A Broad Genomic Survey Reveals Multiple Origins and Frequent Losses in the Evolution of Respiratory Hemerythrins and Hemocyanins

José M. Martín-Durán[1,*,†], Alex de Mendoza[2,3,†], Arnau Sebé-Pedrós[2,3,†], Iñaki Ruiz-Trillo[2,3,4], and Andreas Hejnol[1]

[1]Sars International Centre for Marine Molecular Biology, University of Bergen, Bergen, Norway

[2]Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta, Barcelona, Spain

[3]Departament de Genètica, Universitat de Barcelona, Barcelona, Spain

[4]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, Barcelona, Spain

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: chema.martin@sars.uib.no.

Accepted: July 1, 2013

Data deposition: GenBank/DDBJ/EMBL accession numbers of the sequences used in the phylogenetic analyses are indicated in the supplementary data, Supplementary Material online.

Abstract

Hemerythrins and hemocyanins are respiratory proteins present in some of the most ecologically diverse animal lineages; however, the precise evolutionary history of their enzymatic domains (hemerythrin, hemocyanin M, and tyrosinase) is still not well understood. We survey a wide dataset of prokaryote and eukaryote genomes and RNAseq data to reconstruct the phylogenetic origins of these proteins. We identify new species with hemerythrin, hemocyanin M, and tyrosinase domains in their genomes, particularly within animals, and demonstrate that the current distribution of respiratory proteins is due to several events of lateral gene transfer and/or massive gene loss. We conclude that the last common metazoan ancestor had at least two hemerythrin domains, one hemocyanin M domain, and six tyrosinase domains. The patchy distribution of these proteins among animal lineages can be partially explained by physiological adaptations, making these genes good targets for investigations into the interplay between genomic evolution and physiological constraints.

Key words: comparative genomics, hemerythrin, hemocyanin, tyrosinase, respiration, lateral gene transfer.

Introduction

Hemoglobins, hemerythrins, and hemocyanins are three different respiratory proteins present in animals (Terwilliger 1998). Hemoglobins have a Fe–protoporphyrin ring to reversibly bind oxygen and are the most common molecules for oxygen transport and storage in the Bilateria (Weber and Vinogradov 2001). Globin proteins are widespread in the tree of life and, in animals, respiratory globins likely evolved from a membrane-bound ancestor that acquired a respiratory function independently in different lineages (Roesner et al. 2005; Blank and Burmester 2012). In contrast to the widespread hemoglobins, hemerythrins and hemocyanins have been detected in fewer animal groups. Hemerythrins transport oxygen using two Fe^{2+} ions that bind directly to the polypeptide chain and have been described in a cnidarian (*Nematostella vectensis*), priapulids, brachiopods, some annelids, and sipunculans (Terwilliger 1998; Bailly et al. 2008). Recently, it has been shown that regulation of iron homeostasis in vertebrates involves an E3 ubiquitin ligase (*FBXL5* gene) with an iron-responsive hemerythrin domain in its structure (Salahudeen et al. 2009; Vashisht et al. 2009), although it is not clear how this hemerythrin domain-containing protein is related to invertebrate respiratory hemerythrins. Hemocyanins are large proteins that have copper-binding sites to transport oxygen in arthropods and molluscs (Bonaventura and Bonaventura 1980). Despite their shared name, arthropod and molluscan respiratory hemocyanins are considered to have evolved independently from a common ancestral copper

protein based on protein similarities (Burmester 2001; van Holde et al. 2001). A deeper understanding of the evolutionary history of hemerythrins and hemocyanins, and thus of the different respiratory strategies in animals, is limited by the absence of data for many invertebrate groups and, in particular, for unicellular holozoans and other eukaryotes. In this study, we survey a wide phylogenetic distribution set of genomes and RNAseq data to identify new hemerythrin and hemocyanin proteins, and we then reconstruct their evolution within the eukaryote tree of life, with particular focus on animal lineages.

In addition to the respiratory hemerythrin sequences previously characterized in animals (Vanin et al. 2006; Bailly et al. 2008; Meyer and Lieb 2010), we identified respiratory hemerythrins in the priapulid *Priapulus caudatus*, the arthropod *Calanus finmarchicus,* and the bryozoans *Alcyonidium diaphanum* and *Membranipora membranacea* (see supplementary table S1, Supplementary Material online). In bryozoans, no respiratory proteins have been previously described, despite the presence of a circulatory system in these animals (Schmidt-Rhaesa 2007). Differently from other animal hemerythrins, the hemerythrin domain shows a Ca^{2+}-binding EF-hand domain in its N-terminal region in both bryozoan species. Additionally, we identified an E3 ubiquitin ligase containing an F-box domain together with a hemerythrin domain, as in *FBXL5*, in cnidarians and across bilaterally symmetrical animals (see supplementary table S1, Supplementary Material online), suggesting an ancient origin of the iron-sensing system described in vertebrates. Our phylogenetic analyses show two major clades of hemerythrin-containing proteins (fig. 1), one comprising the metazoan *FBXL5* gene and most of the eukaryote hemerythrins (clade A) and the other comprising the metazoan respiratory hemerythrins (including the newly identified sequences from this study) (clade B). As shown in a previous report (Bailly et al. 2008), respiratory hemerythrins are closely related to some *Naegleria gruberi* hemerythrins and a sequence from the amoebozoan *Acanthamoeba castellanii*. To eliminate possible bacterial contamination, we checked the gene structure and confirmed that the *Acanthamoeba* gene has introns within the hemerythrin domain. The two major clades (A, nonrespiratory, and B, respiratory) are separated with high nodal support, which is in agreement with observed structural differences (Histidine 74 being only present in clade B) (Thompson et al. 2012). Interestingly, clade B hemerythrins seem to be more common and highly diversified in prokaryotes (French et al. 2008) than in eukaryotes. Metazoans may have acquired them during an ancient event of lateral gene transfer (LGT), but, according to our phylogeny, it is more likely that clade B hemerythrins are ancient and have been lost in many eukaryotic lineages, so far only present in three extant distant lineages: Amoebozoa, Excavata, and Metazoa. Clade B prokaryote hemerythrins have been shown to bind oxygen as metazoan respiratory hemerythrins (Xiong et al. 2000) but

