Rat pancreatic ribonuclease: agreement between the corrected amino acid sequence and the sequence derived from its messenger RNA

Jaap J. Beintema

Biochemisch Laboratorium, Rijksuniversiteit, Nijenborgh 16, 9747 AG Groningen, The Netherlands

Received 15 June 1983

A corrected amino acid sequence of rat pancreatic ribonuclease is presented which is in agreement with the messenger RNA sequence in [J. Biol. Chem. (1982) 257, 14582–14585]. The corrections do not change the general position of rat ribonuclease in trees which can be constructed for ribonuclease sequences, although they do place the rat sequence somewhat closer to the mouse sequence. The evolutionary rate of ribonuclease in the rodent family of the Muridae (rat and mouse) now has been calculated to be 140 nucleotide substitutions per 10^8 years per 100 codons, and still is one of the highest rates yet observed. The occurrence of 4 additional amino acids at the C-terminus in several mammalian ribonucleases is in agreement with the position of a second stop codon in the 3′ non-coding region of the rat messenger RNA sequence.

Ribonuclease *Evolutionary rate* *(Rat pancreas)* *Signal sequence* *Stop codon*

1. INTRODUCTION

Relatively shortly after the elucidation of the primary structure of bovine pancreatic ribonuclease [1], the amino sequence of rat pancreatic ribonuclease was determined [2,3]. The rat enzyme was found to differ at 44 positions from the bovine enzyme, but residues which were known to be important for structural stability and enzymic activity were identical in both enzymes. This comparative study of bovine and rat ribonuclease was followed later by a more extensive study of the molecular evolution of pancreatic ribonuclease in mammals [4,5]. This demonstrated that ribonucleases from species in different placental mammalian orders differ at 25–40 positions, with the rat enzyme having the most deviating sequence. Yet, in a most parsimonious tree, the rat sequence groups with the sequence of mouse, and these two group with the sequences of two other myomorph rodent species, hamster and muskrat, which is in agreement with accepted zoological classification [4,5]. Consequently, the deviating sequence of rat ribonuclease is caused by an increased evolutionary rate of ribonuclease in the myomorph rodents in general and in the branch leading to rat after the divergence from mouse in particular [4–6].

The sequence of rat pancreatic ribonuclease messenger RNA in [7] demonstrated 5 differences with the protein sequence we have determined. Here, I present a corrected sequence of rat pancreatic ribonuclease, which is in agreement with the sequence derived from the messenger RNA, and discuss several evolutionary aspects arising from new data on the protein and mRNA sequence.

2. MATERIALS AND METHODS

Rat pancreatic ribonuclease was isolated by extraction with 0.125 M sulfuric acid, ammonium sulfate fractionation [8] and affinity chromatography [9]. Trypsin (treated with L-1-tosylamide-2-phenyl ethyl chloromethyl ketone, TPCK) and α-chymotrypsin were from Worthington Biochemicals (Freehold NJ). Other materials were as in [10].

Rat ribonuclease (4 mg) was oxidized with performic acid and digested with 40 μg TPCK-treated

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

trypsin at 37°C in 400 μl of 0.2 M ammonium bicarbonate for 3.5 h. Peptides were isolated by gel filtration on Sephadex G-25 and high-voltage paper electrophoresis at pH 3.5. Only tryptic peptides T10 and T15 [3], which contain the sequences to be reinvestigated, were isolated. The positions of these peptides on the electropherogram were identified by staining for peptides containing histidine (T10) or tyrosine (T15) [11]. Peptide T10 was further purified by high-voltage paper electrophoresis at pH 6.5 and thin-layer chromatography with 1-butanol–acetic acid–water–pyridine (15:3:12:10, by vol.). Peptide T15 (90 nmol) was digested with 10 μg chymotrypsin at 37°C in 100 μl 0.2 M ammonium bicarbonate for 2.5 h and the digest was fractionated by high-voltage paper electrophoresis at pH 3.5 (peptides T15C..).

