Annelid erythrocruorins are extracellular respiratory complexes assembled from 180 subunits into hexagonal bilayers. Cryo-electron microscopic experiments have identified two different architectural classes. In one, designated type I, the vertices of the two hexagonal layers are partially staggered, with one hexagonal layer rotated by about 16° relative to the other layer, whereas in the other class, termed type II, the vertices are essentially eclipsed. We report here the first crystal structure of a type II erythrocruorin, that from *Arenicola marina*, at 6.2 Å resolution. The structure reveals the presence of long continuous triple-stranded coiled-coil “spokes” projecting towards the molecular center from each one-twelth unit; interdigitation of these spokes provides the only contacts between the two hexagonal layers of the complex. This arrangement contrasts with that of a type I erythrocruorin from *Lumbricus terrestris* in which the spokes are broken into two triple-stranded coiled coils with a disjointed connection. The disjointed connection allows formation of a more compact structure in the type I architecture, with the two hexagonal layers closer together and additional extensive contacts between the layers. Comparison of sequences of the coiled-coil regions of various linker subunits shows that the linker subunits from type II erythrocruorins possess continuous heptad repeats, whereas a sequence gap places these repeats out of register in the type I linker subunits, consistent with a disjointed coiled-coil arrangement.

© 2006 Elsevier Ltd. All rights reserved.

**Keywords:** erythrocruorin; hemoglobin; HBL Hb; coiled coils; protein assembly

**Introduction**

In many annelids, oxygen transport relies on giant extracellular respiratory proteins (∼3.6 × 10⁶ Da), known as either erythrocruorins or hexagonal bilayer hemoglobins (HBL Hbs). Such large complexes offer a number of important advantages as oxygen transport vehicles: erythrocruorins are readily retained in the vascular system as freely dissolved entities, each complex possesses large oxygen binding capacity and subunits can be arranged to permit cooperative oxygen binding and additional regulatory features that enhance oxygen transport.

Abbreviation used: HBL Hb, hexagonal bilayer hemoglobin.

E-mail address of the corresponding author: William.Royer@umassmed.edu

Electron microscopic investigations dating back to the 1960 s established the overall shape of the erythrocruorins as consisting of two hexagonal layers. More recent investigations using cryo-electron microscopy with image reconstruction have revealed two distinct forms of erythrocruorins. In one, designated as type I, the vertices of the two layers are partially staggered with one hexagonal layer rotated relative to the other layer by about 16°. In the second form, designated as type II, the two halves of the molecule are essentially eclipsed. The type I architecture appears to be much more widespread, having been observed in erythrocruorins from oligochaetes (earthworm), achaetes (leech) and vestimentiferans (hydrothermal vent tube worm) and from two polychaete chlorocruorins. The type II erythrocruorin architecture has been observed in three polychaete species. The crystal structure of a type I erythrocruorin, that from the common earthworm *Lumbricus*
linker complexes of coiled-coil arrangements.10 The architecture of the two classes of annelid erythrocruorin is probably the most well studied of the known type II erythrocruorins, and recently has been proposed as a potential blood substitute.12 It shows cooperative oxygen binding, with maximum Hill coefficients above 4, and strong sensitivity to divalent cations and protons.13 Mass spectrometry experiments have identified eight distinct hemoglobin subunits, five of which have now been sequenced and two distinct linker subunits, one of which has been sequenced.14,15 As in the case of Lumbricus erythrocruorin, dissociation reveals both disulfide linked hemoglobin trimers and hemoglobin subunits not disulfide linked to other subunits.14

We have undertaken crystallographic analysis of Arenicola erythrocruorin in order to examine the structural determinants underlying the alternate architectures of the two classes of annelid erythrocruorins. The low resolution crystal structure presented here reveals long uninterrupted triple-stranded coiled coils formed from linker chains in type II erythrocruorin. Coupled with alternate coiled coil arrangements are striking differences in packing between hexagonal layers in type I and II architectures.

Results

Structure determination

The crystal structure of Arenicola erythrocruorin was determined by a combination of molecular replacement and non-crystallographic symmetry averaging. Use of a search molecule comprising one-half of Lumbricus erythrocruorin (excluding the long coiled coils) allowed placement of both halves of this 3.6 × 10^6 Da complex in the crystallographic asymmetric unit. This molecular replacement solution permitted calculation of preliminary phases at 8.5 Å resolution and determination of initial non-crystallographic symmetry operators relating the 12 protomers in the whole molecule. Molecular averaging, with phase extension to 6.2 Å, resulted in a readily interpretable map revealing the disposition of 144 hemoglobin subunits and clear rods of density corresponding to the 12 triple-stranded coiled-coil spokes of the linker subunits. In addition, density features for the non-helical linker LDL-A and β-barrel domains are consistent with the structures and disposition of these domains in Lumbricus erythrocruorin.

