

Beatriz G. Guimarães,<sup>a</sup> Djemel Hamdane,<sup>b,c</sup> Christophe Lechauve,<sup>d</sup> Michael C. Marden<sup>d\*</sup> and Béatrice Golinelli-Pimpaneau<sup>b,c\*</sup>

<sup>a</sup>Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, 91190 Gif-sur-Yvette, France, <sup>b</sup>Laboratoire d'Enzymologie et Biochimie Structurales, Centre de Recherche de Gif, CNRS, 1 Avenue de la Terrasse, 91190 Gif-sur-Yvette, France, <sup>c</sup>Laboratoire de Chimie des Processus Biologiques, CNRS, Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France, and <sup>d</sup>Inserm U779, Université Paris XI, 78 Rue du Général Leclerc, 94275 Le Kremlin-Bicêtre, France

Correspondence e-mail:  
michael.marden@inserm.fr,  
beatrice.golinelli@college-de-france.fr

# The crystal structure of wild-type human brain neuroglobin reveals flexibility of the disulfide bond that regulates oxygen affinity

Neuroglobin plays an important function in the supply of oxygen in nervous tissues. In human neuroglobin, a cysteine at position 46 in the loop connecting the C and D helices of the globin fold is presumed to form an intramolecular disulfide bond with Cys55. Rupture of this disulfide bridge stabilizes bi-histidyl haem hexacoordination, causing an overall decrease in the affinity for oxygen. Here, the first X-ray structure of wild-type human neuroglobin is reported at 1.74 Å resolution. This structure provides a direct observation of two distinct conformations of the CD region containing the intramolecular disulfide link and highlights internal cavities that could be involved in ligand migration and/or are necessary to enable the conformational transition between the low and high oxygen-affinity states following S–S bond formation.

Received 11 October 2013

Accepted 2 January 2014

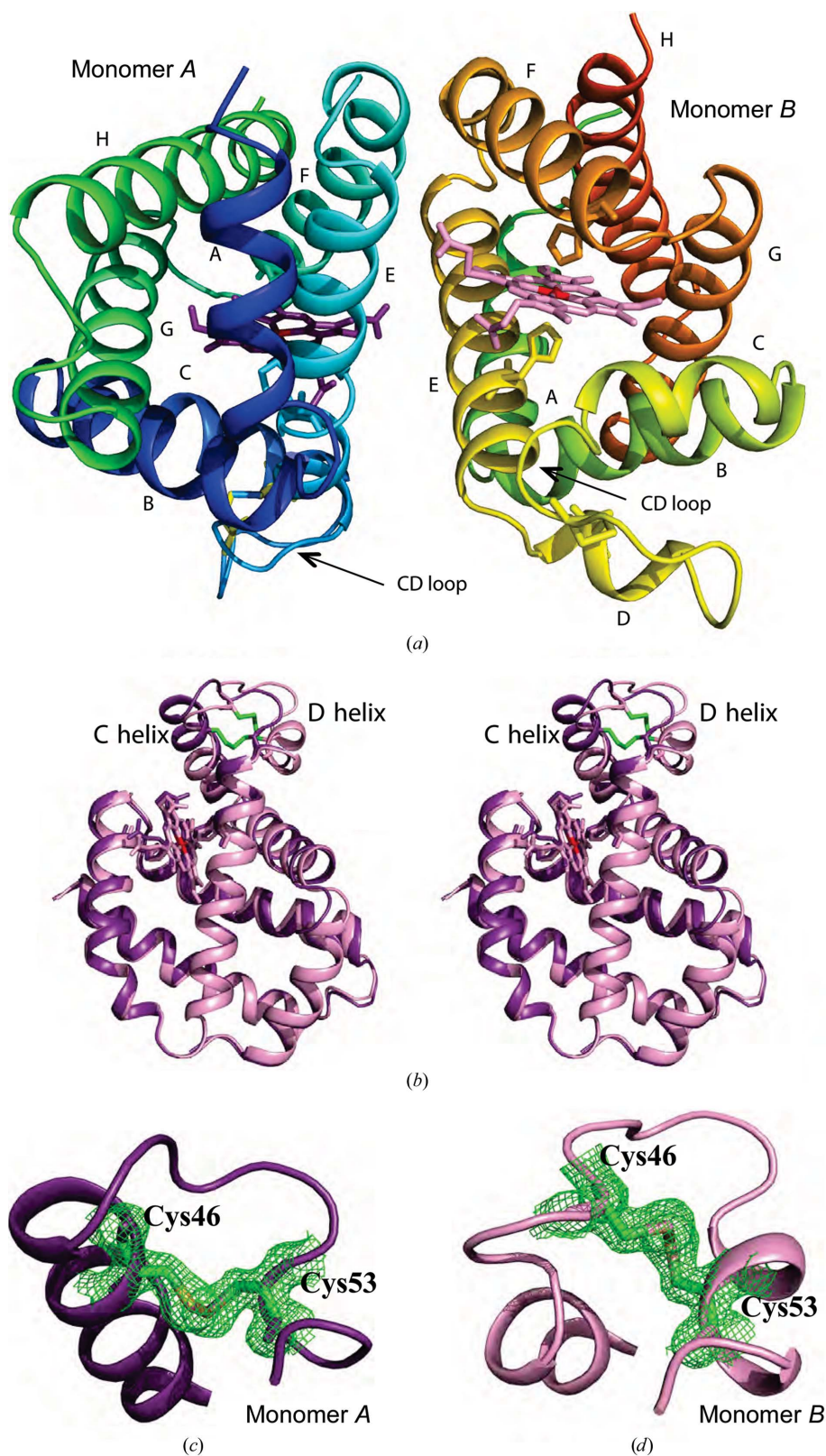
**PDB reference:** neuroglobin,  
4mpm

## 1. Introduction

Neuroglobin (Ngb) is a monomeric 16.9 kDa cytosolic protein that is common to all vertebrates and is expressed in nerve tissues, mainly in the brain and the retina (Burmester *et al.*, 2000). Ngb is up-regulated in the absence of adequate oxygen supply (hypoxic conditions) and has been proposed to act as an endogenous neuroprotective factor in cerebral ischaemia, a disease in which restriction of blood supply results in tissue damage (Sun *et al.*, 2001, 2003; Khan *et al.*, 2008). In addition, Ngb is a nitrite oxidase that modulates intracellular nitric oxide levels (Tiso *et al.*, 2011). However, NO and CO binding are common to haem proteins, and the physiological role of Ngb is not yet completely understood (Burmester & Hankeln, 2009).

The haem-containing Ngb belongs to the vertebrate globin family, although it shares little sequence identity (<25%) with myoglobin, haemoglobin and cytoglobin (Cygb) (Burmester *et al.*, 2000). Ngb adopts the typical globin fold consisting of eight  $\alpha$ -helices denoted A–H (Pesce *et al.*, 2003; Vallone, Nienhaus, Matthes *et al.*, 2004; Supplementary Fig. S1<sup>1</sup>). In contrast to haemoglobin and myoglobin, in which the haem iron is ligated only to the proximal histidine (HisF8), which leaves the distal position accessible for the coordination of O<sub>2</sub>, the haem iron of Ngb is hexacoordinated in both its ferrous and ferric forms, with the distal His64 (E7) being the internal ligand (Dewilde *et al.*, 2001). His64 competes in ligand binding with any potential exogenous diatomic ligand, such as O<sub>2</sub>, CO or NO in the ferrous form (Fe<sup>2+</sup>) of the protein and CN or NO in the ferric form. Hexacoordination has been proposed as a novel

<sup>1</sup> Supporting information has been deposited in the IUCr electronic archive (Reference: WA5062).



