Sequence Analysis of Rat Hypothalamic Corticotropin-Releasing Factor with the *o*-Phthalaldehyde Strategy†

Joachim Spiess,* Jean Rivier, and Wylie Vale

ABSTRACT: Sequence analysis was performed on a 41-residue polypeptide that has been identified as the predominant form of high intrinsic corticotropin-releasing activity of rat hypothalamus. The sequence of residues 1–39 of this corticotropin-releasing factor (CRF) was determined by Edman degradation of a partially purified peptide in a highly sensitive spinning cup sequencer after selective blocking of CRF or its main contaminant with *o*-phthalaldehyde. This approach was validated by peptide mapping of CRF of a highly purified preparation. Peptide mapping was accomplished with reverse-phase high-pressure liquid chromatography of CRF fragments obtained by digestion with clostripain. The identities of the fragments cleaved from CRF were established by chromatographic comparison with synthetic peptides, amino acid analysis, and Edman degradation. On the basis of these experiments, the primary structure of rat hypothalamic CRF was established to be H-Ser-Glu-Glu-Pro-Pro-Ile-Ser-Leu-Asp-Leu-Thr-Phe-His-Leu-Leu-Arg-Glu-Val-Leu-Glu-Met-Ala-Arg-Ala-Glu-Gln-Leu-Ala-Gln-Gln-Ala-His-Ser-Asn-Arg-Lys-Leu-Met-Glu-Ile-Ile-NH_2. It is expected that the *o*-phthalaldehyde strategy will facilitate the sequence analysis of partially purified peptides containing proline residues.

In 1981, we reported the first complete amino acid sequence analysis of a highly potent, large corticotropin-releasing factor (CRF),[1] a 41-residue polypeptide (Spiess et al., 1981a; Vale et al., 1981) that was purified from ovine hypothalamus (Rivier, J., et al., 1982). The primary structure of ovine hypothalamic CRF has recently been confirmed by sequence analysis of cDNA encoding for the amino acid sequence of ovine hypothalamic CRF precursor (Furutani et al., 1983). The 41-residue ovine hypothalamic CRF (or one or several closely related peptides), which plays a key role in the regulation of the adrenocorticotropin (ACTH) and β-endorphin secretion (Vale et al., 1981; Rivier, C., et al., 1982a,b), may also be important in the control of the autonomic nervous system (Brown et al., 1982) and of behavior (Sutton et al., 1982).

On the basis of chromatographic and immunologic evidence, it was suggested that the amino acid sequences of CRF's present in human, dog, and rat hypothalamic extracts deviate significantly from the sequence of ovine CRF (Vale et al., 1983). In view of this observation and the recognition of the rat as the principal laboratory animal, we purified CRF from rat hypothalamus and characterized the predominant form of the high intrinsic corticotropin-releasing activity of the hypothalamic extract. Many aspects of this characterization, including purification, synthesis, and structure–activity relationships, have been described and discussed elsewhere (Rivier et al., 1983). We present here a novel approach in the sequence analysis (of this peptide), which was mainly accomplished by Edman degradation of peptides of an inhomogeneous fraction. This approach was possible on the basis of the *o*-phthalaldehyde strategy of sequence analysis, which allows analysis of partially purified peptides containing proline residues.

Experimental Procedures

Materials. Peptides were synthesized by the solid-phase approach and purified as described (Rivier et al., 1983). Clostripain was obtained from Boehringer, Mannheim, West Germany. Bovine L-1-(tosylamido)-2-phenylethyl chloromethyl ketone–trypsin (TPCK–trypsin) was purchased from Worthington. *o*-Phthalaldehyde (OPA) and 3-sulfophenyl isothiocyanate (3-SPITC) were purchased from Pierce and used without further purification.

Amino Acid Analysis. Analysis was performed with a Beckman 121MB amino acid analyzer. A single column with a ninhydrin detection system and sodium citrate buffers for elution was used. The coefficient of variation was usually between 1% and 10%. More details are published elsewhere (Spiess et al., 1979, 1981b).

Edman Degradation. Peptides (0.7–1.4 nmol) were degraded in a Wittmann-Liebold (1980) modified Beckman 890C spinning cup sequencer equipped with variable speed drive (1000–6000 rpm), automatic converting flask, and microprocessor. Degradation was carried out in the presence of 4–6 mg of Polybrene (Tarr et al., 1978) as described earlier (Spiess et al., 1982a). Phenylthiohydantoin (PTH)-amino acids were determined with reverse-phase high-pressure liquid chromatography (reverse-phase HPLC) (coefficient of variation, 1–6%; minimal detectable amount, <10 pmol) (Spiess et al., 1982a).

Coupling with OPA. In the Edman cycle chosen for OPA treatment, 190 μL of 0.33 M Quadrol buffer (pH 8.9) containing 1 mM OPA was added to the spinning cup (1200 rpm) through the reagent and solvent delivery line. After incubation for 27 min (42 °C) under argon, the Polybrene–peptide film was dried and subjected to the regular degradation procedure.