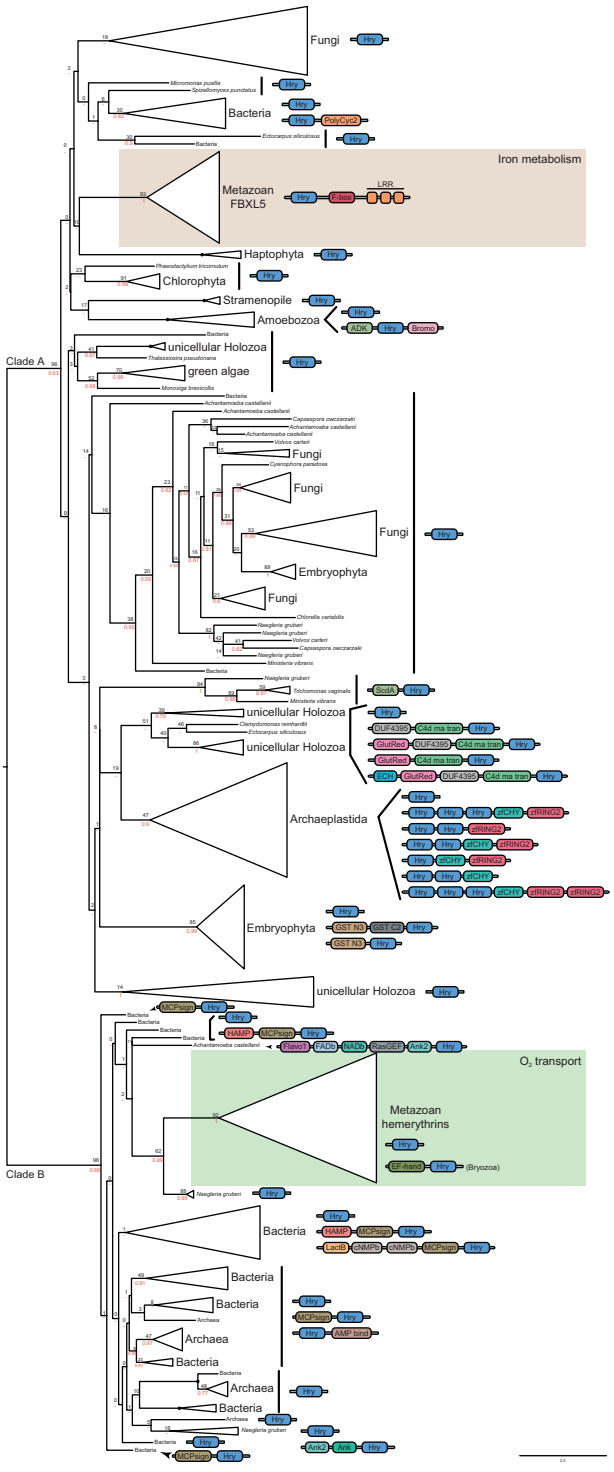

Fig. 1.—Maximum likelihood (ML) phylogenetic tree of the hemerythrin domain as obtained by RAxML. The tree is rooted using the midpoint-rooted tree option. 1,000 replicate bootstrap values (BV, in black) and BPP (in red) are shown for each node. A black dot in the node indicates BV >95% and BPP >0.95. Metazoan hemerythrins are highlighted by colored rectangles. Domain architectures are shown for major lineages (abbreviations and accession numbers of each domain are listed in supplementary table S2, Supplementary Material online).

have been mostly related to oxygen sensing and aerotaxis processes (Xiong et al. 2000; Isaza et al. 2006) and oxygen supply to other metabolic enzymes (Karlsen et al. 2005). This suggests that the oxygen storage and transport function of hemerythrins may have evolved independently in metazoans, given the lack of functional data for the Excavata and the Amoebozoa. Under this scenario, the role of respiratory hemerythrins in iron storage, metal detoxification, and immunity observed in some annelids (e.g., the leeches *Theromyzon tessulatum* and *Hirudo medicinalis* and the polychaete *Neanthes diversicolor*) (Baert et al. 1992; Demuynck et al. 1993; Vergote et al. 2004) are secondary specializations of this type of proteins. Finally, nonrespiratory hemerythrins (clade A) are quite common in eukaryotes and have recruited several companion domains in different lineages.

