Direct Sequence Determination of Ovine Luteinizing Hormone Releasing Factor by Mass Spectrometry†

Nicholas Ling,* Jean Rivier, Roger Burgus, and Roger Guillemin

ABSTRACT: The primary structure of ovine luteinizing hormone releasing factor (LRF), pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2, has been confirmed by mass spectrometry. The method involves treating LRF with: (1) acetylacetone for arginine modification, (2) chymotrypsin digestion, (3) acetylation of the amino groups, (4) N,O-permethylation with sodium methylsulfinylmethide and methyl iodide. The resulting peptide mixture was analyzed directly by fractional vaporization in the mass spectrometer.

A primary structure for porcine luteinizing hormone releasing factor (LRF), pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2, was proposed by Matsuo *et al.* (1971). This structure was based on the results from the combined Edman-dansyl procedure coupled with the selective tritiation method for C-terminal analysis directly on the unresolved digestion products of LRF with chymotrypsin and thermolysin. However, Matsuo *et al.* (1971) did not provide any direct evidence for the pGlu-His linkage as well as the amide grouping of the C-terminal glycine in their proposed structure. Synthetic replicates based on this structure were prepared shortly and found to have biological activity (Matsuo *et al.*, 1971; Monahan *et al.*, 1971). Subsequently, Schally *et al.* (Baba *et al.*, 1971) were able to confirm the proposed structure by performing conventional Edman-dansyl analysis on porcine LRF after cleavage of the N-terminal pyroglutamyl residue with pyrrolidonecarboxylylpeptidase. Shortly thereafter, we independently reported that ovine LRF had the same primary structure (Burgus *et al.*, 1971, 1972) on the basis of hydrolysis of the peptide with chymotrypsin or pyrrolidonecarboxylylpeptidase and analysis of the products by dansyl-Edman method as well as determination of the phenylthiohydantoinyl derivatives by mass spectrometry. However, the limited amount of material available (40 μg or 30 nmol) for our structural work prevented the use of other methods for substantiating the reported structure. Recently, we have isolated another 3.5 mg of ovine LRF from 490,000 fragments of ovine hypothalami and found it most convenient to determine its sequence by mass spectrometry.

The amino acid composition of the new batch of ovine LRF was found to be identical with the previously isolated material (Amoss *et al.*, 1971): His 1, Arg 1, Ser 1, Glu 1, Pro 1, Gly 2, Leu 1, Tyr 1; in addition, one residue of tryptophan was also detected which had not been recovered in the previous analysis (Amoss *et al.*, 1971). From our earlier studies, we had already shown that ovine LRF contained a blocked pyroglutamic acid N-terminus (Amoss *et al.*, 1970, 1971), which rendered the Edman sequencing procedure not directly applicable. Mass spectrometry is therefore the method of choice for determining its structure, provided the peptide can be vaporized into the ion source of the mass spectrometer without decomposition.

The most widely used method to enhance the volatility of a peptide is N,O-permethylation with methyl iodide and sodium methylsulfinylmethide (Vilkas and Lederer, 1968; Thomas, 1968). However, permethylation cannot be performed on a peptide containing arginine without prior derivatization of the guanido group. Two methods commonly used to modify the guanido group of arginine are the condensation with β-dicarbonyl compounds to give a pyrimidyl derivative (Shemyakin *et al.*, 1967; Vetter-Diechtl *et al.*, 1968) or conversion of arginine to ornithine by hydrazinolysis (Shemyakin *et al.*, 1967). But, even with arginine derivatization followed by permethylation, we could not, with the techniques presently available, observe the molecular as well as all the sequence ions from a series of short-chain LRF analogs with more than seven amino acids (Rivier *et al.*, 1973). In order to overcome this problem, we treated LRF by: (1) transforming the guanido group of arginine to a 2,4-dimethylpyrimidyl derivative with acetylacetone, (2) cleavage of the resulting peptide with chymotrypsin, (3) without separation, acetylation of the free amino groups in the peptide mixture, (4) N,O-permethylation with sodium methylsulfinylmethide and methyl iodide in dimethyl sulfoxide. The resulting mixture of products was then analyzed directly by fractional vaporization in the mass spectrometer.

Materials and Methods

The isolation of ovine LRF and its physicochemical constants will be reported in detail in a separate paper.

