C_{18} column (4.6×250 mm, 5 μm) equilibrated with 10% B. A linear gradient from 10 to 80% was applied over a period of 80 min (0.875% B/min) at a flow rate of 1.0 ml/min at 25 °C.

2.3.3. Lys-C endoproteinase digestion

A 160 μg aliquot (10 nmol) of the globin was submitted to digestion with endoproteinase Lys-C at a 1:10 (w/w) enzyme/protein ratio in 0.4 M *N*-ethylmorpholine, pH 8.0, for 22 h at 37 °C. The reaction was stopped with 5% TFA. The peptide mixture was separated by HPLC on a Vydac C_{18} column as described above.

2.4. Molecular modeling

2.4.1. Sequence alignment

The amino acid sequence of the *G. paulistus d*-chain was compared to those of globins of known three-dimensional structure using the BLAST program [26]. For sequence alignment, the following globin chains and their associated Brookhaven identifiers were considered: *Chironomous* hemoglobin (1ECD), *Urechis caupo* hemoglobin (fat innkeeper worm, 1ITH); *Glycera dibranchiata* hemoglobin (2HBG); leghemoglobin (1LH4); horse hemoglobin (2DHB, α and β chains); human hemoglobin (3HHB, α and β chains); *pagothenia bernachii* (fish) hemoglobin (1PBX, α and β chains); *Petromizon marinus* hemoglobin (lamprey, 2LHB); *Scapharca inaequivalis* (clam) dimeric (2SDH) and tetrameric hemoglobins (1SCT, *a* and *b* chains); human myoglobin (5MBN); *Aplysia* myoglobin (1MBA) and *Lumbricus terrestris* (*a*, *b*, *c* and *d* chains). Structurally equivalent residues were initially identified by rigid body superposition of two short sections of main-chain centered around the proximal and distal histidines, as described previously [11]. The *L. terrestris* and *G. paulistus* hemoglobins were subsequently introduced using the MULTALIGN program of the AMPS suite [27]. For this purpose, default values of the program were used employing the PAM250 scoring matrix of Dayhoff [28] corrected by a constant term of 8 in order to turn all matrix elements positive and a gap penalty of 8, but with the restriction that all insertions and deletions were limited to regions outside the common core of α-helices.

2.4.2. Building of the main-chain and side-chains

The previous model from *L. terrestris* (*d* chain) was chosen as the template for homology modeling

of the *G. paulistus* sequence because of its high sequence identity (55%), the functional equivalence of the two molecules, and particularly due to the conservation of Trp33, as discussed below. This model also was selected because it presented only one-residue extensions at the N and C termini with respect to *L. terrestris* and no internal insertions or deletions. The main-chain, therefore, required modification in order to accommodate these two extensions (one at each terminus), which were made by reference to one or more of the remaining globin structures.

A similar procedure was adopted for modeling the side-chains. Side-chain conformations were initially chosen based on those observed in other globin structures or alternatively selected from a backbone-dependent library of rotamers using the graphics programs WHATIF [29] and Insight II [30]. Manual adjustments were made in order to alleviate steric clashes prior to model refinement. In the case of the loss of a glycine, the backbone conformation was determined in order to avoid substitution of a residue into a disallowed region of Ramachandran space.

2.4.3. Model refinement

The resulting model was energy minimized in order to relieve steric strain using the AMBER force field [31] using the Steepest descent algorithm with the solvent modeled implicitly by a distance-dependent dielectric constant. For this purpose, the atoms of the heme group were fixed at their original positions, with the exception of the propionate side-chains and minimization was continued until stereochemical and packing parameters failed to further improve. Manual adjustments were made to regions which failed to improve on minimization prior to subsequent cycles.

2.4.4. Model evaluation

Throughout this procedure, the model was periodically evaluated for its stereochemical quality as well as for its residue packing and atomic contacts. The latter were analyzed using the QUALTY [32] option of the program WHATIF [33]. Residue packing analysis was performed by the method of Lüthy et al. [34], employing a sliding window of 21 residues. Stereochemistry was evaluated with the PROCHECK program [35]. Particular attention was paid to the orientation of Trp33, which forms part of the heme-binding pocket, and in this case the DGROTA option of WHATIF was employed in order to examine alternative conformations consistent with side-chain packing within the protein core.

2.5. Fluorescence lifetimes

Absorbance of the oxymonomer was determined using a Hitachi U 2000 spectrophotometer. For the steady-state fluorescence experiments a Hitachi F 4500 spectrofluorimeter was used. Excitation was carried out at 295 nm in order to avoid interference by tyrosine whose absorption peak has a maximum at 276 nm. The absorbance at 295 nm was kept below 0.1 in order to avoid the inner filter effect. Emission was monitored from 305 to 450 nm with an excitation and emission bandpass of 3 nm. The quantum yield was evaluated at pH 7.0 using a quantum yield of 14.0% for the tryptophan as a reference [36].

Time-resolved experiments were performed using an apparatus based on the time-correlated single photon counting method [37]. The excitation source was a Tsunami 3950 Spectra Physics titanium-sapphire laser, pumped by a 2600 Spectra Physics argon laser. The repetition rate of the 5-ps pulses was set at 400 kHz using a 3980 Spectra Physics pulse picker. The laser was tuned to provide an output at 890 nm, and a third harmonic generator BBO crystal (GWN-23PL Spectra Physics) provided the 295-nm excitation pulse that was directed to an Edinburgh FL900 spectrometer. The right angle configuration of the spectrometer was used. The emission wavelength was selected by a monochromator, and emitted photons were detected with a Hamamatsu R3809U microchannel plate photomultiplier. The FWHM of the instrument response function was typically 45 ps and a time resolution of 12 ps per channel was employed. Software provided by Edinburgh Instruments was used to analyze the decay curves. The adequacy of the exponential decay fitting was evaluated by inspection of the plots of weighted residuals and by the reduced chi-square (χ^2) statistical parame-

ter. The fluorescence decay curves were fitted to the multi-exponential decay model [Eq. (1)]:

$$I(t) = \sum \alpha \exp - \frac{t}{\tau} \qquad (1)$$

2.6. Lifetime computations

The fluorescence lifetimes of the oxymonomer were evaluated employing the Förster theory [38] for the radiationless energy transfer of tryptophan excitation energy to the heme. The orientation factor, κ^2, for each tryptophan–heme pair was calculated from the atomic coordinates resulting from the molecular modeling. The overlap integral was computed from the extinction coefficient spectrum of the oxymonomer and the fluorescence normalized to the total emission spectrum of the heme-free monomer.

Gryczynski and collaborators have performed calculations based on the Förster mechanism of energy transfer in several tryptophan- and heme-containing protein systems [12–14]. It has been assumed that the $^1L_a \rightarrow {}^1S_0$ tryptophan transition is more relevant than the $^1L_b \rightarrow {}^1S_0$ transition for the tryptophan fluorescence emission, as previously established by Ruggiero et al. [39]. The transition moments of these transitions lie in the indole plane and are orthogonal. The $^1L_a \rightarrow {}^1S_0$ transition is polarized at $39 \pm 7°$ clockwise from the pseudo-symmetry axis of indole as deduced from UV and IR linear dichroism of indole spread on stretched polyethylene film [40]. C_{2v} heme symmetry was characterized as a non-symmetrical absorber, implying a rod-like shape; linear dichroism data showed the UV region absorption and Soret band dominated by a non-degenerated linear transition moment oriented in the heme plane [41]. The UV band (deconvoluted with a single Gaussian component from 300 to 380 nm) has been considered to be responsible for the overlap integral, which reflects the tryptophan to heme energy transfer in the heme-containing proteins [23,42]. The heme UV band transition moment is reported to be oriented at approximately $60°$ counter-clockwise from the pseudo two-fold axis of the heme.

