What are the origins and phylogeny of plant hemoglobins?

Serge N. Vinogradov,1,* David Hoogewijs2 and Raúl Arredondo-Peter3
1Department of Biochemistry and Molecular Biology; Wayne State University; School of Medicine; Detroit, MI USA; 2Institute of Physiology and Zürich Center for Integrative Human Physiology (ZIHP); University of Zürich; Zürich Switzerland; 3Laboratorio de Biofísica y Biología Molecular; Facultad de Ciencias; Universidad Autónoma del Estado de Morelos; Cuernavaca, Morelos, México

Land plants and algae are now represented by about 40 genomes. Although most are incomplete, putative globins appear to be present in all the ca. 30 land plant genomes and in all except one algal genomes. The globins have either the canonical 3/3 α-helical fold characteristic of vertebrate myoglobin (Mb) or 2/2 α-helical folds, characteristic of bacterial globins with a truncated Mb-fold. In view of the fairly complete picture of the globin superfamily that is now available from analyses of over 1,000 bacterial genomes and >200 other eukaryote genomes, it is now possible to seek answers to the following twin questions: what is the phylogenetic relationship of plant and algal globins to those of other eukaryotes and what is their likely bacterial origin? We summarize below the available results. Molecular phylogenetic analyses indicate that plant and algal 3/3 globins are related to bacterial flavohemoglobins and vertebrate neuroglobins. Furthermore, they also suggest that plant and algal 3/3 and group 1 2/2 Hbs originated from the horizontal gene transfers that accompanied the horizontal gene transfers that accompanied the two generally accepted endosymbioses of a proteobacterium and a cyanobacterium with a eukaryote ancestor. In contrast, the origin of the group 2 2/2 Hbs unexpectedly appears to involve horizontal gene transfer from a bacterium ancestral to Chloroflexi, Deinococcales, Bacilli and Actinomycetes. We present additional results which indicate that the shared ancestry is likely to be with the Chloroflexi alone.

All the known plant genomes (ca. 30) contain sequences coding for globins (Hbs), which have either a 3/3 or 2/2 α-helical folds. The former comprise symbiotic Hbs, including the legHbs (Lbs) and nonsymbiotic Hbs which are divided into two classes nsHb-1 (nsHbs) and nsHb-2.1-3 The plant 3/3 Hbs appear to be monophyletic, with the symbiotic Hbs likely originating from the nsHb-2 globins.4 The 2/2 Hbs of plants and of the related Chlorophytes and Rhodophytes (ca. 10 genomes) belong to groups 1 and 2.5 The topics we would like to address are the origins of the plant and Chlorophyte Hbs.

To answer this question, we need to provide a brief overview of our current understanding of the globin superfamily phylogeny in the light of the genetic information accrued over the last 15 years. With over a 1,000 bacterial genomes now known, about two thirds contain Hbs that belong to one of three families: flavohemoglobin (F), sensor (S) and truncated Mb-fold (T).6,7 Although the F and S family members evince the canonical 3/3 α-helical fold characteristic of vertebrate Hb,8,9 the T globins have an abbreviated 2/2 α-helical fold and exist in 3 structural classes.10,11 The globin superfamily has been even more successful in eukaryotes, where apart from unicellular parasitic eukaryotes, all eukaryote groups have Hbs, ranging from >80% of the more than 100 fungal genomes to 100% in about 100 metazoan genomes. It is obvious that the present day eukaryote Hbs, including plants, must share one or more common origin with the bacterial globins. A tentative model of Hb evolution proposed that the 3/3 and 2/2 eukaryote globin genes emerged in ancestral eukaryotes via horizontal gene transfer (HGT) of bacterial F...
mitochondria and plastids, respectively.12 Involving an α-globin that led to the formation of generally accepted endosymbiotic events and that the HGTs occurred via the two ν-β-globin families.5,14

Figure 1. Bayesian phylogenetic tree (Mr. Bayes 3.1.2; WAG model; 1 x 10⁵ generations; burnin 500) based on a MUSCLE alignment of 3 concatenated sequences representing plant, Chlorophyte and Stramenopile 3/3 and 2/2 globins with 3/3 vertebrate globins and representatives of bacterial F and T families. Branch support values are bayesian posterior probabilities. FHbs, flavohemoglobin; Fgbs, single domain F globins; Agna, Agnathan; Cygb, cytoglobin; Ngb, neuroglobin.

How do the plant and Chlorophyte Hbs fit into this scenario?

