Published in final edited form as:

J Mol Biol. 2007 January 5; 365(1): 226–236.

Low Resolution Crystal Structure of *Arenicola* Erythrocrucorin: Influence of Coiled Coils on the Architecture of a Megadalton Respiratory Protein

William E. Royer Jr^{*}, Michael N. Omartian, and James E. Knapp

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA

Summary

Annelid erythrocrucorins are extracellular respiratory complexes assembled from 180 subunits into hexagonal bilayers. Cryoelectron 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 erythrocrucorin, that from *Arenicola marina*, at 6.2Å resolution. The structure reveals the presence of long continuous triple-stranded coiled-coil “spokes” projecting towards the molecular center from each one-twelfth 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 erythrocrucorin 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 erythrocrucorins 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.

Introduction

In many annelids, oxygen transport relies on giant extracellular respiratory proteins (~3.6 x 10⁶ Daltons), known as either erythrocrucorins or hexagonal bilayer hemoglobins (HBL Hbs). Such large complexes offer a number of important advantages as oxygen transport vehicles: erythrocrucorins 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.

Electron microscopic investigations dating back to the 1960s established the overall shape of the erythrocrucorins as consisting of two hexagonal layers¹. More recent investigations using cryoelectron microscopy with image reconstruction have revealed two distinct forms of erythrocrucorins. 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

^{*}Corresponding author Email address: William.Royer@umassmed.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

form, designated as “Type II”, the two halves of the molecule are essentially eclipsed². The Type I architecture appears to be much more widespread, having been observed in erythrocrurins from oligochaetes (earthworm)^{3, 4}, achaetes (leech)⁵ and vestimentiferans (hydrothermal vent tube worm)⁶ and from two polychaete chlorocrurins^{7, 8}. The type II erythrocrurin architecture has been observed in three polychaete species^{2, 9, 10}.

The crystal structure of a Type I erythrocrurin, that from the common earthworm *Lumbricus terrestris*, has been reported at 3.5 Å resolution¹¹. *Lumbricus* erythrocrurin assembles from 144 oxygen-binding hemoglobin subunits and 36 structural “linker” subunits. The hemoglobin subunits are arranged into twelve dodecamers, which assemble onto a central core formed from linker subunits. The overall D₆ symmetry of the molecule, and arrangement of the hexagonal layers, is dictated by the central linker complex. Twelve interdigitated heterotrimeric coiled-coil “spokes”, formed from the amino termini of the linker subunits, project towards the center of the complex. A striking break separates a long (~40 Å) from a short (~25 Å) coiled coil in the linker complexes of *Lumbricus* erythrocrurin. Based on electron microscopic analysis, it has been proposed that the alternate Type I and Type II architectures may be coupled with differences in the coiled-coil arrangements¹⁰.

Arenicola marina erythrocrurin is probably the most well studied of the known type II erythrocrurins, and recently has been proposed as a potential blood substitute¹². It shows cooperative oxygen binding, with maximum Hill coefficients above 4, and strong sensitivity to divalent cations and protons¹³. 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* erythrocrurin, dissociation reveals both disulfide linked hemoglobin trimers and hemoglobin subunits not disulfide linked to other subunits¹⁴.

We have undertaken crystallographic analysis of *Arenicola* erythrocrurin in order to examine the structural determinants underlying the alternate architectures of the two classes of annelid erythrocrurins. The low resolution crystal structure presented here reveals long uninterrupted triple stranded coiled coils formed from linker chains in Type II erythrocrurin. 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* erythrocrurin was determined by a combination of molecular replacement and non-crystallographic symmetry averaging. Use of a search molecule comprising one-half of *Lumbricus* erythrocrurin (excluding the long coiled coils) allowed placement of both halves of this 3.6 million Dalton 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 twelve 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 twelve triple-stranded coiled coils 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* erythrocrurin.

Overall structure of *Arenicola* erythrocrurin

The final electron density map at 6.2Å 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 six-fold are dyad axes oriented every 30°. Two unique dyad axes are present; similarly to our earlier description of *Lumbricus* erythrocrurin¹⁶, we designate these axes as “P” and “Q” (Figure 1). These molecular symmetry axes relate twelve protomers, each of which is composed of twelve hemoglobin and three linker subunits.

