Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Engineering the internal cavity of neuroglobin demonstrates the role of the haem-sliding mechanism

G. Avella,[a] C. Ardiccioni,[b] A. Scaglione,[a] T. Moschetti,[c] C. Rondinelli,[a] L. C. Montemiglio,[a] C. Savino,[a] A. Giuffrè,[a] M. Brunori[a] and B. Vallone[a]*

[a]Istituto Pasteur–Fondazione Cenci Bolognetti and Istituto di Biologia e Patologia Molecolari del CNR, Dipartimento di Scienze Biochimiche 'A. Rossi Fanelli', Sapienza Università di Roma, Piazzale A. Moro 5, 00185 Rome, Italy, [b]Departments of Physiology and Cellular Biophysics, Columbia University College of Physicians and Surgeons, Russ Berrie Pavilion, 1150 St Nicholas Avenue, New York, NY 10032, USA, and [c]Biochemistry Department, Cambridge University, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, England

Correspondence e-mail: beatrice.vallone@uniroma1.it

Received 23 January 2014
Accepted 29 March 2014

PDB references: neuroglobin, M144W mutant, 4mu5; F106W mutant, 4o1t; F106W mutant, CO-bound, 4o35; V140W mutant, 4nzi; V140W mutant, CO-bound, 4o2g

Neuroglobin is a member of the globin family involved in neuroprotection; it is primarily expressed in the brain and retina of vertebrates. Neuroglobin belongs to the heterogeneous group of hexacoordinate globins that have evolved in animals, plants and bacteria, endowed with the capability of reversible intramolecular coordination, allowing the binding of small gaseous ligands (O_2, NO and CO). In a unique fashion among haemoproteins, ligand-binding events in neuroglobin are dependent on the sliding of the haem itself within a preformed internal cavity, as revealed by the crystal structure of its CO-bound derivative. Point mutants of the neuroglobin internal cavity have been engineered and their functional and structural characterization shows that hindering the haem displacement leads to a decrease in CO affinity, whereas reducing the cavity volume without interfering with haem sliding has negligible functional effects.

1. Introduction

Members of the globin family are spread among all living organisms, stemming from an ancestor that predates the appearance of eukaryotes. Haemoproteins belonging to this group are ubiquitous in animals and plants, and their function spans from oxygen transport to redox catalysis and signalling (Hardison, 1996, 1998; Vinogradov & Moens, 2008). The structural features typical of globins are the presence of a haem group and a characteristic three-dimensional structure (Bashford *et al.*, 1987; Holm & Sander, 1993) consisting of eight α-helices arranged in the so-called 3-over-3 α-helical fold. Sequence variability is very high, with only two positions that are strictly conserved among all members: His(F8), the so-called proximal histidine, and Phe(CD1) involved in π–π stacking interactions with the haem. A second His at position E7, called the distal histidine, is also often conserved and plays a major role in determining affinity and selectivity for ligands at the sixth iron-coordination position (we have thus used the globin-fold structure nomenclature and numbering; Antonini & Brunori, 1971; Perutz, 1979). The current hypothesis is that oxygen-transporting globins, such as haemoglobin (Hb) and myoglobin (Mb), evolved from an ancestor endowed with redox functions; the haem was most likely hexacoordinate, with the sixth ligand provided by the distal His(E7). Hexacoordination of the haem is indeed energetically favourable, and stabilization of the pentacoordinate state to allow reversible ligand binding represents a 'thermodynamic challenge' (Hoy *et al.*, 2007).

Within the prolific globin family, the best known members adopt the more recent open pentacoordinate form, but over the last decade it has become clear that His-hexacoordinate globins are widespread (in plants, invertebrates, vertebrates

and cyanobacteria). These globins have been more elusive since, being implicated in scavenging, catalysis or signalling functions, they are not as abundant as typical oxygen carriers (Halder *et al.*, 2007). In an elegant paper focusing on nonsymbiotic plant Hbs, Hargrove and coworkers deduced on the basis of structural comparisons that pentacoordination was achieved during evolution by stabilizing the position of the E-helix (Hoy *et al.*, 2007); this would result from increased interactions with other secondary-structure elements and from shortening of the flanking CD and EF loops, eventually preventing the distal His(E7) from binding to the haem iron.

Neuroglobin (Ngb) is a small monomeric globin mainly expressed in the brain and retina of vertebrates (Burmester *et al.*, 2000). The characteristic feature of Ngb is the hexacoordinate state, with (ferrous and ferric) iron being bound to the proximal and distal histidines; as outlined below, this His-hexacoordinate state is the main control in the binding of exogenous ligands such as O_2, CO and NO. Enhanced expression of Ngb was observed in rat neurons exposed to hypoxia and focal ischaemic conditions (Sun *et al.*, 2001), and it was shown to protect neurons after ischaemic insults (Sun *et al.*, 2001, 2003). The molecular basis of neuroprotection by Ngb, however, remains unclear. Several mechanisms have been proposed, including oxygen/redox sensing, scavenging of reactive oxygen and nitrogen species, and inhibition of apoptosis (Burmester & Hankeln, 2009).

The presence of a very large (∼300 Å^3) internal tunnel is another feature peculiar to Ngb, absent in other globins from vertebrates (such as Mb and Hb; Tilton *et al.*, 1984; Savino *et al.*, 2009; Vallone, Nienhaus, Brunori *et al.*, 2004). MD simulation studies showed some analogies between the Ngb tunnel and the Xe binding sites identified in Mb (Anselmi *et al.*, 2007), suggesting this tunnel to somehow be an amplified version of the canonical set of cavities shown to play a role in exogenous ligand binding to the haem (Lutz *et al.*, 2009). The crystallographic structure of Ngb in complex with Xe or Kr showed the presence in the cavity of four and two binding sites, respectively (Moschetti *et al.*, 2009), suggesting that these positions can be transiently occupied by gaseous ligands on their way to and from the distal haem site.

