Known knowns, known unknowns, unknown unknowns and the propagation of scientific enquiry

David C. Logan*

Department of Biology, WP Thompson Building, 112 Science Place, University of Saskatchewan, Saskatoon SK S7N 5E2, Canada


In February 2002, Donald Rumsfeld, the then US Secretary of State for Defence, stated at a Defence Department briefing: ‘There are known knowns. There are things we know that we know. There are known unknowns. That is to say, there are things that we now know we don’t know. But there are also unknown unknowns. There are things we do not know we don’t know.’ As a result, he was almost universally lampooned since many people initially thought the statement was nonsense. However, careful examination of the statement reveals that it does make sense, indeed the concept of the unknown unknown existed long before Donald Rumsfeld gave it a new audience.

Much scientific research is based on investigating known unknowns. In other words, scientists develop a hypothesis to be tested, and then in an ideal situation experiments are best designed to test the null hypothesis. At the outset the researcher does not know whether or not the results will support the null hypothesis. However, it is common for the researcher to believe that the result that will be obtained will be within a range of known possibilities. Occasionally, however, the result is completely unexpected—it was an unknown unknown.

There are many known knowns of intracellular protein targeting and, as with many fields of research, it seems that the number of known unknowns increase in parallel. The key determinants for targeting to mitochondria have been determined, and there are at least four subclasses of mitochondrial-targeted proteins containing different targeting signals that are directed to different sites within the mitochondrion (outer membrane, inter-membrane space, inner membrane, and matrix) by different mechanisms (Bolender et al., 2008; Whelan and Glaser, 2007) (Fig. 1). However, much remains unknown, especially in plants. For example, there are now a number of known unknowns resulting from a previously unknown unknown: the existence of proteins dually targeted to both plastids and mitochondria (Peeters and Small, 2001; Ma and Taylor, 2002; Whelan and Glaser, 2007). We now know that dual targeting to plastids and mitochondria occurs as a result of ambiguous signal sequences, but we do not know how these signals are recognized by both organelles, when other proteins are only recognized by one (Whelan and Glaser, 2007).

The paper by Chatre et al. (2009) in this issue is an excellent example of research uncovering unknown unknowns. Typically, investigations into the mechanics of intracellular protein targeting have been performed using protein biochemistry, but the investigation by Chatre et al. (2009) is not typical. If the study had simply been investigating mitochondrial targeting using in vitro translation of various proteins with altered targeting signals, detected by protein electrophoresis and immunoblotting, the results would have been a combination of ‘yes, the construct targets to mitochondria’ and ‘no, the construct does not target to mitochondria’. However, because a cell biological approach was taken, so much more information was garnered; that is the power of using fluorescent protein fusions in vivo.

By means of a bioinformatics screen, Chatre et al. (2009) identified nucleus-encoded proteins that were predicted, based on the presence of a coiled-coil domain, to be targeted to the secretory pathway. However, one of the proteins identified, and named MITS1, was targeted to mitochondria when a full-length protein fusion was made to the N-terminus of YFP (Fig. 2A). MITS1, a putative actin-binding protein, was also correctly predicted, by various
bioinformatic tools, to be localized to mitochondria. Further experimentation demonstrated that the predicted mitochondrial targeting peptide (mTP), residues 1–39, was sufficient to target YFP to mitochondria (Fig. 2B). In addition, when only the first 11 amino acids of MITS1 (MITS11–11) was fused to YFP, the fusion protein was cytosolic, the same result occurred using the first 31 amino acids (MITS11–31), consistent with the requirement for a positively charged region in mTPs; in MITS1 this is towards the C-terminus of the mTP and is preceded by a hydrophobic region (Fig. 2A, B). So far, so predictable. However, the research entered the world of the unknown unknowns when the team started an in-depth exploration of the targeting determinants of the N-terminal region. When MITS112–39 was fused to YFP, targeting was to both the ER and mitochondria demonstrating that residues 1–11, although not sufficient for targeting to mitochondria, can overcome targeting to the ER. Dual targeting to mitochondria and the ER is a known known in mammalian cells (Huang et al., 1999; Ma and Taylor, 2002), and occurs by means of N-terminal splice variants. It is not known if wild-type MITS1 is dual targeted. In any case, a major difference between MITS1 and the dual targeted cAMP-dependent protein kinase anchoring protein, D-AKAP1, in mammals is that it is the shorter forms of D-AKAP1, those missing an extreme N-terminal region, that target to mitochondria (Huang et al., 1999).

The change in MITS1 targeting resulting from changes in the N-terminal mTP is interesting, but what I think is much more interesting is the finding that a single distal tryptophan residue (W361) affects targeting. Full-length MITS1 fused to YFP is found to be targeted solely to mitochondria, but a single W361A mutation (MITS1W361A) redirects the protein to the ER and Golgi (Fig. 2C). The W361A mutation also redirects MITS112–573 from the ER and mitochondria to the cytosol (Fig. 2C). The obvious questions that arise include: How does this tryptophan residue affect targeting? How does this one mutation lead...
to redistribution of a mitochondrial-targeted protein to the ER and Golgi? Does the mutation lead to increased interaction of the protein with ER signal recognition particles (SRPs)? Is there competition for MITS1W361A between the SRPs and the mitochondrial import machinery?

Once again, the discovery of a previously unknown unknown shows us how little we know and leads to the propagation of a family of known unknowns which can then be tackled by traditional hypothesis forming and testing, occasionally throwing-up another unknown unknown, and so the cycle continues.

References


Fig. 2. MITS1-YFP constructs used by Chatre et al. (2009). (A) Schematic representation of MITS1 and its N-terminal region. (B) Schematic representations of the constructs used to dissect the roles of subdomains of the mTP in MITS1. (C) Schematic representations of the constructs used to investigate the role of distal tryptophan residues. All figures redrawn from Chatre et al. (2009).