The amino acid sequences of the peptides were determined by dansyl-Edman degradation, with the Edman degradation in a 10-times scaled-down version of the method in [12]. Dansylated amino acids were identified by thin-layer chromatography [12], except for dansyl-cysteic acid, which was identified by paper electrophoresis at pH 4.4 [13]. During dansyl-Edman degradations the C-terminal residues of peptides were generally identified by analysis of both hydrolyzed and non-hydrolyzed samples of the dansylated reaction products.

Amino acid analyses were performed with a Kontron Liquimat III amino acid analyzer. Amide positions were assigned by electrophoresis at pH 6.5 [14] of peptides both before and after several Edman degradation steps. During dansyl-Edman degradation these assignments were confirmed by conversion of the aspartic acid and glutamic acid derivatives to their phenylthiohydantoin derivatives and their identification by high-pressure liquid chromatography [15], if these residues had been identified as the dansylated derivatives of aspartic acid or glutamic acid after the previous Edman step.

3. RESULTS AND DISCUSSION

The amino acid sequences of peptides T10 and T15 and the results of the dansyl-Edman degradations are shown in fig. 1. The peptide pattern obtained after gel filtration of the tryptic digest is shown in fig. 2. Table 1 summarizes the fractionation procedures used for the purification of each peptide, the electrophoretic mobilities of the peptides at pH 6.5 and their amino acid compositions.

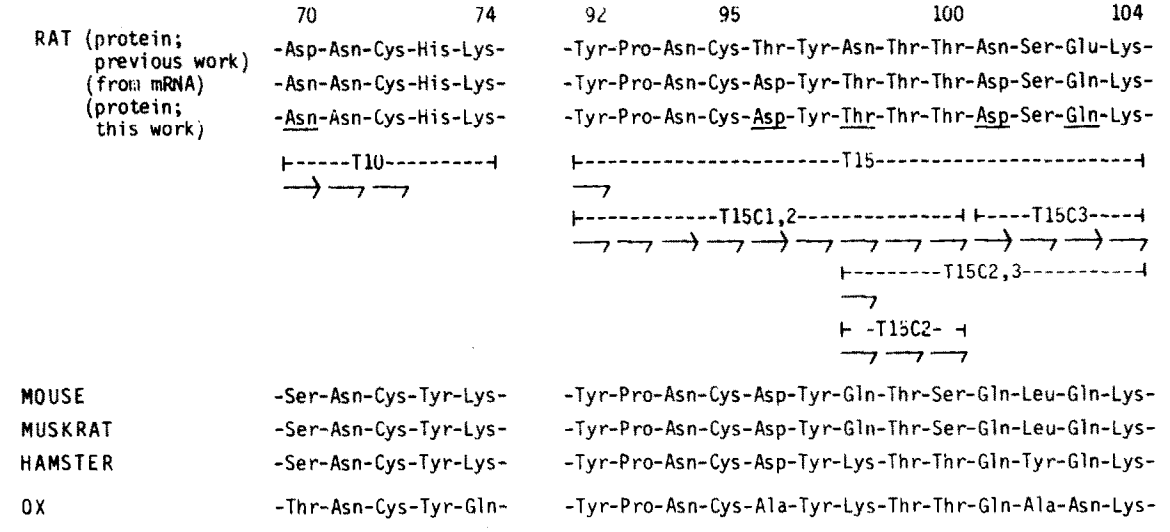

Fig. 1. Amino acid sequence of corrected peptides of rat pancreatic ribonuclease: (——) peptides used in reconstruction of the sequence; (– – –) less satisfactory amino acid analysis, peptide nomenclature in section 2; (→) identified as dansyl amino acid; (→) identified as phenylthiohydantoin derivative. The amino acid residues are numbered according to the sequence of bovine pancreatic ribonuclease (1). Corrected residues are underlined. For comparison the sequences of homologous parts of other myomorph rodent ribonucleases and of bovine ribonuclease are included.