Binding of exogenous ligands to some His-hexacoordinate globins takes place upon displacement of the distal His(E7), either by 'swinging out' or by relocation of the E-helix mediated to a variable extent by the CD and EF loops (Hoy *et al.*, 2007; Trent *et al.*, 2004; Makino *et al.*, 2011). In the case of Ngb only, binding of an external ligand (CO) has a minor effect on the position of the distal His(E7) while being associated with a unique conformational change whereby rupture of the distal His–iron bond is associated with a substantial sliding motion of the haem, which fills part of the pre-existing tunnel into the protein matrix, causing displacement of the CD loop and repositioning of Phe106(G3) (Vallone, Nienhaus, Matthes *et al.*, 2004). As shown in Fig. 1, upon CO binding the haem moves more deeply into its crevice (iron displacement of 2 Å) followed by the proximal His(F8) (1.5 Å displacement of the Nε coordinated to the iron); on the other hand, the distal His(E7) is displaced by only 0.7 Å (Nε). Based on the crystallographic structure of carbomonoxy Ngb (Ngb–CO; PDB entry 1w92), Vallone, Nienhaus, Matthes *et al.* (2004) assigned a role to the large internal cavity also present in this protein in the reduced His-hexacoordinate state. This haem-sliding mechanism has been observed by molecular-dynamics (MD) simulations, which indicated, however, that in Ngb–CO two almost equivalent energy minima could be accessible to the haem: the less populated would correspond to the slid haem

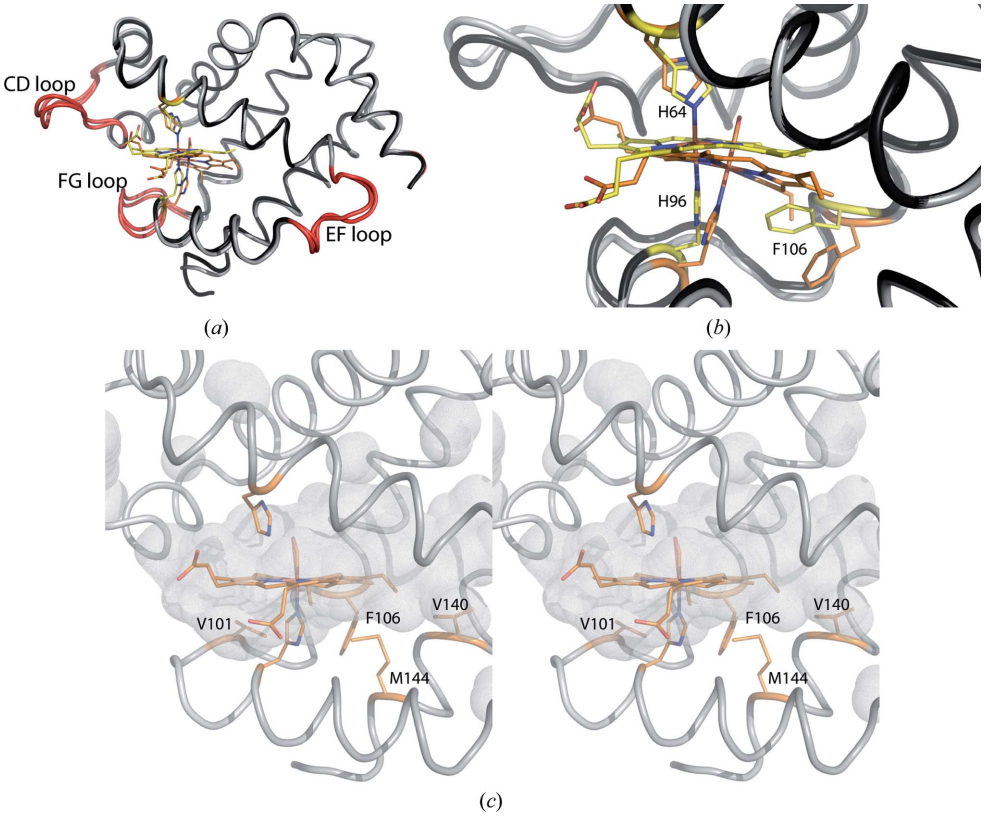

Figure 1
Superposition of wild-type metNgb (dark grey; PDB entry 1q1f) and wild-type Ngb–CO (light grey; PDB entry 1w92). (*a*) The CD, EF and FG loops are shown in red. (*b*) The displacement of Phe106 upon CO binding and haem sliding is shown; the haem group and Phe106 are shown in yellow for the met-hexacoordinate form and in orange for the CO-bound state. (*c*) Stereoview of the haem cavity of wild-type Ngb–CO, showing the residues selected for mutation (Val140, Val101, Met144 and Phe106) in orange.

Table 1
Crystallographic statistics.

Values in parentheses are for the highest resolution shell.

Structure (PDB code)	M144W (4mu5)	F106W (4o1t)	F106W–CO (4o35)	V140W (4nzi)	V140W–CO (4o2g)
Space group	$R3_2$:H	$R3_2$:H	$R3_2$:H	$R3_2$:H	$R3_2$:H
Unit-cell parameters					
$a = b$ (Å)	87.636	87.636	89.155	87.214	85.067
c (Å)	115.217	114.139	108.622	114.399	110.202
Resolution range (Å)	45.89–1.80 (1.864–1.800)	31.67–1.60 (1.657–1.600)	30.00–1.80 (1.86–1.80)	45.6–2.10 (2.175–2.100)	34.94–2.70 (2.795–2.700)
No. of reflections	30608 (3015)	45216 (4448)	155339 (15350)	19925 (1929)	8502 (808)
Unique reflections	15838 (1561)	22608 (2224)	15481 (1535)	9964 (966)	4286 (408)
Completeness (%)	98.89 (98.05)	99.97 (99.96)	99.1 (99.9)	99.75 (97.97)	97.90 (97.37)
$R_{\rm merge}$†	0.0208 (0.186)	0.0205 (0.118)	0.074 (0.215)	0.0203 (0.151)	0.02792 (0.1649)
Multiplicity	1.9 (1.9)	2.0 (2.0)	10.0 (10.5)	2.0 (2.0)	2.0 (2.0)
$\langle I/\sigma(I)\rangle$	22.61 (3.84)	25.38 (5.91)	13.7 (10.1)	28.00 (4.22)	18.79 (3.16)
$CC_{1/2}$	0.999 (0.878)	0.999 (0.955)		0.999 (0.914)	0.999 (0.912)
Wilson B (Å^2)	24.62	14.72	21.05	31.06	63.11
$R_{\rm work}$‡/$R_{\rm free}$§	0.199/0.249	0.168/0.195	0.187/0.198	0.219/0.271	0.208/0.227
No. of residues	148	148	150	148	148
No. of waters	93	160	160	35	0
No. of ligands	1 haem, 1 SO_4^{2-}	1 haem, 1 SO_4^{2-}	1 haem, 1 CO, 1 SO_4^{2-}, 3 1,4-dioxane	1 haem, 1 SO_4^{2-}	1 haem, 1 CO
R.m.s.d., bond lengths¶ (Å)	0.017	0.008	0.012	0.009	0.010
R.m.s.d., angles¶ (°)	1.92	1.04	2.15	1.26	1.48
Average B factors (Å^2)					
All	38.00	21.80	25.60	39.80	63.80
Macromolecule	38.50	21.00	25.30	40.00	64.00
Waters	42.20	33.80	32.40	38.50	0
Ligands	27.00	33.80	21.60	36.80	57.50
Ramachandran plot analysis (%)					
Favoured	97	97	93	98	99
Allowed	3	1	8	2	1
Outliers		2	3		