Arrangement of hemoglobin subunits

The hemoglobin subunits of *Arenicola* erythrocrurin are arranged into dodecamers that are very similar to the dodecamers found in *Lumbricus* erythrocrurin. A molecular three-fold within a hemoglobin dodecamer relates three tetrameric units that, by analogy with *Lumbricus* erythrocrurin¹¹, most likely correspond to a heterotetramer. The arrangement of hemoglobin subunits is similar to that in *Lumbricus* erythrocrurin. The twelve highest electron density peaks in each protomer, all above 6.8σ, correspond to the twelve heme iron atoms. Iron peaks are present in pairs, 19–20Å apart; these pairs correspond to the iron atoms within hemoglobin EF dimeric assemblages that are characteristic of all cooperative invertebrate hemoglobins investigated to date¹⁷.

Although the structure of individual hemoglobin dodecamers from *Lumbricus* and *Arenicola* erythrocrurin are very similar, arrangements of dodecamers in the two whole molecules are different. Our maps of *Arenicola* erythrocrurin confirm earlier electron microscopic results² 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* erythrocrurin are partially staggered. (Density shown for *Lumbricus* erythrocrurin was calculated at 6.2Å resolution, following averaging procedures similar to those used for the *Arenicola* erythrocrurin maps.) Using the orientation shown in Figure 2a, transition from *Arenicola* to *Lumbricus* erythrocrurin involves rotation of the top layer clockwise by 16° and translation of the two hexagonal layers closer together by ~15Å (Figure 2b&c). This brings *Lumbricus* hemoglobin *b* subunits from different hexagonal layers into proximity at the P dyad¹¹. 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.

Protomer structure

The protomer (one-twelfth of a whole molecule) of *Arenicola* erythrocrurin consists of a hemoglobin dodecamer and a trimeric linker complex. Density maps for protomers of *Arenicola* and *Lumbricus* erythrocrurins are shown in Figure 3. Each protomer exhibits a local three-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 three-fold axis relates the globular portion of three linker subunits (correlation coefficient of 0.70–0.71); this quasi three-fold is coincident, within 0.3°, to that of the molecular three-fold of the globin dodecamer threefold. Alignment of the three-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* erythrocrurin.

Since the globular portion of the linker subunits contains little α-helical structure, 6.2Å resolution maps are not sufficient to interpret their structure. However, the known 3.5Å

Lumbricus erythrocrurin structure ¹¹ does allow approximate placement of the LDL-A and β -barrel linker domains within the *Arenicola* erythrocrurin maps. In addition to density at the 1.5σ level that nicely outlines the main-chain β -barrel structure, three density peaks above 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* erythrocrurin that presumably result from their close proximity to the coiled-coil helices (see below).

The major difference between protomers of *Arenicola* and *Lumbricus* erythrocrurins 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* erythrocrurin indicates a continuous ~ 70 Å long coiled coil. In contrast, the coiled-coil density in *Lumbricus* erythrocrurin 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 three-fold axis; however the coiled coil three-fold is inclined in *Arenicola* erythrocrurin 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 three-fold in *Lumbricus* erythrocrurin. (The short coiled coil in *Lumbricus* erythrocrurin possesses a quasi three-fold that is essentially coincident with that of the hemoglobin dodecamer.) As discussed below, the different arrangements of coiled coils in *Arenicola* and *Lumbricus* erythrocrurins 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 six-fold axis (Figure 4a) show the linker globular region of *Arenicola* erythrocrurin to be slightly staggered between layers, but with the local three-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 II architectures can be appreciated from views of the complete linker complex along the P-dyad (Figure 4b&c). The continuous coiled coils of *Arenicola* erythrocrurin limit the extent of contacts between the two hexagonal layers. Introduction of a break in the coiled coils, as observed in *Lumbricus* erythrocrurin, is coupled with rotation of one layer about the sixfold axis and translation of layers closer together (~ 15 Å). This results in a more compact linker complex and greater contacts between layers in *Lumbricus* erythrocrurin. 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* erythrocrurin. It appears, at this resolution, that the bilayer interactions are limited to the coiled coils themselves in *Arenicola* erythrocrurin, with no contacts between bilayers involving the β -barrel linker domain. The less extensive bilayer interactions in *Arenicola* erythrocrurin may contribute its lower stability at alkaline pH compared with *Lumbricus* erythrocrurin ^{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 mass²⁰. This emphasizes the critical role of the *Arenicola* erythrocrucorin 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* erythrocrucorin 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; erythrocrucorins from polychaetes with Type II architecture show continuous heptad repeats, whereas those sequences from Type I erythrocrucorins 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* erythrocrucorin, two distinct triple-stranded coiled coils are formed on either side of the gap¹¹. In contrast, the coiled-coil sequences of Type II linker subunits are fully compatible with the continuous coiled coils observed in the *Arenicola* erythrocrucorin 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 cryoelectron microscopic studies of this vestimentiferan erythrocrucorin⁶ 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 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* erythrocrucorin 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* erythrocrucorin¹¹. Additional distinctions are present in the breakdown of quasi three-fold symmetry in the linker region of the protomer structure. Whereas the density for the LDL-A and β -barrel domains possess a quasi three-fold axis that is aligned with the dodecamer three-