The energy transfer from tryptophan to the heme was predicted employing the Förster dipole–dipole theory [38,43], which permits evaluation of the energy transfer rate factor, $(R_0/r)^6$, calculated by Eq. (2) and based on the distance between the emitter and acceptor, r:

$$R_0^6 = 8.785 \times 10^{-25} \times \frac{1}{n^4} \times \kappa^2 \times Q_0 \times J \qquad (2)$$

where κ^2 is the orientation factor for the dipole–dipole interaction, J is the overlap integral [Eq. (3)], n is the refractive index of the medium between donor and acceptor, and Q_0 is the quantum yield of donor fluorescence in the absence of an acceptor. The parameter R_0 is associated with the distance at which the transfer efficiency is 50%. The overlap integral can be calculated from the experimental spectra [43]:

$$J = \frac{1}{\displaystyle\int_{\tilde{\nu}1}^{\tilde{\nu}2} I(\tilde{\nu})d\tilde{\nu}} \int_{\tilde{\nu}1}^{\tilde{\nu}2} \frac{I(\tilde{\nu})\varepsilon(\tilde{\nu})}{\tilde{\nu}^4} d\tilde{\nu} \qquad (3)$$

where $I(\tilde{\nu})$ is the fluorescence normalized to the total emission spectrum of the donor, $\varepsilon(\tilde{\nu})$ is the absorption spectrum of the acceptor, and $\tilde{\nu}$ is the wavenumber, where $\tilde{\nu}1$ and $\tilde{\nu}2$ define the limits of the experimental fluorescence spectrum. The orientation factor κ^2 is a function of the angle between the donor and acceptor transition moments, α; the angle between the donor moment and the translation vector connecting the two oscillators, β; and the angle between the acceptor moment and the translation vector, γ:

$$\kappa^2 = (\cos\alpha - 3\cos\beta\cos\gamma)^2 \qquad (4)$$

Finally, the fluorescence lifetimes can be estimated from Eq. (5) [43]:

$$\frac{\tau}{\tau_0} = \frac{1}{1 + (R/R_0)^6} \qquad (5)$$

Protein refractive index was calculated by summing the molar refractions of the constituent amino acids [44] of the monomer and assuming the partial specific volume of globin to be 0.738 cm^3 g^{-1} as
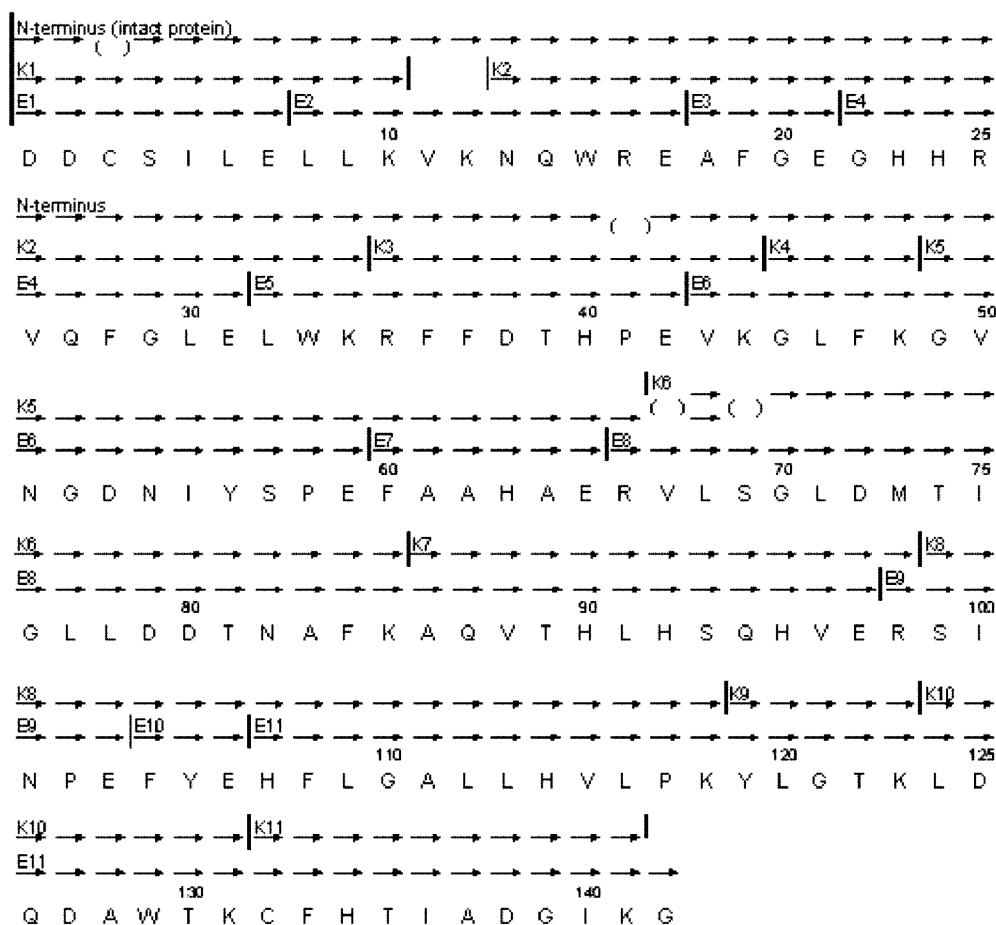
Fig. 1. Amino acid sequence of the hemoglobin monomer *d*-chain of *Glossoscolex paulistus*. Peptides are identified by letters that indicate the enzyme used for their preparation with endoproteinase Glu-C (E) and endoproteinase Lys-C (K). The arrows indicate sequence data obtained by Edman degradation of peptides or intact protein (N-terminus). The parentheses ( ) indicate residues not positively identified in some peptides, but which were identified in others. Quantitative Edman sequence data are given in Table 1.

determined by Martin et al. [5] for the *d*-chain of *Lumbricus terrestris*.

3. Results and discussions

3.1. Sequence determination

Fig. 1 presents the primary structure of the 142 amino acid residues of *G. paulistus* hemoglobin *d*-chain, with most of the residues being assigned on the basis of two or three independent Edman degradation determinations. Edman degradation

data for 11 Lys-C endoproteinase peptides and 11 Glu-C endoproteinase peptides, from which the primary structure was deduced, are presented in Table 1. The yield of these peptides varied from 5 to 93% and the quantitative Edman degradation data were straightforward and unambiguous. The sequence of the first 48 residues was also obtained by Edman degradation of the intact protein with an initial yield of 87.5% and a repetitive yield of 92.1%.

Endo-Glu-C peptide E-11 was assigned to the carboxyl-terminus of the molecule because of the

Table 1
Edman sequence data for endo-Lys-C (K) and endo-Glu-C (E) peptides derived from *Glossoscolex paulistus d*-chain