Searches for the hemocyanin M domain (arthropod hemocyanins) identified this copper-binding protein in amoebozoans, the fungus *Aspergilus niger*, the sponge *Amphimedon queenslandica*, the ctenophore *Mnemiopsis leidyi*, and the hemichordate *Saccoglossus kowalevskii* (see supplementary table S1, Supplementary Material online). Therefore it is likely to be a unikont synapomorphy. Our phylogenetic analyses show the monophyly of all metazoan sequences, as well as of the main arthropod protein families (fig. 2). The relationship of the sponge and fungal sequences may be due to an ancient LGT, though the presence of hemocyanins in the more distant amoeobozoans does not support that idea. Moreover the *A. niger* gene has a N-terminal intron and is located between two other fungal genes, making it less likely to come from a recently incorporated segment of metazoan DNA. Furthermore, a pseudogene with a hemocyanin M domain is present in the fungus *Neosartorya fischeri* (a congeneric species despite the name), but absent from all 6 other *Aspergillus* genomes and also from all the other fungi sequenced to date. The newly identified sequences in this study demonstrate that the N-domain of arthropod hemocyanins and related proteins is a specific molecular signature of the Panarthropoda (Onychophora + Arthropoda), although there is some degree of similarity of these regions in nonarthropod sequences. The presence of hemocyanin-like proteins in the tunicate *Ciona intestinalis* with putative phenoloxidase activity suggested that respiratory hemocyanins evolved from an ancestral prophenoloxidase (Immesberger and Burmester 2004). Given the absence of functional data for the nonbilaterian animals (i.e., the ctenophore *M. leidyi* and the sponge *A. queenslandica*) and our phylogeny (fig. 2), this is still the most parsimonious functional explanation for the evolution of the respiratory properties of arthropod hemocyanins.

An extensive search for the tyrosinase domain (molluscan hemocyanin) demonstrated a wide distribution of this copper-binding protein across metazoan lineages (see supplementary table S1, Supplementary Material online), with the remarkable exception of arthropods. The absence in arthropods could be associated with the expansion and diversification of the hemocyanin M domain in this lineage, which can exhibit similar activities to the tyrosinase domain, for example in melanin biosynthesis (Sugumaran 2002). In contrast to a previous analysis (Esposito et al. 2012), our phylogenetic reconstruction of a broader dataset shows that the animal tyrosinase domains group in six independent clades (clades A–F) (fig. 3), which are further supported by the domain architecture of the proteins nested in each clade (e.g., clade D and clade F). The tyrosinase domain that gave rise to the molluscan hemocyanins is related to brachiopod and tunicate sequences (clade B), and the series of duplications that lead to the typical arrangement of eight tyrosinase domains in tandem (Bonaventura and Bonaventura 1980) is specific to molluscs. With the exception of clade E, which is restricted to nonbilateral animals (fig. 3), the other clades exhibit an extremely patchy distribution across bilaterally symmetrical animals (see supplementary table S1, Supplementary Material online), not only between major animal groups but also within the same group (e.g., in molluscs, *Crassostrea gigas* has only a clade D tyrosinase, *Lottia gigantea* has clade A and D tyrosinases, and *Sepia officinalis* has both clade B and D tyrosinases). Despite the poor resolution of deeper nodes, our phylogenetic scenario at least strongly supports three independent origins of metazoan tyrosinases. Clades A, B, and F are well supported (PP $>$ 0.9) and nested with nonmetazoan sequences. The other three clades (clades C, D, and E) are not robustly supported, but have unique domain architectures and do not significantly cluster with other metazoan groups, therefore they might also come from independent origins. Moreover, our phylogenetic analysis demonstrates that the tyrosinase-containing proteins of plants likely originated due to a LGT event from bacteria, corroborated by these proteins exhibiting the same domain architecture (see fig. 3).

Altogether, our data clarify the origins and evolutionary history of the alternative respiratory strategies observed in animals (fig. 4). Respiratory hemerythrins, arthropod hemocyanins, and molluscan respiratory tyrosinases originated independently from enzymatic domains that were most likely already present in the last common metazoan ancestor. Although their function in early branching lineages that do not possess circulatory systems needs to be elucidated (e.g., the function of hemerythrins in the cnidarian *N. vectensis* or the hemocyanin M domain in sponges and ctenophores), the cooption of these domains for respiratory purposes occurred independently, and most likely took place at the base of the Protostomia (hemerythrin), the (Pan-)Arthropoda (arthropod hemocyanins), and the Mollusca (molluscan tyrosinase "hemocyanins"). Accordingly, the similarities observed between arthropod and molluscan hemocyanins (e.g., use of copper to reversibly bind oxygen as a respiratory strategy, oligomerization, and secretion to the hemolymph) are the result of convergent evolution. The evolutionary history of hemerythrins and hemocyanins is characterized by frequent losses, even after a respiratory function has been acquired when a