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is the mean value of $I_i(hkl)$ for all i measurements. ‡ $R_{\rm work} = \sum_{hkl} \big||F_{\rm obs}| - |F_{\rm calc}|\big| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ is the observed structure factor and $F_{\rm calc}$ is the calculated structure factor. § $R_{\rm free}$ is the same as $R_{\rm cryst}$ except that it as calculated with a subset (5%) of data that were excluded from the refinement calculations. ¶ Engh & Huber (1991).

conformation, while the more probable would have the distal His(E7) out of the haem pocket coupled to a conformational transition of the CD loop (Anselmi *et al.*, 2007). The latter conformation was not detected by crystallography, possibly owing to contacts constraining the CD corner in the crystal lattice or to an actual predominance of the haem displacement upon ligand binding (Anselmi *et al.*, 2008).

The work presented here aims at investigating the molecular mechanism controlling ligand affinity in Ngb using mutagenesis, kinetics and crystallography. Since the haem-sliding mechanism seems to be (to date) unique to Ngb and demands the occupancy of a fraction of the large internal tunnel, we decided to focus on specific experiments to test (i) the role of this cavity in the dynamic mechanism of haem sliding and (ii) the effect of hindering this motion on the binding of CO as the simplest prototype of functionally relevant ligands (O_2 and NO). A recent paper, based on experimental and molecular-dynamics data, indicates that restricting the space available to the haem by systematically varying the size of the distal residue Ile109(G8) might favour the Ngb pentacoordinate state (Nienhaus *et al.*, 2013). Thus, we set out to analyze by kinetics and crystallography the effect of mutations that reduce the volume of the tunnel in order to obtain functional evidence for haem sliding in solution and to assess the role of this cavity in the mechanism of ligand binding. We have identified by modelling four positions on the proximal haem side of the Ngb large cavity and introduced mutations with bulky side chains that are compatible with allowed rotamers and do not clash with neighbouring side chains in the deoxy His-hexacoordinate state. Therefore, in order to sterically hinder and partially oppose haem sliding or just to fill up the cavity (see Fig. 1*c*), we have purified and characterized four mutants at positions 101(FG3), 106(G3), 140(H15) and 144(H19), namely V101F, F106W, V140W and M144W. We have investigated the effects of these mutations on the kinetics of CO binding and determined their crystal structure in deoxy and CO-bound states in order to assess the effect of steric hindrance on haem dynamics.

2. Materials and methods

The gene coding for murine Ngb was synthesized based on the protein sequence reported in the dbEST database and applying the codon frequencies observed in *Escherichia coli*. Expression in *E. coli* and purification of recombinant murine Ngb were carried out as described previously (Moschetti *et al.*,

2009). Point mutations were introduced by means of the QuikChange site-directed mutagenesis kit (Stratagene).

2.1. Kinetic measurements

Absorption spectra were collected using an HP8453 spectrophotometer (path length 1 cm). Stopped-flow experiments were carried out using a thermostated instrument (DX.17MV, Applied Photophysics, Leatherhead, England) with a 1 cm path length. Assays were performed at 293 K in 50 m*M* phosphate pH 7.0, 20 µ*M* EDTA, 260 U ml^{-1} catalase. Anaerobic conditions were obtained by extensive N_2 equilibration and the addition of a minute amount of sodium dithionite. Stock solutions of CO were prepared by equilibrating degassed buffer at 293 K with the pure gas at 101.3 kPa ([CO] = 1 m*M*). Typically, 10 µ*M* deoxy ferrous Ngb (wild type or mutated) was anaerobically mixed with solutions containing CO at a concentration varying from 1 m*M* down to 12.5 µ*M*, and the binding reaction was followed at 429 nm. Observed rate constants and amplitudes (overall absorption changes) were obtained by fitting the kinetic traces to single exponentials. For each protein sample, the half-saturating concentration of CO (c_{50}) was obtained by plotting the fraction of CO-bound protein (Y) as a function of the concentration of free unbound CO ($[CO]_{free}$) and fitting the data to the equation $Y = [CO]_{free}/(c_{50} + [CO]_{free})$. Data were analyzed using the *MATLAB* software (MathWorks, Natick, Massachusetts, USA).

2.2. Crystallization and crystal handling

Ferric Ngb (wild type and mutated) was crystallized by the hanging-drop vapour-diffusion method using a 1:1 mixture of protein (10 mg ml^{-1}) and reservoir solution [1.6 *M* ammonium sulfate, 0.1 *M* MES pH 6.5, 10%(*v*/*v*) dioxane] at 293 K. Crystals of Ngb–CO mutants were obtained by soaking the crystals into a CO-saturated solution after protein reduction by sodium dithionite as described previously (Vallone, Nienhaus, Matthes *et al.*, 2004).

Cryoprotection was achieved by dipping the crystals into mother liquor containing 25% glycerol. X-ray diffraction data were all collected at 100 K at either the ESRF (beamline ID14.1), Grenoble, France or BESSY (beamline ID14.1), Berlin, Germany.

