Amino Acid Sequence of the Nonsecretory Ribonuclease of Human Urine†

Jaap J. Beintema,*,‡ Jan Hofsteenge,§ Masanori Iwama,∥ Toru Morita,∥ Kazuko Ohgi,∥ Masachika Irie,∥ Ruth H. Sugiyama,⊥ Gary L. Schieven,⊥ Charles A. Dekker,⊥ and Dohn G. Glitz#

Biochemisch Laboratorium, Rijksuniversiteit, Nijenborgh 16, 9747 AG Groningen, The Netherlands, Friedrich Miescher Institut, P.O. Box 2543, 4002 Basel, Switzerland, Department of Microbiology, Hoshi College of Pharmacy, Ebara, Shinagawa-ku, Tokyo 142, Japan, Department of Biochemistry, University of California, Berkeley, California 94720, and Department of Biological Chemistry, UCLA School of Medicine, University of California, Los Angeles, California 90024

Received September 11, 1987; Revised Manuscript Received January 27, 1988

ABSTRACT: The amino acid sequence of a nonsecretory ribonuclease isolated from human urine was determined except for the identity of the residue at position 7. Sequence information indicates that the ribonucleases of human liver and spleen and an eosinophil-derived neurotoxin are identical or very closely related gene products. The sequence is identical at about 30% of the amino acid positions with those of all of the secreted mammalian ribonucleases for which information is available. Identical residues include active-site residues histidine-12, histidine-119, and lysine-41, other residues known to be important for substrate binding and catalytic activity, and all eight half-cystine residues common to these enzymes. Major differences include a deletion of six residues in the (so-called) S-peptide loop, insertions of two, two, and nine residues, respectively, in three other external loops of the molecule, and an addition of three residues at the amino terminus. The sequence shows the human nonsecretory ribonuclease to belong to the same ribonuclease superfamily as the mammalian secretory ribonucleases, turtle pancreatic ribonuclease, and human angiogenin. Sequence data suggest that a gene duplication occurred in an ancient vertebrate ancestor; one branch led to the nonsecretory ribonuclease, while the other branch led to a second duplication, with one line leading to the secretory ribonucleases (in mammals) and the second line leading to pancreatic ribonuclease in turtle and an angiogenic factor in mammals (human angiogenin). The nonsecretory ribonuclease has five short carbohydrate chains attached via asparagine residues at the surface of the molecule; these chains may have been shortened by exoglycosidase action. This characteristic plus the resistance of the protein to denaturation and proteolytic degradation suggests that it originated in lysosomes and escaped digestion by lysosomal proteinases due to its high net positive charge and compact structure.

Sierakowska and Shugar (1977) classified the pyrimidine-specific mammalian neutral and alkaline ribonucleases into two classes: secretory and nonsecretory ribonucleases. The secretory ribonucleases, found mostly in secretory organs, show optimal enzymic activity with RNA at around pH 8.0; bovine pancreatic ribonuclease is a typical enzyme of this class. The sequence, three-dimensional structure, and many properties of this enzyme have been extensively studied (Blackburn & Moore, 1982). Amino acid sequences and a number of enzymatic and structural features have been determined for about 40 other ribonucleases from the pancreas of other mammals, from turtle pancreas, and from bovine semen. These results have allowed the study of the molecular evolution of this enzyme (Beintema et al., 1986; Beintema, 1987). The nonsecretory ribonucleases have a lower preference for poly(C) than the secretory ribonucleases. They cleave RNA most effectively at pH 6.5–7.0 and occur predominantly in liver, lung, spleen, and leucocytes. Ribonucleases of this class may be lysosomal enzymes (Morita et al., 1986). Several isolations of ribonucleases from lysosome-rich sources have been described (Delaney, 1963; Ohta et al., 1982; Niwata et al., 1985).

Both secretory and nonsecretory ribonucleases occur in body fluids including urine, serum, cerebral spinal fluid, and semen, and both enzymes have been purified from human urine (Cranston et al., 1980; Iwama et al., 1981; Sugiyama, 1981). The secretory ribonuclease (also known as RNase C, RNase U_L, and band A in the references mentioned before) was found to be identical with the pancreatic enzyme, except in the glycosylation characteristics and except for the presence of an additional C-terminal residue (Beintema et al., 1985a). Primary structure investigations of the nonsecretory ribonuclease (also known as RNase U, RNase U_s and band D) have been initiated in order to learn more about the structural and evolutionary relationships between the two classes of ribonucleases. Preliminary sequence information already indicated homology in the region of the active-site histidines (Beintema et al., 1986). In this paper we present the complete amino acid sequence of the nonsecretory ribonuclease, present some information about its glycosylation pattern, and give evidence that both classes of enzyme belong to a superfamily of ribonucleases, which also includes turtle pancreatic ribonuclease (Beintema et al., 1985b) and human angiogenin (Strydom et al., 1985).

Materials and Methods

Nonsecretory Ribonuclease from Human Urine. The batch of protein used for the digestions with trypsin and chymotrypsin after reduction and carboxymethylation (see Table I) was isolated following the procedure developed by Sugiyama (1981):

† This work has been supported by a travel grant from the Collaborative Research Grant Programme of the NATO Scientific Affairs Division to enable J.J.B. to visit Berkeley.

* Address correspondence to this author.
‡ Rijksuniversiteit.
§ Friedrich Miescher Institut.
∥ Hoshi College of Pharmacy.
⊥ University of California, Berkeley.
University of California, Los Angeles.

0006-2960/88/0427-4530$01.50/0

Human urine (50 L), collected from healthy volunteers, was concentrated tenfold by rotatory evaporation (<25 °C) or by dialysis against poly(ethylene glycol) (Carbowax 6000) at 4 °C. The concentrated urine was dialyzed against water and made 5 mM in sodium phosphate, pH 6.7. Sodium azide (0.05%) was added as a preservative. Phosphocellulose fractionation (Yamanaka et al., 1977) was used for further concentration and partial purification. Application of the sample in 0.01 M sodium phosphate, pH 6.7, and elution with 2.0 M NaCl in the same buffer gave fractions enriched in the nonsecretory ribonuclease as ascertained by polyacrylamide gel electrophoresis and activity staining (Blank & Dekker, 1981).

(*A*) *Sephadex G-75 Chromatography.* The pooled phosphocellulose fractions (666 mL; 4.5×10^6 units) were dialyzed against 5 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) buffer, pH 7.4, lyophilized, and taken up in 2 mL of 2.5 M NaCl–0.02 M Tris-HCl buffer, pH 7.4. The sample was loaded on a Sephadex G-75 column (1.4 × 100 cm) preequilibrated with 0.5 M NaCl–0.02 M Tris-HCl buffer, pH 7.4, and eluted with equilibration buffer. Fractions with the highest ribonuclease activities (88–114 mL) were pooled.

