Table 1 Distribution of PDI polypeptides between PDI dimers and prolyl-4-hydroxlase tetramers

	% Total cell protein	
	PDI	β-subunit of prolyl-4-hydroxylase
Embryo chick tendon (day 16)	0.08	0.12
Adult mammalian liver	0.35–0.4	~0.01
Ig-secreting lymphoma cell	0.5–0.8	<0.1*

Ratios from ref. 4 Table VI except for * which uses the value quoted for chick embryo spleen and is probably an overestimate; absolute values derived from original papers cited in ref. 13.

in embryo chick tendon — the source most active in procollagen synthesis — free PDI is comparable in amount to β-subunits incorporated into prolyl-4-hydroxylase. In contrast to Pain's claim, the level of total PDI subunits does not correlate across tissues specifically with the rate of synthesis of procollagen but with the total rate of synthesis of disulphide-bonded proteins.

This point underlines the primary role of PDI. Its location within the lumen of the endoplasmic reticulum[5,6] implies that it is not required for protein synthesis *per se*, but for the synthesis of secretory and cell-surface proteins. Such extracellular proteins characteristically contain disulphide bonds which are absent from cytosolic proteins. In the folding *in vitro* of reduced denatured secretory proteins, the rate determining step is disulphide bond isomerization. Extensive *in vitro* studies have established that pure PDI can catalyse disulphide bond formation, reduction or isomerization in over 20 disparate proteins (ref. 7, Table 2). Detailed analysis of the action of PDI in catalysing the reoxidation and refolding of bovine pancreatic trypsin inhibitor shows that all the steps involving thiol: protein disulphide interchange and conformational change are catalysed[8]. PDI has never been observed to catalyse the folding of a protein lacking disulphide bonds (see discussion in ref. 8). The sequence of PDI[9] indicates a homology with thioredoxin, a protein which functions through a redox-active disulphide/dithiol couple, and we have subsequently established a functional homology in the chemical properties of the PDI active-site disulphide/dithiol couples[5].

In the circumstances it seems perverse of Pain to propose a quite new hypothetical mechanism for PDI, based on *cis–trans* isomerization of prolyl-peptide bonds (which are, after all, found indifferently in cytosolic and secreted proteins, in proteins with and without disulphide bonds). The stimulus for this suggestion is clearly the participation of PDI as β-subunit of prolyl-4-hydroxylase, but there is already evidence that it is the α-subunit of that enzyme which binds the prolyl peptide substrate[10]. If Pain's proposal was ever tenable it is contradicted by the work of Koivu and Myllylä[11], who have shown that PDI accelerates procollagen assembly and folding *in vitro* to give rates comparable to those *in vivo*; however, it is only the assembly process, in which three polypeptides are linked into a disulphide-bonded trimer, that is catalysed, while the folding of the central proline-rich domain of the molecule into the characteristic triple helix occurs at the same rate in absence or presence of PDI (for further discussion see ref. 12).

Thus, in contrast to Pain's claim that the recent work obviates "the previously held view that PDI is required in collagen biosynthesis", the results show that PDI is essential for rapid formation of native procollagen through its catalysis of native disulphide bond formation. This confirms in dramatic fashion the conventional view of PDI action[13]. The PDI dimer contains four domains homologous to thioredoxin each with an essential dithiol/disulphide couple with a characteristic low pK thiol (our unpublished observations). This allows it to catalyse the breaking and forming of protein disulphide bonds at physiological pH, and hence to catalyse isomerization to the set of disulphides compatible with the most stable folded protein conformation. Hence PDI participates in the biosynthesis of secretory proteins by ensuring that they fold to the native disulphide-linked conformation on a timescale consistent with the secretory process as a whole.

The recent results reveal not "a case of mistaken identity" as stated in your headline, but exciting evidence for an additional role, in a polypeptide with a well-established function.