Peptide Mapping. Peptide (1.5 nmol) was digested for 3.5 h at 37 °C in 50 μL of aqueous 0.05 M Quadrol buffer, 5 mM $CaCl_2$, 1 mM dithiothreitol, and 6% (v/v) 1-propanol (adjusted to pH 8.0 with trifluoroacetic acid) containing 6.8 μg of clostripain/mL. The substrate to enzyme weight ratio was approximately 20:1. In some experiments, peptides (0.5–5.0 nmol) were acetylated with acetic anhydride (Spiess et al., 1981a) before digestion with TPCK–trypsin in 50 μL of buffer composed as described above (except for dithiothreitol, which

† From the Peptide Biology Laboratory, The Salk Institute, San Diego, California 92138. *Received April 8, 1983.* This research was supported in part by National Institutes of Health Grants AM-26741, AM20971, AA03504, and HD13527 and by grants from The Rockefeller Foundation and The March of Dimes–Birth Defects Foundation. This research was conducted in part by The Clayton Foundation for Research, California Division. J.S. and W.V. are Clayton Foundation Investigators.

[1] Abbreviations: CRF, corticotropin releasing factor; rCRF, rat CRF; oCRF, ovine CRF; OPA, *o*-phthalaldehyde; TPCK, L-1-(tosylamido)-2-phenylethyl chloromethyl ketone; 3-SPITC, 3-sulfophenyl isothiocyanate; PTH, phenylthiohydantoin; PTH-N^{ϵ}-SPTC-lysine, PTH derivative of 3-sulfophenylthiocarbamyllysine.

0006-2960/83/0422-4341$01.50/0 © 1983 American Chemical Society

Table I: Amino Acid Analysis of Rat Hypothalamic CRF

amino acid	acid hydrolysis[a] A[b]	B[c]	C[d]	sequence analysis
Asx	9.4 (9)	3.0 (3)	2.0 (2)	2
Thr	3.2 (3)	1.1 (1)	0.8 (1)	1
Ser	3.7 (4)	3.3 (4)	3.2 (3)[e]	3
Glx	15.3 (15)	9.0 (9)	9.0 (9)	9
Pro	6.8 (7)	2.4 (2)	1.8 (2)	2
Gly	4.4 (4)	1.1 (1)	0.4 (0)	0
Ala	7.0 (7)	4.4 (4)	3.8 (4)	4
Cys	<0.3 (0)	<0.4 (0)		0
Val	6.5 (7)	0.9 (1)	0.9 (1)	1
Met	2.0 (2)	1.6 (2)	1.9 (2)[e]	2
Ile	6.4 (6)	1.9 (2)	2.8 (3)	3
Leu	14.1 (14)	6.7 (7)	7.2 (7)	7
Tyr	4.0 (4)	<0.4 (0)	<0.1 (0)	0
Phe	2.1 (2)	1.2 (1)	1.0 (1)	1
Lys	6.1 (6)	0.9 (1)	1.1 (1)	1
His	3.4 (3)	1.8 (2)	2.0 (2)	2
Trp	1.8 (2)	<0.4 (0)	<0.1 (0)	0
Arg	4.5 (5)	3.4 (3)	2.9 (3)	3
total	100	43	41	41

[a] Amino acid ratios are presented; the nearest integer is given in parentheses. [b] Peptide (1.7 μg) of batch 1 (total: 66 μg of peptide; approximately 5.7 nmol of CRF) was hydrolyzed in 25 μL of 4 M methanesulfonic acid and 0.2% tryptamine (110 °C, 24 h). Norleucine (1 nmol) was added as internal standard. [c] Peptide (0.6 μg) of batch 2 (total: 7.7 μg of peptide; 1.6 nmol of CRF) was hydrolyzed as described for A. [d] Peptide (1.4 μg) of batch 3 (total: approximately 8.1 μg of peptide; 1.8 nmol of CRF) was oxidized with hydrogen peroxide and hydrolyzed in 50 μL of constant boiling HCl and 0.03% (v/v) β-mercaptoethanol (140 °C, 9 h). Norleucine (1 nmol) was added as internal standard. [e] Values were corrected for losses during hydrolysis.

was omitted). The reaction was carried out at 37 °C for 4 h with a substrate to enzyme weight ratio of 7:1. All digests were immediately applied to reverse-phase HPLC or frozen prior to chromatography. Reverse-phase HPLC was carried out with a Hewlett-Packard 1084B liquid chromatograph equipped with an autosampler and a programmable variable wavelength detector.

Results

Three batches of CRF purified with gel filtration and reverse-phase HPLC from extracts of 100000 rat hypothalami (Rivier et al., 1983) were available to sequence analysis.

Most of the sequence data were obtained from 66 μg of peptide (based on amino acid analysis) that was provided by purification of CRF from 60000 rat hypothalami (batch 1). The purity of CRF of this preparation was examined by amino acid analysis and Edman degradation. The amino acid ratios of the hydrolysate (Table I, A) deviated significantly from integers and the corresponding values derived from the amino acid composition of ovine CRF. From these data, it was concluded that the CRF fraction under investigation was probably inhomogeneous. This interpretation was confirmed by sequence analysis of 15.8 μg of peptide of this fraction. On the basis of the PTH-amino acid yields, the presence of approximately equal amounts of two main polypeptides was suggested. The PTH-amino acids could not be unambigously assigned to either polypeptide, because the differences between the PTH-amino acid yields in the same Edman cycle were not significant. However, the sequence data were compatible with the assumption that the analyzed fraction contained a polypeptide closely related to ovine CRF.