Arginine Derivatization. For condensing the guanido function with acetylacetone, we applied the method published by Bacon *et al.* (1969). A 200-μg sample (150 nmol) of ovine LRF was dissolved in 100 μl of dry EtOH in a 12 × 75 mm test tube and 30 μl of redistilled $CH_3COCH_2COCH_3$ was added. The solution was adjusted to pH 7 with bicarbonate form Bio-Rad AG 1-X8 resin (about 40 mg). A few pieces of Linde Molecular Sieves (type 3A, size $^1/_{16}$) were added and the mixture was stoppered and left overnight at room temperature. The supernatant liquid was withdrawn and the residue was washed with 3 × 100 μl of dry EtOH. The solid resulting from the evaporation of the combined EtOH extract was taken up in 100 μl of H_2O in a 13 × 100 mm centrifuge tube and the self-condensation product of $CH_3COCH_2COCH_3$ extracted with CH_2Cl_2. The aqueous phase was then lyophi-

† From the Salk Institute, La Jolla, California 92037. *Received July 17, 1973.* Supported by the Agency for International Development (Contract No. AID/csd 2785), the Ford Foundation, and the Rockefeller Foundation.

lized. To minimize the loss of material, the following reactions were done in the same centrifuge tube and the solvents were transferred by syringes.

Chymotrypsin Digestion. The lyophilized material was incubated with 10 μg of chymotrypsin A (Worthington) in a 100-μl solution of 0.1 M NH_4OAc at pH 8.2 for 17 hr at 37°. Two drops of HOAc was added to stop the digestion and the solution was lyophilized. The resulting residue was lyophilized once more from 100 μl of 0.1% HOAc.

Acetylation.[1] The lyophilized sample was taken up in 100 μl of H_2O and 3 μl of 50:50 (v/v) $(CH_3CO)_2O$ and $(CD_3CO)_2O$ was added, followed by 2 μl of 4 M NaOAc. The mixture was agitated at 0° for 30 min and then lyophilized. The excess acetate ion was removed by addition of 100 μl of MeOH saturated with HCl and evaporation to dryness. The residue was further dried under vacuum at 60° for 2 hr.

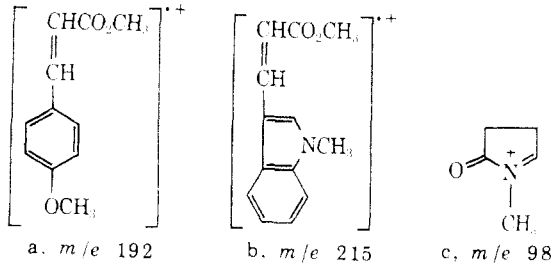
Permethylation.[2] A 0.5 M solution of $NaCH_2SOCH_3$ in Me_2SO was prepared by heating 200 mg of NaH (50% oil suspension, Gallard-Schlesinger Chemical Corp.), which had been washed with five 2-ml portions of anhydrous Et_2O, in 8 ml of redistilled Me_2SO under argon at 65° for 40 min. After letting the reaction mixture cool down to room temperature, a pale yellow liquid with a small amount of gelatinous material at the bottom was obtained. Sixty microliters (30 μmol) of this solution was added to the acetylated peptide mixture, agitated for 5 min, followed by 1.9 μl (30 μmol) of CH_3I, and the mixture was agitated for 10 min. The reaction was terminated by the addition of 0.5 ml of H_2O and the permethylated peptides were extracted with three 100-μl portions of $CHCl_3$. After evaporation, the residue was transferred with $CHCl_3$ in small portions to a gold crucible used for the direct inlet system of the mass spectrometer and the solvent was evaporated under a stream of argon.

Mass Spectrometry. Mass spectra were obtained from a Varian Mat CH-5 single focusing mass spectrometer with the direct inlet system. The electron current was set at 1 mA with an ionizing potential of 70 eV. The accelerating voltage was maintained at 3 kV and the ion source temperature at 250°. The spectra were recorded at a resolution of 1200 by a Varian 620/i computer and plotted on a Varian Statos I plotter. The temperature of the probe was raised gradually from 80 to 300° with spectra collected at approximately every 4° intervals.

Results and Discussion

The mass spectra of N-acetylated and N,O-permethylated peptides exhibit a series of very intense peaks A_1, A_2, A_3, etc., deriving from cleavage of the CO—NCH_3 bonds with charge retention on the carbonyl portion of the molecule (Thomas *et al.*, 1968). These acylium (sequence) ions, in turn, may or may not lose carbon monoxide to give another series of iminium ions N_1, N_2, N_3, etc. As a result, the sequence of a peptide can be deduced if all the acylium and/or iminium ions from the N-terminus are identified in the mass spectrum.

In the following discussion, the primary structure of ovine LRF is assumed unknown except for its amino acid composition and that it possesses a pyroglutamic acid N-terminus. Since LRF contains a tyrosine and a tryptophan residue, after chymotrypsin digestion it should yield two or three fragments depending on the position of these two amino acids. Furthermore, since these two amino acids will be at the C-terminus of the cleaved fragments, after derivatization it is very easy to recognize them in the mass spectrum by their characteristic rearrangement ions a and b, respectively, which are very intense peaks (Grützmacher and Heyns, 1966; Pfaender, 1967; Morris *et al.*, 1971).



a, *m/e* 192 b, *m/e* 215 c, *m/e* 98

In the direct sequencing of peptide fragments by mass spectrometry, it is essential that the N-terminal amino acid of each fragment can be recognized. In the present case, since we already knew that LRF has a pyroglutamic acid N-terminus (Amoss *et al.*, 1970, 1971), after enzyme digestion one of the fragments recovered should still have this amino acid at its N-terminus. In the mass spectrometer, pGlu peptides after permethylation give a very intense ion at *m/e* 98 corresponding to structure c (Lenard and Gallop, 1969; White and Desiderio, 1971; Morris *et al.*, 1971).