fold axis, the quasi three-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 viruses²². One can speculate that evolution of at least two unique linker paralogs, as present in *Arenicola* erythrocrurin¹⁴, 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). Erythrocrurin was purified from the blood by twice repeated ultracentrifugation (120,000 x g) to pellet the 3.6×10^6 Da complex. This was followed by resuspension in 10mM Hepes pH 7.5, 100mM NaCl, 10mM CaCl₂, to obtain a concentration of ~100mg/ml for crystallization.

Crystallization and Data collection

Crystals of carbon monoxide (CO) liganded *Arenicola* erythrocrurin were grown at 4°C by vapor diffusion against a solution of 30% PEG 400, 13mM CaCl₂, 150mM NaCl, 0.1M Hepes, pH 7.5, and were successfully flash frozen from this solution on a cryoloop 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 6Å at best, despite screening many crystals. Data from 195 frames, each with an oscillation angle of 0.5°, were processed with HKL2000 and scaled with SCALEPACK²³. 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* erythrocrurin¹¹, but leaving out coordinates for the long coiled coil. Rotation and translation calculations against the *Arenicola* erythrocrurin diffraction data to 8.5Å using CNS²⁴ were readily successful at locating the two halves of the molecule in the asymmetric unit.

Initial calculated phases at 8.5Å resolution were improved by six cycles of 12-fold averaging with the RAVE package of programs²⁵ 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 8.5Å to 6.2Å by small incremental increases in resolution using six cycles of 12-fold averaging at each of 22 resolution steps. Based on this 6.2Å 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 6.2Å 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 erythrocrurorin*. Phases, a molecular envelope and non-crystallographic symmetry operators were obtained from the refined atomic coordinates of *Lumbricus erythrocrurorin*¹¹. 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 (with the NCBI protein codes in parentheses) were used: L1 (**AAF99389**), L2 (**ABB71122**), L3 (**ABB71123**), and L4 (**ABB71124**) from *Lumbricus terrestris*; L1 (**BAC82449**) from *Macrobdella decora*; L1 (**CAB38536**) and L3 (**CAC37413**) from *Sabella spallanzanii*; AV1 (**P16222**) from *Lamellibrachia*; L2 (**CAJ00866**) from *Arenicola marina*; L1 (**CAJ00867**) and L2 (**CAJ00868**) from *Alvinella pompejana*; L1 (**P18207**) and L2 (**P18208**) from *Tylorrhynchus heterochaetus*; L2 (**BAA09580**) from *Neanthes (Nereis) diversicolor*; LX (**CAJ00870**) and LY (**CAJ00871**) fragments from *Riftia pachyptila*. Alignments were made by hand, using the 3.5 Å crystal structure of *Lumbricus erythrocrurorin* 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 *Riftia* 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 chains¹¹. The resulting alignment (Figure 6) showed that both Type I and 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 erythrocrurorin* and the BioCARS staff for assistance with data collection. This work was supported by NIH grant DK43323 to WER. Use of the Advanced Photon Source was supported by the U.S. 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