Peptide	Sequence position[a]	Recovery (%)[b]	Peptide sequencing[c]																			
K-1	1-10	26	D 100.9	D 95.5	C 52.1	S 20.6	I 35.9	L 30.4	E 23.8	L 18.9	L 16.7	K 12.1											
K-2	13-34	18	N 15.9	Q 10.2	W 3.8	R 12.2	E 8.4	A 8.6	F 9.1	G 8.6	E 5.5	G 6.8	H 3.6	H 5.7	R 5.1	V 3.8	Q 2.4	F 2.9	G 4.4	L 2.1	E 1.8	L 2.5	W 1.9	K 1.4
K-3	35-44	39	R 70.9	F 60.2	F 59.2	D 44.6	T 18.0	H 35.2	P 28.9	E 17.6	V 20.4	K 17.9												
K-4	45-58	38	G 20.8	L 20.2	F 21.3	K 22.9																		
K-5	49-68	17	G 65.9	V 68.5	N 8.3	G 18.5	D 11.0	N 11.2	I 15.1	Y 10.7	S 3.7	P 9.2	E 7.2	F 11.8	A 13.5	A 16.9	H 4.7	A 13.6	E 4.4	R 3.9	V 5.9	L[d] 9.5		
K-6	67-85	8	X	L 13.4	X	G 12.8	L 9.8	D 10.3	M 5.5	T 3.9	I 5.8	G 7.5	L 5.6	L 4.2	D 4.3	D 4.6	T 2.7	N 2.2	A 2.0	F 2.5	K 1.4			
K-7	86-98	7	A 16.9	Q 31.3	V 23.2	T 9.5	H 12.8	L 15.9	H 14.4	S 2.4	Q 13.8	H 9.4	V 9.3	E 4.2	R[d] 5.6									
K-8	99-118	10	S 15.2	I 33.0	N 23.9	P 23.1	E 18.9	F 27.1	Y 24.4	E 15.0	H 14.8	F 23.7	L 20.2	G 18.1	L 18.2	L 23.3	H 13.1	V 13.4	L 12.9	P 9.1	K 6.6			
K-9	119-123	64	Y 46.1	L 50.7	G 36.7	T 9.5	K 11.8																	
K-10	124-131	36	L 52.5	D 60.2	Q 52.6	D 50.9	A 43.9	W 11.8	T 14.8	K 25.6														
K-11	132-141	31	C 85.6	F 138.5	H 84.2	T 60.6	I 106.1	A 99.8	D 112.2	G 65.3	I 68.3	K[e] 60.4												
E-1	1-7	84	D 183.4	D 204.6	C 116.7	G 56.4	I 156.3	L 139.5	E 89.5															
E-2	8-17	63	L 156.2	L 171.5	K 157.2	V 152.6	K 148.1	N 114.4	Q 107.1	W 62.4	R 128.5	E 36.2												
E-3	18-21	93	A 122.4	F 135.3	G 117.5	E 81.1																		
E-4	22-31	75	G 307.9	H 160.5	H 242.2	R 205.9	V 255.2	Q 204.6	F 208.2	G 159.4	L 178.4	E 70.0												
E-5	32-42	93	L 159.5	W 99.8	K 115.7	R 95.4	F 138.4	F 148.0	D 83.7	T 60.3	H 30.4	P 31.5	E 33.6											
E-6	43-59	11	V 222.8	K 177.9	G 138.8	L 83.5	F 195.9	K 159.4	G 154.9	V 176.6	N 129.12	G 119.4	D 82.6	N 104.2	I 111.9	Y 95.4	S 23.4	P 63.7	E 23.2					
E-7	60-65	60	F 225.7	A 188.9	A 188.7	H 90.1	A 141.0	E 60.7																

Table 1 (Continued)

Peptide	Sequence position[a]	Recovery (%)[b]	Peptide sequencing[c]																					
E-8	66-97	27	R	V	L	S	G	L	D	M	T	I	G	L	L	D	D	T	N	A	F	K	A	Q
			90.6	131.7	127.6	33.8	97.9	109.1	83.9	114.7	47.2	80.1	72.4	77.9	80.9	59.9	68.2	31.3	49.6	49.9	57.2	42.6	51.5	49.9
			V	T	H	L	H	A	Q	H	V	E												
			46.1	19.2	32.4	31.0	35.8	12.3	20.7	22.1	16.8	5.7												
E-9	98-103	6	R	S	I	N	P	E																
			2.1	19.3	57.4	41.0	39.7	60.7																
E-10	104-106	47	F	Y	E																			
			82.2	99.4	48.4																			
E-11	107-142	5.8	H	F	L	G	A	L	L	H	V	L	P	K	Y	L	G	T	K	L	D	Q	D	A
			127.2	182.6	164.3	140.3	139.3	154.7	160.7	109.2	152.8	139.3	96.4	99.9	120.9	108.1	84.4	51.7	71.1	75.4	58.5	62.3	56.4	49.6
			W	T	K	C	F	H	T	I	A	D	G	I	K	G								
			7.14	20.9	30.1	21.3	31.3	22.2	12.4	21.2	18.3	14.4	15.1	17.2	12.4	9.6								

[a] Position of E- and K-peptides at the chain-d (Fig. 1).

[b] Recovery yields of E- and K-peptides after C18-HPLC.

[c] The number below each amino acid is the yield of the PTH-amino acids in pmol.

[d] Peptides obtained by non-specific cleavages.

[e] The expected Gly142 in the K-11 was released by endo-Lys. Then, the Gly142 was identified in the E-11.

absence of glutamic acid (Table 1) and the failure to identify any peptide that extends beyond the C-terminus, Gly142. Two apparently unexpected cleavages were obtained in the endo-Lys-C peptides at residues Arg66 and Arg98. In both cases, glutamic acid preceded the arginine residues.

3.2. Molecular modeling

3.2.1. Model quality

Model evaluation by PROCHECK [35] indicates that 90.4% of the residues are in the most favorable regions of ϕ/ψ space. Residue environment evaluation performed with the VERIFY_3D program, which is an implementation of the method of Lüthy et al., resulted in a quality index of 59.28, which is well above the minimum value for correctly-folded structures (28.90) and close to the average expected value for experimentally determined structures of proteins of 142 residues (64.23). The atomic contact quality score for the final model was −0.589, which compares favorably to the minimum acceptable value of −1.2 for well-determined structures. All analyses, therefore, suggest that the model presents overall sequence-structure compatibility.

The model is compatible with the classical globin fold lacking the D helix, as seen, for example, in mammalian α-chains. Despite the model's overall quality, the VERIFY_3D results identified the first 10 residues (corresponding to the A helix) to be in an unusual chemical environment. This was also observed in our previous analysis of the *d*-chain from *L. terrestris* and is readily explained by the fact that the analysis is performed on the isolated monomer, out of its natural structural context within the full hemoglobin particle. Specifically, in the tetrameric structure predicted by Viana et al. [11] and borne out by Royer et al. [20], the A helix is buried towards the center of the tetramer. This would be expected to result in the burial of several hydrophobic residues exposed on its surface and responsible for the low VERIFY_3D score in this region.

3.2.2. Alignment

Fig. 2 compares the alignment of the *d*-chain of *G. paulistus* and *L. terrestris* [10]. The *Glossos-*

colex paulistus monomer presents an additional amino acid at both its N- and C-terminal, as compared with the *Lumbricus terrestris d*-chain. Besides these, no other insertion or deletion was necessary during model construction. In order to detect uncommon substitutions and evaluate their structural consequences, the amino acid sequence was analyzed with reference to Bashford's [45] systematic study of more than 200 globins. The most relevant substitutions, which characterize unusual amino acids in the given positions, include a tryptophan at position 33 (B10), a histidine at 92 (F5) and a tyrosine at 105 (G5). The model suggests that His95 (F8) may form a novel salt-bridge with the C-terminal. The Trp33 and Tyr105 are described more fully below.

3.2.3. Tryptophan environment

Fig. 3 shows the relative orientations of the indole rings of the three tryptophan residues relative to the heme. The orientation of the indole rings of Trp15 and Trp129 can be readily assigned as they are present in many other globin structures, and yield good atomic contact scores of 1.0 and −0.5, respectively. Although unusual, the Trp33 at position B10 has already been seen in our previous modeling study [11], as well as in other more recently determined sequences of giant hemoglobin chains. It can be accommodated within the heme pocket by slight adjustment to the relative position of helix B. Trp33 can be modeled in three acceptable conformations, which we term Rot1, Rot2 and Rot3, identified by systematic rotamer searches. Rot1 corresponds to the conformation as modeled in *L. terrestris* and characterized by $\chi_1 = -80.5°$ and $\chi_2 = 108.8°$. Rot 2 was characterized by $\chi_1 = -111.0°$ and $\chi_2 = -108.8°$, and Rot3 by $\chi_1 = -73.3°$ and $\chi_2 = -87.7°$. Although the side-chain torsion angles of Rot2 correspond to a staggered configuration, all three rotamers correspond to conformations which are free from steric hindrance. Alternative options could be eliminated as they gave rise to unacceptable inter-atomic contacts.