2.3. Structural data collection and analysis

All of the crystals were isomorphous with wild-type Ngb crystals. Data were indexed and processed using the programs *MOSFLM* and *SCALA* (Leslie, 2006) and *XDS* (Kabsch, 2010) in space group *H*32. Initial phases were derived from the deposited structures of wild-type metNgb and Ngb–CO (PDB entries 1q1f and 1w92; Vallone, Nienhaus, Brunori *et al.*, 2004; Vallone, Nienhaus, Matthes *et al.*, 2004). The models were subjected to iterative rounds of refinement and model building. Refinement was carried out with *PHENIX* (Adams *et al.*, 2010) followed by model adjustment and water addition with *Coot* (Emsley & Cowtan, 2004). The quality of the final model was assessed using *PROCHECK* (Laskowski *et al.*, 1993). Solvent molecules were added to the $F_o - F_c$ electron-density map contoured at 3σ with the solvent-picking procedure of *PHENIX*. Figures were prepared using *PyMOL* (v.1.5.0.4; Schrödinger). Relevant crystallographic statistics and geometric details of the refined models were extracted from the programs used in the analysis and are reported in

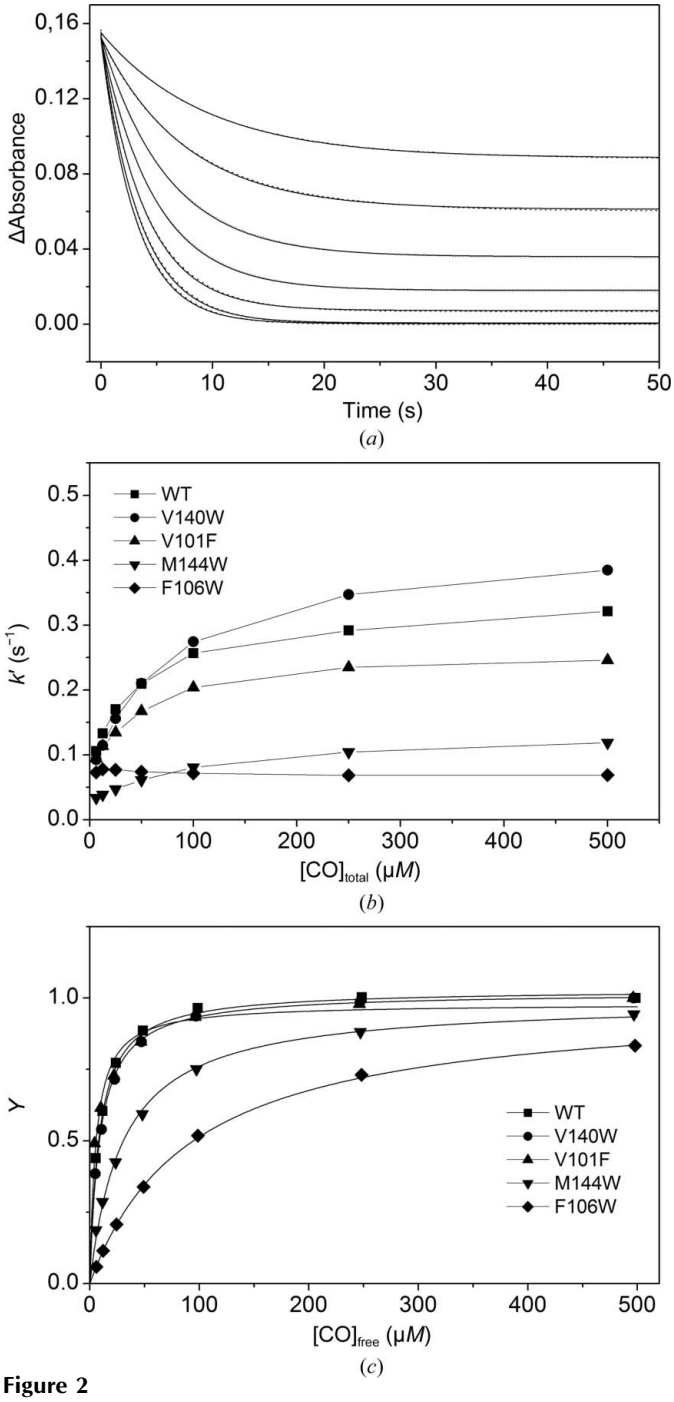

Figure 2
(*a*) Absorption changes at 429 nm after anaerobically mixing 10 µ*M* deoxy ferrous wild-type Ngb with CO. T = 293 K. CO concentrations after mixing (from top to bottom) are 6.25, 12.5, 25, 50, 100, 250 and 500 µ*M*. Traces are shown together with their best fit to a single exponential decay. (*b*) Rate constant measured for CO binding to wild-type and mutated neuroglobins, plotted as a function of the total CO concentration. (*c*) Fraction of Ngb–CO as a function of the concentration of free CO.

Table 2
CO binding to wild-type and mutated deoxy ferrous Ngb.

The rate constant was measured at 500 μ*M* CO, together with the c_{50} values determined for each protein sample. T = 293 K.

Ngb sample	k' (s^{-1}) at 500 μ*M* CO	c_{50} (μ*M*)
Wild type	0.32	8.5 ± 0.3
V140W	0.38	9.1 ± 0.4
V101F	0.25	5.6 ± 0.8
M144W	0.12	30.0 ± 1.7
F106W	0.07	92.7 ± 2.2

Table 1. Well defined electron densities were detected at the mutated residue sites. The electron-density maps allowed a description of the whole molecule. The final models did not contain residues 1 and 2 at the N-terminus and residues 149, 150 and 151 at the C-terminus. The N- and C-terminal regions, as well as the CD, EF and FG loops, proved to be more mobile than the rest of the structure, showing higher *B*-factor values.

3. Results and discussion

3.1. CO affinity and binding kinetics

The effect of V140W, V101F, M144W and F106W mutations on the binding kinetics and on the overall affinity of Ngb for CO was investigated by stopped-flow absorption spectroscopy. As shown in Fig. 2(*a*), after mixing deoxy ferrous wild-type Ngb with CO at concentrations between 12.5 μ*M* and 1 m*M* the formation of the Ngb–CO adduct was monitored as a decrease in absorption at 429 nm. Within experimental error, the reaction followed a single-exponential time course and the rate constant had a modest dependence on the CO concentration, tending to a plateau at ∼0.3 s^{-1} for the wild-type protein (Fig. 2*b*, square symbols). Such kinetic behaviour is expected since, owing to endogenous hexacoordination, CO binding to Ngb is rate-limited by the slow dissociation of His64(E7) from the haem iron. The affinity of the protein for CO was then evaluated from the fraction of Ngb–CO formed at each CO concentration, calculated from the amplitude of the CO-binding reaction. As shown in Fig. 2(*c*) (square symbols), CO binds to wild-type ferrous Ngb with a c_{50} of ∼8.5 μ*M*.