1. Roche J, Bessis M, Thierry JP. [Study of the plasmatic hemoglobin of some Annelidae with the electron microscope]. *Biochim Biophys Acta* 1960;41:182–4. [PubMed: 14438064]
2. Jouan L, Taveau JC, Marco S, Lallier FH, Lamy JN. Occurrence of two architectural types of hexagonal bilayer hemoglobin in annelids: comparison of 3D reconstruction volumes of *Arenicola marina* and *Lumbricus terrestris* hemoglobins. *J Mol Biol* 2001;305:757–71. [PubMed: 11162090]
3. Schatz M, Orlova EV, Dube P, Jager J, van Heel M. Structure of *Lumbricus terrestris* hemoglobin at 30 Å resolution determined using angular reconstitution. *J Struct Biol* 1995;114:28–40. [PubMed: 7772416]
4. Taveau JC, Boisset N, Vinogradov SN, Lamy JN. Three-dimensional reconstruction of *Lumbricus terrestris* hemoglobin at 22 Å resolution: intramolecular localization of the globin and linker chains. *J Mol Biol* 1999;289:1343–59. [PubMed: 10373371]
5. de Haas F, Biosset N, Taveau JC, Lambert O, Vinogradov SN, Lamy JN. Three-dimensional reconstruction of *Macrobdella decora* (leech) hemoglobin by cryoelectron microscopy. *Biophys J* 1996;70:1973–84. [PubMed: 8785356]
6. de Haas F, Zal F, Lallier FH, Toulmond A, Lamy JN. Three-dimensional reconstruction of the hexagonal bilayer hemoglobin of the hydrothermal vent tube worm *Riftia pachyptila* by cryoelectron microscopy. *Proteins* 1996;26:241–56. [PubMed: 8953646]

7. de Haas F, Taveau JC, Boisset N, Lambert O, Vinogradov SN, Lamy JN. Three-dimensional reconstruction of the chlorocruorin of the polychaete annelid *Eudistylia vancouverii*. *J Mol Biol* 1996;255:140–53. [PubMed: 8568862]
8. Lanzavecchia S, Wade RH, Ghiretti Magaldi A, Tognon G, Bellon PL. A two-exposure technique for ice-embedded samples successfully reconstructs the chlorocruorin pigment of *Sabella spallanzanii* at 2.1 nm resolution. *J Struct Biol* 1999;127:53–63. [PubMed: 10479617]
9. de Haas F, Zal F, You V, Lallier F, Toulmond A, Lamy JN. Three-dimensional reconstruction by cryoelectron microscopy of the giant hemoglobin of the polychaete worm *Alvinella pompejana*. *J Mol Biol* 1996;264:111–20. [PubMed: 8950271]
10. Jouan L, Marco S, Taveau JC. Revisiting the structure of *Alvinella pompejana* hemoglobin at 20 Å resolution by cryoelectron microscopy. *J Struct Biol* 2003;143:33–44. [PubMed: 12892724]
11. Royer WE Jr, Sharma H, Strand K, Knapp JE, Bhyravbhata B. Lumbricus erythrocrurorin at 3.5 Å resolution: Architecture of a megadalton respiratory complex. *Structure* 2006;14in press
12. Rousselot M, Delpy E, Drieu La Rochelle C, Lagente V, Pirow R, Rees JF, Hagege A, Le Guen D, Hourdez S, Zal F. *Arenicola marina* extracellular hemoglobin: a new promising blood substitute. *Biotechnol J* 2006;1:333–45. [PubMed: 16897713]
13. Ochiai T, Weber RE. Effects of magnesium and calcium on the oxygenation reaction of erythrocrurorin from the marine polychaete *Arenicola marina* and the terrestrial oligochaete *Lumbricus terrestris*. *Zoolog Sci* 2002;19:995–1000. [PubMed: 12362052]
14. Zal F, Green BN, Lallier FH, Vinogradov SN, Toulmond A. Quaternary structure of the extracellular haemoglobin of the lugworm *Arenicola marina*: a multi-angle-laser-light-scattering and electrospray-ionisation-mass-spectrometry analysis. *Eur J Biochem* 1997;243:85–92. [PubMed: 9030725]
15. Chabasse C, Bailly X, Rousselot M, Zal F. The multigenic family of the extracellular hemoglobin from the annelid polychaete *Arenicola marina*. *Comp Biochem Physiol B Biochem Mol Biol* 2006;144:319–25. [PubMed: 16766219]
16. Royer WE Jr, Strand K, van Heel M, Hendrickson WA. Structural hierarchy in erythrocrurorin, the giant respiratory assemblage of annelids. *Proc Natl Acad Sci U S A* 2000;97:7107–11. [PubMed: 10860978]
17. Royer WE Jr, Zhu H, Gorr TA, Flores JF, Knapp JE. Allosteric hemoglobin assembly: diversity and similarity. *J Biol Chem* 2005;280:27477–80. [PubMed: 15932877]
18. Daniel E, Lustig A, David MM, Tsfadia Y. Towards a resolution of the long-standing controversy regarding the molecular mass of extracellular erythrocrurorin of the earthworm *Lumbricus terrestris*. *Biochim Biophys Acta* 2003;1649:1–15. [PubMed: 12818185]
19. Rousselot M, Le Guen D, Zal F. Novel dissociation mechanism of a polychaetous annelid extracellular haemoglobin. *Febs J* 2006;273:1582–96. [PubMed: 16689943]
20. Snyder GK. Blood viscosity in annelids. *J Exp Zool* 1978;206:271–277.
21. Woolfson DN, Alber T. Predicting oligomerization states of coiled coils. *Protein Sci* 1995;4:1596–607. [PubMed: 8520486]
22. Caspar DL, Klug A. Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* 1962;27:1–24. [PubMed: 14019094]
23. Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* 1997;276:307–326.
24. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 1998;54(Pt 5):905–21. [PubMed: 9757107]
25. Kleywegt, GJ.; Jones, TA. From First Map to Final Model. In: Bailey, S.; Hubbard, R.; Waller, DA., editors. *Halloween ... Masks and Bones*. Science and Engineering Research Council Daresbury Lab; Warrington, UK: 1994. p. 59–66.
26. DeLano, WL. The PyMOL Molecular Graphics System. 2002. <http://www.pymol.org>
27. Suzuki T, Riggs AF. Linker chain L1 of earthworm hemoglobin. Structure of gene and protein: homology with low density lipoprotein receptor. *J Biol Chem* 1993;268:13548–55. [PubMed: 8514788]