Table 1

Procedures used for purification of peptides from rat pancreatic ribonuclease, charge determinations of peptides at pH 6.5 and amino acid compositions of peptides

Peptide	T10	T15	T15C1,2	T15C3	T15C2,3	T15C2
Sephadex G-25 pool (fig. 2)	E	D				
Paper electrophoresis pH 3.5 (m_{Lys})	0.69	0.22	0.02	0.54	0.50	0.43
pH 6.5 (m_{Asp})	+0.30		−0.51	0	0	
pH 6.5 charge	0 to +1		−2	0	0	
after 4 (T15C1,2) or 1 (T15C3) (m_{Asp})		−0.4 to −0.6	+0.56			
after 4 (T15C1,2) or 1 (T15C3) charge		−1	+1			
Edman steps						
Thin-layer chrom. (R_f)	0.05					
Amount (nmol)	9	110	32	31	9	5
Position in sequence	70–74	92–104	92–100	101–104	98–104	98–100
Amino acid:						
Cysteic acid	0.9 (1)	1.0 (1)	1.0 (1)			
Aspartic acid	1.8 (2)	3.1 (3)	2.2 (2)	1.0 (1)	1.1 (1)	0.8
Threonine		3.1 (3)	2.1 (3)		2.6 (3)	3.0 (3)
Serine		1.1 (1)		1.0 (1)	1.3 (1)	1.0
Glutamic acid		1.1 (1)		1.1 (1)	1.1 (1)	0.6
Proline		1.0 (1)	1.0 (1)			
Glycine				0.2	0.3	0.8
Tyrosine		1.8 (2)	2.0 (2)			
Lysine	1.1 (1)	0.9 (1)		0.9 (1)	0.9 (1)	
Histidine	1.1 (1)					

Mobilities from the origin relative to lysine (m_{Lys}) and from the neutral band to aspartic acid (m_{Asp}) are given for peptides isolated by paper electrophoresis at pH 3.5 and 6.5, respectively. Mobilities from the origin relative to the front (R_f) are given for peptides isolated by thin-layer chromatography. Amino acid analysis of peptides were performed on samples hydrolyzed in 0.25 ml 6 M HCl at about 110°C in evacuated sealed tubes. Values were not corrected for hydrolysis losses. The composition of each peptide as derived from the amino acid sequence determinations is shown within parentheses.

3.1. *Peptide T10 (residues 70–74)*

This peptide co-purified with peptide T11 (Ser-Ser-Ser-Thr-Leu-Arg; residues 75–80) during gel filtration and paper electrophoresis at pH 3.5. Partially purified peptide T10 contaminated with 0.3 equivalents of T11 was obtained after electrophoresis at pH 6.5 and was used for dansyl-Edman degradation. Its final purification was achieved by thin-layer chromatography; this preparation was used for amino acid analysis.

In [3] we presented the N-terminal sequence Asp–Asn for this peptide; a form with N-terminal Asp–Asp was also obtained. However, the mRNA sequence indicates the sequence Asn–Asn [7]. Here, we confirm the latter result. Staining for histidine-containing peptides on the electropherogram (pH 3.5) demonstrated that no deamidated forms of the peptide were present. Previously we used column chromatography on Dowex 1 at 37°C, with a starting buffer of pH 9, for purification of the peptides. Probably these conditions have caused a rather specific deamidation of this Asn–Asn sequence.