Overall structure of Arenicola erythrocruorin

The final electron density map at 6.2 Å resolution for one whole molecule is shown in Figure 1. Hemoglobin subunits, whose density is colored dark violet, occupy the surface of the molecule, with linker subunits (orange and blue) forming a central core. The molecule exhibits overall D₆ symmetry. Perpendicular to the molecular 6-fold symmetry axes are present; similarly to our earlier description of Lumbricus erythrocruorin,16 we designate these axes as P and Q (Figure 1). These molecular symmetry axes relate 12 protomers, each of which is composed of 12 hemoglobin and three linker subunits.

Arrangement of hemoglobin subunits

The hemoglobin subunits of Arenicola erythrocruorin are arranged into dodecamers that are very similar to the dodecamers found in Lumbricus erythrocruorin. A molecular 3-fold within a hemoglobin dodecamer relates three tetrameric units that, by analogy with Lumbricus erythrocruorin,11 most likely correspond to a heterotetramer. The arrangement of hemoglobin subunits is similar to that in Lumbricus erythrocruorin. The 12 highest electron density peaks in each protomer, all above 6.8σ, correspond to the 12 heme iron atoms. Iron peaks are present in pairs, 19–20 Å apart; these pairs correspond to the iron atoms within hemoglobin EF dimeric assemblies that are characteristic of all cooperative invertebrate hemoglobins investigated to date.17

Although the structure of individual hemoglobin dodecamers from Lumbricus and Arenicola erythrocruorin are very similar, arrangements of dodecamers in the two whole molecules are different. Our maps of Arenicola erythrocruorin confirm earlier electron microscopic results,2 indicating that the hemoglobin dodecamers in the two hexagonal layers are in an eclipsed orientation (Figure 2). In contrast, hemoglobin dodecamers from the two hexagonal layers of Lumbricus erythrocruorin are partially staggered. (Density shown for Lumbricus erythrocruorin was calculated at 6.2 Å resolution, following averaging procedures similar to those used for the Arenicola erythrocruorin maps.) Using the orientation shown in Figure 2(a), transition from Arenicola to Lumbricus erythrocruorin involves rotation of the top layer clockwise by 16° and translation of the two hexagonal layers closer together by ∼15 Å (Figure 2(b) and (c)). This brings Lumbricus hemoglobin b
subunits from different hexagonal layers into proximity at the P dyad. Thus, the type I architecture is more compact than the type II architecture. Despite the closer approach of layers, hemoglobin subunits from one layer do not contact hemoglobin or linker subunits from the apposing layer in either type I or type II architecture.
Figure 2. Hemoglobin dodecamers in Arenicola and Lumbricus erythrocruorin. Density at 6.2 Å resolution is displayed at the 1.5σ level for Arenicola erythrocruorin and 1.3σ level for Lumbricus erythrocruorin. (a) Views along the molecular 6-fold axis. Note the eclipsed arrangement of dodecamers in Arenicola erythrocruorin compared with the partially staggered arrangement in Lumbricus erythrocruorin. (b) Stereo view of Arenicola hemoglobin dodecamers along the P axis. (c) Stereo view of Lumbricus hemoglobin dodecamers along the P-dyad axis.
Protomer structure

The protomer (one-twelfth of a whole molecule) of Arenicola erythrocruorin consists of a hemoglobin dodecamer and a trimeric linker complex. Density maps for protomers of Arenicola and Lumbricus erythrocruorins are shown in Figure 3. Each protomer exhibits a local 3-fold axis of symmetry within the hemoglobin portion, with density essentially identical (correlation coefficient of 0.95) for three tetramers within a hemoglobin dodecamer. A quasi 3-fold axis relates the globular portion of three linker subunits (correlation coefficient of 0.70–0.71); this quasi 3-fold is coincident, within 0.3°, to that of the molecular 3-fold of the globin dodecamer 3-fold.

Alignment of the 3-fold axes assures similar interactions between the hemoglobin subunits and each linker subunit. Electron density is consistent with the LDL-A and β-barrel domains interacting with the hemoglobin subunits in much the same way as their counterparts do in Lumbricus erythrocruorin.