The relative orientation and distance of the indole ring with respect to the heme are important parameters for the evaluation of radiationless energy transfer. All three Trp33 rotamers were approx-
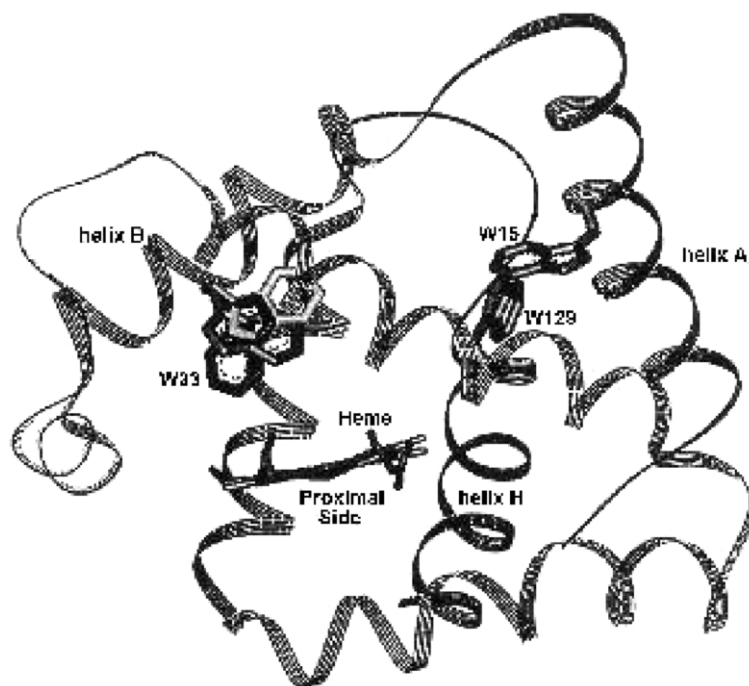
```
G. p. monomer         0            1            2              3          4          5              6          7
Sequence Numbering  1 2 3  4 5 6 7 8 9 0 1 2 3 4 5 6 7 8  9 0   1 2 3  4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  0 1 2 3 4 5 6  7 8 9 0 1 2 3 4  5 6      7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6
Helical notation            A + + + + + + + + + + + + A15          B + + + + + + + + + + + + B16 C + + + + + 7          D + + + + + D7 E + + + + + + + + + + + + + + + + + + + E20
G.p. heme contacts                                              d       d              d       d                                d       d   d   d   d

G. paulistus Monomer  D D C  S I L E L L K V K N Q W R E A  F G - - - E G H  H R V Q F G L E L W K R F F D T  H P E V K G L  F K G V N G D N  I Y - - - - -  S P E F A A H A E R V L S G L D M T I G
L. terrestris Lt-d (d1')  E C  L V T E S L K V K L Q W A S A  F G - - - H A H  E R V A F G L E L W R D I I D D  H P E I K A P  F S R V R G D N  I Y - - - - -  S P E F G A H S Q H V L S G L D I T I S


                          8                9                 100          110            120                130            140
                        7 8 9 0  1 2 3 4 5 6 7  8 9 0 1 2 3 4 5 6    7 8 9  0  1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9  0 1 2 3 4    5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2
                            F + + + + + + F9          G + + + + + + + + + + + + + + + + G19          H + + + + + + + + + + + + + + + + + + + + H21
                            p     p p        p       p        p p     d

G. paulistus Monomer  L L D D - - T N A F K A Q  V T H L H S Q H V  - - - E R S - I  N P E F Y E H F L G A L L H V L P K Y  L G T K L  - - - D Q D A W T K C F H T I A D G I K G
L. terrestris Lt-d (d1')  M L D T - - P D M L A A Q  L A H L K V Q H V  - - - E R N - L  K P E F F D I F L K H L L H V L G D R  L G T H F  - - - D F G A W H D C V D Q I I D G I K -
```

Fig. 2. Comparison of the primary structure of *G. paulistus* monomer *d*-chain and *L. terrestris* monomer globin. The α-helices are identified by the usual A–H notation. The lower case 'd' and 'p' are used to identify the residues localized at *distal* and *proximal* sides of the heme.

Fig. 3. Line ribbon diagram illustrating the globin-like folding of *Glossoscolex paulistus* hemoglobin monomer. Residues W15 (A12), W33 (B10-Rot1, Rot2 and Rot3), and W129 (H8) and heme are rendered as sticks.

imately orthogonal to the heme plane and the distances to the heme center (as measured with respect to the average coordinates for the $C_{\varepsilon2}$ and $C_{\delta2}$ atoms of the indole ring system) were 8.3, 8.4 and 9.1 Å for Rot1, Rot2 and Rot3, respectively. The other two tryptophans were essentially parallel to the heme and the distances were 18.8 and 18.5 Å for Trp15 and Trp129, respectively. Although the 90° orientation between the rings is unfavorable for energy transfer, the Trp33 rotamers are localized within the heme pocket at a distance almost half that of Trp15 and Trp129. Since energy transfer falls with the sixth power of the distance, this factor turns out to be decisive.

Solvent accessibility is also an important parameter for understanding fluorescence behavior as it influences the dielectric relaxation route [46]. All three tryptophans of the *G. paulistus d*-chain are relatively buried. The accessibility values calculated for Trp15, Trp33 and Trp129 by the method of Lee and Richards [47] using a water molecule approximated by a rolling sphere of radius 1.4 Å

were 14.06%, 7.06% and 0.41%, respectively, considering the area for a tryptophan in the extended conformation of a Gly–Trp–Gly tripeptide to be 271 Å² [48]. The model further suggests that the ε-NH$_3$ groups of Lys10, Lys12 and Lys131 may act as specific quenchers for Trp129. Arg25 and Tyr105 may play similar roles for Trp15 and Trp33, respectively.

3.2.4. Heme pocket

Besides the distal and proximal histidines [His63 (E7) and His95 (F8), respectively], which are present as expected, the heme environment can be characterized by 10 amino acid residues on its distal side (Fig. 4a) and seven residues on its proximal side (Fig. 4b) as shown in the alignment scheme (Fig. 2). In general, the hydrophobicity of the heme pocket is well conserved in all globins. For example, Phe47, Val67, Leu71, Leu91 and Ile100 of the *G. paulistus* sequence are representative of well conserved hydrophobic amino acids. The most common lysine residues are substituted
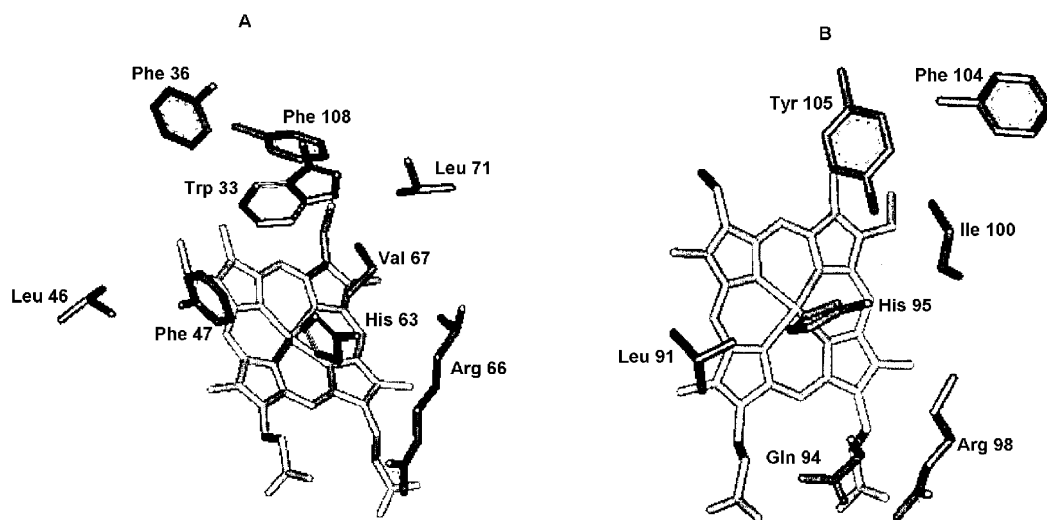
Fig. 4. Heme environment—proximal side (a) and distal side (b).

by Arg66 (E10) and Arg98 (FG2), but without alteration of their stabilizing role of the propionate groups of the heme. At positions F3 (His90) and F7 (Gln94) the residues observed in all chains of the *L. terrestris* hemoglobin are conserved in *G. paulistus*. It has been suggested that these residues form cross-bridges with the propionate side-chains of the heme group from a neighboring monomer [11]. Their conservation in the *G. paulistus* sequence, together with the recent crystallographic data on the full particle, which suggest that inter-subunit contacts are made via the E and F helices [20], are consistent with this hypothesis.