With *m/e* 98 as the starting point in the mass spectrum, the rest of this particular fragment can be deduced on the basis of the observed acylium and/or iminium ions as shown later. However, this will only solve the sequence of one of the fragments. Since we are working on the assumption that we do not know the sequence of LRF, the rest of the nine amino acids (we have not yet solved the sequence of the pGlu fragment) have equal probability to be the N-terminus of the other remaining fragment(s). In order to facilitate the recognition of these N-terminal acylium ions (see Table I, row 1) as well as the sequencing work later, the digested peptide was acetylated with a 50:50 (v/v) mixture of $(CH_3CO)_2O$ and $(CD_3CO)_2O$. By this modification, all the acylium ions containing the acetyl group will have a peak of approximately equal intensity but 3 mu higher.

As the temperature of the probe was gradually raised above 80°, an ion at *m/e* 192 began to appear in the mass spectra, indicating that one of the peptide fragments with a tyrosine C-terminus was being evaporated into the ion source. This ion reached its maximum intensity at 128° and is presented in Figure 1. The rest of Figure 1, however, does not show a doublet of equal intensity at *m/e* 234 and 237, which would signify an N-terminal tyrosine. Consequently, this fragment must have at least another LRF amino acid besides tyrosine. In view of the low vaporization temperature, this fragment is probably no bigger than a di- or tripeptide. Since there is no *m/e* 98, its N-terminus cannot be pyroglutamic acid. Searching the rest of this spectrum, it will be found that the only doublets of about equal intensity but 3 mu apart are at *m/e* 158, 161,

[1] This method of acetylation does not acylate the secondary amino group in N^{δ}-2-(4,6-dimethylpyrimidyl)ornithine.

[2] We found that by using equimolar ratios of $NaCH_2SOCH_3$ and CH_3I as well as shortening the permethylation time to 10 min, a clean product containing the expected number of methyl groups was obtained. Apparently, no quaternization of histidine and N^{δ}-2-(4,6-dimethylpyrimidyl)ornithine occurred under these conditions.

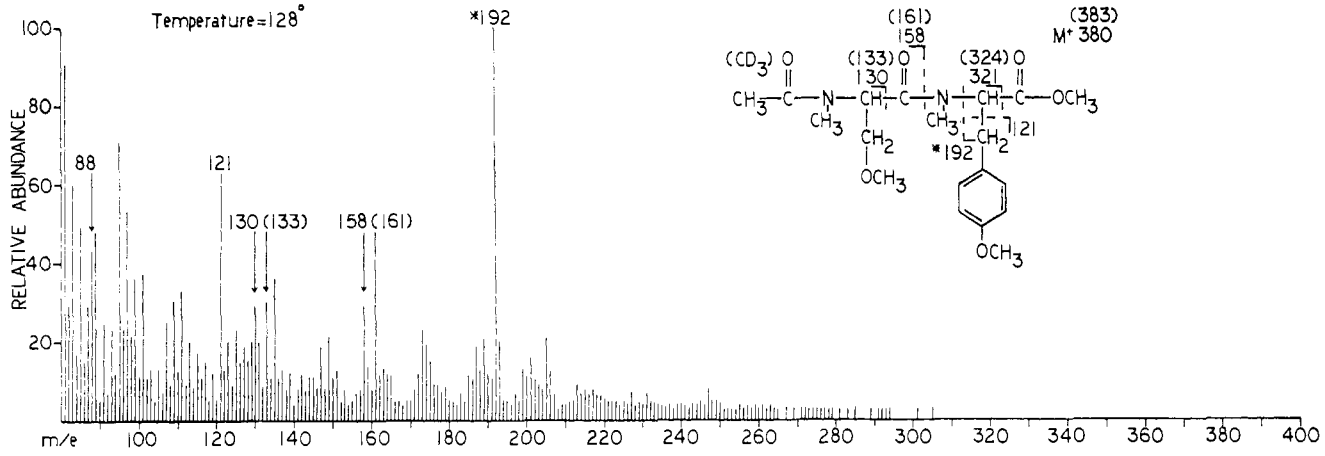


FIGURE 1: Mass spectrum (above *m/e* 80) of modified LRF, obtained by fractional vaporization. The peptide detected has the sequence H-Ser-Tyr-OH. * denotes rearrangement peak.