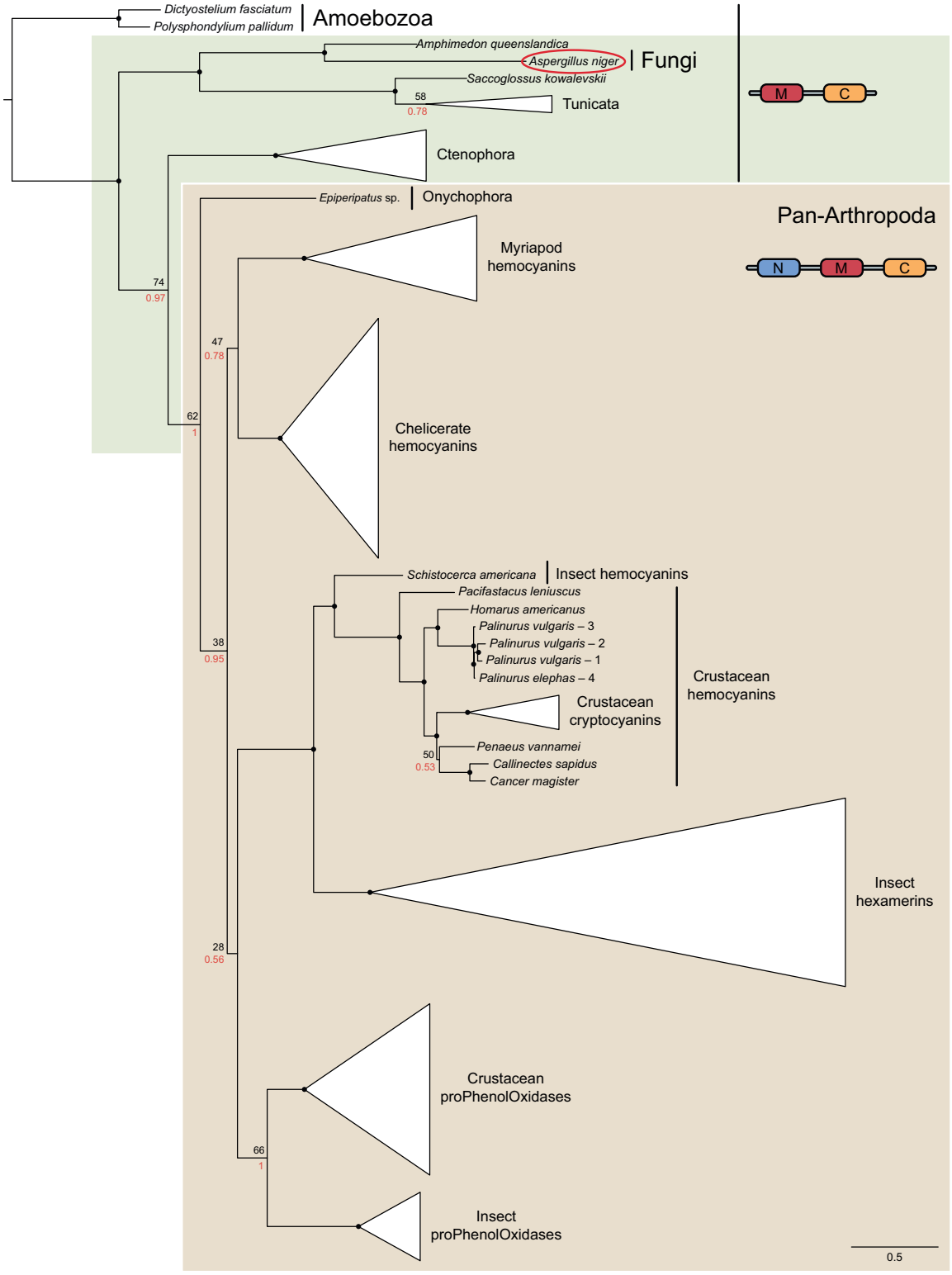

Fig. 2.—Maximum likelihood (ML) phylogenetic tree of the hemocyanin M domain as obtained by RAxML. The tree is rooted using the amoebozoan hemocyanin genes as outgroup. 1,000 replicate bootstrap values (BV, in black) and Bayesian posterior probabilities (BPP, in red) are shown for each node. A black dot in the node indicates BV >95% and BPP >0.95. Metazoan and panarthropod hemocyanins are highlighted by colored rectangles. The *Aspergilus niger* hemocyanin sequence is enclosed by a red circle. Domain architectures are shown for major lineages (abbreviations and accession numbers of each domain are listed in supplementary table S2, Supplementary Material online).