Since the globular portion of the linker subunits contains little α-helical structure, 6.2 Å resolution maps are not sufficient to interpret their structure. However, the known 3.5 Å Lumbricus erythrocruorin structure11 does allow approximate placement of the LDL-A and β-barrel linker domains within the Arenicola erythrocruorin maps. In addition to density at the 1.5σ level that nicely outlines the main-chain β-barrel structure, three density peaks above

---

**Figure 3.** Stereo views of one-twelfth protomers. Surface density representation is shown for Arenicola erythrocruorin at 1.5σ and Lumbricus erythrocruorin at 1.3σ. Density corresponding to the hemoglobin dodecamer is depicted in dark violet, that corresponding to the LDL-A and β-barrel domains is shown in blue and that for the coiled coils are in yellow. An approximate local 3-fold (vertical) relates the hemoglobin subunits along with the LDL-A, β-barrel domain and, for Lumbricus, the short coiled coil. The similar structures for the two erythrocruorins break down at the coiled coils (yellow).
5.5σ clearly pinpoint the calcium ions bound to the LDL-A domains. These peaks suggest slightly different orientations of the LDL-A modules in *Arenicola* and *Lumbricus* erythrocruorins that presumably result from their close proximity to the coiled-coil helices (see below).

The major difference between protomers of *Arenicola* and *Lumbricus* erythrocruorins occurs in the triple-stranded coiled coils (shown in yellow in Figure 3). In both molecules, each coiled coil projects towards the molecular center from the globular portion of the protomer. Density in *Arenicola* erythrocruorin indicates a continuous ∼70 Å long coiled coil. In contrast, the coiled-coil density in *Lumbricus* erythrocruorin is broken up into a long (∼45 Å) coiled coil and a shorter (∼20 Å) coiled coil within the globular portion of the linker head region.

The density for the coiled-coil helices also possesses a quasi 3-fold axis; however, the coiled-coil 3-fold is inclined in *Arenicola* erythrocruorin relative to those of the dodecamer and linker globular portion by 15°. This inclination angle is larger than, but oriented in a similar direction to, the 9° angle observed between the long coiled coil and hemoglobin dodecamer 3-fold in *Lumbricus* erythrocruorin. (The short coiled-coil in *Lumbricus* erythrocruorin possesses a quasi 3-fold that is essentially coincident with that of the hemoglobin dodecamer.) As discussed below, the different arrangements of coiled coils in *Arenicola* and *Lumbricus* erythrocruorins have profound implications for the overall architecture of these complexes.

Central linker complex

The differences between type I and II architectures in arrangement of the hemoglobin dodecamers (Figure 2) are evident in, and indeed dictated by, the arrangement of the 36 linker subunits at the core of these complexes. Views along the 6-fold axis (Figure 4(a)) show the linker globular region of *Arenicola* erythrocruorin to be slightly staggered between layers, but with the local 3-fold oriented such that bound hemoglobin dodecamers attain an eclipsed arrangement. (The local 3-fold axes of the linker head region and dodecamers do not project radially from the center of the complex). The effect of the linker coiled coils on the overall type I and type II architectures can be appreciated from views of the complete linker complex along the P-dyad (Figure 4(b) and (c)). The continuous coiled coils of *Arenicola* erythrocruorin limit the extent of contacts between the two hexagonal layers. Introduction of a break in the coiled coils, as observed in *Lumbricus* erythrocruorin, is coupled with rotation of one layer about the 6-fold axis and translation of layers closer together (∼15 Å). This results in a more compact linker complex and greater contacts between layers in *Lumbricus* erythrocruorin. Contributing greatest to the difference is *Lumbricus* linker chain L1, whose connection between long and short coiled-coil helices exhibits the longest inter-helical segment and whose β-barrel domains participate in extensive contacts at the Q-dyad of the bilayer interface.11 These bilayer interactions likely provide greater stability for the central linker complex in *Lumbricus* erythrocruorin. It appears, at this resolution, that the bilayer interactions are limited to the coiled coils themselves in *Arenicola* erythrocruorin, with no contacts between bilayers involving the β-barrel linker domain. The less extensive bilayer interactions in *Arenicola* erythrocruorin may contribute to its lower stability at alkaline pH compared with *Lumbricus* erythrocruorin.18,19

Discussion

Design of extracellular oxygen carriers requires a balance between large size, permitting a large number of oxygen binding sites and vascular retention, and adequate solution properties. It was shown many years ago that the blood viscosity in *Arenicola marina* is less than 25% of that predicted for a linear polymer of comparable molecular mass.20 This emphasizes the critical role of the *Arenicola* erythrocruorin shape, elucidated here, which permits the binding of 144 oxygen molecules to a single 3.6 MDa protein complex.