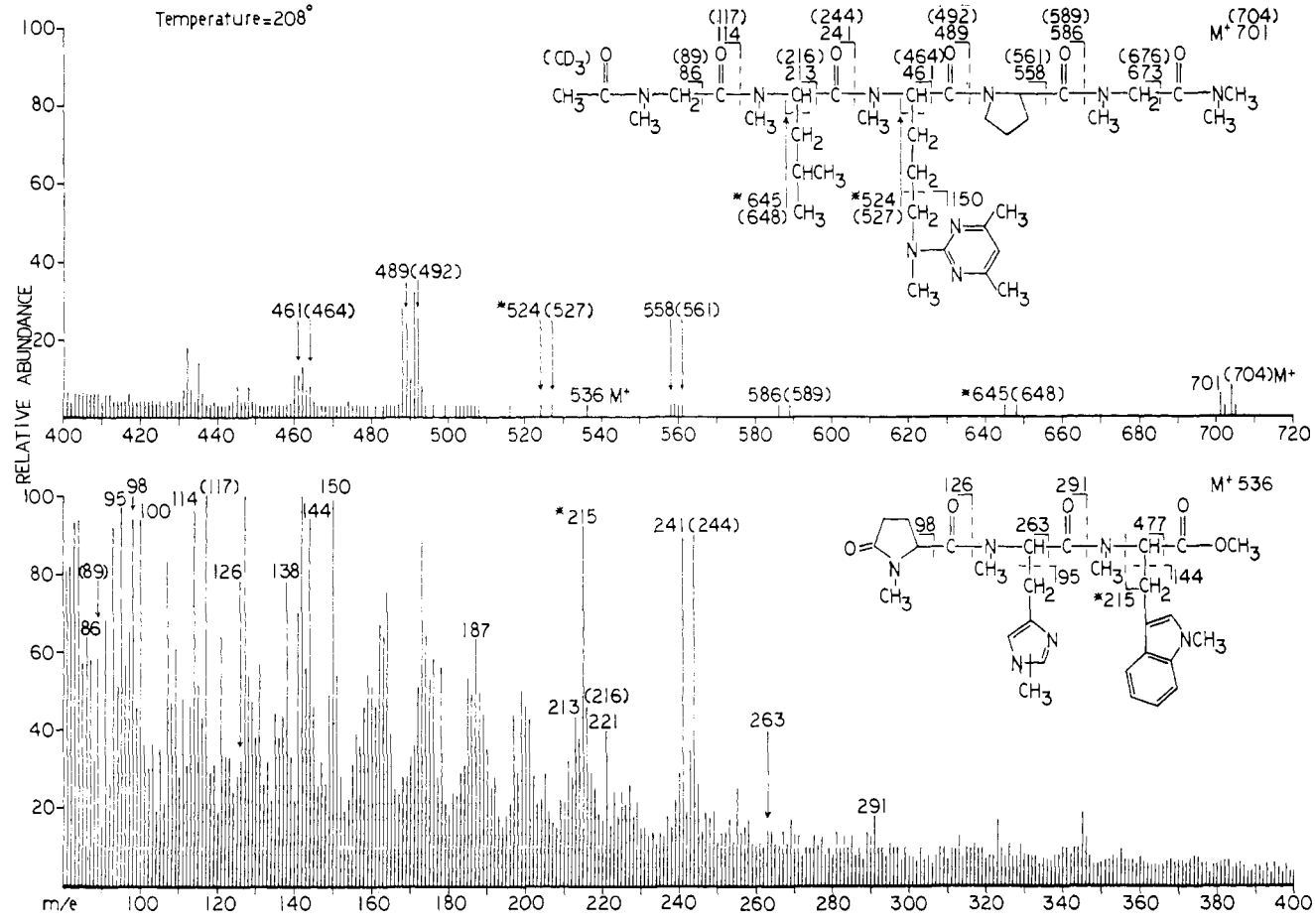


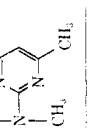
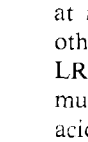
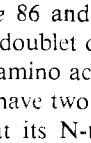
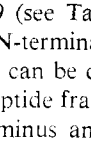
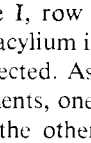
FIGURE 2: Mass spectrum (above *m/e* 80) of modified LRF obtained by fractional vaporization. The peptides detected have the sequence pGlu-His-Trp-OH and H-Gly-Leu-Arg-Pro-Gly-NH_2. * denotes rearrangement peak.

and 130, 133, which correspond to the N-terminal acylium and iminium ions of serine.

That this spectrum indeed contains serine is substantiated by the immonium ion at *m/e* 88 (Wipf *et al.*, 1973) as calculated in Table I, row 7. As a result, this fragment must have serine at its N-terminus and tyrosine at its C-terminus with or without another amino acid in between. However, because there is no other doublet separated by 3 mu in the mass region higher than *m/e* 161, this fragment is most probably a H-Ser-Tyr-OH dipeptide. Since no ion can be detected beyond *m/e* 305, the structure of this dipeptide cannot be confirmed by its molecular ions which should be at *m/e* 380 and 383. The ion at *m/e* 121 corresponds to the side-chain fragmentation of tyrosine (Shemyakin *et al.*, 1971).

