Minireview

Truncated Hemoglobins: A New Family of Hemoglobins Widely Distributed in Bacteria, Unicellular Eukaryotes, and Plants*§

Published, JBC Papers in Press, November 5, 2001, DOI 10.1074/jbc.R10058200

Jonathan B. Wittenberg‡§, Martino Bolognesi†, Beatrice A. Wittenberg‡, and Michel Guertin¶

From the ‡Department of Physiology and Biophysics, Albert Einstein College of Medicine, Bronx, New York 10461, ¶Department of Physics, INFN and Advanced Biotechnology Center, IST, University of Genova, Largo Rosanna Benzi 10, 16132 Genova, Italy, and §Département de Biochimie et Microbiologie, Pavillon Marchand, Université Laval, Québec G1K 7P4, Canada

Truncated hemoglobins (trHbs)1 (1) constitute a family of small oxygen-binding heme proteins distributed in eu-bacteria, cyanobacteria, protozoa, and plants (Table I, Supplemental Material) forming a distinct group within the hemoglobin (Hb) superfamily (2). They are nearly ubiquitous in the plant kingdom, occur in many aggressively pathogenic bacteria, and are held to be of very ancient origin. None have been detected in the genomes of archaea or metazoa. Characteristically, trHbs occur at nanomolar to micromolar intracellular concentration, hinting at a possible role as catalytic proteins.

Many trHbs display amino acid sequences 20–40 residues shorter than non-vertebrate hemoglobins to which they are scarcely related by sequence similarity. Crystal structures (1, 3) show that trHB tertiary structure is based on a 2-on-2 α-helical sandwich, which represents an unprecedented editing of the highly conserved globin fold. Moreover, an almost continuous hydrophobic tunnel, traversing the protein matrix from the molecular surface to the heme distal site, may provide a path for ligand diffusion to the heme.

**trHbs Are Phylogenetically Distinct within the Hb Superfamily**

More than 40 putative trHB genes have been identified (Table I, Supplemental Material). A phylogenetic analysis (Fig. 1) indicates that trHbs form a distinct family separate from bacterial flavohemoglobins, Vitreoscilla Hb, plant symbiotic and non-symbiotic hemoglobins, and animal globin. Notably, trHB genes and flavohemoglobin genes frequently coexist in the same bacterium (Table I, Supplemental Material), suggesting distinct functions for each. Three distinct groups (groups I, II, and III) can be distinguished within the trHB family with four clusters within group II. The extent of amino acid identity between members of the different groups, e.g. Mycobacterium tuberculosis trHBo (group II) and trBHO (group I), can be low (18%) (Fig. 1, Supplemental Material). However, identity rises dramatically when the M. tuberculosis trHBo sequence is compared with the orthologue sequences from Mycobacterium avium (84%), Mycobacterium leprae (83%), Corynabacterium diptheriae (64%), and Streptomyces coelicolor (66%) or when the M. tuberculosis trHBn is compared with its M. avium orthologue (79%). The presence of trHB genes from groups I, II, and III in M. avium (Gram-positive) and in Methylococcus capsulatus (proteobacteria) may indicate that the different trHB lineages diverged from their last common ancestor before the separation of the main prokaryotic lineages (Fig. 1).

An interesting progression is found in the genus Mycobacterium. The genome of the opportunistic pathogen, M. avium, contains one trHB from each group of the family, trHBp (group III), trHBo (group II), and trBHO (group I) (Fig. 1). The facultative intracellular pathogen, M. tuberculosis, that infects man, has two, trBHO and trHBn, and the obligate intracellular pathogen, M. leprae, which has undergone extensive reductive evolution and is thought to have only a minimal gene set for a pathogenic mycobacterium (4), retains solely trHBo, which accordingly may play an essential role.

**The 2-on-2 α-Helical Fold Characterizes trHbs**

Crystal structures of trHbs from Chlamydomonas eugameto-s, Paramaecium caudatum, and M. tuberculosis show that their three-dimensional fold is based on a subset of the classical globin fold (the so-called 3-on-3 α-helical sandwich). In trHbs the antiparallel helix pairs β/E and G/H are the main secondary structure elements arranged in a 2-on-2 sandwich (1, 3) (Fig. 2). Within the trHB fold the N-terminal A helix is almost completely deleted, and the whole CD-D region is trimmed to about 3 residues, possibly the minimum polypeptide stretch to bridge between C- and E-helices. Moreover, most of the hemoglobin F-helix is substituted by a polypeptide segment (pre-F) in extended conformation, followed by the one-turn F-helix that properly supports the HisF8 residue, allowing heme iron coordination. Thus, the trHB polypeptide chain is not simply a truncated version of a conventionally folded globin. Rather, it owes its conformational stability to residue deletions and substitutions at specific sites, as compared with non-vertebrate globins (1, 3).

Specific sequence motifs support attainment of the compact trHB fold. Among these, three glycine motifs (present in groups I and II), located at the AB and EF interhelical corners and immediately before the one-turn F-helix, help the pre-F segment to build a properly structured heme crevice within a very short polypeptide chain. Heme isomerism has been reported in trHbs (1, 5).

Very few amino acids are strictly conserved throughout the known trHB sequences, the proximal HisF8 being the only invariant residue. A Phe-Tyr pair is strongly conserved at the B9-B10 sites, where Tyr-B10 participates in heme ligand stabilization. Site CD1, invariably Phe in non-vertebrate Hbs, hosts Phe, Tyr, or His, whereas at least six different residue types occupy the distal E7 position. Moreover, the almost invariant Phe-E14, located along the heme distal face, may be related to a heme/solvent shielding role together with apolar residues of the pre-F segment.
Heme Coordination

Distal pocket residues may ligate the heme iron to form 6-coordinate, low spin, structures. Residue Tyr-B10 has been suggested to play this role in ferric *C. eugametos* trHb (6); His-E10 of *Synechocystis* sp. trHb may ligate the iron of ferric and ferrous species (7–9). These structures are in equilibrium with 5-coordinate or aquoferric forms with which exogenous ligands react. At high ligand concentration, where conversion of the 6- to the 5-coordinate form becomes rate-limiting, the rates of ligand combination give a measure of the rate of conversion (9, 10). If this rate is low, about 30 s \(^{-1}\) as in *Synechocystis* sp. trHb (7, 9), the 6-coordinate form is strongly favored.

In the instance of *C. eugametos* trHb, where the rate of this conversion in the ferrous alkaline form is 5-fold higher, 6- or 5-coordinate species prevail at different pH (10). Five-coordinate, high spin ferrous species of *P. caudatum* (11), *M. tuberculosis* trHbN (12), and *M. tuberculosis* trHbO \(^2\) predominate at all pH. A 6-coordinate form of *Arabidopsis thaliana* deoxy-trHb is short-lived and reverts to a 5-coordinate species (13).

In the known trHb three-dimensional structures the proximal His-F8 imidazole ring is markedly staggered relative to the heme pyrrole N atoms (1, 3, 14). This together with the high value of His-F8 NE2-Fe stretching frequencies (220–232 cm \(^{-1}\)) (10, 11, 15) indicates an unstrained proximal His, whose coordination to the heme iron facilitates a heme in-plane location of the iron atom that in turn supports fast oxygen association and electron donation to the bound distal ligand (16). Moreover, the O-O stretching frequencies measured for oxygenated *C. eugametos*, *Synechocystis* sp., and *M. tuberculosis* trHbO (1136, 1133, and 1140 cm \(^{-1}\), respectively) (17) are consistent with a ferric superoxide character of the heme-ligand pair.

2 M. Guertin and S. R. Yeh, unpublished observations.
3 Although *M. tuberculosis* trHbN was originally proposed to be a cooperative dimer with a cooperativity index of 2 (12), ongoing investigations indicate that oxygenation of this trHb may not be cooperative (M. Guertin, unpublished observations).
The heme distal site of trHb is characterized by the nearly invariant Tyr-B10, the main residue providing direct hydrogen bonding to the heme-bound ligand (as observed in aquo-Met \( P. caudatum \) trHb, cyano-Met \( C. eugametos \) trHb, and \( M. tuberculosis \) oxy-trHbN) (1, 3, 11) (see Fig. 3). A distal network of H-bonds has been shown to stabilize the ligated \( O_2 \) in \( M. tuberculosis \) oxy-trHbN (with Leu at the E7 site) through direct interaction with the Tyr-B10 phenolic oxygen atom and hydrogen bonding of this oxygen atom to Gln-E11 (3). Analysis of the crystal structures shows that distal site networks based on Tyr-B10 ligand and Tyr-B10-Gln/Thr-E11 hydrogen bonds are conserved in ferric cyano-Met \( M. tuberculosis \) and \( C. eugametos \) trHbs, respectively (1). Additionally, the latter ferric trHbs display hydrogen bonds between Gln-E7 and the heme ligand, indicating that polar E7 residues are effective in ligand stabilization and that hydrogen bonding between E7 and Tyr-B10 may also contribute to effective structuring of the distal site residues. In this respect, it should be noted that the occurrence of small apolar residues (Ala, Gly) at the E7 site may be compensated by CD1 Phe-224 Tyr/His mutations such as observed in different trHbs of group II (Fig. 1, Supplemental Material).

The role played by residues Tyr-B10 and Gln-E7 in ligand stabilization is further stressed by mutational studies (10, 12, 15) and by results of Das et al. (17) on the simultaneously observed O-O and Fe-O stretching frequencies in \( C. eugametos \) and \( Synechocystis \) sp. trHbs. It should also be noticed that the trHb hydrogen-bonded network centered on Tyr-B10 is strongly reminiscent of that observed for the very high oxygen affinity Hb from the nematode \( Ascaris suum \) (18, 19). In these cases, extraordinarily low oxygen dissociation rates (0.004–0.0014 s\(^{-1}\)) result in a high oxygen affinity (0.004–0.005 mm Hg) (9, 10, 20).

The Fe-OO stretching frequency of \( P. caudatum \) oxy-trHb detected by resonance Raman spectroscopy indicates strong polar interactions (including hydrogen bonding) between the bound ligand and the nearby Tyr-B10 residue, implying slow dissociation and high ligand affinity (11). In addition, the two Fe-CO stretching frequencies observed in the resonance Raman spectra of both \( M. tuberculosis \) trHbN-CO and \( Ascaris \) Hb-CO derivatives indicate two conformers expected to display slow and rapid ligand dissociation rates, respectively (15, 21). The actual dissociation of bound \( O_2 \) or \( CO \) is attributed to a fast equilibrium between two conformers, with the ligand off-rate being determined either by the rate of conformer interconversion or by ligand dissociation from the conformer with the higher rate.

### Ligands Can Enter the Distal Heme Pocket through a Protein Matrix Tunnel

Because of the orientation of the CD-D region, the E-helix of trHb falls close to the distal face of the heme. Crowding by distal residues and their interactions with the heme block access to the distal site cavity through the classical E7 residue gate, typically achieved in vertebrate Hbs by His-E7 (22–25). Remarkably, however, a different route for ligand diffusion to/from the heme appears to be coded in trHb structures as a cavity network or tunnel through the protein matrix. In \( C. eugametos \) trHb and \( M. tuberculosis \) trHbN, the tunnel is composed of two roughly orthogonal branches converging at the heme distal site from two distinct protein surface access sites. On one hand, a 20-Å long tunnel branch connects the globin region nestled between the AB and GH corners to the heme distal site. On the other, a path of about 8 Å connects an opening in the protein structure between G- and H-helices to the heme. The tunnel branches display inner diameters in the 5–7 Å range for a ligand-accessible volume of 330–360 Å\(^3\) (1, 3). Residues lining the tunnel branches are hydrophobic and are substantially conserved throughout the trHb family, suggesting that the tunnel plays a functional role and is suited for small nonpolar ligand diffusion or storage. A study of ligand rebinding following photolysis of CO in \( C. eugametos \) or \( P. caudatum \) trHbs\(^4\) suggests that the tunnel/cavity network in these proteins may indeed act as a CO store whose filling strikingly affects ligand rebinding kinetics.