This fraction was not further purified to avoid possible losses of CRF. Instead, a strategy was followed which allowed sequence analysis of the peptide mixture after selective blocking of CRF or its contaminant. This strategy was based on the distribution of PTH-proline over Edman cycles 1, 4, 5, 10, and 30 and the availability of *o*-phthalaldehyde (OPA) as a blocking reagent (Machleidt & Hofner, 1982) not affecting proline residues.

Whether OPA would be useful in generating quasi-homogeneous conditions for the sequence analysis of partially purified peptides containing proline residues was tested by Edman degradation of synthetic ovine CRF (2 nmol) in the presence of five other synthetic polypeptides: porcine somatostatin-28, human ACTH, corticotropin-like intermediate lobe peptide (CLIP), angiotensin II, and β-endorphin (2 nmol of each). In cycle 4 (corresponding to proline-4 of ovine CRF), all peptides were successfully blocked with OPA except CRF, which was the only peptide remaining accessible to Edman degradation.

The strategy of sequence analysis utilizing OPA for selective blocking was employed to the characterization of rat hypothalamic CRF in two experiments. In the first experiment, approximately 7.9 μg of peptide (containing 0.7 nmol of rat CRF) was applied to the cup and treated with OPA in the first cycle. On the basis of the PTH-amino acid yields, the N-terminal sequence of one of the two peptides was recognized to be H-Pro-Leu-Gln-Gly-Thr-Thr-Gly-Leu-Ile-Pro-Leu-Leu-Gly-Ile-. Since this sequence did not show any significant homology with ovine CRF, it was assigned to the contaminant. Accordingly, the N-terminal sequence of rat hypothalamic CRF could be deduced by taking the results of the sequence analyses of CRF and its contaminant with and without OPA blocking in the first cycle into account (Figure 1). In a second experiment, approximately 15.8 μg of peptide (containing 1.4 nmol of CRF) was applied to the cup, reacted in the first cycle with 3-SPITC and in the fourth cycle with OPA. The modification with 3-SPITC was performed as described (Spiess et al., 1982b) to introduce sulfonic acid groups into the peptides, thereby facilitating their binding to the positively charged peptide carrier Polybrene. After the addition of OPA, only one major PTH-amino acid was identified per cycle, although in some cycles additional PTH-amino acids were detected. However, the yields of these PTH-amino acids never exceeded 10% of the yields of the major PTH-amino acids. The relative yields of PTH-amino acids of consecutive cycles corresponded to the values expected on the basis of sequence analysis of homogeneous fractions of synthetic peptides. Therefore, the PTH-amino acids could unambigously be assigned to residues 4–39 (Figure 2) of a polypeptide that was suggested to be rat hypothalamic CRF on the basis of its structural resemblance to ovine CRF (Figure 4).

Confirmation of the sequence of rat CRF (1–39) and determination of the C-terminal sequence were accomplished by peptide mapping supplemented with amino acid and sequence analysis. These experiments were mainly performed with another preparation of rat hypothalamic CRF (batch 2) purified from 20000 rat hypothalami. On the basis of amino acid analysis (Table I, B), it was estimated that the purity of this CRF preparation was 80–90%. CRF of this preparation was digested with clostripain, which is known to cleave peptides preferentially at the C-terminal side of arginine residues (Mitchell, 1977). By reverse-phase HPLC (Figure 3), eight major products were observed on the basis of their absorbance at 210 nm. Amino acid analysis revealed that products A, C, and D (Figure 3) did not consist of protein, whereas products B, E, F, G, and H represented 1.1, 0.58, 0.54, 1.21, and 0.48 μg of protein, respectively. The amino acid ratios found in the analysis of products B, E, F, and G (Table II) did not significantly deviate from integer numbers. It was thus sug-

