Expanding the repertoire of the eukaryotic selenoproteome

Robert J. Stillwell and Marla J. Berry*

Department of Cell and Molecular Biology, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, HI 96822

Selenium, an essential trace element known primarily for its antioxidant properties, exerts a diversity of functions through its presence in selenoproteins, proteins in which selenium is incorporated into the nascent polypeptide cotranslationally by means of the amino acid selenocysteine (Sec). Although the function of many selenoprotein families is unknown, the importance of selenoprotein synthesis has been clearly demonstrated in animals, in which mutations that render the cell unable to incorporate selenium into proteins result in embryonic lethality (1). With few exceptions, characterized selenoprotein families are known to function enzymatically and are typically involved in redox reactions. In these selenoprotein families, the Sec residue is found in the active site of the enzyme. Because nearly all eukaryotic selenoproteins are represented within mammalian genomes, mammals have been thought to recapitulate the eukaryotic selenoproteome, the set of all selenoproteins. However, by using bioinformatic approaches, Castellano *et al.* (2) have identified and characterized a fish selenoprotein family, SelJ, that challenges these paradigms. Their findings, published in this issue of PNAS, reinforce the seemingly universal importance of selenium by showing the vast and diverse phylogenetic distributions of selenoproteins across taxa.

Evolution of Sec Usage

Selenoprotein translation relies on the concerted involvement of several cis- and trans-acting factors that effectively recode an in-frame UGA from the canonical stop codon to that of Sec (see refs. 3–5 for detailed reviews). As suggested by its unique mechanism of synthesis and incorporation, Sec is likely a late addition to the genetic code. However, Sec is used in all three domains of life, inviting speculation that it appeared early on in protein evolution and that the origin of the Sec machinery is common, evolving early and only once. To date, the glutathione peroxidases and selenophosphate synthetases are the only selenoprotein families that appear to be shared between domains (6). This finding implies that there has been an independent rise of selenoproteins in each lineage and that different selenoproteomes consist of selenoprotein families of widely different evolutionary ages.

It has been shown that fish proteins have diverged significantly faster than homologs found in mammals. The nucleotide divergence per year between two pufferfish, *Tetraodon* and *Takifugu*, is approximately twice the level of divergence between human and mouse or between rat and mouse, and comparison of orthologs between these species indicates that protein divergence is also greater in fish (7). Recent findings reveal that a fish selenoprotein family, SelU, has a widespread eukaryotic distribution of Cys-containing homologs (8). Only fish incorporate Sec into this protein, and it is unclear whether Sec or Cys incorporation is the ancestral state. Interestingly, for most selenoprotein families, homologs in which the active-site Sec residue is replaced by Cys can be found. Mutational studies of several selenoproteins have concluded that Sec and Cys versions have analogous functions, but Sec versions have increased catalytic efficiency (9–11). The discovery of SelJ in fish is an example of the existence of ancient and new selenoprotein families within a selenoproteome and highlights the possibility for dynamic lineage-specific patterns of selenoprotein evolution. Thus far, no known Cys incorporating homologs of SelJ have been discovered in mammals.

Each selenoprotein family within any specific taxa has a distinct evolutionary history.

SelJ: A Structural and Catalytic Selenoprotein?

A limited set of selenoprotein families have been characterized. For example, the glutathione peroxidases, thioredoxin reductases, and methionine sulfoxide reductases are involved in antioxidant protection and cellular redox balance (12–14). Iodothyronine deiodinases catalyze thyroid hormone activation and inactivation (9, 15). In contrast, Castellano *et al.* (2) suggest that SelJ may have a structural, rather than enzymatic, role. Cys-containing homologs to SelJ were found only in cnidarians (the jellyfish and anemones), and these proteins have been previously characterized as being structural crystallins. Further analysis revealed that SelJ and the jellyfish J1-crystallins form a subfamily within a large family of ADP-ribosylation enzymes. The selenoprotein phospholipid hydroperoxide glutathione peroxidase is known to function alternately as a glutathione-dependent hydroperoxide reductase and as a structural protein at different stages during sperm maturation (16). Whether or not SelJ proves to function structurally and enzymatically, it is an interesting addition to the ever-growing repertoire of the eukaryotic selenoproteome.