No significant doublets or *m/e* 98 and 215 appeared on the spectra until the probe temperature reached 200°. At 208° the peaks at *m/e* 98 and 215 exhibited their highest intensity and are presented in Figure 2. Looking at the low mass region of this spectrum, one will notice that there is an equal intensity doublet at *m/e* 114 and 117, which corresponds to an N-terminal glycine acylium ion (see Table I, row 1). This finding is strengthened by its corresponding N-terminal iminium ions

TABLE I: Integral Mass Units (mu) Corresponding to Various Derivatives of Gly, Pro, Ser, Leu, His, Tyr, Trp, and Arg, Respectively.

									
R = CH_3CO–CD_3CO	114 (117)	140 (143)	158 (161)	170 (173)	208 (211)	234 (237)	257 (260)	291 (294)	
R′ = CO									
R = CH_3CO–CD_3CO	86 (89)	112 (115)	130 (133)	142 (145)	180 (183)	206 (209)	229 (232)	263 (266)	
R′ = no substitution									
R = no substitution	71	97	115	127	165	191	214	248	
R′ = CO									
R = no substitution	43	69	87	99	137	163	186	220	
R′ = no substitution									
R = no substitution	102	128	146	158	196	222	245	279	
R′ = $COOCH_3$									
R = no substitution	115	141	159	171	209			292	
R′ = $C(=O)N(CH_3)_2$									
R = H	44	70	88	100	138	164	187	221	
R′ = no substitution									

at *m/e* 86 and 89 (see Table I, row 2). Besides glycine, no other doublet of N-terminal acylium ions from the remaining LRF amino acids can be detected. As a result, this spectrum must have two peptide fragments, one having a pyroglutamic acid at its N-terminus and the other glycine. Also, one of these two fragments must have tryptophan at its C-terminus because of the *m/e* 215 peak. Since it is easier to recognize equal intensity doublets that are 3 mu apart in a mass spectrum, the sequence of the fragment with the N-terminal glycine will be deduced first.

By adding successively the amino acid derivatives in row 3 of Table I to *m/e* 114 and 117 and checking the resulting sum of doublets in the mass spectrum, it will be found that the second amino acid from glycine matches leucine because of the peaks at *m/e* 241 and 244. As a further proof of the H-Gly-Leu sequence, the corresponding pair of iminium ions can be observed at *m/e* 213 and 216. Using the same procedure, one will discover that the next two amino acids match arginine and proline, respectively, because of the doublets at *m/e* 489, 492, and 586, 589. The corresponding pairs of iminium ions show up at *m/e* 461, 464 and 558, 561. However, no match is obtained after H-Gly-Leu-Arg-Pro. As a result, the next amino acid is probably the C-terminus which can either be a methyl ester or a dimethylamide. By adding successively all the C-terminal derivatives in rows 5 and 6 of Table I to *m/e* 586 and 589, one will find that the sum resulting from glycine dimethylamide matches the doublet at *m/e* 701 and 704. Hence, this fragment must have the sequence H-Gly-Leu-Arg-Pro-Gly-NH_2. That *m/e* 701 and 704 are molecular ions is substantiated by the doublet at *m/e* 645 and 648 which corresponds to the loss of the leucine side-chain with a hydrogen transfer back to the parent molecule (Bricas *et al.*, 1965). An analogous pair of rearrangement ions from arginine can be observed at *m/e* 524 and 527. The immonium ions (Wipf *et al.*, 1973) at *m/e* 100 and 221 are from leucine and arginine as calculated in Table I, row 7, respectively, while the *m/e* 150 ion is derived from the pyrimidyl derivative of arginine (Shemyakin *et al.*, 1971). As a result, it can be concluded that a fragment with the sequence H-Gly-Leu-Arg-Pro-Gly-NH_2[3] was cleaved from LRF by chymotrypsin and this fragment must be at the C-terminal portion of the parent molecule because of the amide grouping of the C-terminal glycine.

Having deduced the structure of one of the peptide fragments in Figure 2, the remaining fragment can be readily sequenced. From the spectrum, it is apparent that this fragment must have a pyroglutamic acid N-terminus and a tryptophan C-terminus because of the ions at *m/e* 98 and 215. Furthermore, the ions at *m/e* 95 and 138 indicate that histidine is also present (Shemyakin *et al.*, 1971); this accounts for all the ten amino acids of LRF. Starting with *m/e* 126 which is the N-terminal acylium ion of pyroglutamic acid and adding to it 165 mu which corresponds to the histidine derivative (see Table I, row 3), one obtains a mass of 291 which is present in Figure 2. This finding indicates that the residue next to pyroglutamic acid is histidine. And by adding 245 mu which corresponds to the C-terminal methyl ester derivative of tryptophan (see Table I, row 5) to mass 291, one arrives at a mass of 536 which is observed in the spectrum. This result shows that the remaining fragment has the sequence pGlu-His-

[3] A peptide containing the Arg-Pro sequence after derivatization with acetylacetone and permethylation always gives a more intense ion which is 1 mu lower than the corresponding arginine acylium ion (see *m/e* 488 and 491 in Figure 2).