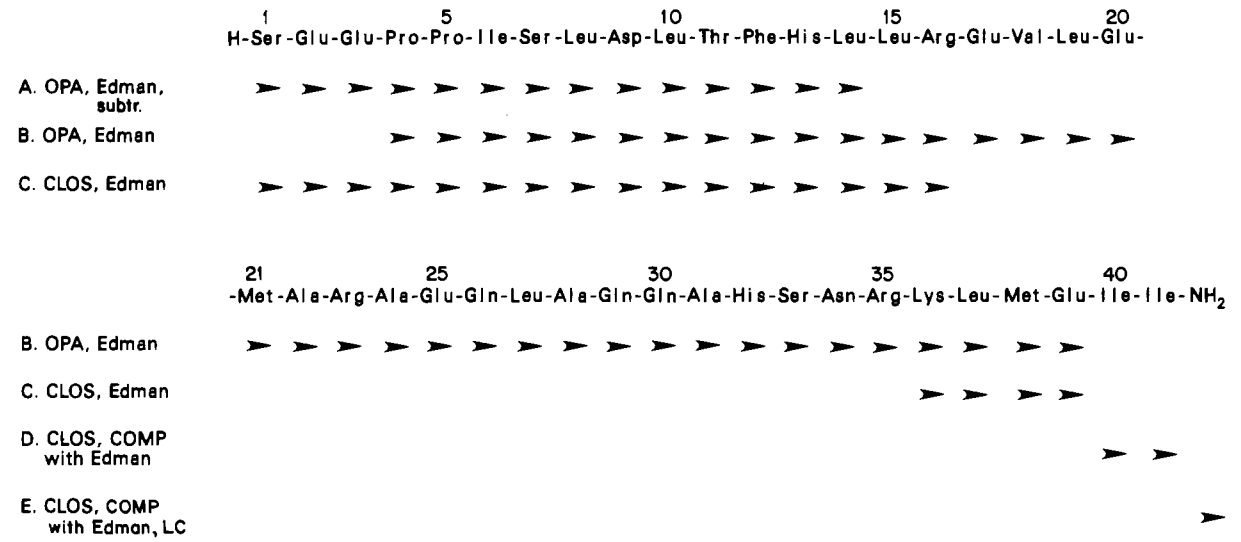

FIGURE 1: Primary structure of rat hypothalamic CRF and strategy of sequence analysis. Arrowheads indicate identified residue or chemical group. (A) Automated Edman degradation of 0.7 and 1.4 nmol of partially purified CRF in two experiments with and without OPA blocking in cycle 1. Residues were identified by subtraction. (B) Automated Edman degradation of 1.4 nmol of partially purified CRF after OPA blocking in cycle 4. (C) Automated Edman degradation of 0.4 nmol of CRF(1–16)-OH (product G, Figure 3) and 0.5 nmol of CRF(36–41)-NH_2 (product F, Figure 3). (D) Residues 40 and 41 were identified on the basis of the composition of CRF(36–41)-NH_2 (product F) and the data presented at line C. (E) C-Terminal amide was determined by comparison of the reverse-phase HPLC retention times of CRF(36–41)-NH_2 and synthetic analogues. CLOS, clostripain; COMP, amino acid composition; subtr, subtraction.

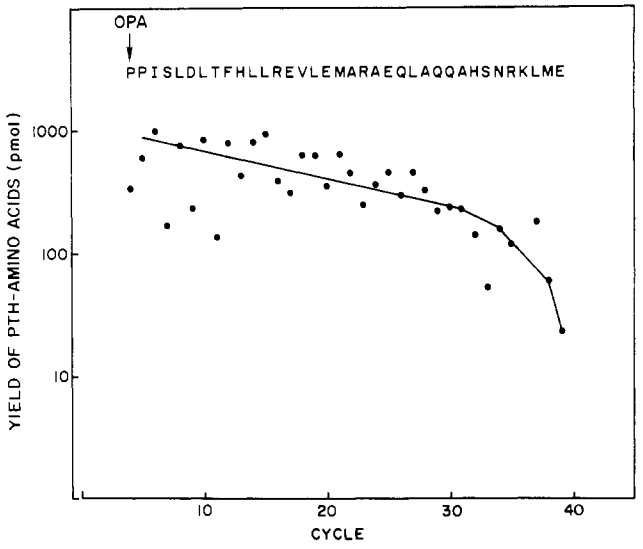

FIGURE 2: Edman degradation of native rCRF(4–39) after OPA blocking. Partially purified rat hypothalamic CRF (batch 1, Table I) (15.8 μg; approximately 1.4 nmol of CRF) was incubated for 80 min at 42 °C in 30 μL of 40 mM 3-SPITC dissolved in 0.33 M Quadrol buffer (pH 9.2) (Spiess et al., 1982a). The mixture, which was protected by a layer of argon, was subsequently transferred to the spinning cup, dried, and subjected to two cleavages with heptafluorobutyric acid before application of the regular degradation program. In cycle 4, OPA was employed as described under Experimental Procedures. In cycles 4 and 5, a second cleavage with heptafluorobutyric acid was carried out. One-third of every cycle fraction was analyzed with reverse-phase HPLC. Serine and threonine were determined as described in Table III. PTH-N^{ϵ}-SPTC-lysine, which was identified but not quantified, behaved on reverse-phase HPLC similarly as did PTH-glutamine (Heil and J. Spiess, unpublished results). Total PTH-amino acid yields are presented. A, alanine; D, aspartic acid; E, glutamic acid; F, phenylalanine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine.

gested that the corresponding fractions contained peptides of a purity that was estimated to be greater than 95%.

On the basis of the amino acid ratios, it was concluded that products B, E, and G were identical with the rat CRF fragments CRF(24–35)-OH, CRF(17–23)-OH, and CRF(1–16)-OH, respectively. This interpretation was supported by the finding that synthetic rat CRF was cleaved with clostripain to fragments that behaved in reverse-phase HPLC like the described products derived from native rat CRF (Figure 3).