Kinetic analysis was then extended to the aforementioned protein variants. Based on the data reported in Fig. 2 and Table 2, in the V140W and V101F mutants both the binding kinetics and the overall affinity for CO were not significantly affected compared with wild-type Ngb. Indeed, while binding CO with slightly different rate constants (possibly reflecting slight variations in the rate of His64 dissociation), both mutants were characterized by a similar CO affinity ($c_{50} \simeq 9.1$ and 5.6 μ*M* for V140W and V101F, respectively). On the other hand, mutants containing substitutions at positions 144 (M144W) and 106 (F106W) bound CO not only with lower rates (Fig. 2*b*) but also with significantly decreased affinities ($c_{50} \simeq 30$ and 92.7 μ*M* for M144W and F106W, respectively; see Fig. 2*c*) compared with wild-type Ngb. In each of the latter mutants, both the rate-limiting value for the binding rate constant (as estimated at 500 μ*M* CO) and the overall affinity decreased by similar factors (3–4-fold for M144W and 5–11-fold for F106W; see Table 2). Based on these results, we deduce that dissociation of haem iron from the distal histidine, the rate-limiting step for CO binding to iron(II) Ngb, is the main (if not the only) step that is affected upon the introduction of bulky side chains into the the protein cavity. These results are therefore fully consistent with a mechanism whereby hampering the haem sliding into the cavity results in a slower rupture of the haem Fe–distal histidine bond and thus in a slower formation of the pentacoordinate species competent for ligand binding.

The kinetic data described above show that Phe106 plays a key role in ligand-binding affinity modulation, in agreement with the finding that this residue undergoes a notable conformational change upon haem sliding, as observed by comparing the structure of the wild-type protein in the ferric and CO-bound states (Figs. 1*a* and 1*b*). Based on this observation, substitution of Phe106 with a bulkier residue such as Trp is expected to hamper the haem sliding inside the cavity, resulting in a decreased affinity for CO. Consistently, a Trp at position 144 should also constitute an obstacle to repositioning of the haem in the cavity. In summary, these kinetic data indicate that the haem displacement observed in the CO–Ngb crystal structure (PDB entry 1w92) upon ligand binding is likely to also take place in solution; hampering the repositioning of the haem by increasing the size of residues at positions 106 and 144 would have no effect on the affinity if

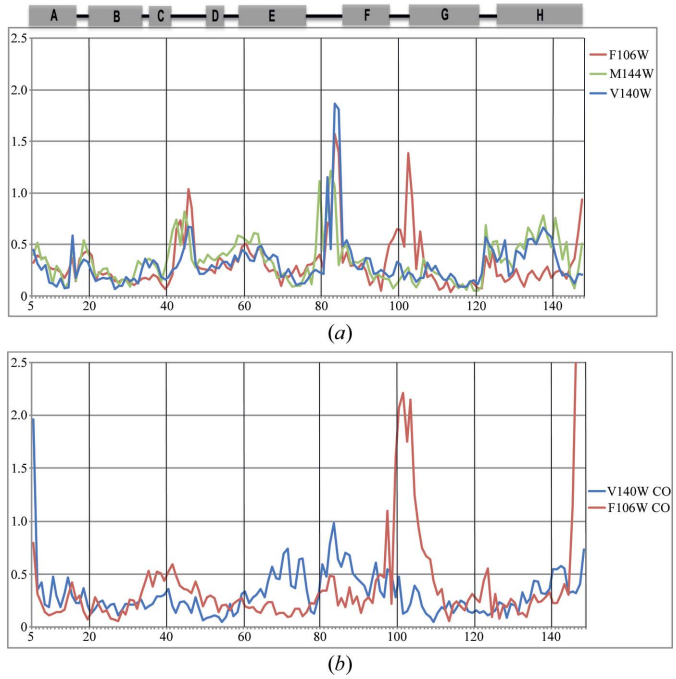

Figure 3
Plots of r.m.s.d. *versus* amino-acid sequence. (*a*) R.m.s.d. between the C^{α} atoms of wild-type metNgb and mutants in the met form (V140W Ngb, F106W Ngb and M144W Ngb). (*b*) R.m.s.d. between C^{α} atoms of wild-type Ngb–CO and mutants in the CO-bound form (V140W Ngb and F106W Ngb). Bars and capital letters indicate the helical segments in murine Ngb.

ligand binding was controlled primarily by swinging out of the distal His(E7). Therefore, haem sliding appears not to be unique to the crystal state but may be a functional feature of Ngb.

3.2. Crystal structures of Ngb mutants

We have determined the structures of the F106W and M144W mutants, in which a major effect on CO binding was observed, and of the V140W mutant, in which ligand binding was almost unperturbed.

Analysis of the structures in the deoxy hexacoordinate state showed that the introduction of additional bulk into the cavity has no major effects on the unliganded state. For the mutants, the average r.m.s.d. (root-mean-square deviation) calculated on the main chain with respect to wild-type Ngb ranges between 0.51 and 0.41 Å; therefore, it is only slightly higher than that observed for two independent determinations of the structure of deoxy wild-type Ngb (0.35 Å; PDB entries 2vry and 1q1f). On the other hand, comparison of the r.m.s.d. between the structures of wild-type Ngb and its mutants in both the met and CO-bound states (Figs. 3*a* and 3*b*) suggests that decreasing the volume of the internal Ngb cavity induces local effects at the mutated sites and some loop rearrangement, whereas helices tend to behave as rigid bodies with only slight displacements, as described below for each mutant (Figs. 4*a*–4*e*).

In the V140W(H15) mutant in the deoxy hexacoordinate state, the presence of a tryptophan induces a modest displacement of helix H, where the mutation is located (0.70 Å for the C^{α} of Trp140), that extends to the GH loop and a slight displacement in the CD loop (Fig. 3*a*). This region has already been shown to be very sensitive to changes in the internal cavity (Moschetti *et al.*, 2009). A larger displacement is observed for the EF loop and, to a lesser extent, for the first two turns of the F helix (Fig. 3*a*), likely owing to the bulky Trp at position 140 that pushes the side chain of Leu89 away by about 0.5 Å (Figs. 4*a* and 4*b*).

