The Amino Acid Sequence of Pike-Whale (Lesser-Rorqual) Pancreatic Ribonuclease

By MARCHIENUS EMMENS, GJALT W. WELLING and JAAP J. BEINTEMA
*Biochemisch Laboratorium, Rijksuniversiteit, Zernikelaan, de Paddepoel, Groningen,
The Netherlands*

(Received 13 November 1975)

Pancreatic RNAase (ribonuclease) from the pike whale (lesser rorqual, *Balaenoptera acutorostrata*) was isolated by affinity chromatography. The protein was digested with different proteolytic enzymes. Peptides were isolated by gel filtration, preparative high-voltage paper electrophoresis and paper chromatography. The amino acid sequence of peptides was determined by the dansyl-Edman method. Although we do not have an amino acid composition for the whole protein, all peptide bonds were overlapped by one or more peptides. Residues 85–96 are bridged by a peptide of unsatisfactory composition and the sequence here depends, at least in part, on homology for its confirmation. Another region in which a similar situation obtains is residues 39–40. This pancreatic RNAase differs at 24–33% of the positions from all other mammalian pancreatic RNAases sequenced to date, except for pig RNAase, from which it differs by 19%. This indicates that whale RNAase has evolved independently during the larger part of the evolution of the mammals. Lesser-rorqual pancreatic RNAase is partially glycosidated (30%) at asparagine-76 in an Asn-Ser-Thr sequence (residues 76–78). Pig RNAase also has carbohydrate attached to asparagine-76 and is identical with lesser-rorqual RNAase in residues 76–98. Detailed evidence for the sequence has been deposited as Supplementary Publication SUP 50066 (11 pages) at the British Library Lending Division, Boston Spa, Wetherby, W. Yorkshire LS23 7BQ, U.K., from whom copies may be obtained on the terms given in *Biochem. J.* (1976) 153, 5.

To date many pancreatic RNAases* have been sequenced (Welling *et al.*, 1975a). Because of the relatively high number of differences in primary structure, the comparative study of ribonucleases may give valuable information on structure–function relationships but also for phylogenetic studies of RNAases (Welling *et al.*, 1975b).

In the present paper we describe the determination of the primary structure of pancreatic RNAase of the pike whale (lesser rorqual, *Balaenoptera acutorostrata*).

Experimental

Materials and Methods

Pancreatic tissue from the lesser rorqual was collected during a whaling cruise in the Barents Sea and was kept frozen until use. All materials and methods were the same as described by Welling *et al.* (1975a) unless otherwise stated.

* Abbreviations: RNAase, ribonuclease; RNAase A, ribonuclease without carbohydrate; RNAase B, ribonuclease with carbohydrate.

Isolation of RNAase

Several preparations of lesser-rorqual RNAase were made from individual samples of pancreatic tissue by extraction with H_2SO_4 , $(NH_4)_2SO_4$ precipitation (Zwiers *et al.*, 1973) and affinity chromatography as described by Wierenga *et al.* (1973). The RNAase activity was measured as described by Campagne & Gruber (1962). A separation of the glycosidated and non-glycosidated components of the lesser-rorqual RNAase was achieved by chromatography on CM-cellulose (CM-32, W. and R. Balston, Maidstone, Kent, U.K.) as described by Åqvist & Anfinsen (1959), with minor modifications (see the legend of Fig. 1).

The number of normally and abnormally titrating tyrosine residues in bovine and lesser-rorqual RNAase was determined spectrophotometrically as described by Tanford *et al.* (1955).

Cleavage methods and peptide nomenclature

Hydrolysis with proteolytic enzymes was carried out with three independent preparations of lesser-rorqual RNAase. RNAases A (35 mg) and B (17 mg)

were separately aminoethylated and digested with trypsin as described by Zwiers *et al.* (1973). The numbers of the peptides obtained from this digest are prefixed by the letter T. The other digests were prepared from mixtures of RNAase A and B. Lesser-rorqual RNAase (44mg) was oxidized with performic acid at 0°C (Hirs, 1956) and digested with thermolysin as described by Ambler & Meadway (1968). The numbers of the fragments obtained from this digest are prefixed by the letter H. Another portion of oxidized lesser-rorqual RNAase (40mg) was digested with a proteinase from *Staphylococcus aureus* V8 (Glu-enzyme) at pH 7.8 as described by Houmard & Drapeau (1972). The numbers of the fragments obtained from this digest are prefixed by the letter G. Whale RNAase was also cleaved with CNBr as described by Welling *et al.* (1975a). Only one peptide of this digest was used for sequencing and it is denoted by the letters CN.