(*B*) *SP-Sephadex C-50 Chromatography.* The pooled Sephadex G-75 fractions were lyophilized, and the volume was brought to 5 mL with water. The sample was desalted by gel filtration on a Sephadex G-25 column and loaded on an SP-Sephadex C-50 column (1.5 × 25 cm), previously equilibrated with 0.05 M sodium acetate buffer, pH 5.2. Elution of ribonuclease activity was achieved with a salt gradient (0–0.6 M NaCl) in 0.05 M sodium acetate buffer, pH 5.2, as described by Cranston et al. (1980). The fractions with the highest activity (240–260 mL) were pooled. Polyacrylamide gel electrophoresis followed by activity staining revealed slight contamination with the secretory ribonuclease components HU_B and HU_C (Sugiyama et al., 1981).

(*C*) *Preparative Isoelectric Focusing.* Sephadex G-75 was allowed to swell in water and washed until the conductivity of the wash water was less than 1 $\mu\Omega^{-1}$. The gel used for preparing the electrophoresis bed was made 2% in arginine (extends pH range to 11.4) and 2% in carrier ampholytes, pH range 3–10. The pooled SP-Sephadex column fractions were lyophilized, taken up in 5 mL of water, and twice dialyzed against 3 L of water. The sample was again lyophilized, brought to 4.0 mL with water, and made 2% in ampholytes. It was loaded on the gel bed and focused in the electrophoresis cell for 13.5 h at an average power of 5.9 W. The gel bed was cut into 1-cm sections, each of which was poured into a small column and eluted with three column volumes of 0.02 M Tris-HCl, pH 7.4. The wicks under the cathode were soaked in the same buffer. Each eluate was assayed for ribonuclease activity and an equivalent number of units (0.6) loaded on each of two sodium dodecyl sulfate (SDS)–12.5% polyacrylamide gels. Following electrophoresis, one gel was subjected to activity staining and the other to silver staining. Uncontaminated nonsecretory ribonuclease was found in the first 1-cm section (nearest cathode) of the isoelectric focusing gel and in the cathode wicks. The appropriate solutions were pooled and dialyzed against three changes (4 L each) of 0.01 M Tris-HCl buffer, pH 7.5. After removal of traces of solids by centrifugation, the nonsecretory ribonuclease was lyophilized and stored at −15 °C: yield, 6 mg; recovery, 18%; purification, 200-fold.

Another batch of nonsecretory ribonuclease was isolated as described by Iwama et al. (1981) and was used for the three digests of the reduced and pyridylethylated protein (see Table I).

Other Materials. Sephadex G-25, G-50, and G-75 and SP-Sephadex G-50 were obtained from Pharmacia. Phosphocellulose (Selectacel phosphate, type 40, 1.02 mequiv/g) was purchased from Brown Co. Trypsin (treated with N^{α}-tosylphenylalanine chloromethyl ketone) and α-chymotrypsin were from Worthington Biochemical Corp. (Freehold, NJ), and thermolysin was supplied by Sigma Chemical Co. (St. Louis, MO). 4-(*N*,*N*-Dimethylamino)azobenzene 4′-isothiocyanate (DABITC) was obtained from Pierce (Rockford, IL). 4-Vinylpyridine (Aldrich) was distilled under reduced pressure and stored at −20 °C. Tri-*n*-butylphosphine (Aldrich) was used as such; it was also stored at −20 °C. Other materials were as described previously (Gaastra et al., 1978; Vereijken et al., 1984) or, otherwise, were of reagent grade.

Analytical Procedures. Ribonuclease activity was assayed as previously described by Blank and Dekker (1981) but with the Tris-HCl buffer replaced by 0.035 M imidazole hydrochloride, pH 6.5. SDS–polyacrylamide gel electrophoresis followed by activity staining was conducted according to the procedure of Blank and Dekker (1981). Silver staining for protein was carried out according to Morrissey (1981). Isoelectric focusing was conducted in a Bio-Rad preparative unit, Model 1405, using Pharmalyte, an ampholyte obtained from Pharmacia. Procedures followed were those contained in the proprietary literature of Bio-Rad and Pharmacia.

Amino Acid Sequence Determination. Reduction and S-carboxymethylation of ribonuclease were done as described by Crestfield et al. (1963). Alternatively, certain batches of 2.5 mg of ribonuclease were pyridylethylated by treatment with 5 μL of tributylphosphine plus 5 μL of 4-vinylpyridine in 500 μL of 50% pyridine in water (v/v) for 3 h at 45 °C (Amons, 1987), followed by lyophilization. The protein was then desalted on a small column of Sephadex G-25 (Beintema et al., 1984) in 30% acetic acid and again lyophilized. This treatment with tributylphosphine and 4-vinylpyridine was repeated after digestion with CNBr (Table I) before fractionation of the CNBr peptides by gel filtration (Figure 1A of the supplementary material; see paragraph at end of paper regarding supplementary material).

The amino acid sequence of human nonsecretory ribonuclease was derived from five separate digests. The procedures for cleavage and the nomenclature of the cleavage products are summarized in Table I. Peptides, numbered according to their position in the polypeptide chain, were isolated by gel filtration, following by high-performance liquid chromatography (HPLC) on a Nucleosil $10C_{18}$ column (300 × 4.6 mm) with a linear gradient of 0–67% acetonitrile in 0.1% trifluoroacetic acid (TFA) over 60 min at a flow rate of 1.0 mL/min. Some peptides were subjected to another purification step by HPLC on the same column, with a linear gradient of 0–67% acetonitrile in 0.1% ammonium acetate (pH 6.0) over 60 min at a flow rate of 1.0 mL/min. Peptide elution was monitored by the absorbance at 214 nm.

Automatic Edman degradation was performed on a gas-phase sequenator (Applied Biosystems, Model 470A) (Hewick et al., 1981), and identification of phenylthiohydantoin derivatives was accomplished by HPLC according to Hunkapillar and Hood (1983), with tetrahydrofuran added to buffer A. During later stages of the work an on-line phenylthiohydantoin analyzer (Applied Biosystems) was used. Procedures for dansylation, Edman degradation combined with dansylation, and identification of dansylated amino acids by thin-layer chromatography have been described by Hartley (1970).