**Figure 1**  
 Overview of the wild-type hNgb structure. (a) Overall view of the two molecules in the asymmetric unit. The haems in molecules *A* and *B* are coloured violet and pink, respectively. Helices are labelled from *A* to *H* and are distinguished by rainbow colours. (b) Stereoview of the superposition of the two monomers of wild-type hNgb. Regions 3–35 and 60–149 were superimposed. Molecules *A* and *B* are coloured violet and pink, respectively. The disulfide bonds are shown in green. The main differences occur in the dynamic loop region. (c, d) Detailed view of the S–S bonds in both monomers with OMIT maps contoured at  $1\sigma$ .

mechanism to regulate ligand affinity in haem proteins because the rate-limiting step in exogenous ligand binding is the cleavage of the covalent bond between the haem iron and the distal histidine. Moreover, in Ngb and other hexacoordinated haemoglobins, the competitive internal ligand leads to low temperature dependence of the external ligand affinity, since the difference in equilibrium energy for the two ligands is much lower than that of ligand binding to pentacoordinated haemoglobin (Uzan *et al.*, 2004).

The crystal structures of Ngb from mouse (Vallone, Nienhaus, Brunori *et al.*, 2004; Vallone, Nienhaus, Matthes *et al.*, 2004) and of a mutant from human brain (Pesce *et al.*, 2003) have previously been determined. Comparison of the structures of unliganded ferric and CO-bound ferrous Ngb from mouse showed haem sliding and a topological reorganization of the internal cavities upon CO binding, with almost no rearrangement of the distal histidine (Vallone, Nienhaus, Brunori *et al.*, 2004; Vallone, Nienhaus, Matthes *et al.*, 2004). Therefore, it was suggested that the binding kinetics of the exogenous ligand were controlled by the haem displacement. Theoretical studies also proposed that breakage of the bond between His64 (E7) and the iron occurs through haem sliding instead of rotation of the distal histidine in the human protein (Bocahut *et al.*, 2009).

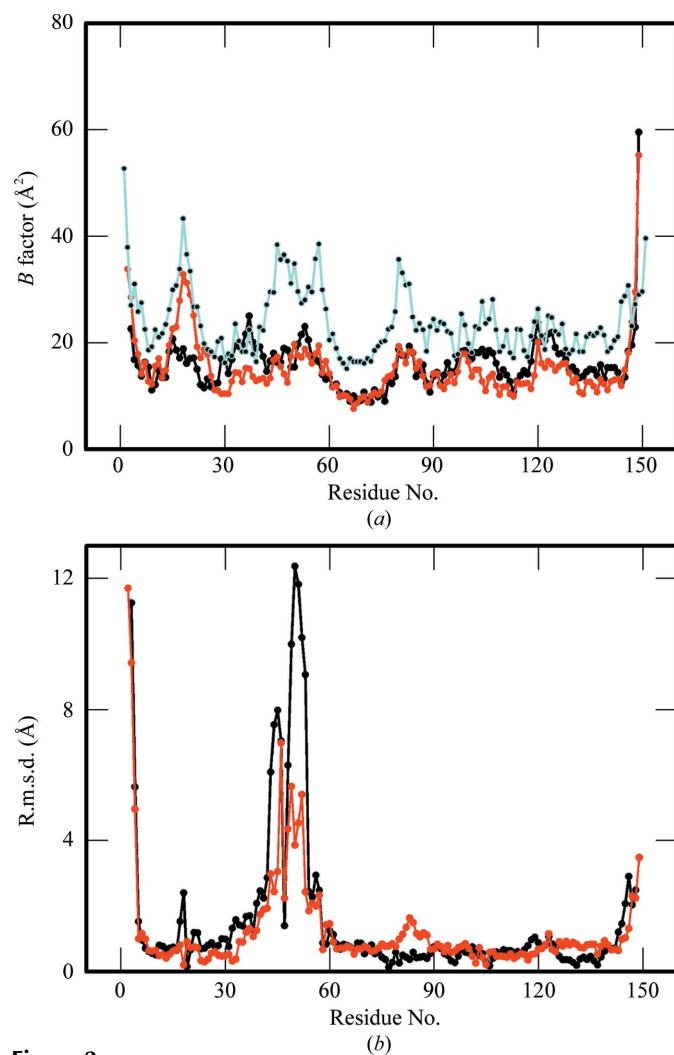
In human Ngb (hNgb), the formation of a disulfide bridge between Cys46 (CD7) and Cys55 (D5) located in the CD loop/D helix is known to affect the haem environment, decrease the His64 (E7) binding affinity for the ferrous protein (Hamdane *et al.*, 2003, 2004) and increase nitrite reductase activity (Tiso *et al.*, 2011). Mutations of the two cysteines had a major impact on the ligand-binding kinetics, thermodynamics and stability of hNgb (Bocahut *et al.*, 2013). This suggested a novel mechanism for oxygen binding, in which reduction of the disulfide bond would promote oxygen release (Hamdane *et al.*, 2003).

The structure of hNgb in the unliganded ferric form has previously been determined (Pesce *et al.*, 2003).

However, because oxidized hNgb always revealed a variable mixture of several forms (Dewilde *et al.*, 2001; Kriegl *et al.*, 2002), Cys46, Cys55 and Cys120 were mutated to Gly, Ser and Ser, respectively, to hinder the formation of disulfide bonds and achieve crystallization. In the mutant structure, Gly46 and Ser55 were positioned 18–18.7 Å apart, and thus could not provide any evidence about the structural importance of the intramolecular disulfide bond. To shed light on the impact of the internal disulfide bond on the conformational state of hNgb and on the mechanism that modulates ligand binding, we have solved the crystal structure of wild-type hNgb.

## 2. Experimental procedures

hNgb was purified and crystallized as described by Hamdane *et al.* (2009). The previous data set (Hamdane *et al.*, 2009),



**Figure 2**  
Plots of  $B$  factors and r.m.s.d.s versus primary-sequence position. (a)  $B$  factors. The  $B$  factors of molecule  $A$  of wild-type hNgb, molecule  $B$  of wild-type hNgb and molecule  $B$  of the mutant are shown in red, black and cyan, respectively. (b) R.m.s.d.s between  $C^\alpha$  atoms of pairs of structures. The superposition of molecule  $A$  of hNgb and molecule  $B$  of the mutant is in red and that of molecule  $B$  of hNgb and molecule  $B$  of the mutant in black.

**Table 1**

Data-collection and refinement statistics for wild-type human neuroglobin crystals.

Values in parentheses are for the outer resolution shell.

Data collection	
Space group	$P2_12_12$
Unit-cell parameters (Å)	$a = 139.15, b = 48.49, c = 39.28$
Resolution	30.0–1.74 (1.85–1.74)
$\langle I/\sigma(I) \rangle$	10.4 (2.4)
$R_{\text{meas}}$ (%)	15.3 (81.5)
Completeness (%)	99.2 (95.5)
Multiplicity	6.9 (6.4)
Refinement	
$R_{\text{work}}$	0.185
$R_{\text{free}}$	0.225
R.m.s.d., bond lengths (Å)	0.009
R.m.s.d., angles (°)	0.92
No. of atoms	
Protein	2330
Water	277
Ligand	363
Ramachandran plot ( <i>MolProbity</i> )	
Favoured (%)	98.3
Outliers (%)	0

re-treated with a later version of *XDS* (released on 26 September 2012; Kabsch, 2010), led to an extended unit cell with a doubled volume in space group  $P2_12_12$  (Table 1). The two monomers in the asymmetric unit were positioned by molecular replacement with *MOLREP* (Vagin & Teplyakov, 2010) using the atomic coordinates of the hNgb mutant (PDB entry 1oj6; Pesce *et al.*, 2003) as a search model. Refinement was carried out with *BUSTER* (Bricogne *et al.*, 2011) alternating with manual model building using *Coot* (Emsley *et al.*, 2010). The final model was evaluated using the *Coot* validation tools and *MolProbity* (Chen *et al.*, 2010). Because the CD-loop region is disordered in molecules  $A$  and  $C$  of the mutant, the wild-type protein was compared with molecule  $B$  of the mutant. The cavity surfaces and volumes were calculated using *CASTp* (Dundas *et al.*, 2006) with a probe of 1.4 Å radius and were visualized with *HOLLOW* (Ho & Gruswitz, 2008).