ROBERT B. FREEDMAN
Biological Laboratory,
University of Kent,
Canterbury, Kent CT2 7NJ, UK

1. Pain, R.H. *Nature* **328**, 298 (1987).
2. Pihlajaniemi, T. *et al. EMBO J* **6**, 643–649 (1987).
3. Koivu, J. *et al. J. biol. Chem.* **262**, 6447–6449 (1987).
4. Kivirikko, K.I. & Myllylä, R. in *The Enzymology of Post-translational Modification of Protiens, Vol. 1* (eds Freedman, R.B. & Hawkins, H.C.) 53–104 (Academic, London, 1980).
5. Lambert, N. Freedman, R.B. *Biochem. J.* **228**, 635–645 (1985).
6. Edman, J.A. *et al. Nature* **317**, 267–270 (1985).
7. Freedman, R.B., Brockway, B.E. & Lambert, N. *Biochem. Soc. Trans.* **12**, 929–932 (1984).
8. Creighton, T.E., Hillson, D.A. & Freedman, R.B. *J. molec. Biol.* **142**, 43–62 (1980).
9. Hawkins, H.C. *et al. Biochem. Soc. Trans.* **14**, 256 (1986).
10. de Waal, A. *et al. Biochemistry* **24**, 6439–6499 (1985).
11. Koivu, J. & Myllylä, R. *J. biol. Chem.* **262**, 6159–6164 (1987).
12. Freedman, R.B. *Nature* **329**, 196–197 (1987).
13. Freedman, R.B. *Trends biochem. Sci.* **9**, 438–441 (1984).

The last ribo-organism

SIR—There has been much discussion, in these pages[1-3] and elsewhere[4,5], about the relevance of RNA catalysis to the origin of life. Three assumptions are generally made: (1) RNA splicing is a vestige of catalytic processes in the first forms of life; (2) the geological time during which RNA served as the principal catalyst in living systems was short; and (3) the first step the primitive ribo-organism took was to develop translation machinery to make proteins.

Each of these assumptions is problematic. Models postulating an early role for RNA catalysts are grounded on three facts of modern biochemistry[3,6,7]: the role of the RNA in ribosomes, the intermediacy of messenger RNA between DNA and proteins and the structure of cofactors that contain RNA (for example, ATP, NAD, FAD, adenosyl methionine, CoA). These facts, discussed in the 1970s[8,9], and the discovery by Usher and McHale that RNA catalyses simple reactions[10] remain the compelling arguments for this model.

RNA catalysis of RNA splicing and processing[11,12] potentially supports this model if we assume that introns are vestiges of the 'RNA world'. This assumption is not by itself secure although some introns appear to be ancient[13], others appear not to be. For example, none of the eight introns in mammalian genes for alcohol dehydrogenase (ADH) coincide in position with the nine introns in plant ADHs[14,15]. Further, the rate at which introns can be gained or lost suggests that they are in many cases recently derived[16]. Finally, introns (or, for that matter, intron positions) in modern proteins cannot be viewed as homologous to introns in an RNA world. This view implies that RNA and protein exons are homologous, and the chemical differences between RNA and proteins all but rule this out.

Nevertheless, in the context of a model supported on other grounds, it is possible to view self-splicing and RNA processing as vestigial. But in recent literature, molecular biologists have come to view self-splicing, introns and RNA processing as the primary grounds for regarding RNA catalysis as important in early life, overlooking the chemical, biological and logical requirements of a model that comprehensively interprets what we know about modern biochemistry. This has led to a series of proposals that are quite unreasonable.

First, what we know about chemistry suggests that the evolution of catalysts for metabolic reactions is far easier than the evolution of translation machinery. Therefore, organisms using RNA as catalysts should have developed a complex metabolism before they developed translation[17]. Contrary views that propose that translation arose in a primitive metabolic

background appear to be grounded in the mistaken belief that RNA intrinsically needs the help of proteins to catalyse interesting reactions. In fact, many chemical mechanisms for RNA-catalysed reactions are just as plausible as analogous mechanisms for protein catalysis, even considering the shortage of functional groups in the four common RNA bases[17]. Further, if transfer RNA and ribosomal RNA are presumed to be vestiges of the RNA world, them it seems natural to view the modified bases in tRNA and rRNA as vestigial also. Given these modified bases, RNA catalysts have essentially the same range of functional groups as proteins, and appear, at least at first glance, to have a catalytic potential comparable to that of proteins.