Trp-OH. The ions at *m/e* 144 and 187 are derived from the tryptophan residue (Shemyakin *et al.*, 1971).

With the sequence of the three peptide fragments from chymotrypsin cleavage of ovine LRF deduced as pGlu-His-Trp-OH, H-Ser-Tyr-OH, and H-Gly-Leu-Arg-Pro-Gly-NH_2, the only complete sequence that can be assembled together is pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2, bearing in mind that LRF has a pyroglutamic acid N-terminus. This result agrees with the previous finding which was obtained by the dansyl-Edman method (Burgus *et al.*, 1971, 1972).

While this work was in progress, Nair and Schally (1972) reported that they could observe the molecular as well as all the sequence ions from permethylated porcine LRF and its methyl ester. They also performed a chymotrypsin digestion on porcine LRF and separated the three peptide fragments, each of which was derivatized and its mass spectrum was determined. However, two of the fragments which were expected to be identical with ours after derivatization (the third fragment was derivatized differently from ours), namely, H-Ser-Tyr-OH and pGlu-His-Trp-OH, do not exhibit the same mass spectra as ours. Their H-Ser-Tyr-OH spectrum showed a very strong molecular ion but no *m/e* 192 and 121. Furthermore, they observed a loss of methanol from the N-terminal acylium ion of serine, whereas we only detected the loss of carbon monoxide. In the pGlu-His-Trp-OH spectrum, they did not detect the *m/e* 215 and 144 ions which had an intensity of 92 and 94%, respectively, of the base peak in our spectrum. It is quite unlikely that the *m/e* 192, 121 and 215, 144 ions observed by us could be artifacts because in the background spectra, which were obtained from the same derivatization procedure but without ovine LRF, the peaks in these positions had an intensity of less than 25%. Furthermore, these ions are present in the fragmentation pattern of pure peptides containing C-terminal tyrosine or tryptophan in instruments different from ours (Grützmacher and Heyns, 1966; Pfaender, 1967; Morris *et al.*, 1971; Shemyakin *et al.*, 1971).

In presenting their mass spectra, Nair and Schally illustrated the "correctly calculated" sequence (acylium) ions as having the highest intensity with many intense satellite peaks that were 1, 2, 3, 4, etc., mu both higher and lower than the sequence ions. The lower ions would have implied a loss of either one, two, three, four, etc. hydrogens from the sequence ions; and this has never been observed by others nor by us. The higher ions could be explained in terms of the [^{14}C]methyl iodide and [1-^{14}C]acetic anhydride that they used in the derivatization. However, the [^{14}C]methyl iodide that they used was 25 Ci/mol, which is equivalent to 40% carbon-14. A simple isotope peak calculation (Biemann, 1962) reveals that any sequence ion that has two [^{14}C]methyl groups should have the sequence ion (without carbon-14) plus 2 mu as the highest peak and this peak shifts continuously higher by 2 mu as more [^{14}C]methyl groups are added. Yet, in all of Nair and Schally's spectra, the sequence peaks containing no carbon-14 were presented with the highest intensity, no matter how many [^{14}C]methyl groups they contained.

To further complicate matters, it is not clear which of the several permethylation methods quoted by them was used in the derivatization. In their earliest paper on permethylation (Nair *et al.*, 1970), they mentioned the preparation of silver oxide so we presumed they used the silver oxide method. However, it is well known that this method causes problems in the permethylation of histidine (Thomas *et al.*, 1968).

We can offer no explanation as to why Nair and Schally's spectra for porcine LRF are different from ours and, also, why these authors did not observe the expected shift of the intensity of the sequence ions due to the ^{14}C-labeled isotope.

In conclusion, our mass spectral data fully support the primary structure of ovine LRF as pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2.

Acknowledgment

The authors are grateful for the excellent technical assistance offered by M. Butcher, C. Garcia, R. Wolbers, and V. Yang.

References

Amoss, M., Burgus, R., Blackwell, R., Vale, W., Fellows, R., and Guillemin, R. (1971), *Biochem. Biophys. Res. Commun. 44*, 205.

Amoss, M., Burgus, R., Ward, D. N., Fellows, R. E., and Guillemin, R. (1970), *Progr. 52nd Meeting Endocrine Soc.*, 61.