3.2. *Peptide T15 (residues 92–104)*

Analysis of subpeptides obtained after cleavage with chymotrypsin confirmed that the sequence predicted from the mRNA sequence [7] is correct

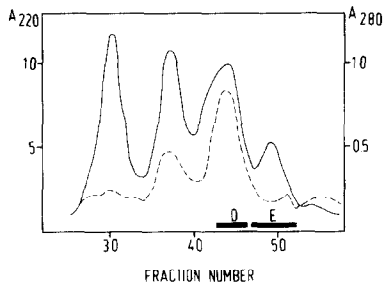

Fig. 2. Gel filtration of the tryptic digest of performic acid-oxidized rat ribonuclease on a column (0.5 × 100 cm) of Sephadex G-25 (superfine). Elution was carried out with 0.2 M acetic acid at 2 ml/h; 0.3 ml fractions were collected. Fractions were pooled as indicated by bars (see table 1): (——)A_{220}; (– – –) $A280$.

and that two interchanges should be made in the sequence proposed in [3]. Our present results indicate that chymotryptic cleavage of the Thr–Asp (100–101) bond was more extensive than that of the Tyr–Thr (97–98) bond. Some dansyl-tyrosine (residue 97) was identified without hydrolysis after the fifth step of the Edman degradation of peptide T15C1,2, indicating that peptide T15C1 (residues 92–97) was a minor contaminant of peptide T15C1,2. This explains the low threonine content in the amino acid analysis of the latter.

In [3], we could not obtain direct evidence for the sequence of this peptide because of lack of material and the less sophisticated techniques existing at that time. A tentative sequence was proposed from incomplete data, with several residues being positioned by homology with ribonucleases from non-related species.

3.3. *Comparison with other ribonucleases*

In addition to the sequence evidence, fig. 1 also shows the homologous parts of the sequences of three other myomorph rodent ribonucleases [6,16,17] and of bovine ribonuclease [1]. The corrected rat sequence differs at the same number of amino acid residues from bovine ribonuclease, as the non-corrected one. However, the corrections at position 96 (Asp) and 103 (Gln) place the rat sequence closer to those of the other myomorph rodents. The correction of Asp to Asn at position 70 does not lower the number of amino acid replacements at this position, but asparagine is one nucleotide substitution closer to serine, observed in the other myomorph rodents, and to threonine, in most non-myomorph rodent sequences.

The corrections in the rat sequence do not influence the topology of several kinds of trees derived for ribonucleases [4,5]. There are 5 fewer nucleotide substitutions in the branch to rat after the divergence from mouse. This results in slightly lower evolutionary rates for ribonuclease as compared with earlier calculations [5]: for instance, 140 instead of 150 nucleotide replacements per 10^8 years per 100 codons in the *Muridae* (rat and mouse) [5]. This value still is very high compared with evolutionary rates calculated both for other substitutions changing amino acids and for silent substitutions [6]. Knowledge of the mouse messenger ribonuclease sequence will be interesting to compare the number of silent substitutions with the high number of amino acid-changing substitutions in rat and mouse ribonuclease.

3.4. *Messenger RNA sequence*

In [7], the signal peptide of rat pre-ribonuclease was shown to have the same length as that predicted for bovine pre-ribonuclease [18]; the addition of 3 residues at the N-terminus of the mature enzyme might be the result of a small duplication event [7]. Since this addition is not observed in mouse and other ribonucleases, this duplication must have taken place after the divergence of rat and mouse, which occurred about 10^7 years ago [5].

The amino acid composition of the signal peptide part of rat pre-ribonuclease as derived in [7] differs in two respects from the composition of the mature enzyme: The signal peptide contains 8 leucines among 25 residues, while the mature protein contains only 3 leucines among 127 residues. A low leucine content is typical of all pancreatic ribonucleases sequenced to-date. The other feature is the presence of a tryptophan residue in the signal peptide. This amino acid residue is not present in the mature protein and has not been found in any other pancreatic ribonuclease. The presence of tryptophan is essential for the in vitro synthesis of ribonuclease by slices of rat pancreas [19]. This observation can now been explained by the presence of tryptophan in the pre-enzyme.