28. Kao WY, Qin J, Fushitani K, Smith SS, Gorr TA, Riggs CK, Knapp JE, Chait BT, Riggs AF. Linker chains of the gigantic hemoglobin of the earthworm *Lumbricus terrestris*: Primary structures of linkers L2, L3, and L4 and analysis of the connectivity of the disulfide bonds in linker L1. *Proteins* 2006;63:174–187. [PubMed: 16425180]
29. Suzuki T, Vinogradov SN. Globin and linker sequences of the giant extracellular hemoglobin from the leech *Macrobdella decora*. *J Protein Chem* 2003;22:231–42. [PubMed: 12962323]
30. Pallavicini A, Negrisolo E, Barbato R, Dewilde S, Ghiretti-Magaldi A, Moens L, Lanfranchi G. The primary structure of globin and linker chains from the chlorocruorin of the polychaete *Sabella spallanzanii*. *J Biol Chem* 2001;276:26384–90. [PubMed: 11294828]
31. Suzuki T, Takagi T, Ohta S. Primary structure of a linker subunit of the tube worm 3000-kDa hemoglobin. *J Biol Chem* 1990;265:1551–5. [PubMed: 2295645]
32. Suzuki T, Takagi T, Gotoh T. Primary structure of two linker chains of the extracellular hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. *J Biol Chem* 1990;265:12168–77. [PubMed: 2373687]
33. Suzuki T, Ohta T, Yuasa HJ, Takagi T. The giant extracellular hemoglobin from the polychaete *Neanthes diversicolor*. The cDNA-derived amino acid sequence of linker chain L2 and the exon/intron boundary conserved in linker genes. *Biochim Biophys Acta* 1994;1217:291–6. [PubMed: 8148374]

Figure 1.

Arenicola erythrocyruorin. Electron density at 6.2Å resolution is represented at the 1.5σ level. Density corresponding to the oxygen-binding hemoglobin chains is shown in dark violet, density for linker subunits is shown in orange for the top hexagonal ring and blue for the bottom hexagonal ring. a) Whole molecule viewed along the molecular six-fold axis. b) Whole molecule viewed along the P-dyad axis, with the molecular six-fold axis vertical and the Q-dyad axis horizontal. c) Pair of one-twelfth protomer subunits in the same orientation as b. Figures 1–4 were produced with PYMOL ²⁶.