Secondary cleavages

Large peptides of the thermolysin and Glu-enzyme digest were digested with trypsin. Peptides (100–400nmol) were dissolved in 250µl of 0.2M-*N*-ethylmorpholine acetate, pH 8.1, and 10µl of trypsin solution (1 mg/ml in 0.001M-HCl) was added. The mixture was incubated at 37°C for 2h, and the resulting peptides were separated by preparative paper electrophoresis at pH 3.5 as described by Ambler (1963).

Peptide studies

The first purification step of the peptides from the different digests was gel filtration on Sephadex G-25 (superfine grade) in 0.1M-acetic acid. The H and T peptides were fractionated on a column (1cm×200cm) at flow rates of 16 and 6ml/h respectively, 1.8ml fractions being collected. The G peptides were fractionated on a column (1.2cm×300cm) at a flow rate of 5ml/h, 1.8ml fractions being collected. Determination of the peptide positions in column eluates, detection of peptides on paper, elution of peptides, dansylation, sequence determination and positions of amide groups were done as described by Welling *et al.* (1975a). Automatic sequencing was done in a JEOL sequence analyser or in a Beckman 890C sequencer. The ratio of monosaccharide residues in glycopeptide TA was determined after methanolysis by g.l.c. essentially as described by Clamp *et al.* (1971).

Results

The elution pattern of the affinity chromatography was comparable with that of giraffe RNAase (Gaastra *et al.*, 1974). From three samples of pancreatic tissues (500–700g each) 300mg of RNAase was isolated in

a yield of about 60% of the initial activity. After polyacrylamide-gel electrophoresis, three enzymically active bands were visible (see Fig. 1). The activity of lesser-rorqual RNAase on RNA at pH 7.5 is about 30% on a molar basis compared with that of bovine RNAase A.

From the tyrosine titration (Tanford *et al.*, 1955), a ratio of normal to abnormal tyrosine residues in lesser-rorqual RNAase of 2:3 was determined, which means that lesser-rorqual RNAase contains one normal tyrosine residue less than does bovine RNAase.

Determination of the *N*-terminal residue showed only arginine. By automatic Edman degradation of lesser-rorqual RNAase with a JEOL sequence analyser, the first 16 amino acid residues were determined, and with a Beckman automatic sequencer we could determine the sequence up to residue 34.

Sequence determination

The amino acid analyses of the isolated peptides are given in Tables 1, 2, 3 and 4 of Supplementary Publication SUP 50066. The results of the sequence determination are summarized in Fig. 2.

Specificity of the proteolytic enzymes

Thermolysin cleaved as expected from its known specificity (Ambler & Meadway, 1968), except for a Thr-Ser (residues 99–100) cleavage. Digestion with the staphylococcal proteinase resulted in cleavage next to glutamic acid (Houmard & Drapeau, 1972). However, cleavage next to aspartic acid residues 14, 53, 83 and 121 also occurred. Cleavage after residue 121 with this proteinase was also found in dromedary RNAase (Welling *et al.*, 1975a) and in kangaroo RNAase (B. van Dijk, G. W. Welling, W. Gaastra & J. J. Beintema, unpublished work). More remarkable was the cleavage of the Thr-Phe bond (residue 45–46).