One particularly notable feature of the *Glossoscolex paulistus* monomer is the presence of Tyr105 at position G5 rather than the highly conserved Phe or Leu observed in other globins. Tyrosine was not observed in this position in any of the 226 globin sequences analyzed by Bashford et al. [45], but has subsequently been observed in several globin sequences from giant hemoglobins [49]. Its presence presents the problem of satisfying the hydrogen-bonding potential of the tyrosine hydroxyl group within the hydrophobic environment of the heme cavity. The model suggests that the most viable solution is an inversion of the orientation of the imidazole of the proximal histidine, maintaining the $N_{\varepsilon 2}$ coordinated to the iron. This places

the $N_{\delta 1}$ in a position suitable to donate a hydrogen bond to the tyrosine hydroxyl (Fig. 5). A similar orientation for the histidine imidazole group has been previously seen in the crystal structure of leghemoglobin [50]. In this case, its conformation is stabilized via the side-chain sulfur of Met73. In neither case is the rotation of the χ_2 torsion angle expected to significantly alter the overall stability of the molecule, as the normal hydrogen bonding partner of $N_{\delta 1}$ is the main-chain carbonyl of Leu91 (F4), which is already satisfied by a conventional i to $i+4$ helical hydrogen bond to the main-chain amide of the proximal histidine itself. This tyrosine could be an important quencher of the Trp33 fluorescence in the apomonomer.

3.3. Spectroscopy

Fig. 6a shows the absorption spectrum of the *d*-chain (solid line) and the emission spectrum of the apo-*d*-chain (dashed line) of *Glossoscolex paulistus* at pH 7. The result of treating the spectral data with Eq. (3) are presented. The overlap integral was computed by integrating the area under this curve. This function was fitted as a sum of three Gaussian curves. For the overlap integral only the UV band (peak 3 in Fig. 6b) was selected, neglecting the other two bands. The area under
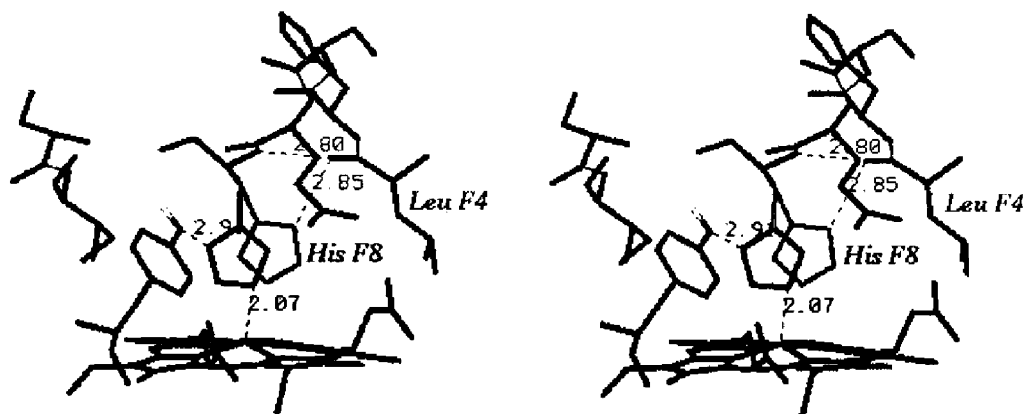
Fig. 5. The figure shows the two permitted positions of the imidazole ring of the proximal histidine (HisF8) and the hydrogen binding with Tyr105. This satisfies the hydrogen bonding potential of the phenol group of Tyr105 within the hydrophobic heme cavity.

peak 3 (centered at 338.1 nm) was equivalent to 72% of the total overlap integral. The other two peaks contribute the remaining 28%, peak 2 (400.59 nm) contributes 23% to the J and peak 1 (413.68 nm, Soret band) contributes 5%. These values show that the contribution of peak 2 is not negligible. However, on the basis of the linear dichroism of heme, only the transition dipole which corresponds to the UV component couples with the indole transition [41], thus only peak 3 was considered in the calculation.

3.4. Tryptophan fluorescence heme quenching analysis

The refractive index of globin d-chain was calculated to be 1.64, based on the amino acid composition, the molar refraction of amino acids reported in the literature [44], and assuming the partial specific volume of the d-chain of *Lumbricus terrestris* determined by Martin et al. [5]. The quantum yield (QY) of the apo-d-chain [Q_0, used in Eq. (2)] was determined to be 10% using 14% as the reference of tryptophan at pH 7 [36]. This QY represents 71% of tryptophan emission at pH 7. In comparison with the oxymonomer, which presented a 0.8% QY, the heme extraction led to a 13-fold increase of QY for the apomonomer. In addition, with the extraction of heme, the indole rings remained in an essentially hydrophobic envi-
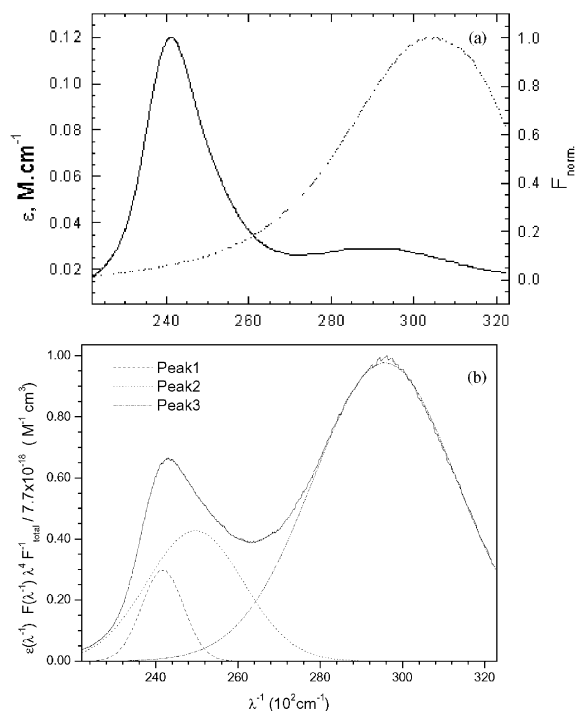
Fig. 6. Electronic transition spectra of d-globin at pH 7. Absorption spectrum (solid line), and normalized fluorescence spectrum (dashed line) (a). The product of absorptivity and normalized emission intensities as a function of wave number which was fitted by three Gaussian components: peak1—dashed line; peak2—dotted line; peak3—dash–dotted line (b).

Table 2
Förster parameters of *d*-globin of *G. paulistus* hemoglobin

Tryptophan	Rotamer	κ^2	R_0, Å	r, Å	τ/τ_0	τ(ns)[*]
W15		0.191	20.6	16.8	0.225	0.675
W129		2.903	32.5	18.5	0.033	0.099
W33	Rot1	0.403	26.2	8.39	0.001	0.003
	Rot2	0.945	29.9	8.43	0.0005	0.0015
	Rot3	0.125	21.3	9.19	0.006	0.018

[*] $\tau_0 = 3$ ns.

ronment, as indicated by the wavelength at the maximum of emission of the apomonomer at approximately 31 000 cm^{-1} (or 320 nm) (Fig. 6a). Furthermore, the lower QY in comparison with the tryptophan in solution is an indication of the presence of intrinsic quenchers in the protein structure as observed for human serum albumin (HSA). These quenchers changed the fluorescence features in the presence of detergents and have been employed to characterize the interaction of detergent molecules with HSA [51,52].

Table 2 summarizes the results of the analysis of radiationless energy transfer according to the Förster approach employing the atomic coordinates obtained from the model. It is clear the importance of both parameters on fluorescence quenching due to radiationless Förster energy transfer. Since Trp33, independent of its conformation, is located inside the heme hydrophobic core, it is expected that its emission must be quenched by the heme very efficiently as indicated by the observed picosecond lifetimes (Table 2) and the peculiarity of each conformer can be characterized by the orientation factor (κ^2). Compared to Trp33, both Trp15 and Trp129 are approximately twice as distant from the heme. The effect of orientation is more noticeable in these cases. Despite the fact that Trp129 is further away, the κ^2 effect super-

imposes and Trp129 should be quenched more efficiently than Trp33. The measured lifetime remains in the sub-nanosecond range (Table 2) even for Trp15, the residue that presented the lowest energy transfer.

The results of lifetime assays of the oxymonomer at pH 7 are presented in Table 3. The lifetimes can be divided into nanosecond and picosecond ranges. The decay curves were adjusted by three or four exponential components. The best model was selected by a combination of the random distribution of residual values along the channels and the optimization of the statistical χ^2 parameter. In Fig. 7, the results of the decay analysis for the oxymonomer are presented. It can be seen that the residues distribution at short times, 5–10 ns, is not totally random with three exponentials model (Fig. 7b), which is consistent with a χ^2 of 1.53. When a fourth component of 33 ps was introduced, the randomness of the residuals was improved (Fig. 7d), consistent with a χ^2 of 1.12, indicating improvement of the statistical fit. Then, the four-exponential model for oxymonomer fluorescence decay at pH 7 was assumed. From the average lifetime, at pH 7, of 3.0 ns, it is clear that the contribution of the long lifetimes to the total fluorescence is predominant.

Based on the same considerations as discussed in relation to the oxymonomer, the fluorescence decay for the apomonomer was best described by a three-exponential model (Fig. 8). The residues were distributed randomly along the channels, Fig. 8b, and the χ^2 value of 1.12 was not improved by introducing a fourth lifetime. Generally, the lifetimes of the apomonomer resemble those of the oxymonomer, lacking, however, the shortest picosecond component (Table 3). The average lifetime for the apomonomer of 2.9 ns is quite close to that of the oxymonomer (3.0). Measurements per-

Table 3
Time-resolved fluorescence analysis

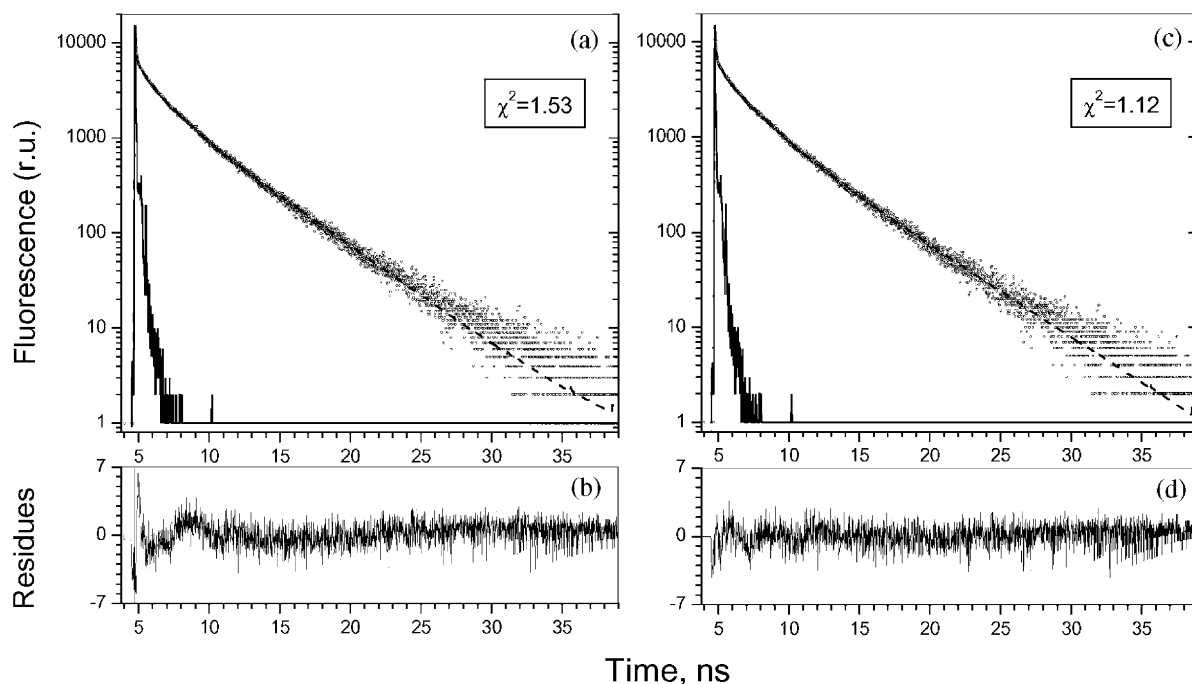| Sample	1		2		3		4		$\langle\tau\rangle$, ns
	τ, ns	α	τ, ns	α	τ, ns	α	τ, ns	α	
Oxymonomer, pH 7	0.033	0.33	0.3	0.02	2.1	0.029	4.6	0.015	3.0
Apomonomer, pH 7			0.25	0.07	1.8	0.046	4.7	0.016	2.9

Fig. 7. Time-resolved fluorescence decay for oxymonomer at pH 7. Analysis with three-exponentials ($\chi^2 = 1.53$). (a) Experimental decay curve for oxymonomer (circles); fit to the model (dashed line); laser pulse (solid line). (b) Residues. Analysis with four-exponentials ($\chi^2 = 1.12$). (C) Decay curve for oxymonomer (circles); fit to the model (dashed line); laser pulse (solid line). (d) Residues.

formed at different emission wavelengths for apo-monomer showed that, at pH 7.0, the contribution of the two shorter lifetimes to the total fluorescence is approximately 12% at 330 nm, changing to 10% at 360 nm emission. The contribution of the two longer nanosecond lifetimes to the total fluorescence is approximately 88% at 330 nm, changing to 90% at 360 nm emission.

The heterogeneity of fluorescence decay of the apomonomer may be associated with some residual heme; however, it is also worth mentioning that the Trp129 is in the proximity of Lys10, Lys12 and Lys131 whose ε-NH$_3^+$ moieties may modify its fluorescence, and the fact that Trp33 is within a hydrophobic cavity close to Tyr105 [53]. The environments of these two tryptophan residues may be important in shortening their fluorescence lifetimes.