The estimated purity and identity of product G as CRF-(1–16)-OH were confirmed by sequence analysis, which was carried out with 0.81 μg of product G (corresponding to 0.44 nmol of peptide). All residues including the C-terminal ar-

Table II: Amino Acid Compositions of Rat CRF Fragments Generated with Clostripain

	acid hydrolysis[a]				
amino acid	G,[b] rCRF-(1–16)	E,[b] rCRF-(17–23)	B,[b] rCRF-(24–35)	F,[c] rCRF-(36–41)	H,[d] rCRF
Asx	1.1 (1)		1.1 (1)		3.2 (2)
Thr	1.0 (1)				2.3 (1)
Ser	1.9 (2)		0.9 (1)		3.9 (3)
Glx	2.0 (2)	2.0 (2)	4.0 (4)	1.0 (1)	9.0 (9)
Pro	1.9 (2)				2.6 (2)
Gly	0.2 (0)	0.1 (0)	0.3 (0)	0.1 (0)	2.7 (0)
Ala		0.9 (1)	3.0 (3)		4.3 (4)
Val		0.9 (1)			2.4 (1)
Met		1.0 (1)		0.8 (1)	1.2 (2)
Ile	0.9 (1)			1.6 (2)	2.5 (3)
Leu	3.9 (4)	1.0 (1)	1.0 (1)	0.9 (1)	6.6 (7)
Tyr					0.8 (0)
Phe	1.1 (1)				1.6 (1)
Lys				0.9 (1)	1.8 (1)
His	1.0 (1)		1.0 (1)		1.5 (2)
Trp					
Arg	0.9 (1)	0.9 (1)	0.9 (1)		3.4 (3)

[a] Ratios of detected amino acids are presented; theoretical values derived from the amino acid compositions of rat CRF or its fragments that were used as markers in reverse-phase HPLC (Figure 3) are given in parentheses. [b] Approximately 0.40, 0.19, and 0.20 μg, respectively, of products G, E, and B (Figure 2) were hydrolyzed for 22 h at 110 °C in 50 μL of constant boiling HCl and 0.03% (v/v) β-mercaptoethanol containing 1 nmol of norleucine as internal standard, dried, and subjected to amino acid analysis. [c] Product F (Figure 3; 0.18 μg) was hydrolyzed for 9 h at 140 °C in constant boiling HCl as described above. [d] Product H (Figure 3; 0.48 μg) was hydrolyzed for 72 h at 110 °C in constant boiling HCl as described above.

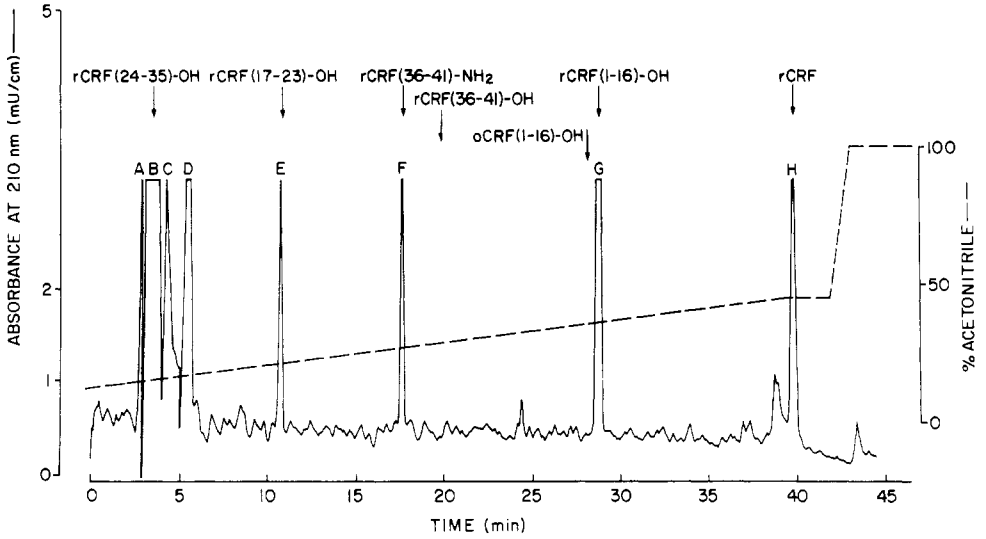

FIGURE 3: Reverse-phase HPLC of fragments cleaved from native rat hypothalamic CRF (1.5 nmol) with clostripain. The sample was applied to a Vydac C_{18} column (0.46 × 25 cm; particle size, 5 m; pore size, 330 Å) and eluted with 0.1% (v/v) trifluoroacetic acid in a mixture of water and acetonitrile. Products detected on the basis of their absorbance are labeled A–H. Arrows mark the retention times of synthetic peptides [rCRF, rCRF(36–41)-OH] and fragments generated with clostripain from synthetic rCRF or oCRF.

ginine were unambigously identified in this experiment (Table III). No evidence for contaminating peptides was provided on the basis of PTH-amino acid analysis.

The retention time of product F did not deviate significantly from the retention time of the synthetic hexapeptide H-Lys-Leu-Met-Glu-Ile-Ile-NH_2. However, product F was eluted significantly earlier (by more than 2 min) than the same hexapeptide with a free C terminus (Figure 3) and significantly later (by 2–5 min) than the synthetic pentapeptides H-Lys-Leu-Met-Glu-Ile-OH and H-Lys-Leu-Met-Glu-Ile-NH_2. On the basis of these data, and in agreement with the results of amino acid analysis (Table II), it was assumed that the C-terminal fragment cleaved from native rat CRF with clostripain was H-Lys-Leu-Met-Glu-Ile-Ile-NH_2. In the sequence analysis of 0.36 μg of product F (corresponding to 0.48 nmol of peptide), residues 1–4 of the C-terminal fragment were confirmed (Table III). No evidence for contaminating peptides was provided. The C-terminal sequence of -Ile-Ile-NH_2 was concluded on the basis of the sequence data in conjunction with the results of amino acid analysis and reverse-phase HPLC. Probably because of its hydrophobic character, the C-terminal dipeptide could be directly determined neither in the analysis of native rat CRF after blocking with OPA nor in the analysis of the C-terminal CRF fragment.