The structure of the CO-bound mutant V140W indicates that the protein can also accommodate the bulky Trp side chain at position 140 when the haem is 'slid' inside the cavity, consistent with the lack of effect on the kinetics and on the affinity. In more detail, in CO–Ngb V140W we observed a flipping of Trp140 owing to the proximity of the haem that slides upon ligand binding (from $\chi_1 = -173^\circ$ to $\chi_1 = -62^\circ$; both angles correspond to stable conformers; Fig. 4*c*). Upon flipping, Trp140 exerts pressure on Leu82 in the EF loop (Fig. 3*b*), which undergoes a displacement of about 1 Å, and extends to the E and F helices, the region that is most affected by CO binding in Ngb V140W.

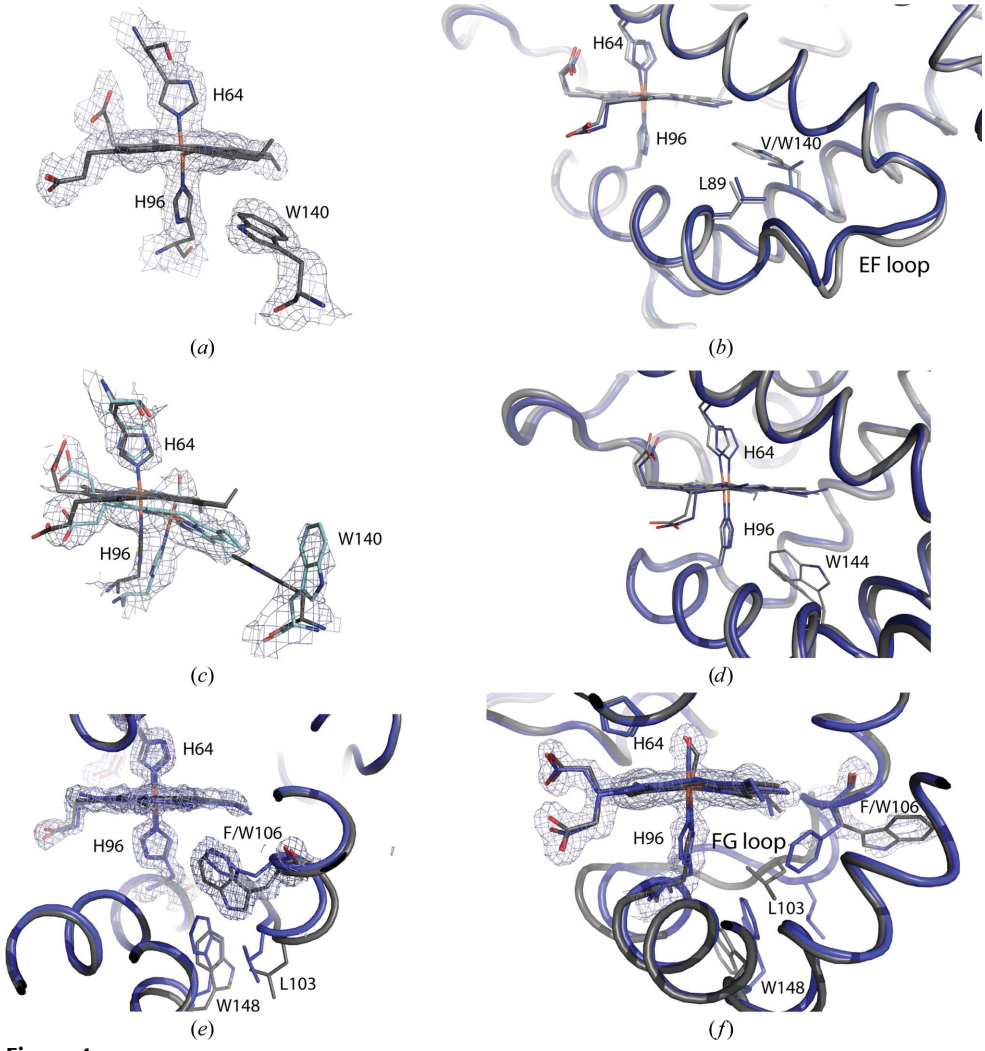

Figure 4
Crystal structure of the V140W and F106W Ngb mutants in met and CO-bound forms. The haem group, the mutated side chains and the proximal and distal histidines are shown, as well as the electron density around them contoured at 1σ. (*a*) For the V140W mutant the residue Trp140, the distal and proximal histidines and the haem group are shown. (*b*) Superposition of wild-type metNgb (dark blue) and met V140W Ngb (grey): Trp140 causes a small displacement of Leu89, and a conformational change of the EF loop and of the first turn of helix F occurs. (*c*) Superposition of met (grey) and CO-bound V140W (cyan) Ngb structures. The conformations adopted by Trp140 in met V140W Ngb and in V140W Ngb–CO are shown. (*d*) Superposition of metNgb (dark blue) and met M144W Ngb (grey). The double conformation of Trp144 is shown. (*b*) Superposition of metNgb (dark blue) and met F106W Ngb (grey). Trp106 causes a displacement of Leu103 and Trp148. (*f*) Superposition of Ngb–CO (dark blue) and F106W Ngb–CO (grey). The main conformations adopted by Trp106, Leu103 and Trp148 upon haem sliding are shown.

In the met hexacoordinate state of the F106W(G3) mutant, the Cα at position 106 is displaced by 0.62 Å: the new side chain induces a double conformation of Trp148 and a displacement of the last visible residues at the C-terminus of the H helix: Asp149 and Gly150. Similarly to the other mutants, the CD, EF and FG loops are affected (Fig. 3*a*). In this mutant, the residue with the largest displacement is Leu103, a residue in the FG loop, which is pushed towards the bulk by the side chain of Trp106 (Fig. 4*e*).

In the F106W mutant, the Trp inserted by mutagenesis is found in a part of the proximal side of the internal pocket that the haem occupies once it slides into the cavity upon rupture of the sixth coordination bond of the iron. In the wild type, this hindrance induces a 3.8 Å displacement of Phe106 (Fig. 1). In the CO-bound structure of the F106W mutant the bulk introduced by a tryptophan upon haem sliding is mainly resolved by a 180° flip, exposing Trp106 to the solvent, with a minor conformer (about 30%; not shown) still in the internal cavity (Fig. 4*f*). This event affects the conformation of the F–FG region, which is displaced and becomes more flexible, adopting a double conformation (not shown). The destabilization and displacement of the FG loop are owing to the collapse of Leu103 in the void created by the movement of Trp106 into the solvent. The rearrangement owing to the departure of Trp106 into the solvent does not extend beyond the proximal His96(F8), since it acts as a constraint owing to the bond between the haem iron and His(F8)Nε. Because of this, only the last two turns of helix F take part into the conformational change triggered by Trp106, and the proximal histidine itself is displaced by about 1 Å (Cα). The new position of Leu103 also induces exposure to the solvent of Trp148, the last residue of the C-terminus that is visible in the electron-density map (Fig. 4*f*). Finally, in the F106W mutant the internal space provided by its repositioning onto the protein surface allows the haem to slide 0.57 Å more deeply into the protein.

In the M144W(H19) mutant the side chain of Trp144 in helix H is found in two conformations with occupancies of 40 and 60%. The position of the Cα atom is changed slightly, with a 0.45 Å displacement with respect to the wild type in both conformers (Fig. 4*d*). This movement extends to the H helix and the GH loop (Fig. 3*a*). Conformer *A* (40% occupancy) is rotated about 120° with respect to the position of conformer *B* and occupies the same position as the methionine found in wild-type Ngb. The effect of this mutation is felt at the level of helix E, where a key residue in haem coordination, the distal histidine His64(E7), is displaced by about 0.66 Å (C^{α}; Fig. 4*d*) owing to lifting of the haem by Trp144 at the proximal side that affects the distal side of the protein. As already observed for the V140W mutant, the CD loop is particularly sensitive to changes in the internal cavity and the EF loop also changes its conformation (Fig. 3*a*).

It was not possible to determine the CO-bound structure of the M144W mutant since crystals grown in the met form and then reduced and soaked with CO dissolved upon treatment, and the protein in the CO-bound state did not yield crystals.

The *in silico* substitution of Met144 with a Trp in the CO-bound state of Ngb leads to a steric clash with the haem that cannot be avoided by adopting different rotamers (see Supplementary Fig. S1[1]). We can not exclude that in the latter mutant binding of CO is achieved by His64(E7) gating, as proposed by molecular dynamics (Anselmi *et al.*, 2008); alternatively, a repositioning upon haem sliding of the Trp at position 144 could take place owing to main-chain rearrangements that might extend to other regions of the protein; in both cases, crystal disruption upon soaking with CO may be understood.

4. Concluding remarks

Dissociation of the distal His(E7) is a crucial event for determining the ligand-binding rate and affinity in His-hexacoordinate globins, and it implies substantial instability of this coordination bond to the haem iron with respect to other hexacoordinate proteins such as cytochrome *c*. Fast transient spectroscopy of external ligand (L) binding to the ferrous haem iron of Ngb has shown (Trent & Hargrove, 2002) that the kinetic mechanism involves three states and two steps as indicated below, where Ngb_P is the pentacoordinate species, Ngb_H is the hexacoordinate species and Ngb·L represents the species bound to an external ligand,

$$Ngb_H \underset{k_H}{\overset{k_{-H}}{\longleftrightarrow}} Ngb_P \underset{k_L}{\overset{k_{-L}}{\longleftrightarrow}} Ngb{\cdot}L. \qquad (1)$$

Such a scheme implies that the state competent for binding an external physiologically relevant ligand (L = CO, NO or O_2) is the pentacoordinate intermediate with a vacant distal site (Ngb_P). This state is populated starting from Ngb_H by dissociation of the distal His(E7), and its equilibrium concentration is generally very low. The CO combination rate for such a scheme predicts that formation of the Ngb–CO complex will be rate-limited at infinite [L] by the dissociation of the distal His(E7) to make the sixth coordination position available. The equilibrium constant for the first step, K_H, varies widely, from 0.45 for rice Hb2 to up to ∼1000 for Ngb, in favour of the His-hexacoordinate species. This variability, which may reflect the different functions of these globins, is mirrored by different structural mechanisms to control the stability of the bond between the distal histidine and the haem iron (Hoy *et al.*, 2007; Halder *et al.*, 2007; Vallone, Nienhaus, Matthes *et al.*, 2004).

Within the family of His-hexacoordinate globins, the case of Ngb is peculiar since the binding of CO is associated with sliding of the haem into a pre-existing tunnel within the protein matrix rather than with displacement of the distal His(E7) alone. In the absence of the constraint imposed by the bond to distal His(E7), the most stable position of the haem within the protein is the slid conformation when bound to an exogenous ligand. Variations in the size and/or shape of the cavity in which the haem is hosted upon sliding are therefore

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: WA5068).

expected to control the equilibrium between the hexacoordinate and pentacoordinate forms, and in turn the rate and the affinity towards exogenous ligands.

The aim of this work was to experimentally test how the internal cavity modulates the ligand-binding rate and affinity by producing and characterizing four single mutants, namely V140W, V101F, M144W and F106W. We have analysed the effect of these mutations on the CO-binding kinetics and have determined their structure both in the ferric His-hexacoordinate state (V140W, M144W and F106W) and in the CO-bound ferrous form (V140W and F106W). None of the mutations appreciably modified protein expression in *E. coli* nor the overall stability during purification or in crystallization experiments (apart from M144W, which was difficult to crystallize reproducibly). The mutants therefore did not display major perturbations in the overall structure, folding or haem incorporation during protein biosynthesis. The results are clear-cut and allow some simple general conclusions.