Figure 2. Hemoglobin dodecamers in *Arenicola* and *Lumbricus* erythrocrurin. Density at 6.2Å resolution is displayed at the 1.5σ level for *Arenicola* erythrocrurin and 1.3σ level for *Lumbricus* erythrocrurin. a) Views along the molecular six-fold axis. Note the eclipsed arrangement of dodecamers in *Arenicola* erythrocrurin compared with the partially staggered arrangement in *Lumbricus* erythrocrurin. b) Stereoview of *Arenicola* hemoglobin dodecamers along the P axis. c) Stereoview of *Lumbricus* hemoglobin dodecamers along the P-dyad axis.

Figure 3.

Stereoviews of one-twelfth protomers. Surface density representation is shown for *Arenicola* erythrocrucorin at 1.5σ and *Lumbricus* erythrocrucorin 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 shown in yellow. An approximate local three-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 erythrocrucorins break down at the coiled coils (yellow).

Figure 4.

Central linker complex of *Arenicola* and *Lumbricus* erythrocrurin. Density at 6.2Å resolution is displayed at the 1.5σ level for *Arenicola* erythrocrurin and 1.3σ level for *Lumbricus* erythrocrurin, with density for linkers in one hexagonal layer shown in orange and other layer in blue. a) Views along the molecular six-fold axis. Note the rotation of layers between *Arenicola* and *Lumbricus* complexes, which is coupled with the introduction of a break in the *Lumbricus* coiled coils. b) Stereoview of *Arenicola* linker complex viewed along the P axis. In *Arenicola* erythrocrurin, the continuous coiled coils limit inter-layer contacts involving the globular portion of the linker subunits. c) Stereoview of *Lumbricus* linker complex viewed

along the P-dyad axis. In *Lumbricus* erythrocrucrin, the break in the coiled coils, coupled with rotation and translation of layers, allows tight packing of β -barrel domains of linker subunit L1 between hexagonal layers.

	a	d	a	d	a	d	a	d	a	d	a	d	Type																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
<i>Lumbricus</i> L1	...YQERRFQYL	VKNQNLH	IDYLAKK	LHD	IEEEY	NK	LTH	--D	VDKKT	IRQLKAR	ISN	LEE	HH	C	I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
<i>Lumbricus</i> L2	...LFSDQLDPR	LGANAF	LIRLDRI	IEKLRTK	LDEAEK	--	ID	PEHFV	SEIDAR	VT	KIEG	TH	C		I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
<i>Lumbricus</i> L3	...PPAQSHDEI	IDKLIERT	KNKITS	ISHVESL	LDDRLD	--	PK	RIRKAGS	LRHR	VEE	LED	P	S	C	I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
<i>Lumbricus</i> L4	A	A	E	D	N	R	A	R	D	I	S	E	R	I	D	K	L	T	A	E	A	F	K	L	G	R	N	L	D	A	R	L	D	--	P	I	R	I	K	K	A	G	T	L	K	A	R	V	D	A	I	A	E	P	T	C		I																																																																																																																																																																																																																																																																																																																																																																																																																																																															
<i>Sabella</i> L1	...WGSSLTTAR	VDSQEM	RLDQ	LAKQL	DS	LEGS	LEG	DL	--	G	I	P	A	S	M	N	S	L	S	L	R	L	D	K	I	E	G	N	D	C		I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
<i>Sabella</i> L3	...PATMXKLDEN	AARLNRL	LVARAV	DVSD	TIKR	RQSE	EYE	--	D	A	E	V	Q	A	Q	N	L	Q	R	L	E	K	R	N	G	V	L	C		I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
<i>Macrobdella</i> L1	...NGSDLEDFK	LKL	RGGL	LLHV	D	K	G	L	E	K	L	E	K	H	S	K	--	F	H	P	D	D	I	Y	Q	A	L	A	Q	R	V	Q	D	A	E	E	--	H	C		I																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
<i>Lamellibrachia</i> AV1	...SVSDAMGAR	VDAQAWR	VDRL	T	K	Q	F	Q	A	I	S	D	A	A	D	T	S	I	G	--	A	A	K	S	G	G	D	I	A	R	H	M	L	N	S	H	L	D	D	H	W	C		I*																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
<i>Riftia</i> LY [†]	...VQVESDASR	KDV	N	D	T	R	L	N	D	L	Q	D	K	V	D	E	L	Q	G	I	L	E	E	R	G	A	I	R	D	Q	R	L	K	E	F	S	E	L	Q	S	R	V	A	K	L	K	G	S	R	C		I																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
<i>Alvinella</i> L1	...AVQVESHER	R	D	V	N	D	M	R	L	N	D	L	Q	G	K	I	E	E	L	Q	Q	T	L	D	E	R	A	K	T	R	D	Q	R	L	E	F	A	E	L	T	A	R	V	H	K	L	E	S	H	C		II																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
<i>Alvinella</i> L2	...AALVRAGAR	F	D	E	I	G	E	E	I	A	E	L	T	A	E	V	V	F	E	N	M	I	E	D	R	I	E	Q	A	E	A	R	H	A	Q	W	H	V	F	S	S	R	I	D	A	V	E	G	S	G		II																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
<i>Tylorrhynchus</i> L1	...ADPHIAEGR	L	A	N	Q	D	A	R	L	G	T	L	E	E	E	V	D	K	L	Q	H	K	Y	D	D	F	I	A	G	K	T	A	R	R	E	R	F	A	Q	L	K	D	R	V	W	G	L	E	A	H	H	C		II																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
<i>Tylorrhynchus</i> L2	...REWASVASK	A	D	S	Q	E	A	R	V	N	R	L	A	G	R	V	E	A	L	A	E	N	L	R	A	G	G	D	R	L	S	H	Y	K	S	E	F	R	E	L	F	R	V	D	E	L	E	G	N	G	C		II																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
<i>Neanthes</i> d. L2	...RQWAGIASK	A	D	A	Q	E	A	R	I	N	R	L	A	G	R	V	E	A	V	A	E	K	L	R	V	G	E	K	V	K	H	F	L	G	D	F	R	E	L	F	R	V	D	E	L	E	G	N	G	C		nd																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
<i>Arenicola</i> L2	...PAGNAETSAD	R	E	Q	S	H	T	L	D	E	L	T	H	Q	L	H	M	L	Q	Q	A	Y	D	T	G	M	G	R	V	D	D	V	M	E	D	M	D	L	S	H	R	I	A	D	H	E	K	E	H	C		II																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a