Second, the assumption that metabolic complexity preceeded the invention of translation is required by the logic of the model. If the 'ribo-structures' of ribo-cofactors are to be explained as vestiges of the RNA world, then the RNA world most plausibly contained RNA enzymes that catalysed reactions using these cofactors, including methyl transfers, phosphorylations and aldol and Claisen condensations. This suggests metabolic complexity. But if metabolic complexity is presumed to post-date translation, the problem of explaining why these metabolically important cofactors all contain fragments of RNA is left unsolved.

Further, biological intuition is contradicted by many recent models, especially those based on a view of RNA viruses as 'living fossils' of early forms of life. Viruses are known to adapt at tremendous rates to survive in specific hosts[18] making it unlikely that RNA viruses living in modern hosts contain many non-functional vestiges of an RNA world that vanished 2,000 million years ago. Conversely, if we interpret the biochemistry of modern viruses as vestiges of an RNA world, we must radically alter our view of the adaptability of viruses.

Thus, if we believe in an 'RNA world', we must believe that translation originated in a 'breakthrough organism' with many riboenzymes catalysing many metabolic reactions. Models for the metabolism of this breakthrough organism can be suggested by modern biochemical data, and the origin of traslation is best viewed in the context of such models. Thus, rather than constructing models for the origin of translation in a primitive metabolic background (a difficult task at best), models for the origin of translation may assume that translation arose in an organism where aminoacylated RNAs were already used metabolically. Some metabolic roles for such RNAs may have survived. A glutamate-RNA ester is the first intermediate in the biosynthesis of chlorophyll[19,20]. A glycine-RNA ester is functionally important in cell wall biosynthesis[21].

Three speculative but interesting suggestions follow from this picture. First a metabolically complex ribo-organism would be expected to evolve divergently to produce many species of ribo-organisms. Most would become extinct after the breakthrough[17]. This presumed extinction implies a limit to our ability to extrapolate from the biochemistry of modern organisms back to the biochemisty of the first life form. To model the biochemistry of an ancient organism, one must normally compare the biochemistries of many of its descendants; examination of a single descendant does not provide sufficient information to distinguish between primitive traits (those present in the ancient organism) and derived traits (those that evolved subsequently). If the picture presented is correct therefore, the breakthrough organism is the most ancient organism that we can hope to reconstruct purely by examining the biochemistry of modern organisms.

Even here, model construction is difficult. But efforts to model the breakthrough organism themselves may be interesting. For example, in one model, the role for RNA in chlorophyll biosynthesis is interpreted as a vestige of a photosynthetic breakthrough organism. The origin of photosynthesis can be approximately dated, and organisms that lived earlier than this date are believed to have left fossils[22]. A conclusion drawn from this model worthy of examination is that these ancient fossils are of ribo-organisms[17].

Steven A. Benner
Andrew D. Ellington

Laboratory for Organic Chemistry, Swiss Federal Institute of Technology, Zurich, Switzerland

