Allosteric Hemoglobin Assembly: Diversity and Similarity*

Published, JBC Papers in Press, June 2, 2005, DOI 10.1074/jbc.R50006200

William E. Royer, Jr.§, Hao Zhu‡, Thomas A. Gorr‡, Jason F. Flores**, and James E. Knapp††

From the ‡Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, §Hematology Division, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts 02115, ‡Institute of Veterinary Physiology, Vetsuisse Faculty of the University of Zurich, CH-8057 Zurich, Switzerland, and **Department of Biology, Pennsylvania State University, University Park, Pennsylvania 16802

Nature is remarkably opportunistic in molding existing proteins to meet the needs of an organism. Hemoglobins provide a fascinating case study of this process. It is now clear that the traditional role ascribed to hemoglobins, that of oxygen transport by circulating hemoglobins, is but one of the functions carried out by members of this superfamily that are united by a similar fold. Recent studies provide strong evidence in favor of primordial hemoglobins functioning as nitric-oxide dioxygenases (1), an activity of considerable importance in present day hemoglobins and myoglobins (2–4). Conversion of such heme-based enzymes to those specialized in the delivery of oxygen proceeded as molecular oxygen became more plentiful because of ongoing photosynthesis and was an especially important step for the development of larger organisms. Although some invertebrates use hemocyanins (5) or hemerythrins (6), hemoglobins are, by far, the most widespread oxygen-carrying molecules with examples in all five kingdoms of life (7). The purpose of this review is to explore the diversity and similarities of invertebrate allosteric hemoglobins and the implications of their architecture on regulation and evolution.

Tertiary Structure of Hemoglobins

All hemoglobins share a common tertiary structure, known as the globin fold (Fig. 1). Our understanding of variations of this fold, as well as the broad diversity of function and widespread nature, has been increasing in recent years with the discovery of truncated hemoglobins (8), hexahemoglobin (9), and even the discovery of two new human hemoglobins, neuroglobin and cytoglobin (10, 11). Despite marked amino acid sequence diversity, tertiary structural similarities are evident, which indicate that all hemoglobins are evolutionarily related (Fig. 1). Hemoglobin residues are conventionally designated by the homologous helical (A-H) or corner (AB-GH) position in sperm whale myoglobin, the first globin structure determined. The only point of direct ligation between the heme and protein moiety is the coordination between the heme iron and the proximal histidine at position F8. Ligands, between the heme and protein moiety is the coordination between

Vertebrate Hemoglobins

The most familiar hemoglobins are, of course, those found in mammalian erythrocytes. Mammalian circulating hemoglobins, assembled into tetramers from two copies each of α and β subunits, played a central role in the history of molecular biology in the 20th century. Pioneering crystallographic studies by Max Perutz provided the methods that revolutionized protein crystallography and also revealed, for the first time, structural transitions that underlie allosteric protein behavior, including large quaternary structural changes (14). This work highlighted roles of the stereochemistry on not only the distal side of the heme but also on the proximal side of the heme in regulating oxygen affinity (14–16). Despite many elegant models and very extensive study, there is no universal agreement concerning the mechanism of cooperativity by human hemoglobin. One intriguing recent finding is that of Ackers and colleagues suggesting cooperativity within each αβ dimer (17). Although controversial, this idea finds support from recent structural results (18) and suggests an intriguing parallel with cooperatively invertebrate hemoglobins, which are assembled from dimeric units that likely possess intrinsic cooperativity.

Invertebrate Allosteric Hemoglobins

Unlike the circulating hemoglobins from higher vertebrates, which invariably display the α2β2 tetrameric form observed in human hemoglobin, invertebrate hemoglobins exhibit remarkable variation in their quaternary assembly. Crystal structures are now available of hemoglobin assemblages ranging in size from dimers to assemblies of 180 subunits (Fig. 2). Generally, hemoglobin assembly into oligomers is coupled with cooperative oxygen binding, but a notable exception is the tetrameric hemoglobin found in the “fat innkeeper” worm Urechis caupo (19).

An intriguing finding from a comparison of allosteric invertebrate hemoglobin assemblages is the recurring presence of a similar dimeric unit, termed an “EF dimer” because of extensive interface contacts involving the E and F helices (13). Isolated EF dimeric hemoglobins have been observed in the mollusk, Scapharca inaequalvis (20), and the echinoderm, Caudiva arenicola (21). The Scapharca HbI homodimer shows significant cooperative ligand binding (22), whereas strong cooperativity requires heterodimeric forms of Caudiva hemoglobin (21, 23). S. inaequalvis also possesses a cooperative tetrameric hemoglobin, which is assembled from two EF heterodimers (24) (Fig. 2).