Rat ribonuclease has no addition or deletion at the C-terminus relative to bovine ribonuclease. In several other ribonucleases additional residues at the C-terminus have been found: 4 residues in two-

toed sloth and several hystricomorph rodent ribonucleases [20–22]; 3 residues in human ribonuclease (unpublished); and 2 residues in horse ribonuclease [23]. These additions can be explained by substitutions of the stop codon at position 125. The frequent occurrence of 4 additional residues can be explained by the additional stop codon at position 129 which has been found [7] in the messenger of rat ribonuclease (AUG). Earlier we noticed that the sequences of the C-terminal additions clearly are not random [20]. Partial conservation of this sequence is also indicated by the fairly good agreement between the sequence of the rat messenger RNA at positions 126–128 (GGC UUC ACG) and the nucleotide sequence at these positions predicted from amino acid sequences of ribonucleases with additions at the C-terminus (VCY UCX ACX; with V = A,G or C; Y = C or U and X = A,G,C or U; unpublished).

ACKNOWLEDGEMENTS

I thank Mr P. Jekel for isolating rat ribonuclease and performing the amino acid analyses and Dr R.N. Campagne for critically reading the manuscript.

REFERENCES

[1] Smyth, D.G., Stein, W.H. and Moore, S. (1963) J. Biol. Chem. 238, 227–234.
[2] Beintema, J.J. and Gruber, M. (1967) Biochim. Biophys. Acta 147, 612–614.
[3] Beintema, J.J. and Gruber, M. (1973) Biochim. Biophys. Acta 310, 161–173.
[4] Beintema, J.J., Gaastra, W., Lenstra, J.A., Welling, G.W. and Fitch, W.M. (1977) J. Mol. Evol. 10, 49–71.
[5] Beintema, J.J. and Lenstra, J.A. (1982) in: Macromolecular Sequences in Systematic and Evolutionary Biology (Goodman, M. ed) pp. 43–73, Plenum, New York.
[6] Lenstra, J.A. and Beintema, J.J. (1979) Eur. J. Biochem. 98, 399–408.
[7] MacDonald, R.J., Stary, S.J. and Swift, G.H. (1982) J. Biol. Chem. 257, 14582–14585.
[8] Beintema, J.J., Campagne, R.N. and Gruber, M. (1973) Biochim. Biophys. Acta 310, 148–160.
[9] Wierenga, R.K., Huizinga, J.D., Gaastra, W., Welling, G.W. and Beintema, J.J. (1973) FEBS Lett. 31, 181–185.
[10] Gaastra, W., Welling, G.W. and Beintema, J.J. (1978) Eur. J. Biochem. 86, 209–217.
[11] Easley, C.W. (1965) Biochim. Biophys. Acta 107, 386–388.
[12] Hartley, B.S. (1970) Biochem. J. 119, 805–822.
[13] Gray, W.R. (1967) Methods Enzymol. 11, 139–151.
[14] Offord, R.E. (1966) Nature 211, 591–593.
[15] Vereijken, J.M., Hofsteenge, J., Bak, H.J. and Beintema, J.J. (1980) Eur. J. Biochem. 113, 151–157.
[16] Van Dijk, H., Sloots, B., Van den Berg, A., Gaastra, W. and Beintema, J.J. (1976) Int. J. Pept. Prot. Res. 8, 305–316.
[17] Jekel, P.A., Sips, H.J., Lenstra, J.A. and Beintema, J.J. (1979) Biochimie 61, 827–839.
[18] Haugen, T.H. and Heath, E.C. (1979) Proc. Natl. Acad. Sci. USA 76, 2689–2693.
[19] Campagne, R.N. and Gruber, M. (1962) Biochim. Biophys. Acta 55, 353–360.
[20] Havinga, J. and Beintema, J.J. (1980) Eur. J. Biochem. 110, 131–142.
[21] Beintema, J.J., Knol, G. and Martena, B. (1982) Biochim. Biophys. Acta 705, 102–110.
[22] Beintema, J.J. and Neuteboom, B. (1983) J. Mol. Evol. 19, 145–152.
[23] Scheffer, A.J. and Beintema, J.J. (1974) Eur. J. Biochem. 46, 221–233.