## 3. Results

### 3.1. Overall structure

The structure of wild-type hNgb contains two independent molecules in the asymmetric unit that form a loosely assembled dimer (Table 1, Fig. 1*a*). Monomer  $B$  adopts the typical globin fold consisting of eight  $\alpha$ -helices denoted A–H (Figs. 1*a* and 1*b* and Supplementary Fig. S1). In contrast, in monomer  $A$  helix C is longer by two helix–turns, whereas helix D is completely absent and is replaced by a coil (Figs. 1*a* and 1*b*). The CD region that contains Cys46 and Cys53 differs considerably among the two protein models (Fig. 1*b*). In both molecules, the haem is hexacoordinated by the distal and proximal histidines His64 and His96, respectively (Supplementary Fig. S2). The N- and C-terminal regions displayed much higher  $B$  factors than the average (Fig. 2*a*), and could not be modelled completely (residues B1, B150, B151, A1, A2, A150 and A151 are missing). In addition, the loop connecting

**Table 2**  
Geometry of the disulfide bonds in molecules *A* and *B* of hNGB.

	Molecule <i>A</i>	Molecule <i>B</i>
Group	− + + + −	− − − − −
S—S distance (Å)	2.05	2.03
C <sup>α</sup> —C <sup>α</sup> distance (Å)	5.79	5.83
C <sup>α</sup> 46—C <sup>β</sup> 46—S46 (°)	114.23	113.79
C <sup>β</sup> 46—S46—S55 (°)	103.94	105.17
S46—S55—C <sup>β</sup> 55 (°)	106.31	103.24
S55—C <sup>β</sup> 55—C <sup>α</sup> 55 (°)	114.35	113.24
χ <sub>1</sub> (N <sup>α</sup> 46—C <sup>α</sup> 46—C <sup>β</sup> 46—S46) (°)	−79.76	−169.09
χ <sub>2</sub> (C <sup>α</sup> 46—C <sup>β</sup> 46—S46—S55) (°)	134.67	−111.45
χ <sub>3</sub> (C <sup>β</sup> 46—S46—S55—C <sup>β</sup> 55) (°)	77.27	−83.60
χ <sub>1</sub> ' (C <sup>α</sup> 55—C <sup>β</sup> 55—S55—S46) (°)	55.93	−57.73
χ <sub>2</sub> ' (S55—C <sup>β</sup> 55—C <sup>α</sup> 55—N55) (°)	−178.68	−53.43
His64 NE2—Fe distance (Å)	2.11	2.11
Dihedral energy strain† (kJ mol <sup>−1</sup> )	13.86	11.63

† Dihedral energy strain (kJ mol<sup>−1</sup>) = 8.37(1 + cos3χ<sub>1</sub>) + 8.37(1 + cos3χ<sub>1</sub>') + 4.18(1 + cos3χ<sub>2</sub>) + 4.18(1 + cos3χ<sub>2</sub>') + 14.64(1 + cos2χ<sub>3</sub>) + 2.51(1 + cos3χ<sub>3</sub>) (Schmidt *et al.*, 2006).

helices *A* and *B* is slightly different between the two monomers.

### 3.2. Disulfide-bond geometry

In both monomers of the asymmetric unit, continuous electron density was observed linking the Cys46 and Cys53 side chains (Figs. 1*c* and 1*d*). The distance between the S atoms of each cysteine (2.0 Å in each monomer) is consistent with the existence of an intramolecular disulfide linkage. The disulfide-bond geometry is characterized by the five χ angles that make up the bond (Table 2). Surprisingly, neither of the disulfide bonds observed in the crystal structure of hNgb belongs to the minus right-handed staple (−RHStaple) group that contains the known allosteric disulfides that regulate protein function in a nonenzymatic way by triggering changes in the intramolecular or intermolecular structure of proteins (Schmidt *et al.*, 2006). In this group, χ<sub>1</sub>, χ<sub>2</sub>, χ<sub>3</sub>, χ<sub>2</sub>' and χ<sub>1</sub>' are −, −, +, − and −, respectively, and the average dihedral strain energy is 18 kJ mol<sup>−1</sup>. The estimated dihedral strain energies of the two disulfide bonds of hNgb are 13.9 and 11.6 kJ mol<sup>−1</sup> for monomers *A* and *B*, respectively, indicating that they are easily cleaved, which is in agreement with their role in controlling protein function. In contrast to most previously identified allosteric disulfide bonds that link adjacent strands in the same antiparallel β-sheets (Schmidt *et al.*, 2006), Cys46 and Cys55 that form the S—S bond in hNgb belong to the CD loop and the D region that forms either an α-helix or a coil, respectively. Therefore, the C<sup>α</sup>—C<sup>α</sup>' distance between the two cysteines is 5.8 Å, much longer than the mean 4.3 Å found in the −RHStaple group. Cys46 and Cys55 are completely buried in monomer *A* [accessible surface areas (ASAs) of 1 Å<sup>2</sup> and 5.9 Å<sup>2</sup>], whereas in monomer *B* Cys55 is buried (ASA of 3.4 Å<sup>2</sup>) but Cys46 is relatively solvent-accessible (ASA of 16.2 Å<sup>2</sup>).

### 3.3. CD loop

Interestingly, monomers *A* and *B* present a different conformation of the region encompassing the two cysteines involved in the intramolecular disulfide bridge (residues 36–

59, corresponding to the CD region and called the dynamic loop), indicating high flexibility of this region despite the presence of the disulfide bond (Fig. 1*b*). Indeed, the two molecules superpose with r.m.s.d. values of 2.59 Å over 147 C<sup>α</sup> atoms and of 0.73 Å over 123 C<sup>α</sup> atoms when this region is excluded. The hinge residues at positions 36 and 59 are two prolines. The presence of the disulfide bond has no major effect on the overall conformation of hNgb outside of the dynamic loop, as shown by the per-residue root-mean-square deviation between the two monomers (Fig. 2*b*). The *B* factors of the dynamic loop in the wild-type protein are around the average (23.6 and 18.9 Å<sup>2</sup> for chains *A* and *B*, respectively) because the loop is involved in crystallographic contacts (Fig. 3; Supplementary Table S1).