Baba, Y., Matsuo, H., and Schally, A. V. (1971), *Biochem. Biophys. Res. Commun. 44*, 459.

Bacon, V., Jellum, E., Patton, W., Pereira, W., and Halpera, B. (1969), *Biochem. Biophys. Res. Commun. 37*, 878.

Biemann, K. (1962), Mass Spectrometry-Organic Chemical Applications, New York, N. Y., McGraw-Hill, p 59.

Bricas, E., van Heijenoort, J., Barber, M., Wolstenholme, W. A., Das, D. C., and Lederer, E. (1965), *Biochemistry 4*, 2254.

Burgus, R., Butcher, M., Amoss, M., Ling, N., Monahan, M., Rivier, J., Fellows, R., Blackwell, R., Vale, W., and Guillemin, R. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 278.

Burgus, R., Butcher, M., Ling, N., Monahan, M., Rivier, J., Fellows, R., Amoss, M., Blackwell, R., Vale, W., and Guillemin, R. (1971), *C. R. Acad. Sci., Ser. A 273*, 1611.

Grützmacher, H.-F., and Heyns, K. (1966), *Justus Liebigs Ann. Chem. 698*, 24.

Lenard, J., and Gallop, P. M. (1969), *Anal. Biochem. 29*, 203.

Matsuo, H., Baba, Y., Nair, R. M. G., Arimura, A., and Schally, A. V. (1971), *Biochem. Biophys. Res. Commun. 43*, 1334.

Monahan, M., Rivier, J., Burgus, R., Amoss, M., Blackwell, R., Vale, W., and Guillemin, R. (1971), *C. R. Acad. Sci., Ser. A 273*, 508.

Morris, H. R., Williams, D. H., and Ambler, R. P. (1971), *Biochem. J. 125*, 189.

Nair, R. M. G., Barrett, J. F., Bowers, C. Y., and Schally, A. V. (1970), *Biochemistry 9*, 1103.

Nair, R. M. G., and Schally, A. V. (1972), *Int. J. Peptide Protein Res. 4*, 421.

Pfaender, P. (1967), *Justus Liebigs Ann. Chem. 707*, 209.

Rivier, J., Vale, W., Burgus, R., Ling, N., Amoss, M., Blackwell, R., and Guillemin, R. (1973), *J. Med. Chem. 16*, 545.

Shemyakin, M. M., Ovchinnikov, Yu. A., and Kiryushkin, A. A. (1971), *in* Mass Spectrometry-Techniques and Applications, Milne, G. W. A., Ed., New York, N. Y., Wiley-Interscience, p 289.

Shemyakin, M. M., Ovchinnikov, Yu. A., Vinogradova, E. I., Feigina, M. Yu, Kiryushkin, A. A., Aldanova, N. A., Alakhov, Yu. B., Lipkin, V. M., and Rosinov, B. V. (1967), *Experientia 23*, 428.

Thomas, D. W. (1968), *Biochem. Biophys. Res. Commun. 33*, 483.

Thomas, D. W., Das, B. C., Géro, S. D., and Lederer, E. (1968), *Biochem. Biophys. Res. Commun. 32*, 199.

Vetter-Diechtl, H., Vetter, W., Richter, W., and Biemann, K.

(1968), *Experientia 24*, 340.
Vilkas, E., and Lederer, E. (1968), *Tetrahedron Lett.*, 3089.
White, P. A., and Desiderio, D. M. (1971), *Anal. Lett. 4*, 141.
Wipf, H.-K., Irving, P., McCamish, M., Venkataraghavan, R., and McLafferty, F. W. (1973), *J. Amer. Chem. Soc. 95*, 3369.

Synthesis and Biochemical Studies of Various 8-Substituted Derivatives of Guanosine 3′,5′-Cyclic Phosphate, Inosine 3′,5′-Cyclic Phosphate, and Xanthosine 3′,5′-Cyclic Phosphate†

Jon P. Miller,* Kay H. Boswell, Kei Muneyama, Lionel N. Simon, Roland K. Robins, and Dennis A. Shuman

ABSTRACT: Several 8-substituted derivatives of guanosine 3′,5′-cyclic phosphate (cGMP), inosine 3′,5′-cyclic phosphate (cIMP), and xanthosine 3′,5′-cyclic phosphate (cXMP) were synthesized and their biochemical properties compared to corresponding 8-substituted adenosine 3′,5′-cyclic phosphate (cAMP) derivatives. cGMP was brominated to give 8-bromo-cGMP which was subsequently used to synthesize *via* nucleophilic reactions, 8-hydroxy-, 8-dimethylamino-, 8-methylamino-, 8-benzylamino-, 8-*p*-chlorophenylthio-, 8-benzylthio-, and 8-methylthio-cGMP. Deamination of cGMP and 8-bromo-cGMP gave cXMP and 8-bromo-cXMP, respectively. 8-Bromo-, 8-azido-, 8-hydroxy-, 8-methylthio-, 8-ethylthio-, 8-benzylthio-, and 8-*p*-chlorophenylthio-cIMP were prepared by deamination of the respective 8-substituted cAMP. Hydrogenation of 8-azido-cIMP gave 8-amino-cIMP. Thiation of 8-bromo-cIMP gave 8-thio-cIMP.