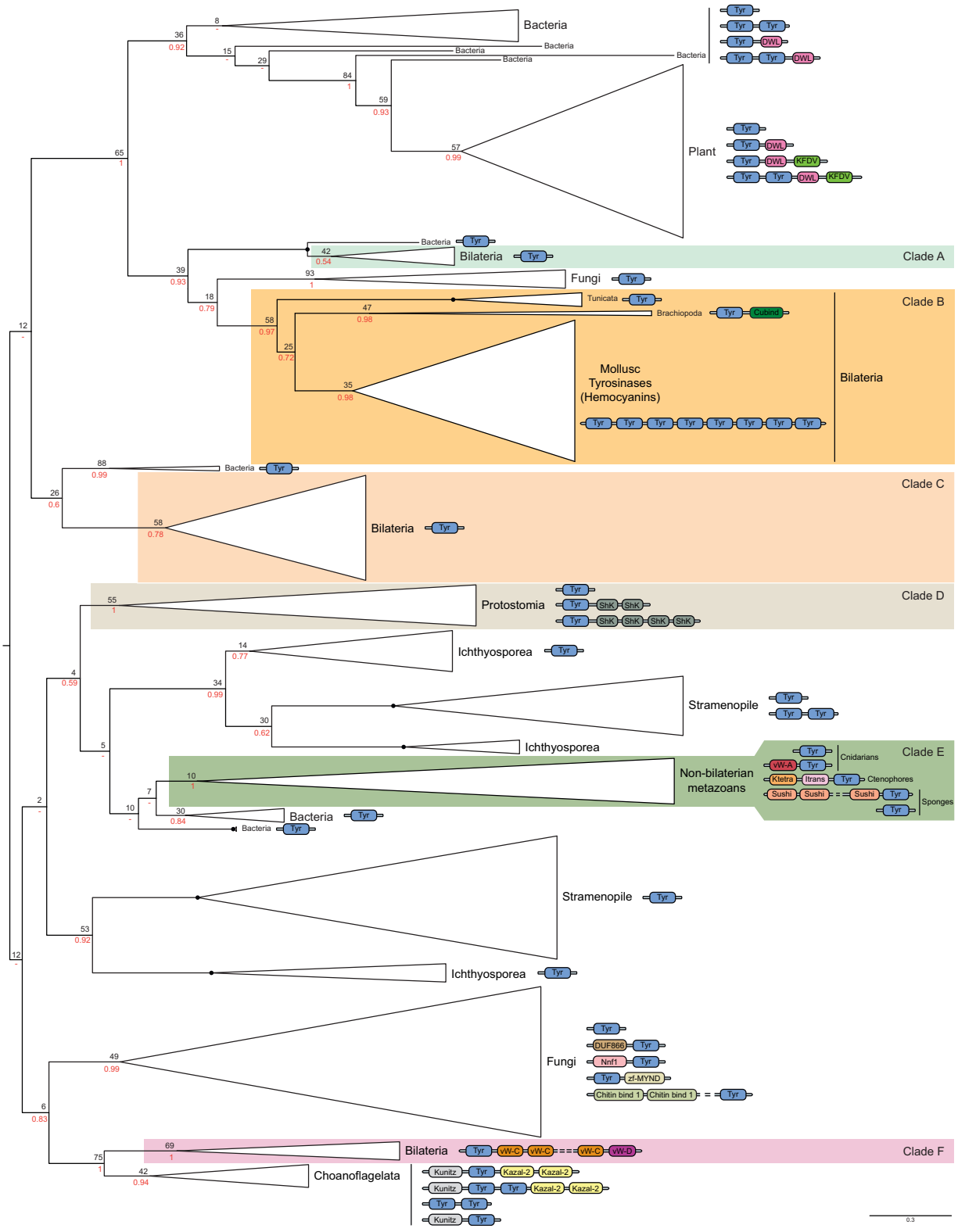

Fig. 3.—Maximum likelihood (ML) phylogenetic tree of the tyrosinase domain as obtained by RAxML. The tree is rooted using the midpoint-rooted tree option. 1,000 replicate bootstrap values (BV, in black) and Bayesian posterior probabilities (BPP, in red) are shown for each node. A black dot in the node indicates BV >95% and BPP >0.95. Metazoan tyrosinase clades are highlighted by colored rectangles. Domain architectures are shown for major lineages (abbreviations and accession numbers of each domain are listed in supplementary table S2, Supplementary Material online).