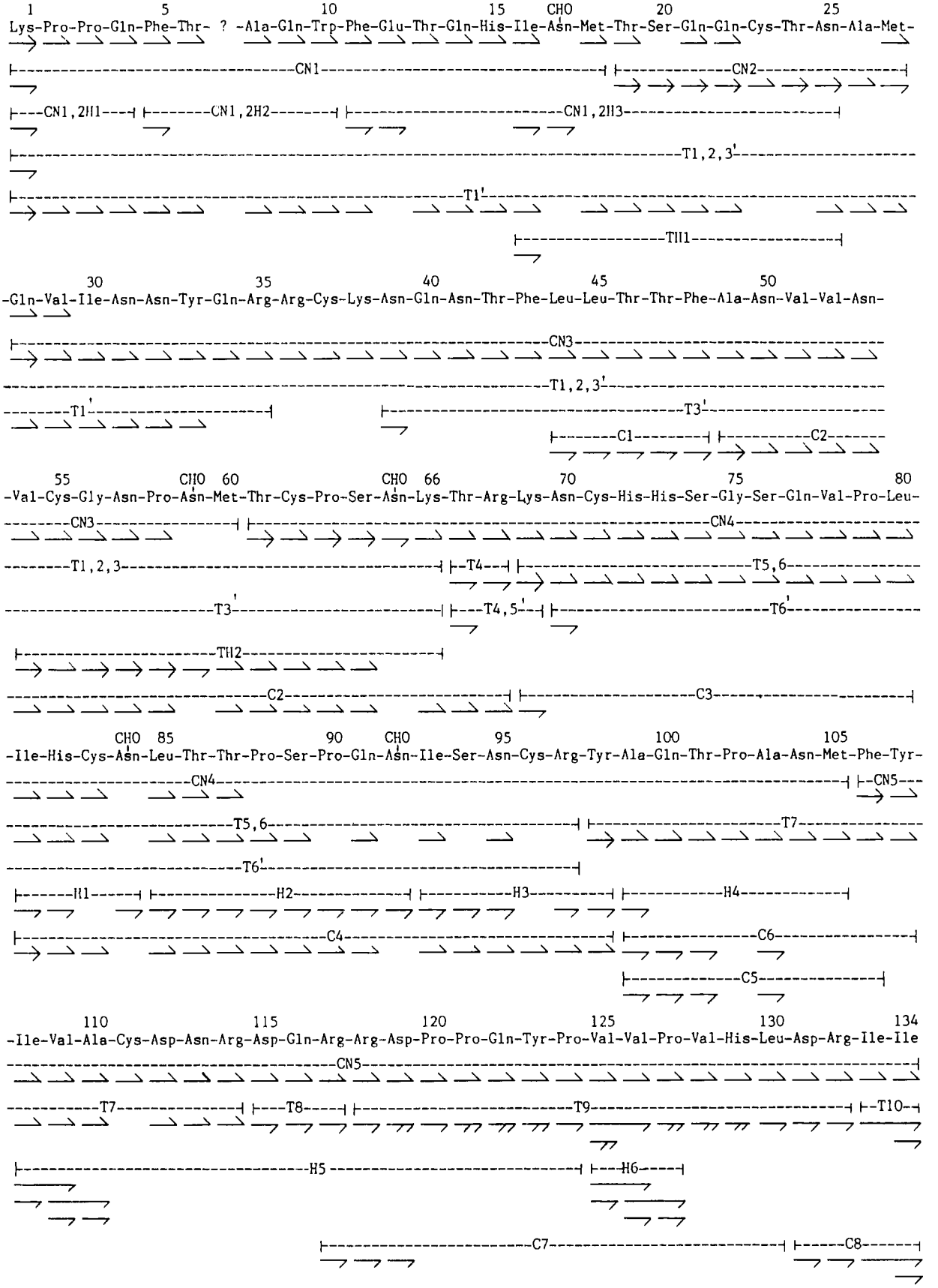


FIGURE 1: Amino acid sequence of human nonsecretory ribonuclease. (|---|) Peptides used as proof for the sequence. The nomenclature of the peptides is presented in Table I. (→) Identified as phenylthiohydantoin during automatic Edman degradation; (⇀) identified as dansyl amino acid; (⟶) identified as dansyl dipeptide; (⇁) identified as (dimethylamino)azobenzene thiohydantoin derivative. CHO, carbohydrate; ?, residue not identified.

Degradation with (dimethylamino)azobenzene isothiocyanate (DABITC) was performed as described by Chang (1983). Amino acid analyses were performed with a Kontron Liquimat III amino acid analyzer.

Table I: Procedures for Cleavage of Peptide Bonds and Nomenclature of the Cleavage Products[a]

protein/peptide treated	quantity treated	cleavage procedure				nomenclature of cleavage products
		enzyme/reagent	quantity	time (h)	vol (μL)	
reduced and carboxymethylated ribonuclease	3 mg	trypsin	25 μg	4	300	T1–T10
reduced and carboxymethylated ribonuclease	2 mg	chymotrypsin	20 μg	3	200	C1–C8
reduced and pyridylethylated ribonuclease	2.5 mg	CNBr	50 mg	20	100	CN1–CN5
reduced and pyridylethylated ribonuclease	2.5 mg	trypsin	40 μg	4	400	T1′–T9′
reduced and pyridylethylated ribonuclease	2.5 mg	(a) trypsin	30 μg	3	300	TH1, TH2 (pool A)
		(b) directly followed by thermolysin	30 μg	3		H1′, H2′ (pool B)
Sephadex G-50 pools of CNBr digest						
A		thermolysin	50 μg	5	500	H1–H6
D	15 nmol	thermolysin	15 μg	5	150	CN1,2H1–CN1,2H3

[a] All enzymic cleavages were done in 0.2 M ammonium bicarbonate buffer, pH 8.0, at 37 °C. Treatment with cyanogen bromide was done in 70% trifluoroacetic acid at room temperature. Peptides with identical sequence (except cysteine modification) but from different digests have the same symbol and number, but differ in the addition of the prime (′) symbol.

Results

The amino acid sequence of human nonsecretory ribonuclease was derived from an automatic Edman degradation run on the reduced and carboxymethylated protein and from structural analyses of peptides derived from five separate digests. The results of the amino acid sequence determination are summarized in Figure 1.

The major portion of the proof for the sequence is formed by the automatic Edman degradations on the intact, reduced, and carboxymethylated protein and on the five CNBr peptides. These peptides formed a complete set and were obtained after reduction with tributylphosphine, followed by pyridylethylation of the protein. The additional positive charges introduction by this treatment facilitated the recovery of peptides during separation on HPLC at low pH. The order of peptides CN1, CN2, and CN3 was obtained from the sequencer run on the intact protein and on the peptide T1′. Peptide TH2 completed the C-terminal sequence of CN3 and established the overlap with CN4. Similarly, the sequence of C4 and T7 yielded the C-terminus of CN4 and the overlap with CN5.

There is only a one-residue overlap of peptides at the position of Tyr-98. Further evidence for this overlap is provided by the amino acid analysis of the rather long CNBr peptide CN4 and by the fact that peptide T7 is the only peptide with N-terminal tyrosine. The three other tyrosine residues in the protein have been identified at other positions in the sequence (positions 33, 107, and 123).

Peptide recoveries from a tryptic and from a chymotryptic digest of the reduced and carboxymethylated protein were incomplete and poor due to rather hydrophobic properties of the protein and the uneven distribution of cleavage sites through the protein. These peptides confirmed the overall amino acid sequence.