Another approach can be used to explain the longer lifetimes, which is based on the equilibrium of heme binding by myoglobins [54,55]. The rate

of the dissociation of the heme can be monitored by following the ligand transfer, for instance, to serum albumin [56]. This dynamic binding of the heme can lead to its reorientation in the cavity, resulting in a disordered orientation [55]. Very interesting features in the single tryptophan-containing W7F and W14F, recombinants of sperm whale myoglobin [57], were reported. Fluorescence lifetimes of hundreds of picoseconds and nanoseconds, together with a very short lifetime, were observed. The longer fluorescence lifetimes were assigned to the disordered hemes, while the nanosecond fluorescence lifetimes were assigned to the apoprotein. However, the fluorescence emissions were dominated by the shortest picosecond lifetime components. In the monomer of *Glossoscolex paulistus*, the components with the longer lifetimes are the predominant ones (Table 3). It has also been reported that the hemes are released from the *Lumbricus terrestris* hemoglobin 2.7-fold faster than from soybean Hb, 2.5-fold faster than
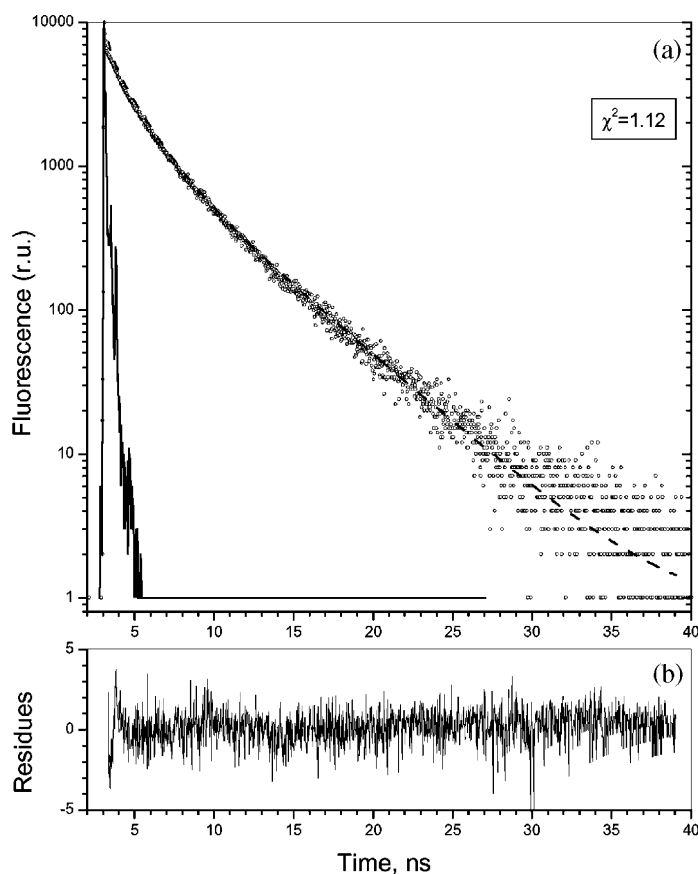
Fig. 8. Time-resolved fluorescence decay for apomonomer at pH 7. Analysis with three-exponentials ($\chi^2 = 1.12$). (a) Experimental decay curve (circles); fit (dashed line); laser pulse (solid line). (b) Residues.

from the human Hb A, and 79-fold faster than from the horseradish peroxidase, and this rate increased further approximately two-fold due to the aging of the hemoglobin [58]. One possibility is that this fast release of heme in the worm hemoglobin can also occur in the monomer of *Glossoscolex paulistus*. The fast release of heme would lead to a significant increase in the contribution of the nanosecond lifetime components. Further studies are needed to better characterize this phenomenon.

Acknowledgments

References

[1] M.C.P. Costa, C.F.S. Bonafé, N.C. Meirelles, F. Galembeck, Sedimentation coefficient and minimum molecular-weight of extracellular hemoglobin of *Glossoscolex-paulistus* (*Oligochaeta*), Braz. J. Med. Res. 21 (1988) 115–118.

[2] N.C. Meirelles, B. Oliveira, A.R. Oliveira, E. de Paula, S. Marangoni, G.M. Rennebeck, Erythrocruorin of *Glossoscolex-paulistus* (*Oligochaeta*, *Glossocolecidae*)— dissociation at alkaline pH and its ligand properties as revealed by chemical, immunochemical and electron-

microscopy studies, Comp. Biochem. Physiol. 88A (1987) 337–379.

[3] T. Svedberg, I.B. Eriksson, The molecular weight of erythrocruorin, J. Am. Chem. Soc. 56 (1934) 2834–2841, 1933.

[4] K. Fushitani, K. Higashiyama, M. Asao, K. Hosokawa, The amino acid sequences of two alpha chains of hemoglobins from Komodo dragon *Varanus komodoensis* and phylogenetic relationships of amniotes, Biochim. Biophys. Acta 1292 (1996) 273–280.

[5] P.D. Martin, A.R. Kuchumov, B.R. Green, et al., Mass spectrometric composition and molecular mass of *Lumbricus terrestris* hemoglobin: a refined model of its quaternary structure, J. Mol. Biol. 255 (1996) 154–169.

[6] A.R. Kuchumov, J.C. Taveau, J.N. Lamy, J.S. Wall, R.E. Weber, S.N. Vinogradov, The roles of linkers in the reassembly of the 3.6 MDa hexagonal bilayer hemoglobin from *Lumbricus terrestris*, J. Mol. Biol. 289 (1998) 1361–1374.

[7] Q. Xie, R.A. Donahue, K. Schneider, et al., Structure of chain d of the gigantic hemoglobin of the earthworm, Biochim. Biophys. Acta 1337 (1997) 241–247.

[8] S.C.M. Agustinho, M.H. Tinto, J.R. Perussi, M. Tabak, H. Imasato, Fluorescence studies of extracellular hemoglobin of *Glossoscolex paulistus* in met form obtained from Sephadex gel filtration, Comp. Biochem. Physiol 118A (1997) 171–181.

[9] K.L. Shukla, T.M. Gund, S.R. Meshnick, Molecular modeling studies of the artemisinin (*Qinghaosu*)–hemin interaction: Docking between the antimalarial agent and its putative receptor, J. Mol. Graphics 13 (1995) 215–222.

[10] C.-C. Chuang, S.-H. Wu, S.-H. Chiou, G.-G. Chang, Homology modeling of cephalopod lens S-crystallin: a natural mutant of sigma-class glutathione transferase with diminished endogenous activity, Biophys. J. 76 (1999) 679–690.

[11] E. Viana, C.H.T.P. Silva, M. Tabak, H. Imasato, R.C. Garratt, A molecular model for the d chain of the giant haemoglobin from *Lumbricus terrestris* and its implications for subunit assembly, Biochim. Biophys. Acta 1383 (1998) 130–142.

[12] I. Pilz, E. Schwartz, S.N. Vinogradov, Small-angle X-ray studies of *Lumbricus terrestris* hemoglobin, Int. J. Biol. Macromol. 2 (1980) 279–283.

[13] O.H. Kapp, S.N. Vinogradov, M. Ohtsuki, A.V. Crewe, Scanning transmission electron microscopy of extracellular annelid hemoglobins, Biochim. Biophys. Acta 704 (1982) 546–548.

[14] O.H. Kapp, M.G. Mainwaring, S.N. Vinogradov, A.V. Crewe, Scanning transmission electron microscopic examination of the hexagonal bilayer structures formed by the reassociation of three of the four subunits of the extracellular hemoglobin of *Lumbricus terrestris*, Proc. Natl. Acad. Sci. USA 84 (1987) 7532–7536.

[15] F. Haas, A. Kuchumov, J.C. Taveau, N. Boisset, S.N. Vinogradov, J.N. Lamy, Three-dimensional reconstruction of native and reassembled *Lumbricus terrestris* extracellular hemoglobin. Localization of the monomeric globin chains, Biochemistry 36 (1997) 7330–7338.

[16] J.C. Taveau, N. Boisset, S.N. Vinogradov, J.N. Lamy, Three-dimensional reconstruction of *Lumbricus terrestris* hemoglobin at 22 Å resolution: intramolecular localization of the globin and linker chains, J. Biol. Chem. 289 (1999) 1343–1359.

[17] F. Mouche, N. Boisset, P.A. Penczek, *Lumbricus terrestris* hemoglobin—the architecture of linker chains and structural variation of the central toroid, J. Struct. Biol. 133 (2001) 176–192.

[18] H. Zhu, M. Hargrove, Q. Xie, et al., Stoichiometry of subunits and heme content of hemoglobin from the earthworm *Lumbricus terrestris*, J. Biol. Chem. 271 (1996) 29999–30006.