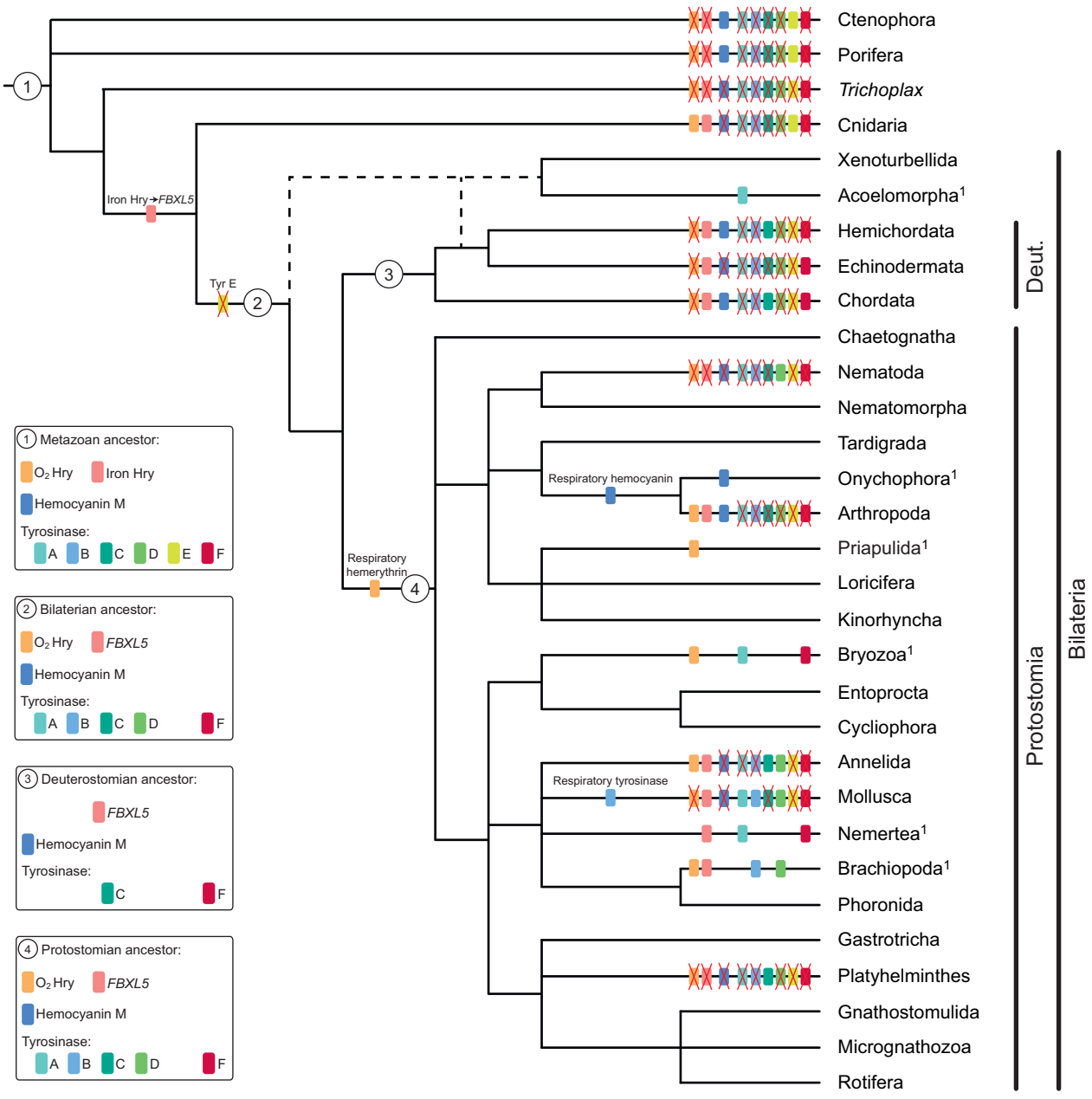

Fig. 4.—Scenario for the evolution of hemerythrins and hemocyanins in metazoans. As shown in this study, the metazoan ancestor likely had two different hemerythrin domains (oxygen subtype and iron-related subtype), one hemocyanin M domain and six independent tyrosinases (A–F subtypes). In the last common ancestor to cnidarians and bilaterians, the iron-related hemerythrin subtype evolved into *FBXL5*, an E3 ubiquitin ligase involved in iron homeostasis. The diversification of bilaterally symmetrical animals was accompanied by extensive independent losses of these proteins in the different lineages, which can be partially related to the colonization of new niches and environments. The adaptation of the oxygen hemerythrin subtype to respiratory purposes occurred at least in the last common ancestor of protostome animals. On the contrary, the adaptation of the hemocyanin M domain and the clade B tyrosinases to oxygen transport and storage occurred independently in panarthropods and molluscs, respectively. Results from groups indicated with a superscript 1 originated from RNAseq data.

higher selective pressure against loss could be expected. For instance, the use of tyrosinase as an oxygen transport molecule seems to be absent in some groups of molluscs, such as solenogasters and pteriomorphids (e.g., *C. gigantea*, as also shown in this study) (Lieb and Todt 2008), in which it was probably replaced by other respiratory proteins that have evolved independently in these lineages. This is the case for gastropods in the group Planorbidae, which lack hemocyanin in their hemolymph and which utilize an extracellular hemoglobin (evolved from an intracellular myoglobin present in the

gastropod radula muscle) as an alternative strategy for oxygen transport. This high molecular mass hemoglobin has a higher affinity for oxygen than the ancestral hemocyanin (Lieb et al. 2006). Similar adaptations are also observed within the Crustacea, such as in branchiopods, ostracods, copepods, cirripeds, and decapods, which lost hemocyanins and evolved hemoglobins as respiratory proteins (Terwilliger and Ryan 2001). In the water flea *Daphnia magna*, for instance, the tandemly duplicated gene cluster of hemoglobin genes shows multiple hypoxia inducible factor (HIF) binding sites, which dramatically induce the expression of hemoglobins when daphnids are exposed to hypoxia (Kimura et al. 1999; Gorr et al. 2004). The extremely patchy distribution of these proteins across the animal phylogeny can be partially understood by the different biochemical properties of their oxygen-binding domains and the changing physiological needs of each particular animal lineage, which make one or the other respiratory protein more effective in their function as oxygen carriers. Recent studies show that many enzymatic genes have complex evolutionary histories, with massive gene losses in most of the eukaryote genomes sampled, but retention in certain tips of the tree of life (Allen et al. 2011; de Mendoza and Ruiz-Trillo 2011; Stairs et al. 2011; Attenborough et al. 2012). In contrast, transcription factors, signaling pathways, and adhesion molecules, for instance, can be traced back in a congruent phylogenetic pattern (Pang et al. 2010; Sebé-Pedrós et al. 2010; Srivastava et al. 2010; Sebé-Pedrós et al. 2011). In some cases, the patchy phylogenetic distribution observed in enzymatic families could be explained by multiple events of LGT, although the phylogenetic signal is often not strong enough. Together with gene structure and synteny analysis we do not find strong evidences of LGT, with the exception of plant tyrosinases (see above). Moreover, the study of the evolution of respiratory proteins emerges as an ideal model to study the interplay between molecular evolution, biochemical constraints, and physiological-ecological needs.

