Complexity and diversity in *c*-type cytochrome biogenesis systems

J.W.A. Allen*[1], M.L. Ginger† and S.J. Ferguson*
*Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K., and †Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, U.K.

Abstract
***c*-type cytochromes contain haem covalently attached to protein by thioether bonds formed post-translationally and requiring a dedicated biogenesis apparatus. Three biogenesis systems, found in different cell types, are well known. Here we discuss emerging evidence for at least one additional system, for unanticipated diversity in the location of the systems and for the co-existence of multiple systems in some cells.**

Introduction
c-type cytochromes are widespread proteins characterized by covalent attachment of haem to the polypeptide through thioether bonds formed between the vinyl groups of the haem and the thiols of, normally, two cysteine residues in a CXXCH motif. The histidine is an axial ligand to the haem iron. Mitochondrial cytochrome *c* is the best known such protein, but there are many other distinct *c*-type cytochrome centres in bacteria that function in electron transfer or occur at the catalytic site of enzymes. Such cytochromes are crucial for the biochemistry of the nitrogen cycle and include cytochrome cd_1 nitrite reductase, cytochrome bc_1, hydroxylamine oxidoreductase, cytochrome c', cytochrome c_4 and nitric oxide reductase.

Remarkably, evolution has produced several distinct biogenesis systems to facilitate the chemically difficult post-translational haem attachment to apocytochromes *c* [1,2]; the three have been characterized. Systems I and II are multicomponent, the former [also called a Ccm (cytochrome *c* maturation) system] is found in α-, some β- and most γ-proteobacteria, deinococci, and the mitochondria of various eukaryotes (see below). System II occurs in δ-, ε- and some β-proteobacteria, at least one γ-proteobacterium (*Acidithiobacillus ferrooxidans*), most Gram-positive bacteria, cytophagales, aquaficales, plant and algal chloroplasts and cyanobacteria. Biogenesis system III, the enzyme haem lyase, is found in the mitochondria of fungi, metazoans and some protozoa [3]. The stereochemistry of the thioether bonds formed by these systems is universally conserved [4]. Both systems I and II mature a wide range of *c*-type cytochromes with varied folds and often multiple haems. The basis for this diversity and extensive functional overlap of cytochrome *c* biogenesis systems is uncertain and intriguing.

Key words: chloroplast, *c*-type cytochrome, haem, post-translational modification, trypanosome.
Abbreviation used: Ccm, cytochrome *c* maturation.
[1]To whom correspondence should be addressed (email allenj@bioch.ox.ac.uk).

A mosaic distribution of cytochrome *c* biogenesis systems in eukaryotes
Within the kingdom Eukaryota, haem lyase, an enzyme that appears to have no ancestral homologue in Archaea or Eubacteria, represents the most widely distributed mitochondrial cytochrome *c* biogenesis system. Exceptions include the land plants, the red alga *Cyanidioschyzon merolae*, the jakaboid protozoan *Reclinomonas americana* and ciliates (*Tetrahymena* and *Paramecium*). In these organisms, which occupy evolutionary positions far away from the fungi and metazoa [5], the Ccm system is at least partially encoded by mitochondrial genes [6,7]. Further components of the Ccm system have been identified in the completely sequenced nuclear genomes of *Arabidopsis thaliana* and *Oryza sativa*, but there is no gene encoding haem lyase [8]. Interestingly, in algae from the Streptophyta, the group that gave rise to land plants, components of the Ccm apparatus have been found to be both absent (e.g. *Chaetosphaeridium globosum*) and present (*Chara vulgaris*) within the mitochondrial genome [9]. However, not only are Ccm components absent from the mitochondria of chlorophyte algae, but cytochrome *c* and c_1 haem lyase genes are encoded in the *Chlamydomonas reinhardtii* nuclear genome [3].

The distribution of eukaryotic cytochrome *c* biogenesis systems is clearly complex. However, if the progenitor of mitochondria was (as often stated) an α-proteobacterium, presumably expressing the Ccm system, then perhaps the early evolution of haem lyase precipitated the disappearance of the multicomponent Ccm apparatus from the vast majority of eukaryotes. In those eukaryotes retaining the Ccm system, either divergence occurred before the arrival of haem lyase or, alternatively, the 'newly evolved' haem lyase was more readily lost. The prevalence of the 'simpler' biogenesis system, haem lyase, does correlate with the observation that in mitochondria there are only two *c*-type cytochromes, *c* and c_1, which have essentially the same fold. The more complicated Ccm system can (in prokaryotes) mature numerous *c*-type cytochromes.

Systems IV and V?