Product H of the clostripain digest (Figure 3) eluted with the retention time of synthetic rat CRF. However, the amino acid composition of this product deviated significantly from the one of rat CRF (Table II). It was additionally observed that product H lacked CRF-like activity in a radioimmunoassay [directed toward oCRF(4–20)], which was shown to read rat hypothalamic CRF (Rivier et al., 1983). Therefore, it was questioned that product H was related to rat CRF but rather suggested that it represented a contaminant of the CRF preparation.

The data of the clostripain fingerprint were in agreement with the results of tryptic peptide mapping of partially purified CRF (batch 1). The tryptic fragments cleaved from 5.8 μg of acetylated peptide were applied to reverse-phase HPLC and identified by amino acid analysis and chromatographic comparison with synthetic peptides. Cleavage occurred only at the C-terminal side of the arginine residues. Most of the contaminant was not digested under these conditions.

Table III: Edman Degradation[a] of Native rCRF(1-16)-OH and rCRF(36-41)-NH_2 Cleaved from rCRF with Clostripain[b]

rCRF(1-16)-OH			rCRF(36-41)-NH_2		
cycle	PTH-amino acid	yield[c] (pmol)	cycle	PTH-amino acid	yield[c] (pmol)
1	Ser	70	1	Lys	360
2	Glu	150	2	Leu	250
3	Glu	170	3	Met	200
4	Pro	140	4	Glu	160
5	Pro	150			
6	Ile	210			
7	Ser	40[d]			
8	Leu	250			
9	Asp	100			
10	Leu	230			
11	Thr	50[e]			
12	Phe	200			
13	His	30			
14	Leu	110			
15	Leu	110			
16	Arg	30			

[a] Edman degradation was carried out with 0.44 nmol of native rCRF(1-16)-OH and 0.48 nmol of native rCRF(36-41)-NH_2. In cycles 4 and 5 of the degradation of rCRF(1-16)-OH, two cleavages with heptafluorobutyric acid were performed. [b] The two fragments, rCRF(1-16)-OH and rCRF(36-41)-NH_2, corresponded to products G and F, respectively, purified with reverse-phase HPLC after digestion with clostripain (Figure 3). [c] Total PTH-amino acid yields are presented. [d] A PTH-serine derivative was obtained as described (Spiess et al., 1981b). [e] Threonine was identified as PTH-dehydrothreonine.

So that the amino acid composition based on sequence data (Table I) could be confirmed, a third preparation of rat hypothalamic CRF (batch 3) was oxidized with hydrogen peroxide (at its methionine residues) (Märki et al., 1981), purified with reverse-phase HPLC (Rivier et al., 1983), and subjected to amino acid analysis. The amino acid ratios (Table I, C) did not deviate significantly from the corresponding values found by sequence analysis.

Thus, the primary structure of rat hypothalamic CRF (Figure 1) was established on the basis of the results of Edman degradation, peptide mapping, and amino acid analysis performed on the native polypeptide and its synthetic replicate. In agreement with the sequence analysis, it was found that

FIGURE 4: Primary structures of ovine and rat hypothalamic CRF. Alignments are marked by shaded areas.

native rat CRF and its synthetic replicate did not differ significantly in their biologic intrinsic activities and potencies to stimulate the secretion of ACTH and β-endorphin from the rat anterior pituitary in vitro and in vivo (Rivier et al., 1983).

Comparison of rat and ovine hypothalamic CRF revealed an 83% sequence homology with conservative replacements (Figure 4) and indistinguishable hypophysiotropic potencies, as determined in the rat anterior pituitary cell culture assay (Rivier et al., 1983). In view of this close chemical and biologic relationship, it appears to be probable that the difference between the two peptides is an expression of species variation.

Discussion

In order to ensure the chemical characterization of rat hypothalamic CRF with the limited amount of peptide available, it was of crucial importance that most of the sequence information be obtained from an inhomogeneous fraction. Thus, losses of CRF by further purification could be avoided, and sequence analysis could proceed while optimal conditions for the purification of rat CRF were established. Difficulties of the CRF purification, which are discussed in more detail elsewhere (Rivier et al., 1983) are, for example, demonstrated by the finding that a highly purified CRF fraction (batch 2) contained a contaminating peptide that was inseparable from native rat CRF by reverse-phase HPLC under the conditions presented in Figure 3.