We carried out molecular phylogenetic analyses of alignments of plant and algal globins with representative vertebrate globins on one hand, specifically the Ngb and Cygb families,13 and with bacterial Hbs representing the F and T families.5,14

In the first instance, we found a close relationship between the plant 3/3 Hbs and the vertebrate Ngbs and of both to the bacterial F globins, a result compatible with the proposed model. Thus, a HGT from an ancestor of modern proteobacteria to the LECA (last eukaryote common ancestor) could represent the origin of vertebrate Ngs and of algal and plant 3/3 Hbs. This finding is underscored by the Bayesian tree shown in Figure 1, based on MUSCLE alignment of three concatenated sequences representing the various globin subfamilies (e.g., the main vertebrate groups are represented by human, bird and fish sequences). This tree also shows the expected similarity of the algal group 1 and 2 T Hbs to their bacterial counterparts. Furthermore, our analysis showed that the Chlorophyte group 1 T Hbs could have originated via HGT from a cyanobacterium to LECA.

Our findings regarding the plant, Chlorophyte and Stramenopile group 2 T Hbs were quite different: their origin unexpectedly appeared to involve HGT from a bacterium ancestral to one of the following bacterial phyla, the Chloroflexi, Deinococcales, Bacillli and Actinomycetes.5 Because of the size of the analyzed datasets, we were never able to obtain trees with significant branch support values. More recently we succeeded to corroborate the earlier results and refine them to show that the six known Chloroflexi group 2 T globins were indeed the closest relatives of the plant, algal and stramenopile T2 Hbs (Fig. 2). The phyllum Chloroflexi, known formerly as “Green Non-Sulfur” bacteria, does not have a clearcut position within the bacterial phylogenetic tree.15,16 The recent suggestion that the root of the tree of life lies close to the Chloroflexi,17 implies that the origin of algal, plant and stramenopile T2 Hbs may be quite ancient.

References
1. Garrocho-Villagés V, Gopalasubramaniam SK, Arredondo-Peter R. Plant hemoglobins: what we know 6 decades after their discovery. Gene Funct Genom 2007; 398:78-85.
2. Hoy JA, Hargrove MS. The structure and function of plant hemoglobins. Plant Physiol Biochem 2008; 46:371-9.
3. Smagghe B, Hoy JA, Persifeld R, Kundu S, Hargrove MS; Sarah G, et al. Correlations between oxygen affinity and sequence classifications of plant hemoglobins. Biopolymers 2009; 91:1083-96.
4. Kakar S, Hoffman FG, Storz JF, Fabian M, Hargrove MS. Structure and reactivity of hexacoordinate hemoglobins. Biophys Chem 2010; 152:1.
5. Vinogradov SN, Fernández I, Hoogewijs D, Arredondo-Peter R. Phylogenetic relationships of plant and algal 3/3 and 2/2 hemoglobins to bacterial and other eukaryotic hemoglobins. Molec Plant 2010; 1:107.
6. Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Guertin M, Gough J, et al. Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. Proc Natl Acad Sci USA 2005; 102:11385-9.
7. Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Gough J, Dewilde S, et al. A phylogenomic profile of globins. BMC Evol Biol 2006; 6:31.
8. Wu G, Wittenwright LM, Poole RK. Microbial globins: Adv Microbiol Physiol 2003; 47-255-310.
9. Freitas TA, Saito JA, Hou S, Alam M. Globin-coupled sensors, protoglobins and the last universal common ancestor. J Inorg Biochem 2005; 99:23-33.
10. Wittenberg JB, Bolognesi M, Wittenberg BA, Guerin M. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes and plants. J Biol Chem 2002; 277:871-4.
11. Yuleitch DA, Lecomte JT. A phylogenetic and structural analysis of truncated hemoglobins. J Mol Evol 2006; 2:196-210.
12. Vinogradov SN, Hoogewijs D, Bailly X, Mizuguchi K, Dewilde S, Moens L, et al. A model of globin evolution. Gene Funct Genom 2007; 398:132-42.
13. Burmester T, Haberkamp M, Mitz S, Roessler A, Schmidt M, Ehner B, et al. Neuroglobin and cytoglobin: genes, proteins and evolution. IUBMB Life 2004; 56:703-7.
14. Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Guertin M, Gough J, Dewilde S, et al. Phylogenetic profile of globins. BMC Evol Biol 2006; 6:31.
15. Vi S, Uwagushi RE, Bourne PE. Structural analysis of polarizing indels: an emerging consensus on the root of the tree of life. Biol Direct 2009; 4:30.
Figure 2. Bayesian phylogenetic tree (Mr. Bayes 3.1.2; WAG model; 5 x 10^6 generations; burnin 1.5 x 10^6) based on concatenating MAFFT and PROBCONS alignments of representative plant, algal and stramenopile T2 globins with representative bacterial T2s and using the Mycobacterium avium and M. marinum T3 globins as outgroups. Branch support values are bayesian posterior probabilities and approximate Likelihood-Ratio Test, aLRT values (italic, if >0.5) of PhyML analysis. FIR, Firmicutes; DEI, Deinococci; BACT, Bacteroidetes; EPR, Epsilon proteobacteria; CHLA, Chlamydia; CHLO, Chloroflexi; ACTI, Actinobacteria.