1. Gilbert, W. *Nature* **319**, 618 (1986).
2. Westheimer, F.H. *Nature* **319**, 534–535 (1986).
3. North, G. *Nature* **328**, 18–19 (1987).
4. Darnell, J.E. & Doolittle, W.F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1271–1274 (1986).
5. Weiner, A.M. & Maizels, N. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
6. Crick, F. *J. molec. Biol.* **38**, 367–379 (1968).
7. Orgel, L.E. *J. molec. Biol.* **38**, 381–393 (1986).
8. White, H.B. III *J. molec. Evol.* **7**, 101 (1976).
9. Visser, C.M. & Kellogg, R.M. *J. molec. Evol.* **11**, 171–187 (1978).
10. Usher, D.A. & McHale, A.H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1149–1153 (1976).
11. Cech, T.R., Zaug, A.J. & Grabowski, P.J. *Cell* **27**, 487–496 (1981).
12. Guerrier-Takada, C. & Altman, S. *Science* **233**, 285–286 (1983).
13. Marchionni, M. & Gilbert, W. *Cell* **46**, 133–150 (1986).
14. Duester, G., Smith, M. Bilanchone, V. & Hatfield, G.W. *J. biol. Chem.* **261**, 2027–2033 (1986).
15. Dennis, E.S. *et al. Nucleic Acids Res.* **12** 3983–4000 (1984).
16. Sogin, M.L., Ingold, A. Karlok, M., Nielsen, H. & Engberg, J. *EMBO J.* **5**, 3625–3630 (1986).
17. Benner, S.A. *et al. Cold Spring Harb Symp. quant. Biol.* **51** (in the press).
18. Domingo, E., Sabo, D., Taiguchi, T. & Weissmann, C. *Cell* **13**, 735–744 (1978).
19. Weinstein, J.D. & Beale, S.I. *Arch. Biochem. Biophys.* **239**, 87–93 (1985).
20. Huang, D.-D., Wang, W.-Y., Gough, S.P. & Kannangara, C.G. *Science* **255**, 1482–1484 (1984).
21. Roberts, R.J. *Nature New Biol.* **237**, 44–45 (1972).
22. Strother, P.K. & Barghoorn, E.H.S. *Origins of Life and Evolution* (eds Halvorsen, H.O. & van Holde, K.E.) 1–18 (Liss, New York, 1980).

Amino-acid sequence similarities

SIR—In a recent News and Views article[1], Lindsay Sawyer proposed that several proteins, including β-lactoglobulin and plasma retinol-binding protein, have common features of amino-acid sequence. He suggests that these features are common to proteins that bind small conjugated molecules that are either sparingly soluble, labile or both and have an associated receptor. This view may be of great importance in the elucidation of the function of the human uterine endometrium during early pregnancy and the molecular interaction between the endometrium and the implanting embryo.

Quantitatively, the main protein product synthesized and secreted *in vitro* in the endometrium from the mid-luteal phase of the menstrual cycle and during the first trimester of pregnancy, is a dimeric glycoprotein with a subunit of relative molecular mass 25,000 (refs 2,3). This protein is a major product of the secretory glandular epithelium. It has been termed endometrial protein 15, pregnancy-associated endometrial α_2-globulin, progestagen-dependent endometrial protein or placental protein 14, and it has been detected in intraluminal uterine fluid and amniotic fluid (see ref. 4 for review). Recent partial sequence analysis reveals strong similarity with the β-lactoglobulin family[5,6] and a 38-residue amino-terminal sequence contains the sequence -Lys-Leu-Ala-Gly-Lys-Trp-His- which conforms to the common sequence proposed by Sawyer[1]: -U-X-X-Gly-X-Trp-Y-(where U is basic, Y aromatic and X any amino acid).

This similarity, and Sawyer's proposals, imply that this endometrial protein may be involved in transport of a ligand and that a receptor system exists. Further sequence and structural analysis may elucidate the function of this protein and the endometrial glandular epithelium in early placental development.

S. C. Bell
J. W. Keyte

Department of Biochemistry, University of Leicester, Leicester LE2 7LX, UK

1. Sawyer, L. *Nature* **327**, 659 (1987).
2. Bell, S.C. *et al. Br. J. Obstet. Gynaec.* **92**, 793–803 (1985).
3. Bell, S.C. *et al. J. reprod. Fertil.* **77**, 221–231 (1986).
4. Bell, S.C. *Hum. Reprod.* **1**, 129–143 (1986).
5. Huhtala, M.-L. *et al. Endocrinology* (in the press).
6. Bell, S.C. *et al. J. clin. Endocr. Metab.* (in the press).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □