Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily

Thomas M. Jermann, Jochen G. Opitz, Joseph Stackhouse & Steven A. Benner

Department of Chemistry, ETH Zürich, CH-8092 Zürich, Switzerland

THE sequences of proteins from ancient organisms can be reconstructed from the sequences of their descendants by a procedure that assumes that the descendant proteins arose from the extinct ancestor by the smallest number of independent evolutionary events ('parsimony')[1,2]. The reconstructed sequences can then be prepared in the laboratory and studied[3,4]. Thirteen ancient ribonucleases (RNases) have been reconstructed as intermediates in the evolution of the RNase protein family in artiodactyls (the mammal order that includes pig, camel, deer, sheep and ox)[5]. The properties of the reconstructed proteins suggest that parsimony yields plausible ancient sequences. Going back in time, a significant change in behaviour, namely a fivefold increase in catalytic activity against double-stranded RNA, appears in the RNase reconstructed for the founding ancestor of the artiodactyl lineage, which lived about 40 million years ago[6]. This corresponds to the period when ruminant digestion arose in the artiodactyls, suggests that contemporary artiodactyl digestive RNases arose from a non-digestive ancestor, and illustrates how evolutionary reconstructions can help in the understanding of physiological function within a protein family[7–9].

The RNase A superfamily includes proteins that display many interesting but poorly understood biological activities, including immunosuppressivity[10], cytostatic activity[11], antitumour activity[12], endothelial-cell-stimulatory activity[13], and lectin-like behaviour[14], many of which have arisen by gene duplication since the time that mammals diverged from reptiles some 300 Myr ago. The abundance of RNase sequences from contemporary artiodactyls allows the reconstruction of the sequences of RNases that were the evolutionary intermediates in the most recent 40 Myr of this evolution (Table 1)[15]. Genes encoding the reconstructed proteins were obtained in the laboratory by site-directed mutagenesis from a synthetic gene for RNase[7]. The genes were then expressed in *Escherichia coli* and the resulting 'ancient' proteins purified to homogeneity using methods reported elsewhere[16–18].

To assess whether reconstruction by parsimony analysis yields proteins plausible as evolutionary intermediates in the evolution of the RNase family, the catalytic activities, substrate specificities, and thermal stabilities of the reconstructed RNases were examined. Most of the reconstructed proteins behave as expected for putative ancestral ruminant digestive RNases. This is particularly apparent when examining their kinetic properties (Table 2). Modern digestive RNases are catalytically active against small RNA substrates and single-stranded RNA[19], so presumably correctly reconstructed ancestral digestive RNases should retain these properties. Consistent with these expectations, the k_{cat}/K_m values for the putative ancestral RNases with UpA as substrate do not differ substantially from those of contemporary bovine digestive RNase (Table 2)[20]. The standard deviation of k_{cat}/K_m with UpA (uridylyl $3'\rightarrow5'$ adenosine) as substrate among the reconstructed ancestral enzymes, is only 25%. With poly(U) as substrate, the deviation is even smaller (18%). Thus, based simply on catalytic power, the sequences reconstructed by parsimony make plausible ancestral pancreatic RNases. Further, if this *in vitro* behaviour alone is accepted as a measure, at least some of the changes in the sequences of ruminant pancreatic RNases over the past 40 Myr appear to have been neutral.

Next, modern digestive enzymes generally are known to be stable to thermal denaturation. To learn whether the putative ancient RNase sequences behaved as digestive enzymes by this criterion, denaturation temperatures were measured (Table 3)[21]. Again, little change was observed in thermal stability back to ancestor **h**. The experimental melting temperatures for these ancient proteins differed with a standard deviation of 1.1 °C when compared with RNase A; typical experimental errors were ±0.5 °C. Therefore, by thermostability data as well as kinetic data, the reconstructions obtained by parsimony analysis are reasonable, at least back to ancestor **h**.

For the more ancient ancestors **i** and **j**, however, thermal stability decreases. The decrease is small, but lies outside experimental error. Of course, this decrease may reflect an incorrect reconstruction, but the change in thermal stability appears in the evolutionary tree at approximately the same time as another change in behaviour, the catalytic power of the reconstructed RNase for the hydrolysis of the duplex RNA, poly(A)·poly(U) (Table 2). Bovine digestive RNase A has only low catalytic activity against duplex RNA under physiological conditions; such activity is presumably not needed for a digestive enzyme. Reconstructed RNases dating back to about 40 Myr behave similarly. This changes markedly, however, in the reconstructed ancestor **h** and its immediate predecessors. With these reconstructed enzymes, catalytic activity against the double-stranded RNA substrate poly(A)·poly(U) is about five times higher than in the RNases that evolved from it (Table 2).

These changes in molecular behaviour correspond to a point in the divergent evolution of mammals where digestive physiology in ungulates also underwent substantial changes, ultimately yielding artiodactyls with 'true ruminant' foregut digestion. In true ruminants (including oxen, deer and giraffe), bacterial fermentation takes place in a stomach (the rumen) preceding the main digestive chambers. This physiology appears to have substantial adaptive value in many herbivorous environments; it may have convergently evolved in marsupial kangaroos, the colobine monkey primates, and more than once within the artiodactyl lineage itself[22]. Ruminants ferment cellulose with

FIG. 1 The evolutionary tree used in this work. Lower-case letters in the nodes in the graph designate putative intermediates in the evolution of the protein family (Table 1). Upper-case letters (D and G) indicate the residue at position 38 in the contemporary and reconstructed RNases. The timescale is approximate. The tree was adapted from ref. 15 with a single alteration. In our tree, the pig and the hippopotamus are joined together in a separate subfamily that branches together from the main line of descent. In the Beintema–Fitch tree, the pig and the hippopotamus diverge from the main line at separate points. Our tree reflects the classical grouping of pig and hippopotamus into a suborder *Suina*, and therefore incorporates some biological information in addition to the RNases themselves. This change has two effects. It reduces the number of putative ancestors on the main line of divergence by one, and helps resolve several ambiguities in the Beintema–Fitch reconstructions (Table 1). The revision does not alter any unambiguously assigned amino acid in the Beintema–Fitch reconstructions. Details of the reconstructions are available (email sab@ezrzl.vmsmail.ethz.ch). The placement of seminal RNase in the tree (as in ref. 15) is retained in light of ref. 30. However, as a rapidly diverging single isolated sequence, its placement on the tree may need future revision.

increased efficiency, and ruminant artiodactyls have been enormously successful in competition with the herbivorous perissodactyls (for example, horses, tapirs and rhinoceroses), which maintain fermentation in digestion in the caecum following the small intestine.

Barnard proposed over twenty years ago that fore-stomach digestion creates a need for especially large amounts of intestinal RNase[23]. Fermenting bacteria deliver large amount of RNA (mRNA, tRNA, rRNA) to the gastric region of the stomach and the small intestine; between 10 and 20% of the nitrogen in the diet of a typical bovid enters as RNA[23]. Consistent with this hypothesis is the fact that fore-stomach digesters have much higher levels of pancreatic RNase than other artiodactyls (such as the pig) and non-ruminant ungulates (such as the horse)[23].

The fact that a ribonuclease emerged with increased stability, decreased catalytic activity against duplex RNA, and increased levels of expression, at the same time as ruminant digestion emerged, may of course be a coincidence; but it may also indicate that ancestor **h** and its predecessors **i** and **j** were not specialized digestive enzymes of the bovine RNase A type, but rather played

TABLE 1 Sequence changes in reconstructed ancient ribonucleases

Bovine RNase A		Ancestral sequences **a**	**b**	**c**	**d**	**e**	**f**	**g**	**h₁**	**h₂**	**i₁**	**i₂**	**j₁**	**j₂**
3	Thr	Thr	Thr	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser*	Thr*	Ser	Ser
6	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Glu	Glu	Lys	Lys
15	Ser	Ser	Ser	Ser	Ser	Pro	Ser	Ser	Ser	Ser	Ser	Ser	Ser	
16	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gly*	Gly*	Gly*	Gly*	Gly	Gly
17	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Ser*	Thr*	Ser	Ser	Ser	Ser
19	Ala	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
20	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Ser	Ser	Ser	Ser	Ser
22	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Asn*	Asn*	Asn	Asn
31	Lys	Lys	Lys	Lys	Lys	Gln	Lys	Lys	Lys	Lys	Lys	Lys	Lys*	Lys*
32	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Arg	Arg	Arg	Arg
34	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn*	Lys*	Lys*	Lys*	Lys*	Asn	Asn
35	Leu	Met	Met	Leu	Leu	Leu	Leu	Leu*	Met	Met	Met	Met	Met	Met
37	Lys	Lys	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln
38	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Gly	Gly	Gly	Gly	Gly	Gly
59	Ser	Ser	Ser	Ser	Ser	Phe	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
64	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Thr	Thr	Thr	Thr	Thr	Thr
70	Thr	Thr	Thr	Thr	Thr	Ser	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr
76	Tyr	Tyr	Tyr	Tyr	Tyr	Asn	Tyr	Asn	Asn	Asn	Asn	Asn	Asn	Asn
78	Thr	Thr	Thr	Thr	Thr	Ala	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr
80	Ser	Ser	Ser	Ser	Ser	His	Ser	Arg*	Arg*	Arg*	His	His	His	His
96	Ala	Ala	Ala	Ala	Ala	Val	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala
100	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Ser	Ser	Ser	Ser
102	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Val	Val	Val*	Val*	Val*	Glu*
103	Asn	Lys	Lys	Lys	Glu	Glu	Glu	Glu	Glu	Glu	Gln	Gln	Gln	Gln

Reconstructed ancient sequences are designated by lower-case bold letters. The ancient sequences were adapted from Beintema *et al.*[15], calculated using the maximum parsimony procedure of Fitch[2] at the amino-acid level. A single modification in the Beintema *et al.* tree (Fig. 1 legend) resolved several ambiguities in the Beintema–Fitch reconstructions without altering any unambiguously assigned amino acids. Those amino acids marked with an asterisk indicate positions in which assignment depends on ambiguous parsimony reconstructions, or might be changed by plausible reorganization of the tree. In several of these cases, multiple sequences were reconstructed; subscripts indicate alternative sequence reconstructions for one node in the tree.

TABLE 2 Kinetic properties of reconstructed ancestral ribonucleases

RNase	Ancestor of	k_{cat}/K_m UpA $\times 10^6$	k_{cat}/K_m as % of RNase A	Poly(U) relative to RNase A	Poly(A)·poly(U) relative to RNase A
RNase A*		5.0	100	100	1.0
RNase A†		4.5	90	97	1.0
a	ox, buffalo, eland	6.1	122	106	1.4
b	ox, buffalo, eland, nilgai	5.9	118	112	1.0
c	b and the gazelles	4.5	91	97	0.8
d	Bovids	3.9	78	86	0.9
e	Deer	3.6	73	77	1.0
f	Deer, pronghorn, giraffe	3.3	67	103	1.0
g	Pecora	4.6	94	87	1.0
$\mathbf{h_1}$	Pecora and seminal RNase	5.5	111	106	5.2
$\mathbf{h_2}$	Pecora and seminal RNase	6.5	130	106	5.2
$\mathbf{i_1}$	Ruminata	4.5	90	96	5.0
$\mathbf{i_2}$	Ruminata	5.2	104	80	4.3
$\mathbf{j_1}$	Artiodactyla	3.7	74	73	4.6
$\mathbf{j_2}$	Artiodactyla	3.3	66	51	2.7

RNase names refer to nodes in the evolutionary tree shown in Fig. 1. All assays were performed at 25 °C. For UpA (Sigma), kinetic values were determined in 100 mM sodium acetate (pH 5.0). For poly(U), kinetic values were determined in 10 mM sodium acetate (pH 5.0) containing 150 NaCl and 20 μg ml^{-1} substrate following change in absorbance at 260 nm over a period of 90 s, using 200–250 ng of RNase per assay. For poly(A) · poly(U) (made by mixing poly(A) and poly(U) from Boehringer Mannheim in equimolar amounts and preannealing)[25], kinetic values were determined in 10 mM Tris-HCl (pH 7.3) containing 150 mM NaCl, 2 mM $MgCl_2$, and 30 μg ml^{-1} substrate following change in absorbance at 260 nm over a period of 5 min, using 1–2 μg RNase per assay[29].

* Expressed in *E. coli*.

† From Boehringer Mannheim.

TABLE 3 Thermal transition temperatures for reconstructed ancient ribonucleases

Enzyme	T_m °C	ΔT_m °C
RNase A*	59.3	0.0
RNase A†	59.7	+0.4
a	60.6	+1.3
b	61.0	+1.7
c	60.7	+1.4
d	58.4	−0.9
e	61.1	+1.8
f	58.6	−0.7
g	59.1	−0.2
$\mathbf{h_1}$	58.9	−0.5
$\mathbf{h_2}$	59.3	0.0
$\mathbf{i_1}$	58.2	−1.1
$\mathbf{i_2}$	58.7	−0.6
$\mathbf{j_1}$	56.5	−2.8
$\mathbf{j_2}$	57.1	−2.2

Melting temperatures (±0.5 °C) were determined according to ref. 21 in 100 mM sodium acetate (pH 5.0).

* Expressed in *E. coli*.

† From Boehringer Mannheim.

a non-digestive role. Of relevance to this hypothesis is the divergence of two non-digestive members of the RNase superfamily at approximately this point on the tree, RNase from brain[24] and RNase from seminal plasma. The physiological significance of catalytic activity against duplex RNA in non-digestive RNases is not yet known. It is interesting to note, however, that bovine seminal RNase has still higher catalytic activity against duplex RNA[25].

Unfortunately, the connectivity of deep branches in the evolutionary tree is not fully specified, either by sequence data or by fossil records, and remains disputed (Table 1)[26]. This makes conclusions that might be drawn from these experiments alone insecure. Therefore, we explored the structural origin of the increased catalytic activity of the ancestral RNases by further site-directed mutagenesis experiments. We found that a variant of h_1 that restores aspartic acid at position 38 (as in RNase A) has a catalytic activity against duplex RNA similar to that of RNase A[27]. Conversely, a variant of RNase A that introduces Gly alone at position 38 has catalytic activity against duplex RNA which is essentially that of ancestor **h**. These results show that substitution at a single position (residue 38) accounts for essentially all of the increased catalytic activity against duplex RNA in ancestor **h**.

As shown in Fig. 1, the reconstructed amino acids at position 38 are unambiguous throughout the tree, even at the most ancient nodes. Thus, it is highly probable that a change in catalytic activity against duplex RNA in fact occurred in RNases at this point. In one interpretation, catalytic activity against duplex RNA was not necessary in the descendent RNases, and therefore was lost. This implies that the replacement of Gly 38 by Asp in the evolution of ancestor **g** from ancestor **h** was neutral. We cannot, however, rule out an alternative model, that Asp 38 confers positive selective advantage on RNases found in advanced ruminants, an interpretation similar to that used to interpret the evolution of lysozymes in ruminants and their evolutionary analogues[22]. Interestingly, an Asp is present at position 38 both in many true ruminants and in the hippopotamus. Although the hippopotamus is not a true ruminant (it does not chew its cud), it does have a complex fore-stomach similar to that found in true ruminants[28]. This suggests the intriguing possibility that this substitution may have an adaptive function in RNases in organisms that have foregut digestion. In either case, these experiments show the value of parsimony analysis as a source of inspiration in experimental biochemistry and as a tool for understanding the physiological role of proteins better, and should encourage a more widespread use of evolutionary reconstruction as an experimental tool to guide site-directed mutagenesis. □

Received 9 May 1994; accepted 6 January 1995.

1. Pauling, I. & Zuckerkandl, E. *Acta chem. scand.* **17** (Suppl. 1), S9–S16 (1963).
2. Fitch, W. M. *Syst. Zool.* **20,** 406–416 (1971).
3. Stackhouse, J., Presnell, S. R., McGeehan, G. M., Nambiar, K. P. & Benner, S. A. *FEBS Lett.* **262,** 104–106 (1990).
4. Malcolm, B. A., Wilson, K. P., Matthews, B. W., Kirsch, J. F. & Wilson, A. C. *Nature* **345,** 86–89 (1990).
5. Carroll, R. L. *Vertebrate Paleontology and Evolution* (Freeman, New York, 1988).
6. Rose, K. D. *Science* **216,** 621–623 (1982).
7. Nambiar, K. P. *et al. Science* **223,** 1299–1301 (1984).
8. Adey, N. B., Tollefsbol, T. O., Sparks, A. B., Edgell, M. H. & Hutchinson, C. A. *Proc. natn. Acad. Sci. U.S.A.* **91,** 1569–1573 (1994).
9. Hillis, D. M., Huelsenbeck, J. P. & Cunningham, C. W. *Science* **264,** 671–677 (1994).
10. Soucek, J., Chudomel, V., Potmesilova, I. & Novak, J. T. *Nat. Immun. Cell Growth Reg.* **5,** 250–258 (1986).
11. Matousek, J. *Experientia* **29,** 858 (1973).
12. Ardelt, W., Mikulski, S. & Shogen, K. *J. biol. Chem.* **266,** 245–251 (1991).
13. Strydom, D. J. *et al. Biochemistry* **24,** 5486–5494 (1985).
14. Okabe, Y. *et al. J. Biochem. (Tokyo)* **109,** 786–790 (1991).
15. Beintema, J. J., Fitch, W. M. & Carsana, A. *Molec. Biol. Evol.* **3,** 262–275 (1986).
16. Nambiar, K. P., Stackhouse, J., Presnell, S. R. & Benner, S. A. *Eur. J. biochem.* **163,** 67–71 (1987).
17. McGeehan, G. M. & Benner, S. A. *FEBS Lett.* **247,** 55–56 (1989).
18. Trautwein, K. & Benner, S. A. in *Proc. Int. Symp. Site Directed Mutagenesis and Protein Engineering* (ed. El-Gewely, M. R.) 53–61 (Elsevier, New York, 1991).
19. Blackburn, P. & Moore, S. *The Enzymes* **15,** 317–434 (1982).
20. Ipata, P. L. & Felicioli, R. A. *FEBS Lett.* **1,** 29–31 (1968).
21. Lang, K. & Schmidt, F. X. *Eur. J. Biochem.* **159,** 275–281 (1986).
22. Jolles, J. *et al. J. molec. Evol.* **28,** 528–533 (1989).
23. Barnard, E. A. *Nature* **221,** 340–344 (1969).
24. Watanabe, H. *et al. J. Biochem. (Tokyo)* **104,** 939–945 (1988).
25. Libonati, M. & Floridi, A. *Eur. J. Biochem.* **8,** 81–87 (1969).
26. Graur, D. *FEBS Lett.* **325,** 152–159 (1993).
27. Ciglic, M. diplomarbeit, ETH Zürich (1994).
28. Stevens, C. E. *Comparative Physiology of the Vertebrate Digestive System* (Cambridge Univ. Press, Cambridge, UK, 1988).
29. Sorrentino, S., Carsaña, A., Furia, A., Doskocil, J. & Libonati, M. *Biochim. biophys. Acta* **609,** 40–52 (1980).
30. Breukelman, H. J. *et al. J. molec. Evol.* **37,** 29–35 (1993).

ACKNOWLEDGEMENTS. We are indebted to S. Leutenegger and F. Dreher for assistance in the early stages of this work, to G. Gonnet for discussion, and to J. Beintema for communicating a set of reconstructed sequences.