### 3.4. Comparison of the structures of the wild-type and mutant proteins

To shed light on the structural changes induced by disulfide-bond formation in the absence of a structure of reduced wild-type hNgb, the structure of the oxidized wild-type protein can be compared with that of the protein in which the cysteines were mutated (Pesce *et al.*, 2003). The dimer of the wild-type protein is similar to the *A/B* and *C/D* dimers observed in the mutant form. Indeed, the structures of the wild-type and mutant dimers superimpose with r.m.s.d. values of 1.18 Å over 246 C<sup>α</sup> atoms (omitting the CD region). The buried surface of 3930 Å<sup>2</sup> and the free energy of assembly dissociation of 16.7 kJ mol<sup>−1</sup> calculated with *PISA* (Krissinel & Henrick, 2007) suggest that the dimer is a stable assembly in solution. However, since the current data indicate that hNgb is a monomer in solution (Burmester *et al.*, 2000), this dimer is thought at present to be crystallographic and not biologically relevant. The main structural changes between wild-type and mutant hNgb lie in the CD region, with Ala55 in the mutant being completely buried (ASA of 1.1 Å<sup>2</sup>) but with Gly46 being completely exposed (ASA of 85.2 Å<sup>2</sup>). Excluding the CD region, the secondary structure of molecule *B* of the wild-type protein is very similar to that of the mutant protein, except that helices *C* and *D* are slightly shifted (Supplementary Fig. S3 and Fig. 4*a*). Monomer *A* of the wild-type protein is more different and superimposes on the mutant with an r.m.s.d. value of 1.34 Å over 147 C<sup>α</sup> atoms, omitting the CD region. The high flexibility of this region observed in the structure of wild-type hNgb had previously been demonstrated in the crystal of the mutant protein, in which the four molecules in the asymmetric unit had either a disordered or an extended conformation of the CD loop (Pesce *et al.*, 2003; Bocahut *et al.*, 2013). Yet, the conformational disorder of this region could be owing to the mutations and in particular to the replacement of Cys46 with a much more flexible glycine residue.

Comparison of the structures of the two copies of the wild-type protein with that of the hNgb mutant highlights critical changes in the CD loop: the side chain of Phe42 slides by 1.5–2 Å and Tyr44, which sequesters the distal pocket from the solvent in mouse Ngb and mutant hNgb, adopts new conformations in wild-type hNgb (Fig. 4*b*). The Tyr44 OH group is

hydrogen-bonded to one propionate O atom in the mutant protein (2.6 Å). In molecule *B* of the wild-type protein, the distance between the two corresponding O atoms is 4 Å, whereas Tyr44 in molecule *A* is directed towards the solvent.

As a consequence of the different conformations of the CD regions, the total accessible surface area for the haem is 99 Å<sup>2</sup>

in molecule *A* and 117 Å<sup>2</sup> in molecule *B*, compared with 140.3 Å<sup>2</sup> in the mutant protein.

Apart from the CD region, the haem pocket is similar in the wild-type and mutant proteins, except that the side chain of Lys102 adopts slightly different positions (Fig. 4*b*). In particular, the side chain of Lys67 (E10) makes a hydrogen bond to one haem propionate group in both the hNgb mutant and

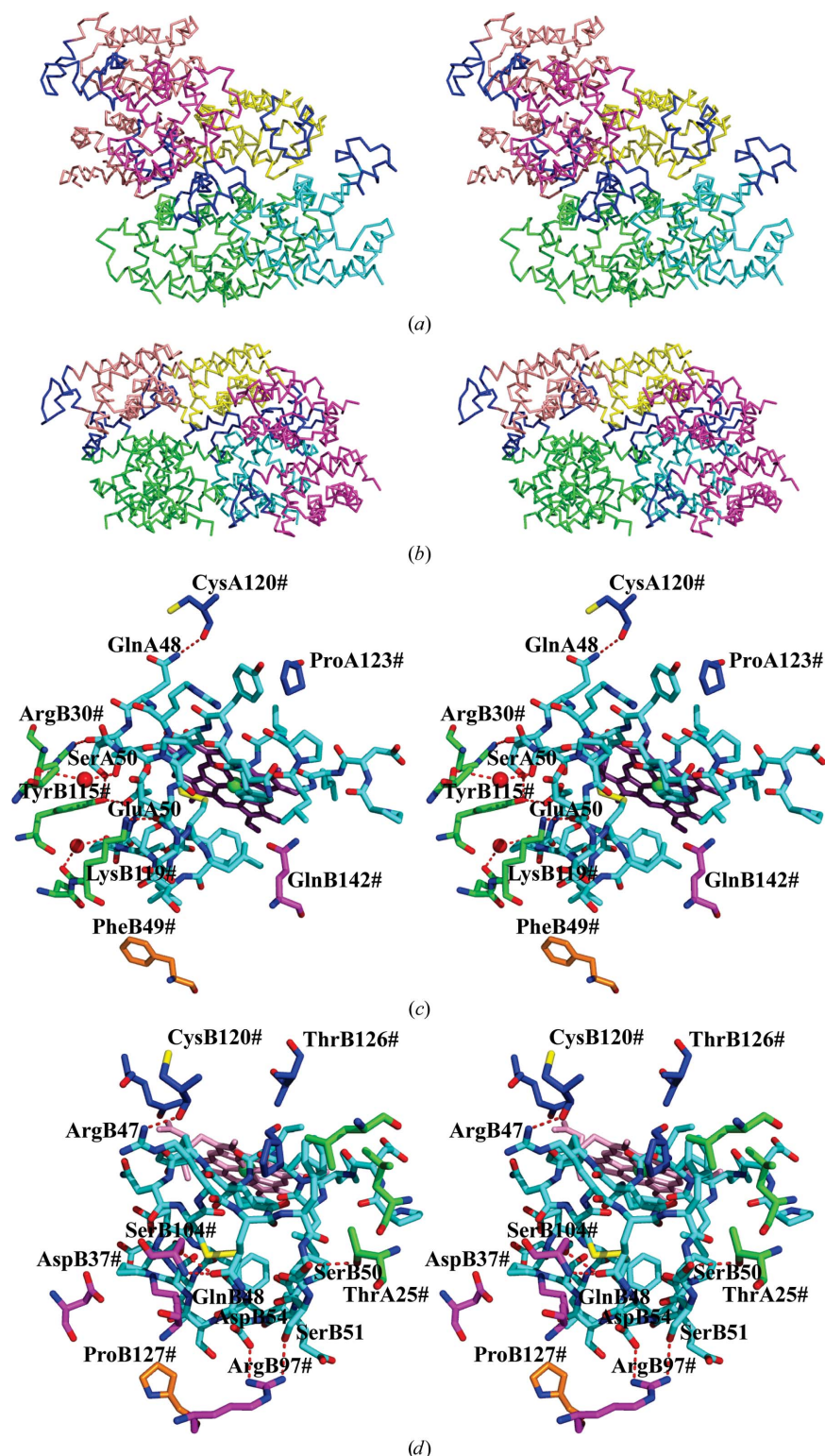
wild-type proteins, an interaction that acts as a barrier to ligand exit and entry and is conserved in wild-type mouse Ngb (Vallone, Nienhaus, Brunori *et al.*, 2004; Vallone, Nienhaus, Matthes *et al.*, 2004). The key residues of the distal site known to control the affinity for exogenous ligands, such as Phe28 (B10) and Val68 (E11), also occupy similar positions. The similarity of the haem pocket in the wild-type and mutant structures agrees with the superimposition of the absorption spectra of oxidized hNgb, DTT-reduced hNgb, the C55S hNgb mutant, rat Ngb and the rat G46S Ngb mutant (Astudillo *et al.*, 2010), which indicates that mutation of the cysteines at positions 46 or 55 does not significantly modify the haem pocket.

### 3.5. Internal cavities

Small globular proteins contain internal cavities and packing defects that may play a role in protein conformational changes or in controlling function by defining pathways for the diffusion of the ligand to the active site. The wild-type hNgb structure shows a predominant solvent-inaccessible cavity enclosing part of the haem distal site (Fig. 5).

**Figure 3**

Crystallographic contacts in the wild-type hNgb structure. Analysis of the intermolecular contacts involving the CD corner shows that the CD loop is trapped by at least four different arrays of short contacts with residues of the neighbouring asymmetric units. The disulfide-bonded conformation of the loops is stabilized by five or six hydrogen bonds between residues in the loop and residues belonging to symmetric molecules for molecules *A* and *B*, respectively. Among these latter residues, Arg30 and Trp148 are conserved in all Ngbs, whereas Arg97, Ser104, Tyr115, Lys119 and Cys120 are mostly conserved (Supplementary Fig. S1). (a) The four neighbouring molecules (coloured cyan, yellow, magenta and tan) of the CD loop (dark blue) of molecule *A* (in green). (b) The four neighbouring molecules of the CD loop of molecule *B*. (c) Enlargement showing the interactions between the CD loop of molecule *A* of wild-type Ngb (in cyan) and the symmetric molecules. (d) Enlargement showing the interactions between the CD loop of molecule *B* of wild-type Ngb (in cyan) and the symmetric molecules.



The cavity is closed to the external solvent and is lined with hydrophobic residues, including Val71, Ile72, Val109, Leu113, Leu136 and Val140. The volume of this cavity is 89 and 225 Å<sup>3</sup> (ASAs of 134 and 272 Å<sup>2</sup>) in molecules *A* and *B* out of totals of 416.5 and 578.8 Å<sup>3</sup>, respectively (Figs. 5*a* and 5*b*). Interestingly, this internal cavity was also present in the mutant protein, but its volume was increased to 274 Å<sup>3</sup> (out of a total of 452.3 Å<sup>3</sup>; Fig. 5*c*). In addition, in both molecules of wild-type hNgb a new internal cavity of around 60 Å<sup>3</sup> is formed in the distal region by the CD loop (Figs. 5*a* and 5*b*). It is inaccessible to solvent but is connected to the distal pocket.

#### 4. Discussion

##### 4.1. The intramolecular disulfide bond regulates the dynamics of the CD loop

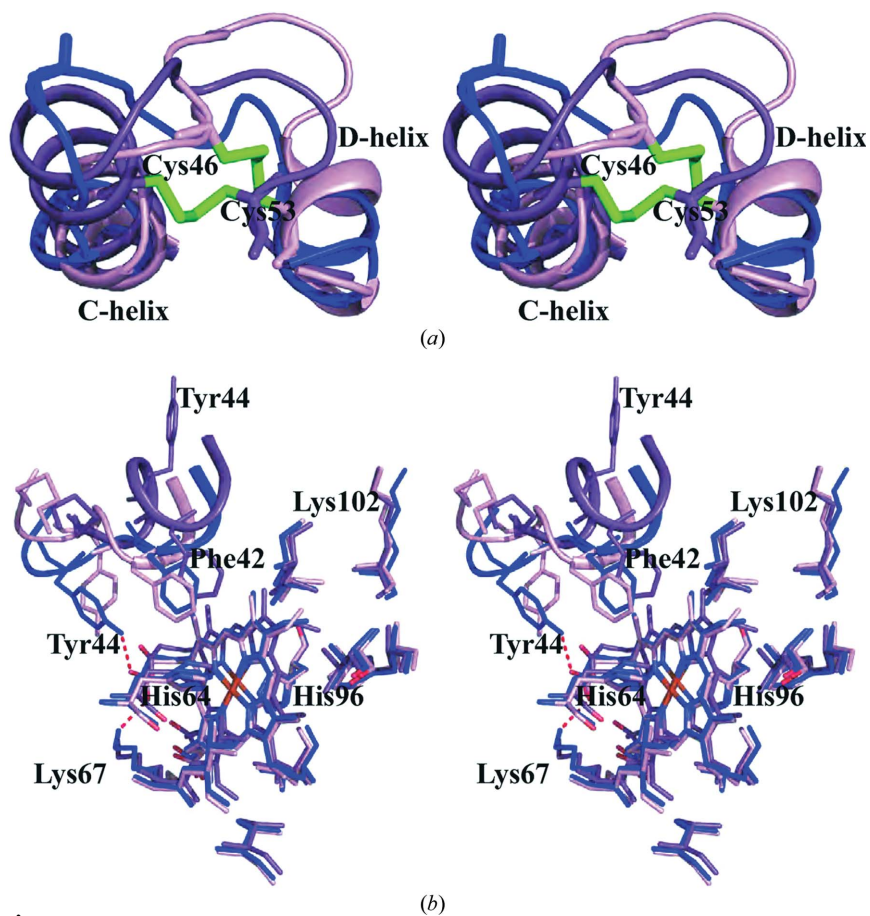
We determined the crystal structure of wild-type hNgb, which shares 94% sequence identity with mouse Ngb with an identical sequence in the CD loop except that Cys at position 46 in human is replaced by Gly in mouse (Supplementary Fig.S1). Similarly to what is observed in the unliganded state

of the mouse protein (Vallone, Nienhaus, Brunori *et al.*, 2004), in wild-type hNgb the haem embedded between the E and F helices is hexacoordinated, probably because the concentration of five-coordinated Ngb is very low at equilibrium (0.1–5%). The two copies in the asymmetric unit display an internal disulfide bond between Cys46 and Cys53, as anticipated for non-rodent mammalian Ngbs. The formation of the disulfide bond in the wild-type protein induces large structural changes in the dynamic loop or CD region compared with the protein in which all cysteines were mutated (Pesce *et al.*, 2003), confirming that the dynamics of the CD–D region is linked to the internal disulfide bond in hNgb. Accordingly, the CD–D region of mutant hNgb is more similar to that of mouse Ngb than to wild-type hNgb (Supplementary Fig. S3). The different mobility in the CD loop may contribute to the distinct functional properties of human Ngb compared with the murine protein. Yet, the existence of at least two different conformations of the CD loop containing the disulfide bridge indicates that the hNgb structure remains flexible even in the presence of the disulfide bridge.

The conformational flexibility of this region was proposed to be crucial for binding of Ngb to the  $\alpha$ -subunit of the  $G\alpha\beta\gamma$  protein (Khan *et al.*, 2008). The interaction sites between Ngb and  $G\alpha\beta\gamma$  (Glu53 and Glu6; Kitatsuji *et al.*, 2007) are located in the CD loop. Under low-oxygen conditions, ferric Ngb would inhibit the rate of exchange of GDP bound to the  $G\alpha\beta\gamma$  protein with GTP by binding to the  $\alpha$ -subunit (Wakasugi *et al.*, 2003). Because the ensuing liberation of  $G\beta\gamma$  would protect the cell against neuronal death, Ngb has been proposed to act as an oxidative stress-responsive sensor for signal transduction in the brain.

##### 4.2. The increased ligand affinity upon disulfide-bond formation is linked to the conformational change of the CD loop

Ligand-binding kinetics are governed by the stability of the iron–distal histidine bond and the ligand diffusion rate to the penta-coordinated iron. Indeed, CO rebinding to hNgb after photodissociation was shown to be a biphasic process, with a fast phase corresponding to CO binding to a population of five-coordinated deoxy-Ngb and a slow phase corresponding to a population in which the distal histidine bound to the Fe atom is replaced by CO. The presence of the internal disulfide bridge in hNgb was shown to be associated with a tenfold decrease in affinity for the internal ligand, owing mainly to a change in the His64 dissociation rate, thus effectively leading to a large increase in affinity for external ligands (Astudillo *et al.*,



**Figure 4** Comparison of the wild-type and mutant hNgb structures. (a) Stereoview of the superposition of the two monomers of wild-type hNgb and the mutant protein. Regions 3–35 and 60–149 of the wild-type protein (molecule *A* is coloured violet and molecule *B* is coloured pink) were superimposed with molecule *B* of the mutant protein (PDB entry 1oj6, coloured dark blue). (b) Superposition of the haem-binding sites.