Materials and Methods

All potential hemerythrin, hemocyanin, and tyrosinase sequences were identified by HMMER searches against the Protein, Genome, and EST databases at the NCBI (National Center for Biotechnology Information) and against completed genome/transcriptome projects databases publicly available or that are being conducted in our laboratories (sequences available in supplementary file S1, Supplementary Material online) with the default parameters and an inclusive *E*-value of 0.05. The retrieved sequences were aligned using MAFFT (Katoh et al. 2002) L-INS-i algorithm, and then manually inspected to remove those hits fulfilling one of the following conditions: 1) incomplete sequences with >99% sequence identity to a complete sequence from the same taxa; 2) sequences that showed extremely long branches in the preliminary maximum likelihood trees; and 3) incorrect gene model predictions. The final alignment was carried out using the MAFFT G-INS-i algorithm (for global homology). Maximum likelihood (ML) phylogenetic trees were estimated by RaxML (Stamatakis 2006) and the best tree from 100 replicates was selected. Bootstrap support was calculated from 1,000 replicates. Bayesian inference analyses were performed with PhyloBayes (Lartillot and Philippe 2004), using two parallel runs for 500,000 generations and sampling every 100. Bayesian posterior probabilities (BPP) were used for assessing the statistical support of each bipartition. The domain architecture of all retrieved sequences was inferred by performing a Pfam scan with the gathering threshold as cut-off value. The domain information was used to assess the reliability of each sequence of the initial dataset, to help define protein families according to their architectural coherence, and to assess the level of functional and structural diversification of hemerythrins, hemocyanins, and tyrosinases across the eukaryote lineages.

Supplementary Material

Supplementary files S1, tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The authors thank J. Ryan for his advice and *M. leidyi* sequences, members of the S9 lab for collecting material and discussion, G. Richards for reading and improving the manuscript, G. Torruella and X. Grau-Boved for their help, and J. Achatz for his support. This work was funded by the Sars International Centre for Marine Molecular Biology to J.M.M.-D. and A.H., and an ICREA contract, an European Research Council Starting Grant (ERC-2007-StG-206883), and a grant (BFU2011-23434) from Ministerio de Economía y Competitividad (MINECO) to I.R.-T. A.S.-P.'s salary was supported by a pregraduate FPU grant and A.d.M.'s salary from a FPI grant, both from MICINN.

Literature Cited

Allen AE, et al. 2011. Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. Nature 473:203–207.

Attenborough RM, Hayward DC, Kitahara MV, Miller DJ, Ball EE. 2012. A "neural" enzyme in nonbilaterian animals and algae: preneural origins for peptidylglycine alpha-amidating monooxygenase. Mol Biol Evol. 29:3095–3109.

Baert JL, Britel M, Sautiere P, Malecha J. 1992. Ovohemerythrin, a major 14-kDa yolk protein distinct from vitellogenin in leech. Eur J Biochem. 209:563–569.

Bailly X, Vanin S, Chabasse C, Mizuguchi K, Vinogradov SN. 2008. A phylogenomic profile of hemerythrins, the nonheme diiron binding respiratory proteins. BMC Evol Biol. 8:244.

Blank M, Burmester T. 2012. Widespread occurrence of N-terminal acylation in animal globins and possible origin of respiratory globins from a membrane-bound ancestor. Mol Biol Evol. 29:3553–3561.