Protein cavities, accessible to xenon atoms in Mb, have been shown to act dynamically as CO secondary docking sites. This has led to the suggestion that protein cavities modulate ligand dynamics and reactivity (Refs. 26 and 27, and references therein). On the other hand, protein matrix tunnels connect the surface to active sites in Ni-Fe hydrogenases (28), in methane monooxygenase hydroxylase (29), and in carbon monoxide dehydrogenase (30). Tunnels serve for internal substrate translocation in some enzymes (31). Different residues may modulate ligand diffusion processes along the trHb tunnel. For instance, Phe-E15 of \( M. tuberculosis \) trHbN is observed in two distinct conformations in the crystal structure and may act as a gate controlling ligand diffusion along the main tunnel branch (3).

### trHbs Serve Diverse Functions

The functional roles of trHbs are little known and may be various. The trHb of the unicellular green alga \( C. eugametos \) is induced in response to active photosynthesis and is localized, in part, along the chloroplast thylakoid membranes (32). The soluble trHb of the cyanobacterium \( Nostoc commune \) is localized on the cytoplasmic face of the cell membrane and is expressed only under anaerobic conditions (33, 34). In addition, the gene encoding this trHb is coexpressed with genes of the nitrogen fixation complex (34). Oxygen supply to the mitochondria of the protozoan \( Paramecium \) is impeded by levels of CO sufficient to block trHb but not cytochrome oxidase (35).

There is a great deal of evidence that NO generated by nitric-oxide synthase II in macrophages controls the develop-

\(^4\) U. Samuni, D. Danskier, A. Ray, L. Moens, M. Guertin, and J. F. Friedman, manuscript in preparation.
ment of \textit{M. tuberculosis} infection in mouse and man and restricts the bacteria to a latent state (36). However, tuberculous infection is in a dynamic balance that teeters for years in a plasticity of the globin fold, presenting us with new general considerations highlight a previously unpredicted structural accessibility but also for local storage of O2 molecules. Notably, although packing defects in Mb and the tunnel in trHb may appear evolutionarily related, their topology and size, at internal ligand diffusion mechanisms different from those based on the E7 distal gate in Mb. Whereas packing defects with the presence of an elongated protein matrix tunnel hints may reflect biological functions distinct from O2 storage or simplified 2-on-2 version of the globin fold observed in trHbs

**Conclusion**

The currently available data indicate that the dramatically simplified 2-on-2 version of the globin fold observed in trHbs may reflect biological functions distinct from O2 storage or transport. In particular, the combination of a closed distal site with the presence of an elongated protein matrix tunnel hints that no longer express trHbN are severely impaired in their ability to metabolize NO \textit{in vitro}.\(^5\) Oxy-trHbN could scavenge NO \textit{in vitro} by promoting dioxygenation as observed in human oxy-Hb, oxy-Mb, and \textit{Escherichia coli} flavohemoglobin, in which NO is converted to nitrate by reaction with the oxygenated heme (37–39).

**REFERENCES**

1. Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., Moens, L., and Bolognesi, M. (2000) \textit{EMBO} J. 19, 2424–2434
2. Moens, L., Vanbeteren, J., Van de Peer, Y., Peeters, K., Kapp, O., Cavusniak, J., Goodman, M., and Vinogradov, S. N. (1996) Mol. Biol. Ecol. 13, 324–333
3. Milani, M., Pesce, A., Ouellet, Y., Ascenzi, P., Guertin, M., and Bolognesi, M. (2001) \textit{EMBO} J. 20, 3892–3909
4. Cole, S. T., Eglmeier, K., Parkhill, J., James, X. D., Thomson, N. R., Wheeler, B. A., Wittenberg, J. B., Ouellet, Y., Guertin, M., and Rousseau, D. L. (2000) Biochemistry 39, 1310–1321
5. M. Guertin, unpublished results.
Truncated Hemoglobins: A New Family of Hemoglobins Widely Distributed in Bacteria, Unicellular Eukaryotes, and Plants
Jonathan B. Wittenberg, Martino Bolognesi, Beatrice A. Wittenberg and Michel Guertin

J. Biol. Chem. 2002, 277:871-874.
doi: 10.1074/jbc.R100058200 originally published online November 5, 2001

Access the most updated version of this article at doi: 10.1074/jbc.R100058200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2002/02/01/277.2.871.DC1

This article cites 39 references, 17 of which can be accessed free at
http://www.jbc.org/content/277/2/871.full.html#ref-list-1