The supporting data are presented as supplementary material and include the gel filtration patterns of the digests, a figure showing the recoveries of phenylthiohydantoins during automatic Edman degradation, and tables with details of the purification procedures used and the amino acid analyses of peptides.

Most amino acid residues were identified at least twice as their phenylthiohydantoin derivatives during automatic Edman degradation. The presence of five glycosylated asparagine residues in the sequence at positions 17, 59, 65, 84, and 92, where no phenylthiohydantoin derivative could be identified during automatic Edman degradation, was demonstrated by separate dansyl–Edman degradations of peptides (Figure 1), by amino acid analysis of several peptides (supplementary material), and by their occurrence in Asn-X-Ser/Thr sequences. No non-glycosylated Asn-X-Ser/Thr sequences were demonstrated in the protein.

At position 7 no phenylthiohydantoin derivative could be identified during automatic Edman degradation. However, in this case the identity of the amino acid residue has not yet been determined. Neither amino acid analysis of acid hydrolyzates of several peptides nor dansyl–Edman degradation of peptide CN1,2H2 provided any indication regarding the identity of the residue at this position. The characterization of the putative posttranslational modification requires further study, a task made difficult because of the small amounts of peptides that can be isolated from this part of the protein.

Several C-terminal residues of peptides were identified by analysis of both hydrolyzed and nonhydrolyzed reaction products during dansyl–Edman degradation after cleavage of the penultimate residue of the peptide. In this way the C-terminal Ile-Ile sequence in tryptic peptide T10 and chymotryptic peptide C8 was identified.

Discussion

The amino acid sequence of the nonsecretory ribonuclease as presented confirms previously published partial sequences of this protein with only a few corrections (Cranston et al., 1980; Niwata et al., 1985; Beintema et al., 1986). The N-terminal amino acid sequence of human liver ribonuclease was also found to be identical with that of this enzyme (D. G. Glitz, unpublished). Amino acid compositions of nonsecretory ribonuclease preparations from human urine, spleen, and liver also are generally in good agreement with the composition calculated from the sequence (Table II). These data, together with immunological studies (Morita et al., 1986), indicate that all of these studies deal with the same or very closely related gene products.

The C-terminal sequence of human nonsecretory ribonuclease is Ile-Ile. Cranston et al. (1980) also found isoleucine as the C-terminal residue in their preparation of the protein by carboxypeptidase digestion. However, they derived a different sequence for other residues in the C-terminal sequence from the results of this experiment.

Gleich et al. (1986) published the N-terminal sequence of human eosinophil-derived neurotoxin; it differs from that of the nonsecretory ribonuclease only in having threonine instead of valine at position 54, the penultimate residue sequenced in this protein. This neurotoxin has a specific activity on RNA similar to that of pancreatic ribonucleases (Slifman et al., 1986), as does the nonsecretory ribonuclease (Iwama et al., 1981). Gleich et al. (1986) also were not able to identify the residue at position 7. These observations indicate that the

Table II: Comparison of Earlier Determined Amino Acid Compositions of Human Ribonucleases from Urine, Spleen, and Liver with That Calculated from the Amino Acid Composition of Human Nonsecretory Ribonuclease

	Delaney (1963)		Cranston et al. (1980), urine	Iwama et al. (1981), urine	Glitz (unpublished), liver	from the sequence (this paper), urine
	spleen	urine				
aspartic acid	19.8	21.4	22.2	21.9	20	21
threonine	9.9	10.8	11.2	10.8	12	12
serine	7.0	5.7	6.6	7.6	8	6
glutamic acid	14.3	15.3	14.3	14.5	14	14
proline	12.1	12.2	12.0	12.3	12	12
glycine	4.5	3.0	4.8	3.2	3	2
alanine	6.4	6.2	5.9	7.0	6	6
half-cystine	6.8	6.8	7.3	5.1	*a*	8
valine	8.0	8.7	7.8	7.3	11	9
methionine	3.3	3.6	4.2	4.6	3	4
isoleucine	6.6	6.6	6.6	4.5	7	7
leucine	5.5	4.9	5.4	5.7	5	5
tyrosine	4.4	3.9	3.7	3.7	4	4
phenylalanine	4.9	4.8	4.9	4.7	5	5
lysine	5.2	4.2	4.1	4.9	4	4
histidine	4.9	4.9	4.9	4.3	5	5
arginine	7.3	7.5	7.9	8.3	8	8
tryptophan	0.8	0.8	nd	1.2	nd	1
unknown						1
total						134

[a] Half-cystine below valine peak.

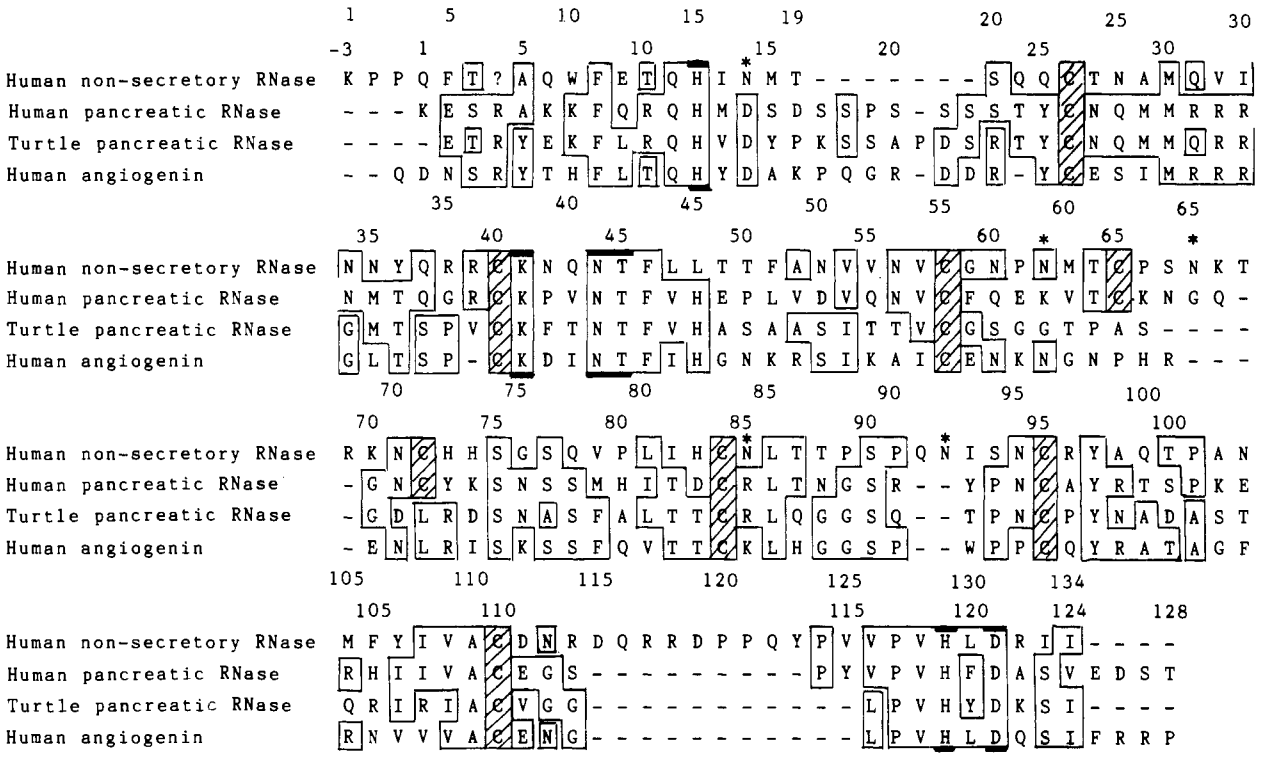


FIGURE 2: Comparison of the amino acid sequences of human nonsecretory ribonuclease, human pancreatic ribonuclease (Beintema et al., 1984), turtle pancreatic ribonuclease (Beintema et al., 1985b), and human angiogenin (Strydom et al., 1985). Positions are numbered according to the residue number of human nonsecretory ribonuclease (upper line) or human pancreatic ribonuclease (lower line). (–) No amino acid residue present; (?) identity unknown. Identical residues in two or more sequences are enclosed in blocks. (Hatched lines) Half-cystine residues; (heavy lines) residues in the active site or involved in substrate binding in bovine pancreatic ribonuclease; (*) carbohydrate attachment sites in human nonsecretory ribonuclease.

neurotoxin may be identical with the nonsecretory ribonuclease.

In Figure 2 we align the sequence of human nonsecretory ribonuclease with the sequences of human pancreatic ribonuclease (Beintema et al., 1984), turtle pancreatic ribonuclease (Beintema et al., 1985b), and human angiogenin (Strydom et al., 1985). If one deletion and three insertions are introduced at external loops of the structure of mammalian pancreatic ribonucleases, a perfect alignment is obtained not only for the active-site residues His-12, His-119, and Lys-41[1] but also for other residues near the active-site and important for substrate binding including Phe-8, Gln-11, Asn-44, Thr-45, Phe-46, Val-116, Pro-117, Val-118, and Asp-121 (Blackburn & Moore, 1982; Beintema & Van der Laan, 1986). Similarly, residues important for formation and conservation of the three-dimensional structure, including the eight half-cystine residues,

[1] In the rest of the Discussion, residues are numbered according to the residue numbers of human and bovine pancreatic ribonucleases, except if specified otherwise.

Met-30, Ser-75, and Tyr-97, are perfectly conserved in human nonsecretory ribonuclease.

Most of these features have also been conserved in the sequences of turtle pancreatic ribonuclease and human angiogenin. However, although many residues have been conserved in the four members of the superfamily, other residues considered important in mammalian secretory ribonucleases have been replaced in human nonsecretory ribonuclease. These include the tyrosine or phenylalanine at position 120 in the pyrimidine-binding site (also replaced in angiogenin), Glu-2 and Arg-10, which form a helix-stabilizing salt bridge (both also replaced in angiogenin), and Asp-14, which may have a function as a helix stop signal.

We have not derived a three-dimensional structure of the nonsecretory ribonuclease (on the basis of its homology with bovine pancreatic ribonuclease and by use of energy-minimization procedures) as has been done for angiogenin (Palmer et al., 1987). However, an impression of the effect of insertions and deletions in the polypeptide chain of these four members of the ribonuclease superfamily can be derived from Figure 3, which represent modifications of the conformation of bovine pancreatic ribonuclease to accommodate the differences in the polypeptide chain. The most extensive change in the nonsecretory ribonuclease is a deletion of six residues between amino acids 16 and 23 in the so-called S-peptide loop. In bovine ribonuclease the distance between the C^{α} atoms of residues 14 and 25 is about 1.3 nm, a gap which can be bridged by a minimum of three amino acids in fully extended conformation. Even then a direct connection is not possible in bovine ribonuclease because of the steric demands of the rather large side chains of Tyr-25 and Met-29. In the nonsecretory ribonuclease these latter two residues have been replaced by glutamine and alanine, respectively, and amino acids 14 and 25 are connected by four residues. Thus, the deletion of six residues can be acommodated without significant rearrangements in other parts of the molecule. The nonsecretory ribonuclease has an addition of three residues at the N-terminus, an insertion of two residues in an external loop near residue 92, and insertions of two and nine residues in two neighboring loops near residues 70 and 114, respectively. The loop near residue 114 has a deletion of two residues in turtle pancreatic ribonuclease and in human angiogenin, and the loop near residue 70 has deletions of three and two residues, respectively, in these two proteins. (In addition, the two half-cystines which form a disulfide bridge between residues 65 and 72 in bovine ribonuclease have been replaced in turtle ribonuclease and in angiogenin.)

In Figure 3 the location of charged residues in these proteins is indicated as well. All four proteins have an excess of positive charges. However, the number of charged residues and their distribution on the molecular surface are quite different. In both pancreatic ribonucleases the negative charges seem to be rather randomly distributed, but the positive charges occur predominantly on the concave side of the molecule where the active site is located, presumably to facilitate the binding of oligonucleotide substrates. Angiogenin contains many positively charged residues, some of which are located near the active-site cleft but others appear to be more randomly distributed on the surface, especially on the lower half of the molecule as presented in Figure 3D. Human nonsecretory ribonuclease only contains 4 negatively charged residues, 12 lysines and arginines, and 5 histidines. Most of the positively charged residues occur in two clusters: one at positions 38, 39, and 41 on the concave surface and the other one in the loops with insertions near residues 70 and 114 at the bottom of the molecule where six of the positive and three of the negative charges are located.

Table III: Comparison of the Number and Percent Differences between Human Nonsecretory Ribonuclease, Human Pancreatic Ribonuclease, Turtle Pancreatic Ribonuclease, and Human Angiogenin[a]

	human nonsecretory ribonuclease	human pancreatic ribonuclease	turtle pancreatic ribonuclease	human angiogenin
human nonsecretory ribonuclease		102	110	110
human pancreatic ribonuclease	71%		77	86
turtle pancreatic ribonuclease	78%	60%		72
human angiogenin	76%	67%	57%	

[a] The number and percent differences are calculated from the four sequences as aligned in Figure 2 according to Dayhoff (1972). All positions where one sequence has an amino acid and the other a deletion were taken into account. Positions where both sequences contain deletions (and position 7 in human nonsecretory ribonuclease) are ignored.