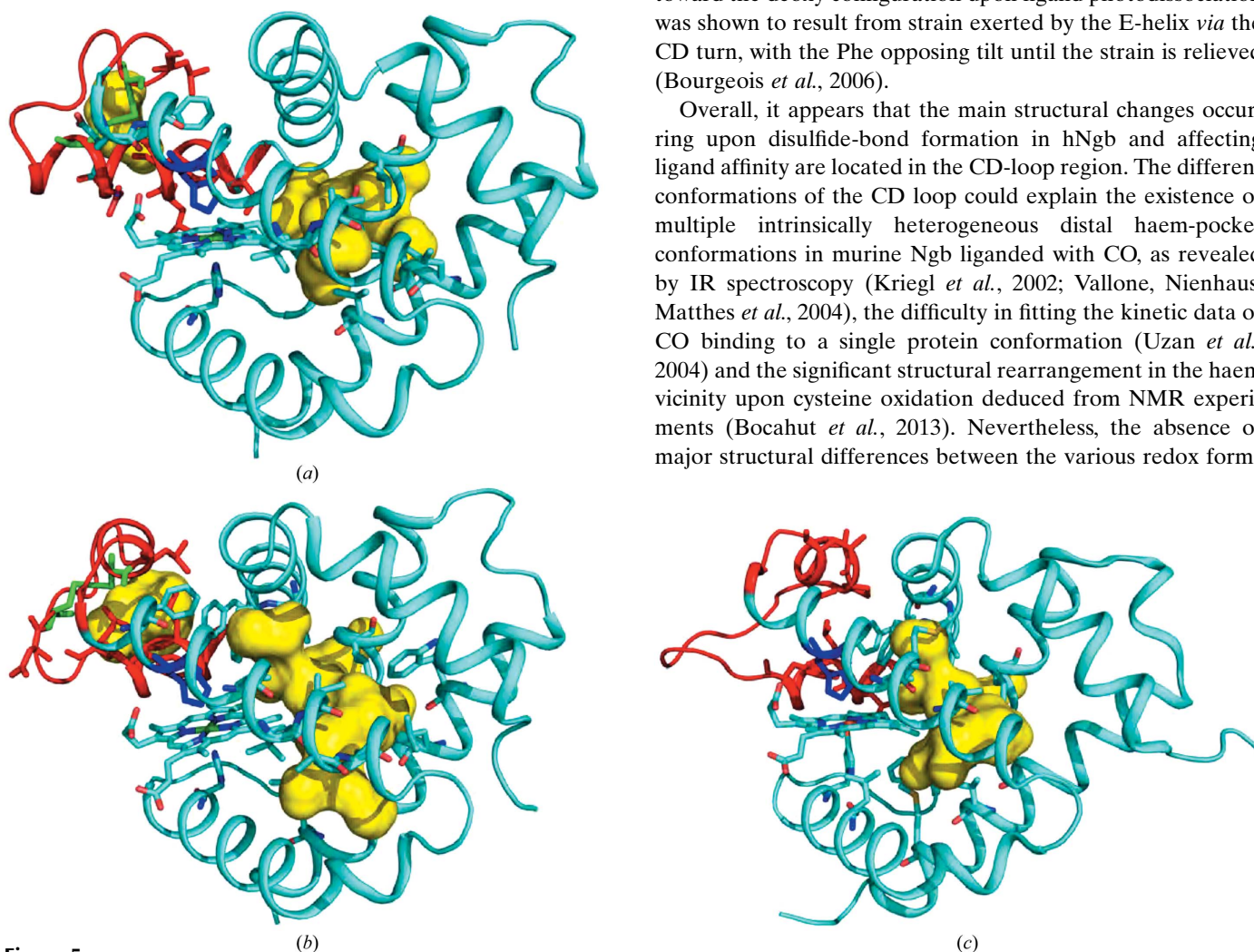
2010; Hamdane *et al.*, 2003; Hamdane, Kiger, Dewilde *et al.*, 2005; Nadra *et al.*, 2008; Bocahut *et al.*, 2013). Moreover, the overall volume change associated with CO photodissociation from the haem active site was smaller when the disulfide bond was present (Astudillo *et al.*, 2010; Hamdane, Kiger, Hoa *et al.*, 2005). Therefore, the internal disulfide bond appeared to modulate the magnitude of the overall structural changes in Ngb upon ligand binding/release.

The flexibility of the CD region was proposed to play a functional role in displacing the distal histidine ligand (His64 in hNgb) upon O<sub>2</sub> binding in hexacoordinated globins (Pesce *et al.*, 2003). Molecular-dynamics (MD) simulations (Anselmi *et al.*, 2007), as well as comparison of unliganded and CO-bound structures of murine Ngb (Vallone, Nienhaus, Matthes *et al.*, 2004), indicated that the flexibility and conformation of the CD loop is affected by ligand binding. In hNgb, it was proposed, based on MD simulations, that in the absence of the disulfide bond an enhanced internal mobility in the CD loop

hinders access of the ligand to the haem group by stabilizing the His64–iron bond and that the formation of the disulfide bridge promotes stabilization of the pentacoordinated deoxy ferrous haem form, facilitating dioxygen release (Bocahut *et al.*, 2013).

Our structural comparison of Ngbs that do or do not contain the disulfide bond indicates that the main elements that could be involved in the propagation of the conformational change of the CD loop to the haem-binding region are the displacements of Tyr44 (CD3) and Phe42 (CD1), which might directly affect the formation of the iron–distal histidine bond. Interestingly, the latter residue in molecule *B* of the wild-type protein, which is positioned slightly nearer to the haem than in the mutant (0.2 Å), is very highly conserved among the myoglobin family (Bashford *et al.*, 1987) and was found to finely affect haem dynamics in myoglobin. Indeed, the equivalent phenylalanine is packed onto the haem in sperm whale Mb complexed with CO, and the haem relaxation toward the deoxy configuration upon ligand photodissociation was shown to result from strain exerted by the E-helix *via* the CD turn, with the Phe opposing tilt until the strain is relieved (Bourgeois *et al.*, 2006).

Overall, it appears that the main structural changes occurring upon disulfide-bond formation in hNgb and affecting ligand affinity are located in the CD-loop region. The different conformations of the CD loop could explain the existence of multiple intrinsically heterogeneous distal haem-pocket conformations in murine Ngb liganded with CO, as revealed by IR spectroscopy (Kriegl *et al.*, 2002; Vallone, Nienhaus, Matthes *et al.*, 2004), the difficulty in fitting the kinetic data of CO binding to a single protein conformation (Uzan *et al.*, 2004) and the significant structural rearrangement in the haem vicinity upon cysteine oxidation deduced from NMR experiments (Bocahut *et al.*, 2013). Nevertheless, the absence of major structural differences between the various redox forms



**Figure 5**

Comparison of the internal cavities. (a) The two main internal cavities of molecule *A* of wild-type hNgb. The cavities are coloured yellow. The small cavity is enclosed by residues of the CD loop: Phe42, Asn45, Cys46, Cys55, Glu60 and Phe61. The large cavity is lined with the side chains of Val71, Ile72, Val109, Leu113, Leu136, Tyr137 and Val140. (b) The two main internal cavities of molecule *B* of wild-type hNgb. The small cavity is enclosed by residues of the CD loop: Phe32, Leu39, Gln43, Asn45, Phe42, Cys55, Leu56 and Phe61. The large cavity is lined with the side chains of Leu27, Val68, Val71, Ile72, Ala75, Val109, Leu113, Trp133, Leu136, Val140 and Met144. (c) The large internal cavity of mutant hNgb (molecule *B*). The large cavity is lined with the side chains of Leu27, Phe28, Leu31, Val68, Val71, Ala75, Val109, Leu113, Trp133, Leu136, Tyr137, Val140, Val141 and Met144.

of the protein indicates that dynamics play a major role in the regulation of the low and high oxygen-affinity states following S—S bond formation. This conclusion is also supported by the greater stability of the reduced protein, which displays a  $T_m$  value of 103°C, 4°C above that of the oxidized form (Hamdane, Kiger, Dewilde *et al.*, 2005).