Extracellular Annelid Allosteric Hemoglobins

Much larger allosteric hemoglobins are found among the annelids, which show striking variability in form. The crystal structures of two very large annelid hemoglobins have been reported. The muscular erythrocrurin, also termed hexagonal bilayer hemoglobin, from the common earthworm, Lumbricus terrestris, is assembled from 144 hemoglobin subunits and 36 non-globin “linker” chains using a hierarchy of symmetrical interactions (25). The hemoglobin subunits are organized into 12 dodecamers, each of which binds to a heterotrimer of linker subunits (25). Each dodecamer is a trimmer of heterotetramers, with each heterotetramer assembled from two distinct EF heterodimers (Fig. 3). The C1 hemoglobin from the deep sea hydrothermal vent tulipworm Riftia pachyptila is assembled from 24 subunits, each half of which forms a hemoglobin dodecamer structure that is very similar to the dodecamers observed in the L. terrestris erythrocrurin structure (26, 27).

The very similar quaternary assembly of dodecamers in 400-kDa vestimentiferan hemoglobin and 3600-kDa erythrocrurins raises interesting issues concerning the evolution of extracellular annelid hemoglobins. The erythrocrurins require assembly of multiple
copies of four different types of hemoglobin subunits and at least three distinct linker subunits\(^1\) organized in a complex hierarchy of symmetry (28, 29). The 400-kDa vestimentiferan hemoglobins show a much more straightforward arrangement, in which six copies of four distinct, but similar, hemoglobin subunits are arranged with D\(_3\) symmetry (26). Moreover, the hollow spherical assembly of the vestimentiferan geometry provides a structural rationale for the irregular, half-spherical shape of the hemoglobin dodecamers first observed in *Lumbricus* erythrocruorin (25, 27).

\(^1\) W.-Y. Kao, J. Qin, K. Fushitani, S. S. Smith, T. A. Gorr, C. K. Riggs, J. E. Knapp, B. T. Chait, and A. F. Riggs, submitted for publication.

---

**Fig. 1. Variability in the myoglobin fold.** A ribbon diagram is shown for each protein with the heme encapsulating E and F helices shown in cyan and the rest of the polypeptide course shown in gray. Also shown are the heme groups (red), oxygen or water ligand (yellow), proximal His(Fe), distal His(E7) or Gln(E7) (blue), and the highly conserved Phe(CD1) (gray). The E and F helices are labeled as are the N and C termini. Note the variability in structure, particularly in the length of the F helices, amino termini, and region just prior to the start of the E helix. A, sperm whale myoglobin (Protein Data Bank (PDB) code 1mbo) (38). B, *Scapharca* dimeric hemoglobin (PDB code 1hbi) (39). Note the longer F helix characteristic of most invertebrate hemoglobins, including those in EF dimer assemblages. C, *Paramecium* hemoglobin (PDB code 1dlw) (40). This truncated hemoglobin is about 75% of the length of myoglobin but still provides a site for the reversible binding of oxygen to the heme iron. Figure was produced with Pymol (46).

**Fig. 2. Quaternary assembly of oligomeric hemoglobins with known three-dimensional structure shown in a phylogenetic arrangement.** Dimers and tetramers are depicted as van der Waals spheres for main chain and heme atoms with heme groups shown in red, E and F helices in cyan, and the rest of the main chain in gray. The 24-subunit *Riftia* C1 hemoglobin is depicted with a main chain trace in color ranging from green to blue according to subunit type and hemes in red, whereas the 180-subunit *Lumbricus* erythrocruorin is shown in a surface rendition of its 5.5-Å electron density map with hemoglobin subunits in magenta and non-globin linker chains in blue and gold. Of the structures shown here, only *Urechis* hemoglobin does not exhibit cooperative oxygen binding. Note the similar assembly of subunits in hemoglobins with EF dimers, including echinoderm, mollusc, and subassemblies of annelid (see Fig. 3) hemoglobins. Structures include: human HbA (PDB code 2hhb) (42), deoxy lamprey HbV (PDB code 3lhb) (43), *Caudina* HbD (PDB code 1hlm) (21), *Lumbricus* erythrocruorin (25), *Riftia* C1 Hb (PDB code 1yhu) (28), *Urechis* Hb (PDB code 1ith) (19), *Scapharca* HbI (PDB code 3dth) (20), and *Scapharca* HbII (PDB code 1set) (24). Molecular images for this figure and Figs. 3 and 4 were produced using MIDAS (44), Ribbons (45), and BOBSCRIPT (41). (Some elements of this figure have been reproduced from Ref. 13 with permission from Elsevier.)
Cooperative Mechanisms within EF Dimers