Analysis of an inhomogeneous fraction of CRF was possible on the basis of the OPA strategy, the rationale of which is to block contaminating peptides with OPA in an Edman cycle, in which the peptide of interest carries a (nonreactive) proline residue as the N terminus. This strategy is based on observations of Bhown et al. (1981), who used fluorescamine in the spinning cup sequence analysis of large polypeptides in order to suppress peptide background built up by incomplete coupling and cleavage reactions. The effect of fluorescamine, which was added to the peptide during Edman degradation when proline (a secondary amine) was the N terminus, was explained as the result of the different reactivity of primary and secondary amines toward this reagent. Since fluorescamine did not remove peptide background during Edman degradation of peptides covalently attached to a solid phase, whereas background suppression was accomplished with OPA (Machleidt & Hofner, 1982), we decided to employ OPA rather than fluorescamine in our experiments. In contrast to Bhown et al. (1981) and Machleidt & Hofner (1982), we did not find any increase of peptide background after blocking. However, the superiority of OPA over fluorescamine as a selective blocking reagent in spinning cup sequence analysis has not been established.

We expect that the OPA strategy, which has been applied here for the first time to the characterization of an unknown peptide, will be useful in the sequence analysis of other proline-containing peptides that are precious and difficult to purify. This approach may be especially advantageous in situations when only partial sequence information is sought, for example, in order to produce antibodies or to synthesize oligonucleotide probes.

Mammalian CRF's and other CRF-like polypeptides are particularly attractive candidates for sequence analysis utilizing OPA, because double proline residues seem to be a characteristic structural element in the N-terminal sequence of these peptides (Spiess et al., 1981a, this work; Lederis et al., 1982; Montecucchi et al., 1979). Should all mammalian CRF's contain the N-terminal sequence X-X-X-Pro-Pro-, it would be highly improbable that a peptide not related to CRF be degraded in the sequencer after blocking with OPA in cycles 4 and 5. The chance to identify a polypeptide not related to CRF would be 1:150 on the basis of currently available data (protein sequence data base of the National Biomedical Research Foundation, Georgetown University Medical Center). In view of the significant sequence homologies between the peptides of the CRF family and on the basis of our experience with the OPA approach, it seems realistic to apply the OPA strategy of sequence analysis to CRF preparations of a purity of 5–10%.

Acknowledgments

The excellent technical assistance of J. Heil, R. Lee, J. Vaughan, G. Yamamoto, L. Keadle, R. Kaiser, R. Galyean, and R. McClintock is gratefully acknowledged. We thank L. Wheatley, D. Weigand, and S. McCall for manuscript preparation.

References

Bhown, A. S., Bennett, J. C., Morgan, P. H., & Mole, J. E. (1981) *Anal. Biochem. 112*, 158–162.

Brown, M., Fisher, L., Spiess, J., Rivier, C., Rivier, J., & Vale, W. (1982) *Endocrinology (Baltimore) 111*, 928–931.

Furutani, Y., Morimoto, Y., Shibahara, S., Noda, M., Takahashi, H., Hirose, T., Asai, M., Inayama, S., Hayashida, H., Miyata, T., & Numa, S. (1983) *Nature (London) 301*, 537–540.

Lederis, K., Letter, A., McMaster, D., Moore, G., & Schlesinger, D. (1982) *Science (Washington, D.C.) 218*, 162–164.

Machleidt, W., & Hofner, H. (1982) in *Methods in Protein Sequence Analysis* (Elzinga, M., Ed.) pp 173–180, Humana Press, Clifton, NJ.

Märki, W., Spiess, J., Tache, Y., Brown, M., & Rivier, J. (1981) *J. Am. Chem. Soc. 103*, 3178–3185.

Mitchell, W. M. (1977) *Methods Enzymol. 47*, 165–174.

Montecucchi, P. C., Henschen, A., & Erspamer, V. (1979) *Hoppe-Seyler's Z. Physiol. Chem. 360*, 1178 (abstract).

Rivier, C., Brownstein, M., Spiess, J., Rivier J., & Vale, W. (1982a) *Endocrinology (Baltimore) 110*, 272–278.

Rivier, C., Rivier, J., & Vale, W. (1982b) *Science (Washington, D.C.) 218*, 377–379.

Rivier, J., Rivier, C., Spiess, J., & Vale, W. (1982) *Anal. Biochem. 127*, 258–266.

Rivier, J., Spiess, J., & Vale, W. (1983) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
Spiess, J., Rivier, J., Rodkey, J. A., Bennett, C. D., & Vale, W. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2974–2978.
Spiess, J., Rivier, J., Rivier, C., & Vale, W. (1981a) *Proc. Natl. Acad. Sci. U.S.A. 78*, 6517–6521.
Spiess, J., Villarreal, J., & Vale, W. (1981b) *Biochemistry 20*, 1982–1988.
Spiess, J., Rivier, J., Rivier, C., & Vale, W. (1982a) in *Methods in Protein Sequence Analysis* (Elzinga, M., Ed.) pp 131–138, Humana Press, Clifton, NJ.
Spiess, J., Rivier, J., Thorner, M., & Vale, W. (1982b) *Biochemistry 21*, 6037–6040.
Sutton, R. E., Koob, G. F., LeMoal, M., Rivier, J., & Vale, W. (1982) *Nature (London) 297*, 331–333.
Tarr, G. E., Beecher, J. F., Bell, M., & McKean, D. J. (1978) *Anal. Biochem. 84*, 622–627.
Vale, W., Spiess, J., Rivier, C., & Rivier, J. (1981) *Science (Washington, D.C.) 213*, 1394–1397.
Vale, W., Rivier, C., Spiess, J., & Rivier, J. (1983) in *Brain Peptides* (Krieger, D., Brownstein, M., & Martin, J., Eds.) Wiley, NY (in press).
Wittmann-Liebold, B. (1980) in *Polypeptide Hormones* (Beers, R. F., Jr., & Bassett, E. G., Eds.) pp 87–120, Raven Press, NY.