Taxa-Specific Preferences for Sec or Cys

In what Castellano *et al.* (2) refer to as "mosaic-like" evolution, the selenoproteomes of different taxa vary widely in the number and types of selenoprotein families found. For instance, the nematode selenoproteomes appear to have been greatly reduced, with the *Caenorhabditis elegans* genome encoding only one selenoprotein, a thioredoxin reductase (17). However, the nematodes have retained many of the known eukaryotic selenoproteins in Cys-containing versions. The alga *Chlamydomonas reinhardtii* is the only eukaryote known to have the protein methionine sulfoxide reductase in Sec form, although its existence is widespread in Cys form. No true Sec-containing selenoproteins have been found in plants, and the Zebrafish selenoproteome, which consists of 19 selenoprotein families, contains all known vertebrate selenoproteins. These examples illustrate the idea that each selenoprotein family within any specific taxa has a distinct evolutionary history. With increasing numbers of selenoproteome descriptions, an apparent preference among taxa and, likewise, among protein families for the use of either

Conflict of interest statement: No conflicts declared.

See companion article on page 16188.

*To whom correspondence should be addressed. E-mail: mberry@hawaii.edu.

Sec or Cys is beginning to be realized. Why this preference exists has yet to be explained.

With the availability of new genomes, ancestral as well as recently acquired selenoproteins are being discovered. Additional taxa-specific selenoproteins are likely to be found. Castellano *et al.* (2) have shown the vast importance and the diversity of functions of selenoprotein families. Among other things, these findings should stimulate further exploration into the functions and the evolutionary histories of selenoprotein families.

1. Bosl, M. R., Takaku, K., Oshima, M., Nishimura, S. & Taketo, M. M. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 5531–5534.
2. Castellano, S., Lobanov, A. V., Chapple, C., Novoselov, S. V., Albrecht, M., Hua, D., Lescure, A., Lengauer, T., Krol, A., Gladyshev, V. N., *et al.* (2005) *Proc. Natl. Acad. Sci.* USA **102,** 16188–16193.
3. Driscoll, D. M. & Copeland, P. R. (2003) *Annu. Rev. Nutr.* **23,** 17–40.
4. Hatfield, D. L. & Gladyshev, V. N. (2002) *Mol. Cell. Biol.* **22,** 3565–3576.
5. Berry, M. J., Tujebajeva, R. M., Copeland, P. R., Xu, X. M., Carlson, B. A., Martin, G. W., III, Low, S. C., Mansell, J. B., Grundner-Culemann, E., Harney, J. W., *et al.* (2001) *Biofactors* **14,** 17–24.
6. Kryukov, G. V. & Gladyshev, V. N. (2004) *EMBO Rep.* **5,** 538–543.
7. Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., *et al.* (2004) *Nature* **431,** 946–957.
8. Castellano, S., Novoselov, S. V., Kryukov, G. V., Lescure, A., Blanco, E., Krol, A., Gladyshev, V. N. & Guigo, R. (2004) *EMBO Rep.* **5,** 71–77.
9. Berry, M. J., Maia, A. L., Kieffer, J. D., Harney, J. W. & Larsen, P. R. (1992) *Endocrinology* **131,** 1848–1852.
10. Lee, S.-R., Bar-Noy, S., Kwon, J., Levine, R. L., Stadtman, T. C. & Rhee, S. G. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 2521–2526.
11. Hazebrouck, S., Camoin, L. Faltin, Z., Strosberg, A. D. & Eshdat, Y. (2000) *J. Biol. Chem.* **275,** 28715–28721.
12. Arthur, J. R. (2000) *Cell. Mol. Life Sci.* **57,** 1825–1835.
13. Zhong, L. & Holmgren, A. (2000) *J. Biol. Chem.* **275,** 18121–18128.
14. Kim, H. & Gladyshev, V. N. (2004) *Mol. Biol. Cell* **15,** 1055–1064.
15. Bianco, A. C., Salvatore, D., Gereben, B., Berry, M. J. & Larsen, P. R. (2002) *Endocr. Rev.* **23,** 38–89.
16. Ursini, F., Heim, S., Kiess, M., Maiorino, M., Roveri, A., Wissing, J. & Flohe, L. (1999) *Science* **285,** 1393–1396.
17. Taskov, K., Chapple, C., Kryukov, G. V., Castellano, S., Lobanov, A. V., Korotkov, K. V., Guigo, R. & Gladyshev, V. N. (2005) *Nucleic Acids Res.* **33,** 2227–2238.