### 4.3. There is no haem conformational disorder in wild-type hNgb

An indication of haem conformational disorder in both mouse and human Ngbs by the presence of two alternative haem orientations within the haem pocket has been provided by X-ray crystallography (Pesce *et al.*, 2003; Vallone, Nienhaus, Brunori *et al.*, 2004) and NMR (Xu *et al.*, 2009; Bocahut *et al.*, 2013). It was proposed that the lack of orientational selectivity is related to the presence of a very large cavity lining the haem, a cavity that is much smaller in wild-type hNgb. In contrast, our structure of the wild-type protein does not reveal the presence of multiple conformations of the haem or the imidazole side chain of the distal histidine ligand. The similarity of the conformations of the two histidine ligands in the wild-type and mutant hNgbs does not support a change in the relative orientations of the HisE7 and HisF8 imidazole planes upon disulfide-bond formation, as was invoked to interpret the difference between the electron magnetic resonance spectra of ferric Ngb (Vinck *et al.*, 2004; Ezhevskaya *et al.*, 2011). In addition, the haem pocket is similar in the wild-type and mutant structures (Fig. 4b), which is in agreement with the superimposition of the absorption spectra of oxidized hNgb, DTT-reduced hNgb, the C55S hNgb mutant, rat Ngb and the rat G46S Ngb mutant (Astudillo *et al.*, 2010), indicating that the mutation of the cysteines in positions 46 or 55 does not significantly modify the haem pocket. However, the different conformations of the CD loop could explain the existence of multiple intrinsically heterogeneous distal haem-pocket conformations in murine Ngb liganded with CO as revealed by IR spectroscopy (Kriegl *et al.*, 2002; Vallone, Nienhaus, Matthes *et al.*, 2004), the difficulty in fitting the kinetic data of CO binding to a single protein conformation (Uzan *et al.*, 2004) and the significant structural rearrangement in the haem vicinity upon cysteine oxidation deduced from NMR experiments (Bocahut *et al.*, 2013).

### 4.4. Remodelling of internal cavities could promote flexibility in wild-type hNgb

Pre-existing internal cavities can serve to control the dynamics and reactivity of reactions. Based on binding experiments of Xe to myoglobin, internal cavities and packing defects were proposed to be the docking sites that were transiently occupied by the ligand during its trajectory to the protein (Brunori & Gibson, 2001). The haems of murine and human Ngbs, like that of myoglobin, are deeply buried in the protein, and the internal cavities can also facilitate ligand migration between the haem pocket and the solvent. A huge internal cavity was observed in ferric Ngb from mouse and human (275 Å<sup>3</sup> for the mutant human protein) between the

distal site and the EF interhelical hinge (Vallone, Nienhaus, Brunori *et al.*, 2004; Pesce *et al.*, 2003). In mouse Ngb, this peculiar feature was proposed to allow the sliding movement of the haem to the preformed docking site upon CO binding, thereby reshaping the cavity connecting the distal and proximal haem sides (Vallone, Nienhaus, Matthes *et al.*, 2004). The formation of the large internal cavity was also thought to be related to the marked mobility of the EF loop which was observed both in mouse Ngb and mutant hNgb (Vallone, Nienhaus, Brunori *et al.*, 2004; Pesce *et al.*, 2003). The loosening of the structural constraints in the EF loop was proposed to be the signal encoding information on low-oxygen conditions. Haem sliding and distortion of the internal cavity were thus proposed to control ligand affinity in mouse Ngb.

The flexibility of the EF loop is not revealed by a *B*-factor analysis of the wild-type hNgb model (Fig. 2a), probably owing to its involvement in packing interactions with the haem propionate groups and the N-terminal part of the other molecule in the asymmetric unit. Thus, despite the low *B* factors of the EF loop in the crystal, this loop could be mobile in solution and its mobility could be correlated with the internal cavity formation, as suggested above. On the other hand, we show here that the disulfide bridge drastically modulates the volume of internal cavities in the human protein, as suggested from MD studies (Bocahut *et al.*, 2009; Nadra *et al.*, 2008). Indeed, the large internal cavity observed in mouse Ngb and mutant hNgb shrinks by 17–30% in wild-type hNgb. Interestingly, our structure instead reveals the creation of a new internal cavity with an approximate volume of 60 Å<sup>3</sup> in the CD loop region upon disulfide-bond formation. This cavity could enable conformational mobility of the distal histidine in hNgb. The reshaping of the internal cavities could thus regulate the dynamics of the equilibrium between the hexacoordinated and pentacoordinated forms or switch the ligand migration pathway, as shown for Cygb (Astudillo *et al.*, 2013). The new distribution of the internal cavities, together with the absence of haem conformational disorder in wild-type hNgb, indicate that a new mechanism for displacing the distal histidine upon ligand binding, distinct from that operating in mouse Ngb or Cygb, may take place in the human protein.

### 4.5. Disulfide bonds acting as regulatory switches

In several other cases, oxidation-mediated disulfide-bond formation within a protein has been shown to induce structural changes with functional consequences (Brandes *et al.*, 2009; Antelmann & Helmann, 2011; Ryu, 2012). Several plant thiol metabolism enzymes such as glutamate–cysteine ligase (Kumar *et al.*, 2013) or adenosine-5'-phosphosulfate kinase (Ravilious *et al.*, 2012) and reductase (Bick *et al.*, 2001), possess cysteines that form disulfide bonds to regulate enzyme activity in response to changes in redox environment. Human succinic semialdehyde dehydrogenase (Kim *et al.*, 2009), the scaffold protein INAD (Liu *et al.*, 2011) and angiotensin (Zhou *et al.*, 2010) are also regulated by redox-switch modulation.



## 5. Conclusion

In hNgb, the breakage or the absence of the disulfide bond was shown to decrease the oxygen affinity by one order of magnitude. Because the overall rates of O<sub>2</sub> binding and dissociation dictate efficient delivery to tissues and thus physiology, the disulfide bond in Ngb is biologically important in humans, in contrast to other organisms such as rodents and amphibia where it is absent given the absence of the relevant cysteines in the sequence.

Because mutation of the cysteine residues in hNgb resulted in important changes in the dynamics of the CD–D region (Bocahut *et al.*, 2009) and haem displacement (Bocahut *et al.*, 2013), the triple cysteine mutant of hNgb only provides part of the description of the wild-type protein. Therefore, it was crucial to obtain the structure of wild-type Ngb. Comparison of this structure with that of the mutant protein allows us to assess the conformational changes occurring upon disulfide-bond formation. Moreover, two distinct conformations of the intramolecular disulfide bond in hNgb were identified by direct structural observation.