[19] H. Zhu, D.W. Ownby, C.K. Rigs, N.J. Nolasco, J.K. Stoops, A.F. Riggs, Assembly of the gigantic hemoglobin of the earthworm *Lumbricus terrestris*—roles of subunit equilibria, non-globin linker chains, and valence of the heme iron, J. Biol. Chem. 271 (1996) 30007–30021.

[20] W.E. Royer, K. Strand, M. van Heel, W.A. Hendrickson, Structural hierarchy in erythrocruorin, the giant respiratory assemblage of annelids, PNAS 97 (2000) 7107–7111.

[21] Z. Gryczynski, C. Fronticelli, T. Tenenholz, E. Bucci, Effect of disordered hemes on energy-transfer rates between tryptophans and heme in myoglobin, Biophys. J. 65 (1993) 1951–1958.

[22] Z. Gryczynski, J. Lubkowski, E. Bucci, Heme–protein interactions in horse heart myoglobin at neutral pH and exposed to acid investigated by time-resolved fluorescence in the picosecond to nanosecond time range, J. Mol. Biol. 270 (1995) 19232–19237.

[23] Z. Gryczynski, S. Beretta, J. Lubkowski, A. Razynska, I. Gryczynski, E. Bucci, Time-resolved fluorescence of hemoglobin species, Biophys. J. 64 (1997) 81–91.

[24] S. Moore, D.H. Spackman, W.H. Stein, Chromatography of amino acids on sulfonated polystyrene resins—an improved system, Anal. Biochem. 30 (1958) 1185–1190.

[25] A. Tsuneshige, T. Yonetani, Preparation of mixed-metal hybrids, Methods Enzymol. 231 (1994) 215–222.

[26] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.

[27] G. Barton, M.J.E. Sternberg, A strategy for the rapid multiple alignment of protein sequences—confidence levels from tertiary structure comparisons, J. Mol. Biol. 198 (1987) 327–337.

[28] R.M. Schwartz, M.O. Dayhoff, in: M.O. Dayhoff (Ed.), Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC, 1979, pp. 353–358.

[29] T.A. Jones, J.-Y. Zon, S.W. Cowan, M. Kjelgaard, Improved methods for building protein models in elec-

tron-density maps and the location of errors in these models, Acta Cryst. A47 (1991) 110–119.

[30] Insight II Modeling Environment, release 98.0, San Diego: Molecular Simulations Inc., 1998.

[31] S.J. Weiner, P.A. Kollman, D.A. Case, et al., A new force-field for molecular mechanical simulation of nucleic-acids and proteins, J. Am. Chem. Soc. 106 (1984) 765–784.

[32] G. Vriend, C. Sander, Quality-control of protein models—directional atomic contact analysis, J. Appl. Cryst. 26 (1993) 47–60.

[33] G. Vriend, What if: A molecular modeling and drug design program, J. Mol. Graph. 8 (1990) 52–56.

[34] R. Lüthy, J.U. Bowie, D. Einsenberg, Assessment of proteins models with three-dimensional profiles, Nature 356 (1992) 83–85.

[35] R.A. Laskowski, M.W. MacArthur, J.M. Thornton, Procheck-A program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.

[36] P. Wiget, P.L. Luisi, Cooplipeptides containing aromatic residues spaced by glycyl residues. 9. Fluorescence properties of tryptophan-containing peptides, Biopolymers 17 (1978) 167–180.

[37] A.L.C.F. Souto, A.S. Ito, Tryptophan fluorescence studies of melanotropins in the amphiphile–water interface of reversed micelles, J. Eur. Biophys. 29 (2000) 38–47.

[38] T. Forster, Zwischenmolekulare Energiewanderung Und Fluoreszenz, Ann. Physik. 2 (1948) 55–75.

[39] A.J. Ruggiero, D.C. Todd, G.R. Fleming, Subpicosecond fluorescence anisotropy studies of tryptophan in water, J. Am. Chem. Soc. 112 (1990) 1003–1014.

[40] B. Albinsson, M. Kubista, B. Nordén, E.W. Thulstrup, Near-ultaviolet electronic-transitions of the tryptophan chromophore—linear dichroism, fluorescence anisotropy, and magnetic circular-dichroism spectra of some indole-derivatives, J. Phys. Chem. 93 (1989) 6646–6654.

[41] Z. Gryczynski, R. Paolesse, K.M. Smith, E. Bucci, Effect of central metal substitution on linear dichroism of porphyrins: evidence of out-of-plane transition moments, Biophys. Chem. 69 (1997) 71–84.

[42] Z. Gryczynski, T. Tenenholz, E. Bucci, Rates of energytransfer between tryptophans and hemes in hemoglobin, assuming that the hemes are a planar oscillator, Biophys.J. 63 (1992) 648–653.

[43] L. Stryer, Fluorescence energy-transfer as a spectroscopic ruler, Ann. Rev. Biochem. 47 (1978) 819–846.

[44] T.L. McMeekin, M. Wilensky, M. Groves, Refractive indices of proteins in relation to amino acid composition and specific volume, Biochem. Biophys. Res. Comm. 7 (1962) 151–156.

[45] D. Bashoford, C. Chotia, A.M. Lesk, Determinants of a protein fold. Unique features of the globin amino acid sequences, J. Mol. Biol. 196 (1987) 199–216.

[46] A.P. Demchenco, J. Gallay, M. Vincent, H.J. Apell, Fluorescence heterogeneity of tryptophans in Na,K-ATPase: evidences for temperature-dependent energy transfer, Biophys. Chem. 72 (1998) 265–283.

[47] B. Lee, F.M. Richards, Interpretation of protein structures—estimation of static accessibility, J. Mol. Biol. 55 (1971) 379–400.

[48] S. Miller, J. Janin, A.M. Lesk, C. Chothia, Interior and surface of monomeric proteins, J. Mol. Biol. 196 (1987) 641–656.

[49] http://bmbsgi11.leeds.ac.uk/bmbknd/promise/ GLOBINS.html.

[50] E.H. Harutyunyan, T.N. Safonova, I.P. Kuranova, et al., The structure of deoxy-leghemoglobin and oxy-leghemoglobin from lupin, J. Mol. Biol. 251 (1995) 104–115.

[51] E.L. Gelamo, M. Tabak, Spectroscopic studies on the interaction of bovine (BSA) and human (HSA) serum albumins with ionic surfactants, Spectrochim. Acta A 56 (2000) 2255–2271.

[52] E.L. Gelamo, C.H.T.P. da Silva, H. Imasato, M. Tabak, Interaction of bovine (BSA) and human (HSA) serum albumin with ionic surfactants: spectroscopy and modeling, Biochim. Biophys. Acta (2001), in press.

[53] S.J. Kroes, G.W. Canters, G. Gilardi, A. Hoek, A.J.W. Viser, Time-resolved fluorescence study of azurin variants: conformational heterogeneity and tryptophan mobility, Biophys. J. 75 (1998) 2441–2450.

[54] G.N. Lamar, H. Toi, R. Krishnamoorthi, Proton NMR investigation of the rate and mechanism of heme rotation in sperm whale myoglobin—evidence for intramolecular reorientation about a heme twofold axis, J. Am. Chem. Soc. 106 (1984) 6395–6401.

[55] W.R. Light, R.J. Rohlfs, G. Palmer, J.S. Olson, Functional-effects of heme orientation disorder in sperm whale myoglobin, J. Biol. Chem. 262 (1987) 46–52.

[56] H.F. Bunn, J.H. Jandl, Exchange of heme among hemoglobins and between hemoglobin and albumin, J. Biol. Chem. 243 (1968) 465–475.

[57] Z. Gryczynski, E. Bucci, Time resolved emissions in the picosecond range of single tryptophan recombinant myoglobins reveal the presence of long range heme protein interactions, Biophys. Chem. 74 (1998) 187–196.

[58] M.L. Smith, J. Paul, P.I. Ohlsson, K.G. Paul, The spontaneous hemin release from *Lumbricus terrestris* hemoglobin, Comp. Biochem. Physiol. 118A (1997) 1241–1245.