

PARTIAL AMINO ACID SEQUENCE OF THE LARGE SUBUNIT OF RIBULOSEBISPHOSPHATE CARBOXYLASE FROM BARLEY

by

CARSTEN POULSEN

Department of Physiology, Carlsberg Laboratory,
Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby

and

BRIAN MARTIN and IB SVENDSEN

Department of Chemistry, Carlsberg Laboratory,
Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby

Keywords: Fraction I, large subunit, ribulose-1,5-bisphosphate carboxylase, cyanogen bromide, amino-terminus, carboxy-terminus, spinach

The large subunit of D-ribulose-1,5-bisphosphate carboxylase (E.C. 4.1.1.39) from barley has been subjected to partial amino acid sequencing. The amino-terminal sequence comprising 46 residues of the reduced and carboxymethylated large subunit polypeptide has been determined. The carboxy-terminal end contained the sequence -Leu-Ala-Val · COOH. Eight of the ten cyanogen bromide fragments have been partially sequenced from the N-terminus. A ninth fragment was found to be blocked to N-terminal sequencing due to cyclization of a glutamine residue after cyanogen bromide cleavage. The eight sequences which have been determined comprise 210 of the approximate 490 residues in the large subunit. The sequences have been compared to five sequences of the large subunit from spinach. Among 37 residues which could be compared four differences were encountered.

Abbreviations: AAA = amino acid analysis, DFP = diisopropylphosphorfluoridate, DMAA = N, N-dimethyl-N-allylamine, HPLC = high performance liquid chromatography, Polybrene = 1,5-dimethyl-1,5-diazaundecamethylene polymethobromide, PTH = phenylthiohydantoin, RuBPCase = D-ribulose-1,5-bisphosphate carboxylase, TFA = trifluoroacetic acid, THEED = N,N,N',N'-tetrakis(2-hydroxyethyl)ethylene-diamine, TLC = thin layer chromatography.

1. INTRODUCTION

In the preceding paper the isolation and purification of the large subunit of D-ribulose-1,5-bisphosphate carboxylase (RuBPCase) of barley and its cyanogen bromide fragments have been presented (12). In the present paper the partial amino acid sequences of eight of the cyanogen bromide fragments are reported. The established sequences which include the 46 N-terminal residues and the C-terminus account for 40% of the complete sequence of this peptide chain coded for by chloroplast DNA. A brief summary of some of the data has been presented earlier (15).

2. MATERIALS AND METHODS

2.1. Materials

DFP-treated carboxypeptidase A was a product of Sigma Chemicals Ltd., St. Louis, U.S.A. Pyroglutamate aminopeptidase was purchased from Boehringer, Mannheim, Germany. Sephadex G-25 Superfine was obtained from Pharmacia Fine Chemicals, Uppsala, Sweden.

For sequence determinations the following chemicals were employed. N,N,N',N'-tetrakis(2-hydroxyethyl)ethylene diamine (THEED) was purchased from ICN-K & K Laboratories Inc., Plainview, New York, U.S.A., and used without further purification. Ethyl acetate (UV solvent grade) was obtained from Merck-Schuchardt, Munich, Germany. Both 0.1M-THEED and ethyl acetate were found to be free of aldehydes as determined by the Tollens reaction (4). Phenylisothiocyanate, heptafluorobutyric acid, 1-propanol, N,N-dimethyl-N-allylamine (DMAA), trifluoroacetic acid, benzene, polybrene and butyl chloride were sequenal grade products from Pierce Chemicals, Rockford, Ill., U.S.A. n-Heptane was a sequencer grade chemical from Fluka, Switzerland. Acetonitrile was an analytical grade chemical from Merck-Schuchardt, Munich, Germany, and was filtered through a 0.5 μ m Millipore filter prior to use with the high performance liquid chromatograph (HPLC).

2.2. Preparation of the large subunit for sequence analysis

S-carboxymethylated large subunit was obtained as described in the preceding paper (12). For

C-terminal sequence analysis increased solubility of the S-carboxymethylated peptide chain was required. Accordingly, 33 mg were citraconylated in 3 ml of 0.5M-NaHCO₃, pH 8.4, 6M in guanidinium chloride. The citraconylation procedure described by ATASSI and HABEED (2) was followed. The modified protein was then dialyzed against 3 \times 5 liters of 0.1M-NH₄OH over a period of 24 hours and freeze dried. SDS polyacrylamide gel electrophoresis and amino acid analysis were performed on this material after decitraconylation (2).