Hydrophilicity plots of the four proteins (Figure 4) are very different. Most similar are the patterns obtained for turtle and human pancreatic ribonuclease (the patterns obtained for other mammalian pancreatic ribonucleases deviate very little from that of the human enzyme). Turtle ribonuclease and human angiogenin are the most closely related sequences of the four members of the superfamily, both in terms of percentage identities (see below) and in terms of pattern of deletions and insertions. However, their hydrophilicity patterns are very different, especially in the region of residues 45–60, which is hydrophobic in turtle ribonuclease but hydrophilic with many positive charges in human angiogenin. The sequence of human nonsecretory ribonuclease is the most hydrophobic, with the three loops with positive charges discussed before as the only hydrophilic segments in the protein. The N- and C-terminal sequences and the loop near residue 92 also are less hydrophobic. But the other parts of the molecule are all hydrophobic, independent of their location in the interior or at the surface of the molecule.

From the comparison of sequences, it is not possible to state with certainty why the three ribonucleases of the superfamily have very similar specific activities on RNA, while the ribonuclease activity of human angiogenin is very small. Notwithstanding its low enzymic activity, angiogenin binds very tightly to the human placental ribonuclease inhibitor, with complete abolition of both enzymic and angiogenin activity (Shapiro et al., 1987). It is likely that a positively charged residue is important for binding the 5′-phosphate of the ribose which carries at its 3′-position the phosphodiester bond which is cleaved by the enzyme (the P_0 site; Iwahashi et al., 1981). In mammalian secretory ribonucleases, Lys-66 plays this role. This lysine is absent in the other three proteins under comparison. We have presented evidence that Lys-122 in turtle ribonuclease can play the same role as Lys-66 (Katoh et al., 1986), and the same explanation can be given for the enzymic activity of human nonsecretory ribonuclease because of the presence of an arginine at position 122. However, the presence of glutamine at this position in angiogenin cannot generate a normal level of ribonuclease activity.

In Table III a difference matrix which summarizes numbers and percent differences between human nonsecretory ribonuclease, human pancreatic ribonuclease, turtle pancreatic ribonuclease, and human angiogenin is shown as derived from the alignment in Figure 2. The nonsecretory ribonuclease

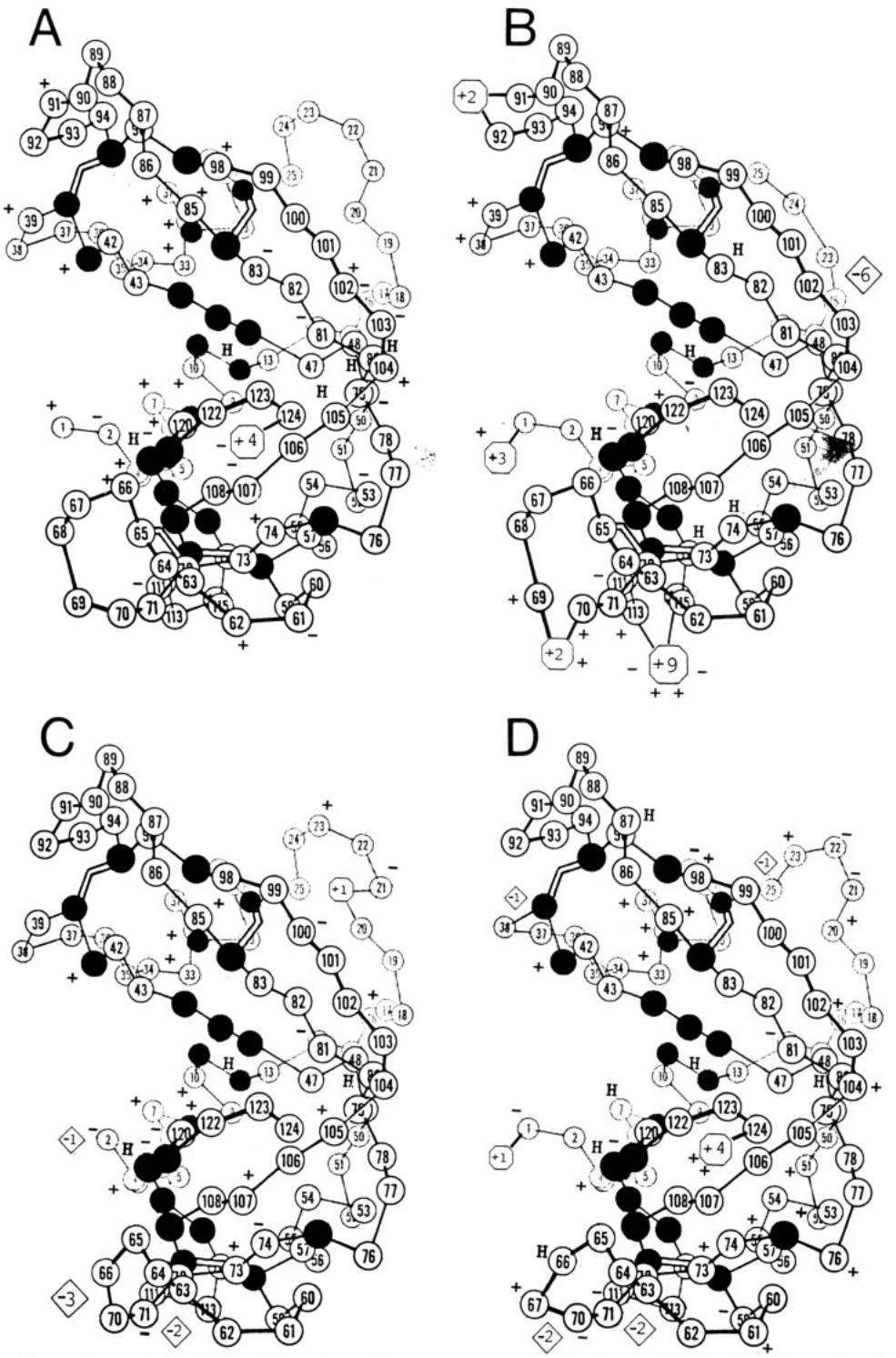


FIGURE 3: Location of insertions and deletions in the amino acid sequences presented in Figure 2 in the three-dimensional structure of bovine pancreatic ribonuclease A. Positions are numbered according to the residue number of bovine ribonuclease. Black circles indicate identical residues in the sequences presented in Figure 2 and all other known sequences of mammalian secretory ribonucleases (Beintema & Van der Laan, 1986). Numbers of deleted residues are presented in diamonds [with a (−) sign] next to the locations of these deletions. Numbers of inserted residues are indicated in polygons [with a (+) sign] inserted in the main chain. Positions with lysine or arginines, with histidines, and with aspartic or glutamic acids are indicated with (+), (H), and (−), respectively. (A) Human pancreatic ribonuclease; (B) human nonsecretory ribonuclease; (C) turtle pancreatic ribonuclease; (D) human angiogenin. Figure adapted from Dickerson & Geis (1969).

sequence deviates most from the other three, while turtle pancreatic ribonuclease and human angiogenin are the two most similar. From these data tentative evolutionary relationships can be proposed: a gene duplication occurred in an ancient vertebrate ancestor; one branch leads to the nonsecretory ribonuclease, while the other branch leads to a second duplication, not much later, with one line leading to the secretory ribonucleases (in mammals) and the second line to a pancreatic ribonuclease in reptiles (turtle) but to an angiogenic factor in mammals (human angiogenin). A more ancient divergence of ribonuclease and angiogenin requires the assumption that turtle ribonuclease and human angiogenin share relatively many identical residues and deletions in the sequence as a result of convergent or parallel evolution.