RNAase A and B

After column chromatography on CM-32, RNAase A and B were obtained in a ratio of 7:3 (see Fig. 1). We could isolate the carbohydrate-containing peptide from RNAase B very easily from the tryptic digest after gel filtration on Sephadex G-25. It was found in the void volume of the column eluate, indicated as pool A (see Fig. 2b of the supplement). This peptide (TA) has the same amino acid composition as peptide T12 from pool C of lesser-rorqual RNAase A; it only differs in the presence of glucosamine. A rough estimate of the glucosamine content of the peptide was obtained by amino acid analysis of samples hydrolysed during 20–24h in 6M-HCl *in vacuo* at about 110°C. This procedure showed 3.9 glucosamine residues. However, 3h hydrolysis in

Fig. 1. Chromatography of purified lesser-rorqual RNAase on a column (1.5 cm $\times$ 40 cm) of CM-cellulose

The column was equilibrated with 0.01 M-sodium phosphate buffer, pH 6.0, at 4°C. After application of the sample a linear gradient was started, generated from 200 ml of the first buffer and 200 ml of 0.15 M-sodium phosphate, pH 7.5. The elution was continued after this gradient with 150 ml of the last buffer. Flow rate, 20 ml/h; fraction size, 4.4 ml; E_{280} , ●; conductivity, ----. Fractions were pooled as indicated by the horizontal bars. Disc-gel-electrophoresis patterns of the pooled fractions are shown over the elution pattern. Electrophoreses in polyacrylamide gels were done as described by Welling *et al.* (1975a). The gels were stained for protein with Amido Black.

4M-HCl *in vacuo* at 100°C of a sample of RNAase B gave a glucosamine content of 4.5 residues/molecule. Carbohydrate analysis of the glycopeptide TA gave the following ratio of sugars (normalized to a glucosamine content of five residues per mole): glucosamine, 5.0; mannose, 5.9; galactose, 4.3; fucose, 2.8; glucose, 0.7; sialic acid, traces.

Supplementary information

Detailed evidence for the amino acid sequence of the lesser-rorqual RNAase has been deposited with the British Library Lending Division, Boston Spa, Wetherby, W. Yorks. LS23 7QB, U.K. (Supplementary Publication SUP 50066). The information comprises: (1) Tables 1, 2 and 3 showing the compositions of the H, G and T peptides; (2) Table 4 showing the amino acid composition of the CN peptide and of the peptides of the secondary cleavages; (3) Table 5 giving the evidence for the presence or absence of amide groups on each aspartic acid or

glutamic acid residue [evidence from the automatic Edman degradation is given for residues up to position 28, and electrophoretic mobilities at pH 6.5 (Offord, 1966) were used in assignments from residues 26 onwards]; (4) information on which pool of collected fractions of the gel filtration each peptide originates from, together with the electrophoretic mobilities, is given in Table 6; (5) Figs. 1 and 2 showing the elution pattern of the gel filtrations of the Glu-enzyme digest and the two tryptic digests, and the 'fingerprint' of the G-peptides made by paper electrophoresis at pH 3.5.

Discussion

Reliability of the sequence determination

The N-terminal amino acid sequence of a mixture of lesser-rorqual RNAases A and B was determined twice by automatic sequencing and resulted unambiguously in one amino acid at each position up to residue 34. Analysis of the peptides of the thermolysin and Glu-enzyme digest of a mixture of RNAases A and B did not show any heterogeneity. The only difference between the peptide 'maps' of the tryptic digest of RNAases A and B after gel filtration is the position of peptide T12, which is glycosidated in RNAase B (denoted as peptide TA). Therefore we may conclude that the only difference in primary structure between lesser-rorqual RNAase A and B is the carbohydrate chain at position 76.

The repetitive features of the sequence Gln-Met-Met-Met-Arg-Arg-Lys (residues 27-34) caused us some concern in view of the fact that it was initially determined by automatic sequencing. However, a number of overlapping peptides from different digests were isolated (see Figs. 2 and 3). The rather long peptide G3 did not give the correct number of methionine sulphone residues, in contrast with peptide H2, which contains five such residues. The sequence was supported by dansyl-Edman degradation of tryptic peptides T4 and T5. All the digests produced peptides from regions 81-91, which have an unexplained excess of aspartic acid. Peptide H8 was a mixture and the residues 85-86 are bridged by a peptide of unsatisfactory composition. However, the extra aspartic acid was not seen when peptides were sequenced, and as this region is usually invariant, it seems most reasonable to propose that the sequence is as shown.

Comparison of the sequence of common- and lesser-rorqual RNAase

The tryptic and chymotryptic peptides isolated from the pancreatic RNAase of common rorqual (*Balaenoptera physalus*) (Yamada *et al.*, 1974) could

Fig. 2. Amino acid sequence of pike-whale (lesser-rorqual) pancreatic RNAase

The numbering of bovine RNAase was used. The sequence was determined by dansyl-Edman degradation (→). The residues determined by automatic sequencing are also indicated (↗). Those amino acids differing from bovine RNAase are indicated in boxes. For peptide nomenclature, see the Experimental section. Attention has been drawn in the synopsis to those regions of the sequence that depend on evidence that is less than totally satisfactory.

Automatic sequencing	-----	Asn-Gln-Met-Met-Met-Arg-Arg-Lys
G3 residues 3-45	-----	Asn, Gln, Met, Met, Met, Arg, Arg, Lys-----
H2 residues 8-35	-----	Asn, Gln, Met, Met, Met, Arg, Arg, Lys, Met
T4 residues 27-32		Asn-Gln-Met-Met-Met, Arg
T5 residues 33-34		Arg-Lys

Fig. 3. Peptides showing the sequence of the residues 27-34

Attention is drawn to the part of the Discussion section 'Reliability of the sequence determination'.

all be placed in our sequence of lesser-rorqual RNAase. Only two parts of the common-rorqual sequence are missing, residues 33-34 and residues 86-91 respectively. Peptide T-I of the tryptic digest of common-rorqual RNAase was the only peptide with a deviating amino acid composition. This amino acid composition is very similar to that of the sequence residues 11-32 of lesser-rorqual RNAase. However, the latter shows one proline residue more and one valine residue less. Yamada *et al.* (1974) could not determine the *N*-terminal residue and suggested a blocked α -amino group of lysine (residue 7). A comparison of the tryptic peptides of common-rorqual RNAase and the amino acid sequence presented here showed that peptide T-XI of Yamada *et al.* (1974) has exactly the same sequence as peptide T1 of lesser-rorqual RNAase and probably represents the first seven amino acid of common-rorqual RNAase.

In conclusion, we suggest that the primary structure of the pancreatic RNAase of common rorqual is identical with that of lesser rorqual, except for a probable Val→Pro replacement in the region residues 11-32. This points to an extreme conservation of primary structure of ribonucleases within the Balaenopteridae. The same kind of conservation has been observed within the subfamily of the Caprinae (Welling *et al.*, 1974) and the tribe Bovini (Muskiet *et al.*, 1976) of the Bovidae.

Activity

At pH 7.5 the activity of lesser-rorqual RNAase on RNA is approximately one-third of that of bovine RNAase A. Irie *et al.* (1966) reported for common-rorqual RNAase A a relative activity of 12.5% and for RNAase B 25% at pH 8.0, also compared with the bovine enzyme.

Kinetic studies using cytidine 2':3'-cyclic phosphate and uridine 2':3'-cyclic phosphate as sub-

strates showed similar activity for lesser rorqual and bovine RNAase (Ronda *et al.*, 1976). Irie (1968) observed the same for common-rorqual RNAase. Another interesting feature of lesser-rorqual ribonuclease is its high activity on double-stranded RNA (M. Libonati, personal communication). These various enzymic properties have also been found for dimeric bovine seminal RNAase and its artificial monomer (Hosokawa & Irie, 1971; Floridi *et al.*, 1972; Libonati *et al.*, 1975).

Comparison with other RNAases

Lesser-rorqual RNAase differs in 30 positions from bovine RNAase (see Table 1). Four replacements involve hydrophobic amino acids; 13 of the substitutions bring about difference in charge. The total net charge of the lesser-rorqual RNAase is 4 units higher than that of the bovine enzyme, irrespective of the additional histidine residue at position 80. Compared with other RNAases, lesser-rorqual and bovine seminal RNAase (Di Donato & D'Alessio, 1973) are the most positively charged RNAases. It is noteworthy that both these RNAases are also rather similar in enzymic properties.

Four substitutions are probably the result of paired mutations (Wyckoff, 1968). Glutamic acid-52 and lysine-55 can form a salt bridge and compensate the change of charge. The same holds for the replacement of Ala-Asn in bovine RNAase by Lys-Glu (residues 102-103).

All hydrophilic replacements, compared with bovine RNAase, are located at or near the surface of the molecule. Two of the hydrophobic replacements, methionine-6 and leucine-64 are also found in RNAases of Hystricomorph rodents (van den Berg & Beintema, 1975). The replacement of lysine by methionine (residue 31) has only been found in the lesser-rorqual RNAase. The fourth hydrophobic substitution at position 35, Leu→Met, is internal.

Table 1. *Difference matrix of pancreatic RNAases*

Data were derived from publications mentioned in the introduction. A deletion was treated as an amino acid. Compared with horse and rat RNAase, the relative differences were calculated on the basis of 126 and 127 amino acid residues respectively. The heterogeneity in dromedary RNAase (Welling *et al.*, 1975a) is treated as half a difference. Latin names for the species investigated are as follows: cow (*Bos taurus*), giraffe (*Giraffa camelopardalis*), reindeer (*Rangifer tarandus*), dromedary (*Camelus dromedarius*), pig (*Sus scrofa*), lesser rorqual (*Balaenoptera acutorostrata*), horse (*Equus caballus*) and rat (*Rattus rattus*).

	Cow	Giraffe	Reindeer	Dromedary	Pig	Lesser rorqual	Horse	Rat
Cow	—	11	14	28	26	30	35	44
Giraffe	9%	—	14	33	28	32	33	47
Reindeer	11%	11%	—	32	25	30	33	45
Dromedary	22%	27%	26%	—	28½	32	33	41
Pig	21%	22%	20%	23%	—	23	30	46
Lesser rorqual	24%	26%	24%	26%	19%	—	31	42
Horse	28%	26%	27%	27%	24%	25%	—	45
Rat	35%	37%	36%	33%	37%	33%	37%	—

Generally it is assumed that the conformation of the main chain is similar in homologous proteins. From the titration behaviour of the tyrosine residues one can obtain an indication of the conservation of the three-dimensional structure. Bovine RNAase contains three abnormally and three normally titrating tyrosine residues, the former at positions 25, 92 and 97, the latter at positions 73, 76 and 115. Because of the Tyr-Asn substitution (at position 76) we expected two normally titrating tyrosine residues, and this proved to be the case in lesser-rorqual RNAase.

However, changes in the main-chain conformation of external loops may occur. Welling *et al.* (1975a) demonstrated that the so-called S-peptide loop of different RNAases shows a different susceptibility to digestion with subtilisin. Lesser-rorqual RNAase is not cleaved by subtilisin (Welling *et al.*, 1975b). The presence of proline residues at positions 19 and 23 probably modifies the course of the main chain in the S-peptide loop, with the result that this loop does not fit into the active site of subtilisin.

In Table 1 a difference matrix of pancreatic RNAases is shown. Lesser-rorqual RNAase differs to the extent of between 24 and 33% of its residues from all other pancreatic RNAases thus far sequenced, except for pig ribonuclease, where the difference is 19%. These findings may be expected from an evolutionary point of view.

Carbohydrate. Irie *et al.* (1966) found the same relative amounts of free carbohydrate (RNAase W1) and bound carbohydrate (RNAase W2) in pancreatic RNAase of common rorqual as we found for lesser

rorqual. The carbohydrate chain in RNAase B is attached to the amide group of the asparagine residue in an Asn-X-Ser/Thr sequence. In the amino acid sequence of lesser-rorqual pancreatic RNAase one carbohydrate attachment site is present, with the amino acid sequence Asn-Ser-Thr (residues 76-78). The attachment site in the common-rorqual RNAase B (RNAase W2) also possesses the sequence Asn-Ser-Thr (Tsuruo *et al.*, 1970) and as the amino acid sequences of common- and lesser-rorqual RNAases are very similar, it may be assumed that asparagine-76 is also the carbohydrate-attachment site in common-rorqual RNAase B.

Pig pancreatic RNAase is so far the only other RNAase with a carbohydrate chain at asparagine-76; however, in this case all molecules are glycosidated. The amino acid sequence around asparagine-76 of pig and lesser-rorqual RNAase is identical for residues 75-98. In other RNAases the amino acid composition of this sequence varies from 10 to 30%. From a three-dimensional model of ribonuclease-S, it is obvious that there is no special reason for this conservation of sequence in whale and pig ribonuclease only.

We thank Dr. A. Jønsgård of the Institutt for Marin Biologi in Oslo and Mr. I. Christensen of the Fiskeridirektoratets Havsforskning sinstitut in Bergen (Norway) for the collection of lesser-rorqual pancreas. We also thank Mr. G. Gerwig, Dr. J. P. Kamerling and Dr. J. F. G. Vliegthart of the Laboratory of Organic Chemistry, University of Utrecht, for performing the carbohydrate analysis.

References

- Ambler, R. P. (1963) *Biochem. J.* **89**, 349–378
- Ambler, R. P. & Meadway, R. J. (1968) *Biochem. J.* **108**, 893–895
- Åqvist, S. E. G. & Anfinsen, C. B. (1959) *J. Biol. Chem.* **234**, 1112–1117
- Campagne, R. N. & Gruber, M. (1962) *Biochim. Biophys. Acta* **55**, 353–360
- Clamp, J. R., Bhatti, T. & Chambers, R. E. (1971) *Methods Biochem. Anal.* **19**, 229–344
- Di Donato, A. & D'Alessio, G. (1973) *Biochem. Biophys. Res. Commun.* **55**, 919–928
- Floridi, A., D'Alessio, G. & Leone, E. (1972) *Eur. J. Biochem.* **26**, 162–167
- Gaastra, W., Groen, G., Welling, G. W. & Beintema, J. J. (1974) *FEBS Lett.* **41**, 227–233
- Hirs, C. H. W. (1956) *J. Biol. Chem.* **219**, 611–621
- Hosokawa, S. & Irie, M. (1971) *J. Biochem. (Tokyo)* **69**, 683–697
- Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3506–3509
- Irie, M. (1968) *J. Biochem. (Tokyo)* **63**, 501–505
- Irie, M., Yamada, T. & Ukita, T. (1966) *J. Biochem. (Tokyo)* **59**, 344–352
- Libonati, M., Malorni, M. C., Parente, A. & D'Alessio, G. (1975) *Biochim. Biophys. Acta* **402**, 83–87
- Muskiet, F., Welling, G. W. & Beintema, J. J. (1976) *Int. J. Peptide Protein Res.* in the press
- Offord, R. E. (1966) *Nature (London)* **211**, 591–593
- Ronda, G. J., Gaastra, W. & Beintema, J. J. (1976) *Biochim. Biophys. Acta* in the press
- Tanford, C., Hauenstein, J. D. & Rands, D. G. (1955) *J. Am. Chem. Soc.* **77**, 6409–6413
- Tsuruo, T., Sudo, K., Terao, T. & Ukita, T. (1970) *Biochim. Biophys. Acta* **200**, 560–563
- van den Berg, A. & Beintema, J. J. (1975) *Nature (London)* **253**, 207–210
- Welling, G. W., Scheffer, A. J. & Beintema, J. J. (1974) *FEBS Lett.* **41**, 58–61
- Welling, G. W., Groen, G. & Beintema, J. J. (1975a) *Biochem. J.* **147**, 505–511
- Welling, G. W., Leijenaar-van den Berg, G., van Dijk, B., van den Berg, A., Groen, G., Gaastra, W., Emmens, M. & Beintema, J. J. (1975b) *BioSystems* **6**, 239–245
- Wierenga, R. K., Huizinga, J. D., Gaastra, W., Welling, G. W. & Beintema, J. J. (1973) *FEBS Lett.* **31**, 181–185
- Wyckoff, H. W. (1968) *Brookhaven Symp. Biol.* **21**, 252–257
- Yamada, T., Nakazawa, Y. & Ukita, T. (1974) *J. Biochem. (Tokyo)* **75**, 153–164
- Zwiers, H., Scheffer, A. J. & Beintema, J. J. (1973) *Eur. J. Biochem.* **36**, 569–574