Characterization of the Activator Site of *Rhodospirillum rubrum* Ribulosebisphosphate Carboxylase/Oxygenase†

Mark I. Donnelly,‡ Claude D. Stringer, and Fred C. Hartman*

ABSTRACT: Carbon dioxide/magnesium ion activated ribulosebisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* was incubated with the transition-state analogue 2-carboxyarabinitol bisphosphate to form the quaternary complex of enzyme·$^{14}CO_2$·Mg^{2+}·carboxyarabinitol bisphosphate, which was then isolated by gel filtration. Despite the relative instability of this complex compared to the analogous one prepared with spinach carboxylase, the $^{14}CO_2$, presumably bound to a lysyl residue as a carbamate, was trapped in high yield by methylation with diazomethane. Following treatment with base to hydrolyze methyl esters of glutamyl and aspartyl residues formed during the trapping procedure, the ^{14}C-labeled protein was digested with chymotrypsin. Two radioactive peptides were purified from the digest by preparative chromatographic procedures. Amino acid analyses, automated Edman degradations, and carboxy-terminal analyses with carboxypeptidases A and Y revealed their sequences to be identical: Leu-Gly-Gly-Asp-Phe-Ile-Lys-Asn-Asp-Glu-Pro-Gln-Gly-Asn-Gln-Pro-Phe. Residue 7, the site of bound $^{14}CO_2$, was identified at the level of its phenylthiohydantoin derivative as N^{ϵ}-(methoxycarbonyl)lysine. Although the observed sequence is similar to that described for the activator-site region of the spinach carboxylase [Lorimer, G. H. (1981) *Biochemistry 20*, 1236–1240], the isoleucyl and asparaginyl residues immediately adjacent to the carbamate-forming lysine of the *R. rubrum* carboxylase are substituted by threonine and aspartic acid, respectively.

Activation of D-ribulose 1,5-bisphosphate (ribulose-P_2)[1] carboxylase/oxygenase involves the sequential binding of CO_2 and divalent metal ion (Lorimer et al., 1976; Laing & Christeller, 1976). For the enzyme from spinach, it has been demonstrated rigorously that activating CO_2 is distinct from CO_2 utilized as substrate (Miziorko, 1979; Lorimer, 1979) and that activation occurs via the formation of a carbamate with Lys-201 of the large subunit (Lorimer & Miziorko, 1980; Lorimer, 1981; Zurawski et al., 1981). This identification entailed isolation of the quaternary complex enzyme·activator CO_2·Mg^{2+}·carboxyarabinitol-P_2. The transition-state analogue carboxyarabinitol-P_2 (Pierce et al., 1980) greatly enhanced the stability of the carbamate and permitted its permanent trapping by esterification with diazomethane; the esterified carbamate survived the conditions used for protein fragmentation and subsequent peptide fractionation.

Universality of reversible carbamylation as a regulatory mode for ribulose-P_2 carboxylases is implied by studies with the enzyme from *Rhodospirillum rubrum*. In contrast to most species of ribulose-P_2 carboxylase which consist of eight large catalytic subunits (M_r = 53 000) and eight small subunits (M_r = 14 000) of ill-defined function, the *R. rubrum* enzyme is a dimer of catalytic subunits (Tabita & McFadden, 1974). Despite this striking difference in quaternary structure, the enzyme exhibits CO_2/Mg^{2+}-dependent activation (Christeller & Laing, 1978; Whitman et al., 1979) which appears to involve the formation of a carbamate between a lysyl ϵ-amino group

† From the Biology Division, Oak Ridge National Laboratory, and the University of Tennessee–Oak Ridge Graduate School of Biomedical Sciences, Oak Ridge, Tennessee 37830. *Received March 31, 1983.* This research was sponsored jointly by the Office of Health and Environmental Research, U.S. Department of Energy, under Contract W-7405-eng-26 with the Union Carbide Corp., by the Sciences and Education Administration of the U.S. Department of Agriculture under Grant 78-59-2474-0-1-161-1, and by the National Science Foundation under Grant PCM-8207516.

* Address correspondence to this author at the Biology Division, Oak Ridge National Laboratory.
‡ Postdoctoral investigator supported by National Science Foundation Grant PCM-8207516.

[1] Abbreviations: ribulose-P_2, D-ribulose 1,5-bisphosphate; carboxyarabinitol-P_2, D-2-carboxyarabinitol 1,5-bisphosphate; Bicine, *N,N*-bis(2-hydroxyethyl)glycine; Quadrol, *N,N,N′,N′*-tetrakis(2-hydroxypropyl)ethylenediamine; TPCK, tosylphenylalanyl chloromethyl ketone; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; Gdn·HCl, guanidine hydrochloride.

0006-2960/83/0422-4346$01.50/0 © 1983 American Chemical Society