Bonaventura J, Bonaventura C. 1980. Hemocyanins—relationships in their structure, function and assembly. Am Zool. 20:7–17.
Burmester T. 2001. Molecular evolution of the arthropod hemocyanin superfamily. Mol Biol Evol. 18:184–195.
de Mendoza A, Ruiz-Trillo I. 2011. The mysterious evolutionary origin for the GNE gene and the root of bilateria. Mol Biol Evol. 28:2987–2991.
Demuynck S, Li KW, Van der Schors R, Dhainaut-Courtois N. 1993. Amino acid sequence of the small cadmium-binding protein (MP II) from *Nereis diversicolor* (annelida, polychaeta). Evidence for a myohemerythrin structure. Eur J Biochem. 217:151–156.
Esposito R, et al. 2012. New insights into the evolution of metazoan tyrosinase gene family. PLoS One 7:e35731.
French CE, Bell JM, Ward FB. 2008. Diversity and distribution of hemerythrin-like proteins in prokaryotes. FEMS Microbiol Lett. 279: 131–145.
Gorr TA, Cahn JD, Yamagata H, Bunn HF. 2004. Hypoxia-induced synthesis of hemoglobin in the crustacean *Daphnia magna* is hypoxia-inducible factor-dependent. J Biol Chem. 279:36038–36047.
Immesberger A, Burmester T. 2004. Putative phenoloxidases in the tunicate *Ciona intestinalis* and the origin of the arthropod hemocyanin superfamily. J Comp Physiol B. 174:169–180.
Isaza CE, Silaghi-Dumitrescu R, Iyer RB, Kurtz DM Jr, Chan MK. 2006. Structural basis for O_2 sensing by the hemerythrin-like domain of a bacterial chemotaxis protein: substrate tunnel and fluxional N terminus. Biochemistry 45:9023–9031.
Karlsen OA, et al. 2005. Characterization of a prokaryotic hemerythrin from the methanotrophic bacterium *Methylococcus capsulatus* (Bath). FEBS J. 272:2428–2440.
Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30:3059–3066.
Kimura S, Tokishita S, Ohta T, Kobayashi M, Yamagata H. 1999. Heterogeneity and differential expression under hypoxia of two-domain hemoglobin chains in the water flea, *Daphnia magna*. J Biol Chem. 274:10649–10653.
Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol. 21:1095–1109.
Lieb B, Todt C. 2008. Hemocyanin in mollusks—a molecular survey and new data on hemocyanin genes in Solenogastres and Caudofoveata. Mol Phylogenet Evol. 49:382–385.
Lieb B, et al. 2006. Red blood with blue-blood ancestry: intriguing structure of a snail hemoglobin. Proc Natl Acad Sci U S A. 103:12011–12016.
Meyer A, Lieb B. 2010. Respiratory proteins in *Sipunculus nudus*—implications for phylogeny and evolution of the hemerythrin family. Comp Biochem Physiol B Biochem Mol Biol. 155:171–177.
Pang K, et al. 2010. Genomic insights into Wnt signaling in an early diverging metazoan, the ctenophore *Mnemiopsis leidyi*. Evodevo 1:10.
Roesner A, Fuchs C, Hankeln T, Burmester T. 2005. A globin gene of ancient evolutionary origin in lower vertebrates: evidence for two distinct globin families in animals. Mol Biol Evol. 22:12–20.
Salahudeen AA, et al. 2009. An E3 ligase possessing an iron-responsive hemerythrin domain is a regulator of iron homeostasis. Science 326: 722–726.
Schmidt-Rhaesa A. 2007. The evolution of organ systems. Oxford: Oxford University Press.
Sebé-Pedrós A, de Mendoza A, Lang BF, Degnan BM, Ruiz-Trillo I. 2011. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan *Capsaspora owczarzaki*. Mol Biol Evol. 28:1241–1254.
Sebé-Pedrós A, Roger AJ, Lang FB, King N, Ruiz-Trillo I. 2010. Ancient origin of the integrin-mediated adhesion and signaling machinery. Proc Natl Acad Sci U S A. 107:10142–10147.
Srivastava M, et al. 2010. The *Amphimedon queenslandica* genome and the evolution of animal complexity. Nature 466:720–726.
Stairs CW, Roger AJ, Hampl V. 2011. Eukaryotic pyruvate formate lyase and its activating enzyme were acquired laterally from a Firmicute. Mol Biol Evol. 28:2087–2099.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.
Sugumaran M. 2002. Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. Pigment Cell Res. 15:2–9.
Terwilliger NB, Ryan M. 2001. Ontogeny of crustacean respiratory proteins. Am Zool. 41:1057–1067.
Terwilliger NB. 1998. Functional adaptations of oxygen-transport proteins. J Exp Biol. 201:1085–1098.
Thompson JW, et al. 2012. Structural and molecular characterization of iron-sensing hemerythrin-like domain within F-box and leucine-rich repeat protein 5 (FBXL5). J Biol Chem. 287:7357–7365.
van Holde KE, Miller KI, Decker H. 2001. Hemocyanins and invertebrate evolution. J Biol Chem. 276:15563–15566.
Vanin S, et al. 2006. Molecular evolution and phylogeny of sipunculan hemerythrins. J Mol Evol. 62:32–41.
Vashisht AA, et al. 2009. Control of iron homeostasis by an iron-regulated ubiquitin ligase. Science 326:718–721.
Vergote D, et al. 2004. Up-regulation of neurohemerythrin expression in the central nervous system of the medicinal leech, *Hirudo medicinalis*, following septic injury. J Biol Chem. 279:43828–43837.
Weber RE, Vinogradov SN. 2001. Nonvertebrate hemoglobins: functions and molecular adaptations. Physiol Rev. 81:569–628.
Xiong J, Kurtz DM Jr, Ai J, Sanders-Loehr J. 2000. A hemerythrin-like domain in a bacterial chemotaxis protein. Biochemistry 39: 5117–5125.

Associate editor: Emmanuelle Lerat