In the V101F and V140W mutants both the binding kinetics and the overall apparent affinity for CO proved to be essentially unchanged compared with the wild-type protein; on the other hand, both the M144W and F106W mutants were characterized by slower binding kinetics and a significantly lower affinity for CO. The asymptotic value for the combination rate extrapolated to 'infinite' CO concentration (Fig. 2) is $\sim$0.32 s^{-1} for wild-type Ngb and decreases to $\sim$0.07 s^{-1} in the M144W mutant. This is consistent with the observation that Phe106 undergoes a substantial conformational change upon haem sliding, as shown by comparing the structure of wild-type Ngb in the ferric and CO-bound states (Figs. 1*a* and 1*b*). Therefore, substitution of Phe106 with the bulkier side chain of Trp hampers the haem sliding inside the cavity, decreasing the CO-binding rate and affinity. We have indeed observed that in the CO-bound state of Ngb F106W the Trp side chain flips out of the cavity and is exposed to the solvent. The mutation M144W is also likely to interfere with the haem sliding, as indicated by modelling this mutation in the CO-bound structure of wild-type Ngb (Supplementary Fig. S1). Owing to crystal cracking upon CO binding, we cannot conclude whether the steric conflict is resolved by rearrangements of the Ngb proximal region of the cavity or whether in this mutant a distal histidine-gating movement takes place to allow occupation of the sixth coordination position.

In conclusion, decreasing the space available in the large Ngb cavity has an effect on ligand rate and affinity when the introduced side chain fills the space occupied by the haem upon sliding (F106W and M144W mutants); this result is fully consistent with expectations based on the crystallographic work on Ngb–CO (Vallone, Nienhaus, Matthes *et al.*, 2004). On the other hand, if the space available in the tunnel is decreased by introducing a bulky side chain in positions that do not interfere with haem sliding (V101F and V140W mutants), the rate and affinity are not significantly affected. The fact that decreasing the volume of the proximal side of the cavity has a negligible effect on affinity (if haem sliding is not affected), in conjunction with the finding that it does not contain any Xe docking site (Moschetti *et al.*, 2009), might indicate that it does not constitute a preferential pathway to and from the haem, in contrast to Mb (Scott & Gibson, 1997).

All in all, the behaviour of the F106W and M144W mutants shows that haem sliding occurs not only in the crystal state of Ngb but also in solution, and that it contributes to modulating the affinity of the protein for exogenous physiologically relevant ligands. Some biologically intriguing questions related to the demonstration of the control mechanism involved in haem sliding in Ngb may be in order, including the following. Are all of the presumed scavenging (*e.g.* of NO) and signalling functions of Ngb maintained in hindered variants such as the F106W and M144W mutants? Is the neuroprotective role of Ngb maintained even in those hindered mutants that reveal a crippled internal tunnel? Given their stability and expression, we believe that these questions may be tackled by suitable *in vivo* experiments.

This work was partially supported by Ministero dell' Istruzione, dell'Università e della Ricerca of Italy (2007B57EAB_004, 20074TJ3ZB_005, RBRN07BMCT_007 and PNR-CNR Aging Program 2012–2014). We acknowledge the Helmholtz-Zentrum Berlin electron-storage ring BESSY II (beamline ID14.1) and the European Synchrotron Radiation Facility (ESRF, Grenoble) for provision of synchrotron radiation (beamline ID14.1). The research leading to these results received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (grant agreement No. 283570) and European Community-Research Infrastructure Action under Contract RII3-CT-2004-506008 from the FP6 'Structuring the European Research Area Programme'.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Anselmi, M., Brunori, M., Vallone, B. & Di Nola, A. (2007). *Biophys. J.* **93**, 434–441.
Anselmi, M., Brunori, M., Vallone, B. & Di Nola, A. (2008). *Biophys. J.* **95**, 4157–4162.
Antonini, E. & Brunori, M. (1971). *Hemoglobin and Myoglobin in their Reactions with Ligands*. Amsterdam: North-Holland.
Bashford, D., Chothia, C. & Lesk, A. M. (1987). *J. Mol. Biol.* **196**, 199–216.
Burmester, T. & Hankeln, T. (2009). *J. Exp. Biol.* **212**, 1423–1428.
Burmester, T., Weich, B., Reinhardt, S. & Hankeln, T. (2000). *Nature (London)*, **407**, 520–523.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Halder, P., Trent, J. T. III & Hargrove, M. S. (2007). *Proteins*, **66**, 172–182.
Hardison, R. C. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 5675–5679.
Hardison, R. C. (1998). *J. Exp. Biol.* **201**, 1099–1117.
Holm, L. & Sander, C. (1993). *FEBS Lett.* **315**, 301–306.
Hoy, J. A, Robinson, H., Trent, J. T. III, Kakar, S., Smagghe, B. J. & Hargrove, M. S. (2007). *J. Mol. Biol.* **371**, 168–179.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Lutz, S., Nienhaus, K., Nienhaus, G. U. & Meuwly, M. (2009). *J. Phys. Chem. B*, **113**, 15334–15343.

Makino, M., Sawai, H., Shiro, Y. & Sugimoto, H. (2011). *Proteins*, **79**, 1143–1153.
Moschetti, T., Mueller, U., Schulze, J., Brunori, M. & Vallone, B. (2009). *Biophys. J.* **97**, 1700–1708.
Nienhaus, K., Lutz, S., Meuwly, M. & Nienhaus, G. U. (2013). *Chemistry*, **19**, 3558–3562.
Perutz, M. F. (1979). *Annu. Rev. Biochem.* **48**, 327–386.
Savino, C., Miele, A. E., Draghi, F., Johnson, K. A., Sciara, G., Brunori, M. & Vallone, B. (2009). *Biopolymers*, **91**, 1097–1107.
Scott, E. E. & Gibson, Q. H. (1997). *Biochemistry*, **36**, 11909–11917.
Sun, Y., Jin, K., Mao, X. O., Zhu, Y. & Greenberg, D. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 15306–15311.
Sun, Y., Jin, K., Peel, A., Mao, X. O., Xie, L. & Greenberg, D. A. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 3497–3500.
Tilton, R. F. Jr, Kuntz, I. D. Jr & Petsko, G. A. (1984). *Biochemistry*, **23**, 2849–2857.
Trent, J. T. III & Hargrove, M. S. J. (2002). *J. Biol. Chem.* **277**, 19538–19545.
Trent, J. T. III, Kundu, S., Hoy, J. A. & Hargrove, M. S. (2004). *J. Mol. Biol.* **34**, 1097–1108.
Vallone, B., Nienhaus, K., Brunori, M. & Nienhaus, G. U. (2004). *Proteins*, **56**, 85–92.
Vallone, B., Nienhaus, K., Matthes, A., Brunori, M. & Nienhaus, G. U. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 17351–17356.
Vinogradov, S. N. & Moens, L. (2008). *J. Biol. Chem.* **283**, 8773–8777.