*The bilayer arrangement of *Lamellibrachia* erythrocrucorin has not been determined, however the closely related *Riftia* erythrocrucorin shows type I architecture.

[†]This sequence is a fragment including the N-terminal coiled coil and LDL-A domains. A partial sequence of linker LX is available that includes the last half of the LDL and the entire β -barrel domain. The lack of identity within the overlap of the LX and LY fragments indicates that these represent different polypeptides.

Figure 5.

Alignment of the coil-coiled domains from available annelid erythrocrucorin linker sequences. Two clearly demarcated groups are evident. Linker sequences from the oligochaete *Lumbricus terrestris*^{27, 28}, the achaete *Macrobdella decora*²⁹, the polychaete *Sabella spallanzanii*³⁰, and the vestimentiferan *Lamellibrachia*³¹ form one group. Alignment of this group with the remaining linker sequences requires introduction of a sequence gap, which has been placed between the two coiled coils observed in the 3.5 Å crystal structure of *Lumbricus* erythrocrucorin (*Lumbricus* linker helical segments are highlighted in light blue). The second group is formed from linkers of polychaetes *Arenicola marina* (**CAJ00866**), *Alvinella pompejana* (**CAJ00867** and **CAJ00868**), *Tylorrhynchus heterochaetus*³², and *Neanthes (Nereis) diversicolor*³³; and the vestimentiferan *Riftia pachyptila* (**CAJ00871**). Lack of a gap in these sequences allows the “a” and “d” positions of the heptad repeat to remain in register for the entire sequence shown, which is necessary for a continuous coiled coil to be formed, as found in the *Arenicola* erythrocrucorin crystal structure. Hydrophobic residues are shown in red and basic residues are shown in blue at the “a” and “d” positions.

Table 1

Statistics for the crystallographic data.

Sample	<i>Arenicola</i> HBL-CO
Space group	P2 ₁ 2 ₁ 2 ₁
Cell Constants	
a (Å)	339.5
b (Å)	362.6
c (Å)	454.6
Resolution range	40.0 – 6.2
(high resolution bin)	(6.42 – 6.2)
Reflection	114201 (9152)
Completeness (%)	88.3 (71.7)
Multiplicity	3.6 (2.3)
I/error (I)	10.2 (1.8)
Rsym	0.128 (0.442)