In only one case, that of the homodimeric hemoglobin from *S. inaequivalvis*, has the structural basis for cooperativity in an invertebrate hemoglobin been investigated in detail. High resolution crystallographic analyses revealed that the ligand-linked transitions involve only small quaternary subunit movements but striking tertiary rearrangements (20). The observed transitions result in substantial functional changes with the high affinity R state estimated to bind oxygen ~300-fold more tightly than the low affinity T state (32). Thus, the EF arrangement of subunits including the heme in the interface permits strong modulation of ligand affinity with limited structural changes. Mutagenesis has confirmed the functional importance of three key aspects of the observed transitions, including residue F4, interface water molecules, and heme group movement. A phenylalanine at position F4 is critical for the functional difference between the low affinity (T) state and high affinity (R) state (33). In the T state, the side chain of Phe(F4) packs in the proximal pocket such that it restricts acquisition of high affinity stereochemistry, but this proximal strain is relieved by its extrusion into the subunit interface in the R state. The ligand-linked movement of Phe(F4) into the interface disrupts a well ordered cluster of interface water molecules that is essential for stabilization of the T state. Stability of the water cluster results from hydrogen bonding by main chain atoms, heme propionates, and Thr(E10). Disruption of the water cluster appears to be at least part of the signal by which one subunit detects the ligand state of its partner subunit (32). Ligand binding also results in a movement of the heme group deeper into each subunit, which is coupled with transitions at the interface involving the heme propionates, water molecules, and Lys(F3). Mutagenesis shows that this heme movement is required for the other ligand-linked transitions suggesting that this movement is the trigger for the allosteric transition (34). Similar ligand-linked heme movements are observed in human hemoglobin β-subunits (15) and hexacoordinate neuroglobin (35) indicating that such heme movement may be a widespread response to ligand binding.

Phylogeny of Hemoglobin Assemblages

Why is the EF dimer hemoglobin assemblage so widespread among invertebrates? It has now been observed in three different phyla (and suggested in a fourth (36)), including not only annelids and molluscs, which are thought to be closely related, but also in the deuterostome phylum of echinoderms (Fig. 2). In fact, there is currently no example available of a cooperative invertebrate hemoglobin that does not exhibit the EF dimer assemblage. Despite conservation of quaternary subunit arrangement, residues in the dimeric interface are remarkably variable. This variation becomes even more striking when considering those residues contributing to allosteric behavior. As discussed above, key players in the cooperative ligand binding behavior of *Scapharca* dimeric hemoglobin include the heme group, ordered interface water molecules, and residues Phe(F4), Lys(F3), and Thr(E10). These residues are conserved among EF dimeric hemoglobins within each phylum but not between different phyla, leaving only the F8-coordinated heme itself as the singular ubiquitous feature of EF dimer and all other hemoglobins (Fig. 4). Position E10 is occupied by an Arg in both annelid and echinoderm EF dimers and...
Minireview: Allosteric Hemoglobin Assembly

is likely to be important in the cooperative mechanisms (27). However, the different conformations of this residue in the liganded forms of annelid and echinoderm hemoglobins (Fig. 4) suggest that its contribution to cooperativity might be rather different in the two systems. Thus, despite a similar quaternary arrangement, the allosteric mechanisms used to modulate oxygen affinity are likely to be quite variable in these hemoglobins comprised of EF dimers.

The observed variability of interface residues among EF dimeric hemoglobins raises the issue of the origin of the EF dimer assemblies. Are these similar assemblies the result of divergent evolution, which all derive from some ancestral dimeric hemoglobin, or might the prevalence of this subunit pairing be an indication that dimerization at the EF face provides an efficient means for gaining selective advantage that has originated on multiple occasions?

To investigate this question, we have carried out phylogenetic analyses of a number of invertebrate and vertebrate hemoglobin sequences. Hemoglobin phylogeny has been a popular endeavor to explore the relationships in this omnipresent molecule; however, to our knowledge it has not been carried out previously to specifically address the relationship among the hemoglobins that exhibit EF dimer assemblers. We have used the program PAUP* Version 4.0 Beta (37) to reconstruct phylogenetic tree(s) with maximum parsimony. Twenty-one amino acid sequences from hemoglobins of known structure, including vertebrate tetramers (human, chicken, and monoc. Twenty-one amino acid sequences from hemoglobins of dimer assemblies. We have used the program PAUP* Version 4.0

address the relationship among the hemoglobins that exhibit EF analyses of a number of invertebrate and vertebrate hemoglobin dimerization at the EF face provides an efficient means for gaining selective advantage that has originated on multiple occasions.