Niwata et al. (1985) have determined the N-terminal amino acid sequence of a ribonuclease from bovine kidneys (30

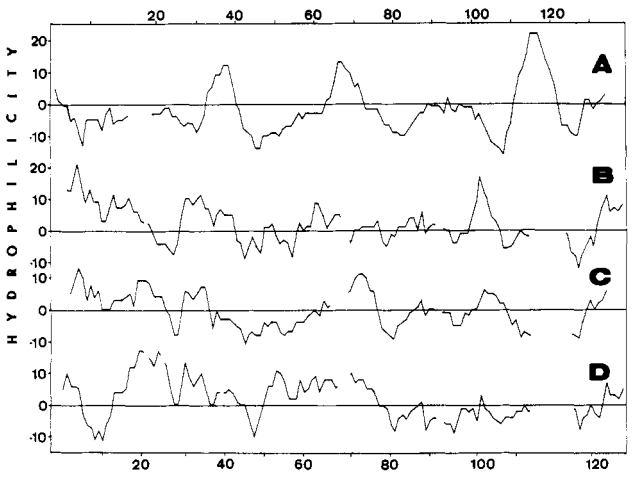


FIGURE 4: Hydrophilicity plots of the amino acid sequences of human nonsecretory ribonuclease (A), human pancreatic ribonuclease (B), turtle pancreatic ribonuclease (C), and human angiogenin (D), obtained by the method of Hopp and Woods (1981) with an average group length of seven amino acid residues. Groups are shifted in steps of one residue. Position 7 in human nonsecretory ribonuclease was given the value 0.0. To obtain values for the N-terminal and C-terminal sequences, three residues with the value 0.0 were added to both termini in the four sequences. The plots were derived for the continuous sequences. Deletions were introduced afterward to obtain the proper alignments. Positions are numbered according to the residue number of human nonsecretory ribonuclease (top) or human pancreatic ribonuclease (bottom).

residues); about half of the residues have been replaced if compared with those of human nonsecretory ribonuclease. Pancreatic ribonucleases of species from different mammalian orders differ at about 25% of the amino acid positions (Beintema et al., 1986).

Gleich et al. (1986) published not only the N-terminal sequence of human eosinophil-derived neurotoxin (which is very similar to and possibly identical with the nonsecretory ribonuclease from urine, spleen, and liver in sequence and specific activity) but also the N-terminal sequence of a human eosinophil cationic protein. The specific activity of the latter protein is about 2% of that of the pancreatic ribonucleases (Slifman et al., 1986). Its sequence differs from that of eosinophil-derived neurotoxin at one-third of the amino acid positions.

Five positions with asparagine-linked carbohydrate have been identified in the sequence of human nonsecretory ribonuclease. However, Cranston et al. (1980) and Iwama et al. (1981) report glucosamine and neutral saccharide contents of only 4.5 and 5.4 and of 2.6 and 1.7 residues per mole of protein, respectively. A carbohydrate analysis of the protein preparation as isolated by Iwama et al. (1981) indicated the presence of 1.2 fucose, 0.2 mannose, and 0.4 glucosamine residues per mole of protein (the latter value has not been corrected for the incomplete cleavage of the N-glycosidic bond between *N*-acetylglucosamine and asparagine during analysis). Also, sugar analysis of glycopeptides after hydrazinolysis indicates that the oligosaccharide chains of this protein are very short and consist of one to three monosaccharides per asparagine residue. This probably is the result of degradation by lysosomal exoglycosidases. Glycosidase action in the urine is unlikely, since the secretory ribonuclease from human urine is very extensively glycosylated (Beintema et al., 1985a).

The carbohydrate attachment sites are located both at the surface of the molecule, i.e., at external loops (positions 65 and 92), and at less exposed sites (positions 17, 59, and 84). A striking feature is the presence of hydrophobic residues at neighboring positions in four of the sites (residues 16 and 18, 60, 83 and 85, and 93; residue numbers of human nonsecretory ribonuclease).

Both the hypothetical three-dimensional structure as proposed in Figure 3B and structural properties encountered during isolation, characterization, and sequence analysis of the protein indicate a very tight and compact structure. Thus, by use of polyacrylamide gel electrophoresis in SDS the protein was shown (Blank & Dekker, 1981) to undergo a heat-induced conformational change resulting in a significant increase in mobility. The protein contains 12 proline residues; several of them are clustered in the N-terminal extension, the external loop near residue 114 where most negatively and positively charged residues are located, and in the less hydrophobic loop near residue 92. Both the compact structure and the high proline content may explain the fact that this protein is very resistant to proteolytic degradation. In addition, the high excess of positive charges, especially in several external loops, will lower the proton concentration near these sites. The effect of this will be an increase of local pH. This will give an additional protection against lysosomal proteinases which have their optimal activity at the low pH values of lysosomes.

We consider that it is very likely that the major function of the nonsecretory ribonuclease is a degradative one, similar to that of other hydrolytic enzymes associated with lysosomes. However, in view of its structural relationship to eosinophil-derived neurotoxin and cationic protein (Gleich et al., 1986), this ribonuclease (and closely related proteins) may have additional extracellular functions.

ACKNOWLEDGMENTS

We thank Robbert Wijtmann and Michael Lee for assisting with the enzyme purification (Berkeley), Peter Jekel for performing amino acid analyses, Piet Wietzes for a preliminary automatic Edman degradation with a Beckman sequencer (Groningen), Renate Matthies for help in performing the automatic Edman degradation with the gas-phase sequenator, and Dr. S. R. Stone for helpful comments (Basel). We also thank Dr. R. Amons (Department of Medical Chemistry, University of Leiden) for giving us advice and reagents for the pyridylethylation reactions and G. J. Gerwig and Dr. J. P. Kamerling (Department of Bio-Organic Chemistry, University of Utrecht) for performing carbohydrate analyses.

SUPPLEMENTARY MATERIAL AVAILABLE

Two figures showing products of enzymic digestion and CNBr cleavage and recoveries of phenylthiohydantoin derivatives during automatic Edman degradation and four tables showing details of the purification procedures used and the amino acid analyses of the peptides (6 pages). Ordering information is given on any current masthead page.