The structure of wild-type hNgb confirms the conformational flexibility of the CD region in the haem vicinity even in the presence of the intramolecular disulfide bond. This flexibility plays a crucial role in the ability of the protein to bind and stabilize exogenous ligands, perhaps by reshaping the internal cavities. In particular, a new internal cavity was created in the CD region upon disulfide-bond formation, which may provide access to ligands and facilitate their diffusion. How the increased backbone flexibility in the CD–D region in the absence of the disulfide bond contributes to stabilizing the iron–His64 bond remains to be explained.

This work was supported by the CNRS, INSERM and the University of Paris XI. We thank Professor T. Burmester and T. Hankeln (University of Hamburg/Mainz) for the Ngb plasmid. We acknowledge SOLEIL for the provision of synchrotron-radiation facilities on the PROXIMA 1 beamline (proposal 20060233).

## References

- Anselmi, M., Brunori, M., Vallone, B. & Di Nola, A. (2007). *Biophys. J.* **93**, 434–441.
- Antelmann, H. & Helmann, J. D. (2011). *Antioxid. Redox Signal.* **14**, 1049–1063.
- Astudillo, L., Bernad, S., Derrien, V., Sebban, P. & Miksovská, J. (2010). *Biophys. J.* **99**, L16–L18.
- Astudillo, L., Bernad, S., Derrien, V., Sebban, P. & Miksovská, J. (2013). *J. Inorg. Biochem.* **129**, 23–29.
- Bashford, D., Chothia, C. & Lesk, A. M. (1987). *J. Mol. Biol.* **196**, 199–216.
- Bick, J. A., Setterdahl, A. T., Knaff, D. B., Chen, Y., Pitcher, L. H., Zilinskas, B. A. & Leustek, T. (2001). *Biochemistry*, **40**, 9040–9048.
- Bocahut, A., Bernad, S., Sebban, P. & Sacquin-Mora, S. (2009). *J. Phys. Chem. B*, **113**, 16257–16267.
- Bocahut, A., Derrien, V., Bernad, S., Sebban, P., Sacquin-Mora, S., Guittet, E. & Lescop, E. (2013). *J. Biol. Inorg. Chem.* **18**, 111–122.
- Bourgeois, D., Vallone, B., Arcovito, A., Sciara, G., Schotte, F., Anfirud, P. A. & Brunori, M. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4924–4929.
- Brandes, N., Schmitt, S. & Jakob, U. (2009). *Antioxid. Redox Signal.* **11**, 997–1014.
- Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonnrhein, C. & Womack, T. O. (2011). *BUSTER v2.11.5*. Cambridge: Global Phasing Ltd.
- Brunori, M. & Gibson, Q. H. (2001). *EMBO Rep.* **2**, 674–679.
- 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.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Dewilde, S., Kiger, L., Burmester, T., Hankeln, T., Baudin-Creuzat, V., Aerts, T., Marden, M. C., Caubergs, R. & Moens, L. (2001). *J. Biol. Chem.* **276**, 38949–38955.
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Res.* **34**, W116–W118.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Ezhevskaya, M., Trandafir, F., Moens, L., Dewilde, S. & Van Doorslaer, S. (2011). *J. Inorg. Biochem.* **105**, 1131–1137.
- Hamdane, D., Kiger, L., Dewilde, S., Green, B. N., Pesce, A., Uzan, J., Burmester, T., Hankeln, T., Bolognesi, M., Moens, L. & Marden, M. C. (2003). *J. Biol. Chem.* **278**, 51713–51721.
- Hamdane, D., Kiger, L., Dewilde, S., Green, B. N., Pesce, A., Uzan, J., Burmester, T., Hankeln, T., Bolognesi, M., Moens, L. & Marden, M. C. (2004). *Micron*, **35**, 59–62.
- Hamdane, D., Kiger, L., Dewilde, S., Uzan, J., Burmester, T., Hankeln, T., Moens, L. & Marden, M. C. (2005). *FEBS J.* **272**, 2076–2084.
- Hamdane, D., Kiger, L., Hoa, G. H. B., Dewilde, S., Uzan, J., Burmester, T., Hankeln, T., Moens, L. & Marden, M. C. (2005). *J. Biol. Chem.* **280**, 36809–36814.
- Hamdane, D., Lechauve, C., Marden, M. C. & Golinelli-Pimpaneau, B. (2009). *Acta Cryst.* **D65**, 388–392.
- Ho, B. K. & Gruswitz, F. (2008). *BMC Struct. Biol.* **8**, 49.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Khan, A. A., Mao, X. O., Banwait, S., DerMardirossian, C. M., Bokoch, G. M., Jin, K. & Greenberg, D. A. (2008). *FASEB J.* **22**, 1737–1747.
- Kim, Y.-G., Lee, S., Kwon, O.-S., Park, S.-Y., Lee, S.-J., Park, B.-J. & Kim, K.-J. (2009). *EMBO J.* **28**, 959–968.
- Kitatsugi, C., Kurogochi, M., Nishimura, S., Ishimori, K. & Wakasugi, K. (2007). *J. Mol. Biol.* **368**, 150–160.
- Kriegel, J. M., Bhattacharyya, A. J., Nienhaus, K., Deng, P., Minkow, O. & Nienhaus, G. U. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 7992–7997.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Kumar, S., Kasturia, N., Sharma, A., Datt, M. & Bachhawat, A. K. (2013). *Biochem. J.* **449**, 783–794.
- Liu, W., Wen, W., Wei, Z., Yu, J., Ye, F., Liu, C.-H., Hardie, R. C. & Zhang, M. (2011). *Cell*, **145**, 1088–1101.
- Nadra, A. D., Martí, M. A., Pesce, A., Bolognesi, M. & Estrin, D. A. (2008). *Proteins*, **71**, 695–705.
- Pesce, A., Dewilde, S., Nardini, M., Moens, L., Ascenzi, P., Hankeln, T., Burmester, T. & Bolognesi, M. (2003). *Structure*, **11**, 1087–1095.
- Ravilious, G. E., Nguyen, A., Francois, J. A. & Jez, J. M. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 309–314.
- Ryu, S. E. (2012). *J. Biochem.* **151**, 579–588.
- Schmidt, B., Ho, L. & Hogg, P. J. (2006). *Biochemistry*, **45**, 7429–7433.
- 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.
- Tiso, M., Tejero, J., Basu, S., Azarov, I., Wang, X., Simplaceanu, V., Frizzell, S., Jayaraman, T., Geary, L., Shapiro, C., Ho, C., Shiva, S., Kim-Shapiro, D. B. & Gladwin, M. T. (2011). *J. Biol. Chem.* **286**, 18277–18289.

- Uzan, J., Dewilde, S., Burmester, T., Hankeln, T., Moens, L., Hamdane, D., Marden, M. C. & Kiger, L. (2004). *Biophys. J.* **87**, 1196–1204.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- 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.
- Vinck, E., Van Doorslaer, S., Dewilde, S. & Moens, L. (2004). *J. Am. Chem. Soc.* **126**, 4516–4517.
- Wakasugi, K., Nakano, T. & Morishima, I. (2003). *J. Biol. Chem.* **278**, 36505–36512.
- Xu, J., Li, L., Yin, G., Li, H. & Du, W. (2009). *J. Inorg. Biochem.* **103**, 1693–1701.
- Zhou, A., Carrell, R. W., Murphy, M. P., Wei, Z., Yan, Y., Stanley, P. L., Stein, P. E., Broughton Pipkin, F. & Read, R. J. (2010). *Nature (London)*, **468**, 108–111.