11. Trent, J. T., III, and Hargrove, M. S. (2002) J. Biol. Chem. 277, 871–874

12. Bolognesi, M., Bordro, R., Rizzi, M., Tarriccone, C., and Ascenzi, P. (1997) Proc. Biophys. Mol. Biol. 68, 29–68

13. Royer, W. E., Jr., Knapp, J. E., Strand, K., and Heaslet, H. A. (2001) Trends Biochem. Sci. 26, 21–23

14. Perutz, M. F. (1970) Nature 228, 726–739

15. van Holde, K. E., and Miller, K. I. (1995) Adv. Protein Chem. 47, 1–81

16. Kavanaugh, J. S., Rogers, P. H., and Arnone, A. (2005) Biochemistry 44, 6101–6121

17. Ackers, G. K., Dalessio, P. M., Lew, G. H., Daugherty, M. A., and Holt, J. M. (1982) J. Biol. Chem. 257, 2753–2760

18. Kavanaugh, J. S., Rogers, P. H., and Arnone, A. (1996) Biochemistry 35, 1073–1084

19. Kolatkar, P. R., Hackert, M. L., and Riggs, A. F. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 14526–14531

20. Yuan, T., Hoang, J. H., Watanabe, H., and Asano, K. (2000) FEBS Lett. 478, 149–152

21. Mitchell, D. T., Kitto, G. B., and Hackert, M. L. (1995) J. Mol. Biol. 251, 421–431

22. Gianccone, E., Vecchini, P., Verzili, D., Ascoli, F., and Antonini, E. (1981) J. Mol. Biol. 152, 577–592

23. Bonaventura, C., Bonaventura, J., Kitto, B., Brunori, M., and Antonini, E. (1976) Biochim. Biophys. Acta 428, 779–786

24. Royer, W. E., Jr., Strand, K., van Heel, M., and Chiancone, E. (1995) J. Mol. Biol. 253, 168–186

25. Flores, J. F., Fisher, C. R., Carney, S. L., Green, B. N., Freytag, J. K., Schaeffer, S. W., and Royer, W. E., Jr. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 159–174

26. Flores, J. F., Fisher, C. R., Carney, S. L., Green, B. N., Freytag, J. K., Schaeffer, S. W., and Royer, W. E., Jr. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 159–174

27. Strand, K., Knapp, J. E., Bhyravbhatla, B., and Royer, W. E., Jr. (2004) J. Mol. Biol. 344, 119–134

28. Zhu, H., Ownby, D. W., Riggs, C. K., Nolasco, N. J., Stoops, J. K., and Riggs, C. K. (1996) Science 271, 25259–25267

29. Esnouf, R. M. (1999) Trends Plant Sci. 4, 159–174

30. Yuasa, H. J., Green, B. N., Takagi, T., Suzuki, N., Vinogradov, S. N., and Takagi, T. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 14526–14531

31. Negrisolo, E., Pallavicini, A., Barbato, R., Dewilde, S., Ghiretti-Magaldi, A., Moens, L., and Laurananchi, G. (1997) J. Biol. Chem. 272, 26391–26397

32. Royer, W. E., Jr., Pardanani, A., Gibson, Q. H., Peterson, E. S., and Friedman, J. M. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2424–2428

33. Pardanani, A., Gibson, Q. H., Colett, G., and Royer, W. E., Jr. (1997) J. Biol. Chem. 272, 13171–13179

34. Knapp, J. E., Gibson, Q. H., Cushing, L., and Royer, W. E., Jr. (2001) Biochemistry 40, 14795–14805

35. Killian, S., Kunz, J. R., and Taylor, M. (1999) J. Biol. Chem. 274, 18813–18818

36. Perutz, M. F. (1970) Nature 226, 555–565

37. Perutz, M. F., Wilkinson, A. J., Pauli, M., and Dodson, G. G. (1998) Annu. Rev. Biophys. Biomol. Struct. 27, 1–34

38. Phillips, S. E. (1980) J. Mol. Biol. 152, 1–34

39. Swofford, D. L. (2001) *Phylogenetic Analysis Using Parsimony (*PAUP*), Version 4, Sinauer Associates, Sunderland, MA

40. Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., Moens, L., and Bolognesi, M. (2000) EMBO J. 19, 2424–2434

41. Esnouf, R. M. (1999) *Acta Crystallographica. Section D. Biological Crystallography*, 55, 938–949

42. Ferrin, G., Perutz, M. F., Shaanan, B., and Fournier, R. (1984) *J. Mol. Biol.*, 175, 1–174

43. Heaslet, H. A., and Royer, W. E., Jr. (1999) *Structure Fold. Des.*, 7, 517–526

44. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics* 6, 13–27

45. Carson, M. (1997) *Methods Enzymol.*, 277, 493–505

46. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA

Acknowledgment—We thank Eric Fontano for help with elements of Figs. 2 and 3.