The crystal structures of *Lumbricus* and *Arenicola* erythrocruorins demonstrate the central role of linker subunits in dictating the relative orientation of the two hexagonal rings. Most important in this regard are the arrangements of the triple-stranded coiled coils. Linker coiled-coil sequences from different annelids (Figure 5) are readily aligned based on the characteristic pattern of seven residue (heptad) repeats in coiled coils in which hydrophobic residues generally occupy the first and fourth positions. These positions, designated a and d,21 line one side of a helix such that hydrophobic contacts form between coiled-coil helices. Two different classes of sequences are evident from the linker sequence alignment; erythrocruorins from polychaetes with type II architecture show continuous heptad repeats, whereas those sequences from type I erythrocruorins generally show a two or three residue gap relative the type II sequences. (The Riftia LY exception is discussed below.) The effect of the type I sequence gap is to place sequences on either side of the gap out of register. As a result, such sequences are incompatible with a continuous coiled coil. Rather, in *Lumbricus* erythrocruorin, two distinct triple-stranded coiled coils are formed on either side of the gap.11 In contrast, the coiled-coil sequences of type II linker subunits are fully compatible with the continuous coiled coils observed in the *Arenicola* erythrocruorin structure presented here.

One exception to the above sequence pattern is evident in Figure 5. The sequence of the Riftia LY subunit does not display the break found in other type I sequences, even though cryo-electron microscopic studies of this vestimentiferan...
Figure 4 (legend on opposite page)
erythrocruorin show it to have type I architecture. While this is the only Riftia pachyptila linker sequence available that encompasses the coiled-coil region, a partial sequence of chain LX is also available. The Riftia LX sequence is very similar to the AV-1 sequence of the closely related vestimentiferan Lamellibrachia, with a sequence identity of 87% for 123 residues in the C-terminal β-barrel domain. Such high homology in the β-barrel domain suggests that the two chains are likely to be similar in remaining portions. Since Lamellibrachia AV-1 exhibits a sequence gap that is similar to that found in other type I hemoglobins, it is appears likely that the Riftia LX subunit will also have a gap within the coiled-coil region. This suggests that the type I architecture might not require that all linker subunits exhibit such a sequence gap within the coiled-coil region. In contrast, formation of uninterrupted coiled coils, as found in type II architecture, probably does require continuous heptad repeats.

The structure and sequences of the type I and type II linker subunits raise the possibility that the type II architecture may represent a more primitive assembly. The disjointed connection between coiled coils in Lumbricus erythrocruorin suggests a greater degree of specialization among linker subunits, since continuous coiled coils could be formed from three identical subunits. Moreover, the β-barrel domains of Lumbricus L1 chains have acquired the capacity to form extensive contacts between hexagonal layers, a property not shared with the other Lumbricus linker subunits. However, the Arenicola linker subunits do have distinctions, even if less striking than those of the Lumbricus linkers. Lateral contacts between protomers likely involve unique interactions between different linker subunits, as
they do in *Lumbricus* erythrocruorin. Additional distinctions are present in the breakdown of quasi-3-fold symmetry in the linker region of the protomer structure. Whereas the density for the LDL-A and β-barrel domains possesses a quasi-3-fold axis that is aligned with the dodecamer 3-fold axis, the quasi-3-fold axis of the coiled coil is offset. This offset presumably results from distinct connections between the coil helices and LDL-A modules. Interestingly, modeling experiments (data not shown) suggest that it would be possible for coiled-coil spokes to project from each protomer with their 3-fold aligned along the dodecamer 3-fold axis without causing interpenetration between spokes. Such an arrangement might be compatible with assembly from just a single type of linker subunit, if subunit flexibility allowed quasi equivalence as is observed in many spherical viruses. One can speculate that evolution of at least two unique linker paralogs, as present in *Arenicola* erythrocruorin, has enhanced the efficiency of subunit assembly. Evolution of additional differences between linker subunits, then, would allow formation of the more compact type I architecture with larger inter-layer contacts, perhaps leading to greater stability of the linker core.

**Materials and Methods**

**Protein isolation and purification**

*Arenicola marina* worms were purchased from the Marine Biological Laboratory (Woods Hole, MA). Worms were cut with a longitudinal slit on the dorsal side and extruded viscera were rinsed with distilled water. Blood was drawn from large dorsal blood vessels using a 1cc tuberculin syringe (26 gauge needle). Erythrocruorin was purified from the blood by repeated ultracentrifugation twice (120,000 g) to pellet the 3.6 × 10^6 Da complex. This was followed by resuspension in 10 mM Hepes (pH 7.5), 100 mM NaCl, 10 mM CaCl_2, to obtain a concentration of ∼100 mg/ml for crystallization.