The substituted cGMP derivatives were specific for lobster muscle cGMP-dependent protein kinase whereas the 8-substituted cAMP derivatives were specific for bovine brain cAMP-dependent protein kinase. The relative potency of the 8-substituted cGMP and cAMP analogs was Br, OH, SR > cyclic nucleotide > NR_1R_2. cIMP and its 8-substituted derivatives exhibited specificity for the cAMP-dependent protein kinase. Only 8-hydroxy-cIMP and 8-*p*-chlorophenylthio-cIMP were more effective than c-IMP at activating either protein kinase. The 8-substituted derivatives were tested for their ability to inhibit the hydrolysis of cAMP and cGMP by phosphodiesterases from beef heart and rabbit lung. Some cGMP derivatives were better inhibitors of cAMP hydrolysis than of cGMP hydrolysis, while some cAMP derivatives were better inhibitors of cGMP hydrolysis than of cAMP hydrolysis. cGMP hydrolysis was inhibited preferentially by cAMP and cIMP derivatives with aromatic 8 substituents, while the inhibition of cAMP hydrolysis did not show this preference. cXMP was hydrolyzed at a rate only 3% that of cAMP by rabbit kidney phosphodiesterase. Of the 8-substituted derivatives, 8-methylthio-GMP and 8-amino-cIMP were substrates for this enzyme, and were hydrolyzed very slowly.

Guanosine 3′,5′-cyclic phosphate[1] (cGMP) has been implicated in the cellular mediation of hormone action (Hardman *et al.*, 1966). Subsequent studies have shown that administration of acetylcholine resulted in the accumulation of cGMP in heart (George *et al.*, 1970), brain (Ferrendelli *et al.*, 1970), and ductus deferens (Schultz *et al.*, 1972). 8-Bromo-cGMP was able to mimic the action of the cholinergic agents in human lung fragments (Kaliner *et al.*, 1972), and a variety of 8-substituted cGMP and cIMP derivatives were able to affect blood glucose and corticosterone levels as well as blood pressure and heart rate in rats (Paoletti *et al.*, 1973).

Results suggesting that the physiological actions of cGMP are inverse to those of cAMP have been reported in smooth muscle (Puglisi *et al.*, 1971; Ball *et al.*, 1970), renal cortex (Goodman *et al.*, 1972; Pagliara and Goodman, 1969, 1970), lysosomes (Ignarro and Colombo, 1973), antigen-sensitive lymphocytes (Strom *et al.*, 1972), psoriatic lesions (Voorhees *et al.*, 1973), cultured heart cells (Krause *et al.*, 1972), and bacterial cell free extracts (Zubay *et al.*, 1970; Emmer *et al.*, 1970). Clonal proliferation of lymphocytes resulted in a 10–50-fold increase in cGMP levels, while cAMP levels were essentially unaffected (Hadden *et al.*, 1972). Also, it is now well established that increased levels of cAMP are associated with reduced growth rate and induction of differentiation (Van Wijk *et al.*, 1972, and references therein). The growth inhibitory effects of cAMP, which may be mediated in part by the inhibition of precursor transport into cells (Kram *et al.*, 1973), is antagonized by cGMP (Kram and Tomkins, 1973). The available data suggest that cAMP limits growth, possibly by promoting differentiation (Kram *et al.*, 1973), while cGMP stimulates growth at the expense of differentiation.

Toward the goal of developing useful biochemical agents for investigating the functions and mechanisms of action of cGMP and developing potential pharmacological agents, we here report the detailed methods of synthesis of 8-substituted derivatives of cGMP, cIMP (**12**), and cXMP (**10**). Previously,

† From ICN Pharmaceuticals, Inc., Nucleic Acid Research Institute, Irvine, California 92664. *Received August 6, 1973.* Portions of these studies were reported at The First International Conference on the Physiology and Pharmacology of Cyclic AMP, Milan, July 1971, and at The Fifth International Congress on Pharmacology, San Francisco, Calif., July 1972.

[1] Abbreviations used are: cAMP (**22**), adenosine 3′,5′-cyclic phosphate; cGMP (**1**), guanosine 3′,5′-cyclic phosphate; cIMP (**12**), inosine 3′,5′-cyclic phosphate; cXMP (**10**), xanthosine 3′,5′-cyclic phosphate. dbcAMP and dbcGMP are N^6,2′-*O*-dibutyryladenosine 3′,5′-cyclic phosphate and N^2,2′-*O*-dibutyrylguanosine 3′,5′-cyclic phosphate, respectively.