Mitochondrial *c*-type cytochromes from euglenids and trypanosomatids (including the medically relevant *Trypanosoma brucei*, *T. cruzi* and *Leishmania*) pose an additional, fascinating puzzle. The cytochromes *c* and c_1 from these protists contain only a single cysteine within the haem-binding motif (i.e. XXXCH), and are thus matured by the formation of only one thioether bond between the haem and apocytochrome. Furthermore, we have recently presented evidence that none of the known *c*-type cytochrome biogenesis systems (I, II or III) are encoded within the mitochondrial or nuclear genomes of any trypanosomatid [10]. Thus, there appears to be another cytochrome *c* biogenesis system, unique to the trypanosomatids (and by inference the closely related euglenids). Alternatively, the biogenesis machinery for the single cysteine *c*-type cytochromes in these organisms may be similar to that for a newly recognized post-translational modification. Crystal structures of the cytochrome b_6f complex from the thylakoid photosynthetic electron transfer pathway reveal that the cytochrome *b* subunit contains an unexpected covalently bound haem, attached by only one thioether bond between a cysteine and a vinyl group [11]. Strikingly, however, the haem iron is not co-ordinated by any amino acid side chains (i.e. there is no histidine corresponding to that of the *c*-type cytochrome CXXCH motif). Four genetic mutants of *C. reinhardtii* that lack a properly assembled cytochrome *b* subunit implicate a requirement for dedicated biogenesis genes for this novel haem attachment [12]. Apparently similar covalent haem attachment is also observed in the *bc* complex of *Bacillus subtilis* [13], so any biogenesis genes should also be found there.

One location, multiple biogenesis systems?

Until recently it had been assumed that no cells had more than one *c*-type cytochrome biogenesis system potentially functioning in the same location. Intriguingly, however, the genomes of some β-proteobacteria, including *Bordetella bronchiseptica*, contain all the components of both systems I and II [2]. In that organism, all the maturation proteins contain every residue that has been experimentally identified as important, so there is no obvious reason why only one maturation machinery would be functional (or preferred). Expression studies are awaited to determine this point, but insight into the evolutionary origin or function could also arise from the genomic locations of each putative biogenesis system. The genes for the Ccm apparatus in *B. bronchiseptica* are adjacent to those for NapB and NapC, *c*-type cytochrome components of a nitrate reductase, whereas system II genes are adjacent to that for a homologue of cytochrome c_{553} (J.W.A. Allen and S.J. Ferguson, unpublished work). Other new genomes (e.g. *Anopheles gambiae*, *Desulfitobacterium hafniense*) indicate additional organisms in which multiple systems are potentially present [2,14]. It remains to be demonstrated how frequently organisms express multiple cytochrome *c* biogenesis systems and to what extent there is functional degeneracy when they do.

Concluding remarks

Recent developments have shown that the complexity and diversity of *c*-type cytochrome biogenesis systems is much greater than had been expected. Whereas three distinct systems were recognized and could be neatly categorized in terms of origin and cellular location, new data question this paradigm. At least one, and possibly two, new maturation systems await characterization in trypanosomatids and thylakoids. Moreover, the biogenesis machinery for *c*-type cytochromes in archaea is not yet clear from genome analyses. Three different biogenesis systems are ostensibly present in various mitochondria, whereas, e.g., both systems I and II are unexpectedly found in both β- and γ-proteobacteria. Systems I and II are apparently quite modular, i.e. the quorum of essential proteins varies between organisms. Genomes also suggest that multiple systems might sometimes operate in the same cellular location. With *c*-type cytochrome biogenesis, the harder one looks, the more complex it becomes.

This work was supported by grants from the U.K. BBSRC (C13343 and BB/C508118/1) and the E.P. Abraham Trust of Oxford University. M.L.G. is a Royal Society University Research Fellow.

References

1 Allen, J.W.A., Daltrop, O., Stevens, J.M. and Ferguson, S.J. (2003) Philos. Trans. R. Soc. Lond. B Biol. Sci. **358**, 255–266
2 Stevens, J.M., Daltrop, O., Allen, J.W.A. and Ferguson, S.J. (2004) Acc. Chem. Res. **37**, 999–1007
3 Bernard, D.G., Gabilly, S.T., Dujardin, G., Merchant, S. and Hamel, P.P. (2003) J. Biol. Chem. **278**, 49732–49742
4 Barker, P.D. and Ferguson, S.J. (1999) Structure Fold. Des. **7**, R281–R290
5 Aravind, L., Iyer, L.M., Wellems, T.E. and Miller, L.H. (2003) Cell (Cambridge, Mass.) **115**, 771–785
6 Burger, G., Zhu, Y., Littlejohn, T.G., Greenwood, S.J., Schnare, M.N., Lang, B.F. and Gray, M.W. (2000) J. Mol. Biol. **297**, 365–380
7 Ohta, N., Sato, N. and Kuroiwa, T. (1998) Nucleic Acids Res. **26**, 5190–5198
8 Spielewoy, N., Schulz, H., Grienenberger, J.M., Thony-Meyer, L. and Bonnard, G. (2001) J. Biol. Chem. **276**, 5491–5497
9 Turmel, M., Otis, C. and Lemieux, C. (2003) Plant Cell **15**, 1888–1903
10 Allen, J.W.A., Ginger, M.L. and Ferguson, S.J. (2004) Biochem. J. **383**, 537–542
11 Stroebel, D., Choquet, Y., Popot, J.L. and Picot, D. (2003) Nature (London) **426**, 413–418
12 Kuras, R., de Vitry, C., Choquet, Y., Girard-Bascou, J., Culler, D., Buschlen, S., Merchant, S. and Wollman, F.A. (1997) J. Biol. Chem. **272**, 32427–32435
13 Yu, J. and Le Brun, N.E. (1998) J. Biol. Chem. **273**, 8860–8866
14 Thöny-Meyer, L. (2003) Biochemistry **42**, 13099–13105

Received 14 September 2004