REFERENCES

Amons, R. (1987) *FEBS Lett. 212*, 68–72.

Beintema, J. J. (1987) *Life Chem. Rep. 4*, 333–389.

Beintema, J. J., & Van der Laan, J. M. (1986) *FEBS Lett. 194*, 338–342.

Beintema, J. J., Wietzes, P., Weickmann, J. L., & Glitz, D. G. (1984) *Anal. Biochem. 136*, 48–64.

Beintema, J. J., Blank, A., & Dekker, C. A. (1985a) *Abstracts of the 13th International Congress of Biochemistry*, August 25–30, Amsterdam, TH-269, Elsevier Science, Amsterdam.

Beintema, J. J., Broos, J., Meulenberg, J., & Schüller, C. (1985b) *Eur. J. Biochem. 153*, 305–312.

Beintema, J. J., Fitch, W. M., & Carsana, A. (1986) *Mol. Biol. Evol. 3*, 262–275.

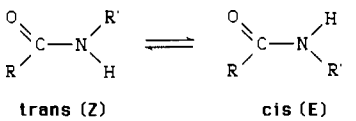
Blackburn, P., & Moore, S. (1982) *Enzymes (3rd Ed.) 15*, 317-433.
Blank, A., & Dekker, C. A. (1981) *Biochemistry 20*, 2261-2267.
Chang, J.-Y. (1983) *Methods Enzymol. 91*, 455-466.
Cranston, J. W., Perini, F., Crisp, E. R., & Hixson, C. V. (1980) *Biochim. Biophys. Acta 616*, 239-258.
Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622-627.
Dayhoff, M. O., Ed. (1972) in *Atlas of Protein Sequence and Structure*, Vol. 5, p D6, National Biomedical Research Foundation, Washington, DC.
Delaney, R. (1963) *Biochemistry 2*, 438-444.
Dickerson, R. E., & Geis, I. (1969) *The Structure and Action of Proteins*, Benjamin, Menlo Park, CA.
Gaastra, W., Welling, G. W., & Beintema, J. J. (1978) *Eur. J. Biochem. 86*, 209-217.
Gleich, G. J., Loegering, D. A., Bell, M. P., Checkel, J. L., Ackerman, S. J., & McKean, D. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3146-3150.
Hartley, B. S. (1970) *Biochem. J. 119*, 805-822.
Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990-7997.
Hopp, T. P., & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3824-3828.
Hunkapiller, M. W., & Hood, L. E. (1983) *Methods Enzymol. 91*, 486-493.
Iwahashi, K., Nakamura, K., Mitsui, Y., Ohgi, K., & Irie, M. (1981) *J. Biochem. (Tokyo) 90*, 1685-1690.
Iwama, M., Kunihiro, M., Ohgi, K., & Irie, M. (1981) *J. Biochem. (Tokyo) 89*, 1005-1016.
Katoh, H., Yoshinaga, M., Yanagita, T., Ohgi, K., Irie, M., Beintema, J. J., & Meinsma, D. (1986) *Biochim. Biophys. Acta 873*, 367-371.
Morita, T., Niwata, Y., Ohgi, K., Ogawa, M., & Irie, M. (1986) *J. Biochem. (Tokyo) 99*, 17-25.
Morrissey, J. H. (1981) *Anal. Biochem. 117*, 307-310.
Niwata, Y., Ohgi, K., Sanda, A., Takizawa, Y., & Irie, M. (1985) *J. Biochem. (Tokyo) 97*, 923-934.
Ohta, T., Ogawa, M., Kurihara, M., Kitahara, T., & Kosaki, G. (1982) *Clin. Chim. Acta 124*, 51-62.
Palmer, K. A., Scheraga, H. A., Riordan, J. F., & Vallee, B. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1965-1969.
Shapiro, R., Fett, J. W., Strydom, D. J., & Vallee, B. L. (1986) *Biochemistry 25*, 7255-7264.
Sierakowska, H., & Shugar, D. (1977) *Prog. Nucleic Acid Res. Mol. Biol. 20*, 59-130.
Slifman, N. R., Loegering, D. A., McKean, D. J., & Gleich, G. J. (1986) *J. Immunol. 137*, 2913-2917.
Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5486-5494.
Sugiyama, R. H. (1981) Thesis, University of California, Berkeley.
Sugiyama, R. H., Blank, A., & Dekker, C. A. (1981) *Biochemistry 20*, 2268-2274.
Vereijken, J. M., De Vlieg, J., & Beintema, J. J. (1984) *Biochim. Biophys. Acta 788*, 298-305.
Yamanaka, M., Akagi, K., Murai, K., Hirano, N., Fujimi, S., & Omae, T. (1977) *Clin. Chim. Acta 78*, 191-201.

Influences of Solvent Water on Protein Folding: Free Energies of Solvation of Cis and Trans Peptides Are Nearly Identical†

Anna Radzicka, Lee Pedersen, and Richard Wolfenden*
Departments of Biochemistry and Chemistry, University of North Carolina, Chapel Hill, North Carolina 27514
Received October 23, 1987; Revised Manuscript Received February 10, 1988

ABSTRACT: Peptide bonds interact so strongly with water that even a modest difference between the free energies of solvation of their cis and trans isomers could have a significant bearing on protein structure. However, proton magnetic resonance studies at high dilution in deuteriated solvents show that *N*-methylformamide exists as the cis isomer to the extent of 8% in water, 10.3% in chloroform, 8.8% in benzene, and 9.2% in cyclohexane. Integrated intensities of proton and carbon resonances show that *N*-methylacetamide exists as the cis isomer to the extent of only 1.5% in water, not changing much in nonpolar solvents. Quantum mechanical calculations using the 6-31G** basis set reproduce these relative abundances with reasonable accuracy and show that there is little difference between the dipole moments of the cis and trans isomers, for either amide. The remarkable insensitivity of cis/trans equilibria to the solvent environment and the heavy preponderance of trans isomers regardless of the polarity of the surroundings (ca. 98.5% for *N*-methylacetamide, whose properties may resemble those of a typical peptide bond) accord with the overwhelming preference of peptide bonds for the trans configuration that is consistently observed in the three-dimensional structures of globular proteins.

The peptide bond prefers to adopt either of two planar configurations (Corey & Donohue, 1950), stabilized by resonance to such an extent that their interconversion is slow on the NMR time scale at room temperature (Drakenberg et al., 1972). Both configurations have been observed in proteins,



but the trans configuration appears to be much more common. Among the rare exceptions are three non-proline cis peptide bonds in carboxypeptidase A (Rees et al., 1981). Steric difficulties would arise if peptide bonds were to adopt the cis

† This work was supported by NSF Research Grant PCM-7823016 (A.R. and R.W.) and by NIH Research Grants HL-27995 and HL-26309 (L.P.).

0006-2960/88/0427-4538$01.50/0