**Crystallization and data collection**

Crystals of carbon monoxide (CO) liganded *Arenicola* erythrocruorin were grown at 4 °C by vapor diffusion against a solution of 30% (w/v) PEG 400, 13 mM CaCl_2, 150 mM NaCl, 0.1 M Hepes (pH 7.5), and were successfully flash frozen from this solution on a cryo-loop in liquid nitrogen. Diffraction data were collected from two crystals at BioCARS beamline 14-BMC at the Advanced Photon Source (APS), Argonne National Laboratory, Argonne, IL. Observable diffraction from these crystals was limited to 0.5°, were processed with HKL2000 and scaled with SCALEPACK. Statistics are provided in Table 1.

**Phasing**

Molecular replacement calculations were carried out using a search molecule created from the atomic model for one hexagonal ring of *Lumbricus* erythrocruorin, but leaving out coordinates for the long coiled coils. Rotation and translation calculations against the *Arenicola* erythrocruorin diffraction data to 8.5 Å using CNS were readily successful at locating the two halves of the molecule in the asymmetric unit.

Initial calculated phases at 8.5 Å resolution were improved by six cycles of 12-fold averaging with the RAVE package of programs using a protein envelope based on the molecular replacement model. This averaged map permitted improved mask definition including portions of the coiled-coil region. The process was iterated several times until a satisfactory envelope was obtained encompassing an entire one-twelfth protomer including the full triple-stranded coiled coil. Phasing was extended from 8.5 Å to 6.2 Å by small incremental increases in resolution using six cycles of 12-fold averaging at each of 22 resolution steps. Based on this 6.2 Å electron density map, the mask envelope and non-crystallographic symmetry matrices were improved and the phase extension procedure was carried out a second time. Final averaging at 6.2 Å led to an averaged map exhibiting an R-factor of 21.6% and a correlation coefficient of 0.946 between the observed structure factors and those calculated from this map.

For comparison, a 6.2 Å averaged electron density map was also obtained for *Lumbricus* erythrocruorin. Phases, a molecular envelope and non-crystallographic symmetry operators were obtained from the refined atomic coordinates of *Lumbricus* erythrocruorin. The electron density map obtained was subjected to six cycles of 24-fold molecular averaging (two whole molecules per asymmetric unit). The map obtained in this way exhibited an R-factor of 20.9% and a correlation coefficient of 0.909 between the observed structure factors and those calculated from this map.

**Alignment of linker sequences**

Linker sequences were obtained from the National Center for Biotechnology Institute (NCBI). The following sequences were used: L1 (AAF99389), L2 (ABB71122), L3 (ABB71123), and L4 (ABB71124) from *Lumbricus terrestris*; L1 (BAC82449) from *Macrobella decora*; L2 (CAB85356) and L3 (CAC57413) from *Sabella spallanzani*; L1 (P16222) from *Lamellibrachia*; L2 (CAJ00866) from *Arenicola marina*; L1 (CAJ00867) and L2 (CAJ00868) from *Alvinella pompejana*; L1 (P18207) and L2 (P18208) from *Tyrrellhynchus heterochaetus*; L2 (BA0100580) from *Neanthes (Nereis) diversicolor*; LX (CAJ00870) and LY (CAJ00871) from *Riftia pachyptila*. Alignments were made by hand, using the 3.5 Å crystal structure of *Lumbricus*
erythrocruorin as a guide. The coiled-coil domains were aligned using the absolutely conserved first cysteine residue of the LDL domain. It was found that the linker chain from *Rifia* LY and all the linker chains from type II complexes could be aligned with their type I counterparts if a two to three amino acid gap was inserted in the type I sequences. This gap was placed into the extended peptide region between the positions of the long and short helices observed in the structure of *Lumbricus* linker chains. The resulting alignment (Figure 5) showed that both type I and type II complexes share homology at the “a” and “d” positions such that one face of the helix is enriched for hydrophobic residues, as commonly found with coiled coils.

Acknowledgements

We thank Sagar Kathuria for his early crystallization experiments on *Arenicola* erythrocruorin and the BioCARS staff for assistance with data collection. This work was supported by NIH grant DK43323 (to W.E.R.). Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-Eng-38. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR07707.

References


*Edited by R. Huber*