2.3. Digestion with carboxypeptidase A

200 nmoles of citraconylated large subunit was dissolved in 2 ml of 0.225M-N-ethylmorpholine acetate buffer pH 8.3, containing 220 nmoles of norleucine as an internal standard. 200 μ l of a 1 mg·ml⁻¹ solution of DFP-treated carboxypeptidase A (6 nmoles) prepared as described by AMBLER (1) was added to the solution in the cold (0° C) as well as to two ml of a similar blank which did not contain protein. From the two digestion mixtures a zero time sample of 100 μ l was immediately removed and added to 50 μ l of glacial acetic acid. The tubes containing the digestion mixtures were sealed and placed in a water bath at 32° C and 100 μ l aliquots were removed from each tube at appropriate time intervals. The aliquots were immediately mixed with 50 μ l of glacial acetic acid to stop further digestion. Subsequently, all samples were dried in an evacuated desiccator and 50 μ l of water were added followed by 5 μ l of a 30% solution of sulphosalicylic acid which precipitated the protein. After centrifugation, 40 μ l of the supernatant was removed and added to 4 μ l of 1M-NaOH to adjust the pH to 2.0. 20 μ l aliquots were analyzed on the Durrum D-500 amino acid analyzer.

2.4. Preparation and purification of cyanogen bromide fragments

The procedure employed for the cyanogen bromide cleavage and subsequent purification of the resulting peptides are described in the preceding paper (12).

Table I.

Automated Edman degradation of the large subunit of RuBPCase and its CNBr-fragments.

Sample code	Quantity (nmoles)	No. of steps	Buffer system	Identification system
LS-1	150	20	0.1M-THEED	HPLC/TLC
CBP-CM2-I	80	34	—	HPLC
CBP-CM2-II	180	46	—	HPLC/AAA
CBP-CM1	100	30	—	HPLC
CBP-CM5X	75	30	—	HPLC
CBCam-IIISP4	150	53	—	HPLC/TLC
CBP-CM2-II	50	20	—	HPLC
CBCam-IIISP2	100	21	THEED + polybrene	HPLC
CBCam-IVSP2	150	30	0.1M-THEED	HPLC
CBCam-IVCM2	125	13	THEED + polybrene	HPLC
CBP-CM8	120	23	DMAA/TFA	AAA/TLC

Fragments obtained by cyanogen bromide cleavage are designated with the letters CB. To this is added the letter P, if the subunit had been S-4-pyridylethylated or the symbol Cam if the subunit had been S-carboxymethylated. Roman numerals indicate the order of appearance of the fragments from gel filtration columns. Fragments separated by ion exchange columns are designated with CM if a CM-cellulose column is used and with SP if a SP-Sephadex column is employed. They are numbered with arabic numerals in order of elution. LS-1 designates a run of S-carboxymethylated large subunit of ribulosebisphosphate carboxylase from barley.

2.5. Amino terminal sequence determination employing the Beckman 890C sequencer

The automated sequence determination was performed using either Beckman programme 102974 (DMAA peptide programme) or 122974 (single cleavage, protein Quadrol). In the latter case the 1.0M-Quadrol was replaced by 0.1M-THEED as described previously (3). In two cases, polybrene was added to the cup prior to the initial step as described by OVERBALLE-PETERSEN et al. (9). Identification of PTH-amino acids in nearly all cases was accomplished by high performance liquid chromatography using a Hewlett-Packard HPLC chromatograph Model 1084B. A Merck reverse phase column (RP-8; 10 μ) was employed with a gradient elution system (10). As additional evidence, identification was made by thin-layer chromatography (6) and/or back-hydrolysis to the free amino acids (8). In the case of peptide CBP-CM8 the identification was made by the latter two methods only. A survey of the CNBr fragments analyzed is presented in Table I.

2.6. Treatment with pyroglutamate aminopeptidase

Peptide CBCam-IIISP1 was treated with pyroglutamic acid aminopeptidase to remove the pyroglutamic acid residue which was thought to have prevented amino-terminal sequencing of this peptide. The method was that described by PODELL and ABRAHAM (11). Forty nanomoles (0.3 mg) of peptide was taken up in 0.5 ml of deblocking buffer and 0.2 mg of the crude enzyme extract was added (10 μ g of active enzyme). The vessel was flushed with N₂, sealed and left to react with mixing for 9 hours at 4° C. Subsequently a second 0.2 mg crude enzyme extract was added, after equilibrating the mixture at room temperature, again flushed with N₂ and sealed. After 14 hours the peptide was passed over a column of Sephadex G-25 superfine (2.5 \times 30 cm) equilibrated and eluted with 0.1M-NH₄OH and then freeze dried prior to N-terminal analysis (12). A 50 μ l control was removed from the assay vessel immediately after the first enzyme addition.

3. RESULTS

3.1. C-terminal sequence

As seen from the SDS polyacrylamide gel electrophoresis pattern in Figure 1 the large subunit sample moves as one band after reduction, S-carboxymethylation, citraconylation and decitraconylation. Treatment of the citraconylated reduced and S-carboxymethylated large subunit with DFP-treated carboxypeptidase A resulted in the identification of the three first amino acid residues released from the C-terminal end: -Leu-Ala-Val-COOH, as can be seen from the release curves in Figure 2. Any further interpretation of the data after three hours was not possible due to a complicated amino acid release pattern.

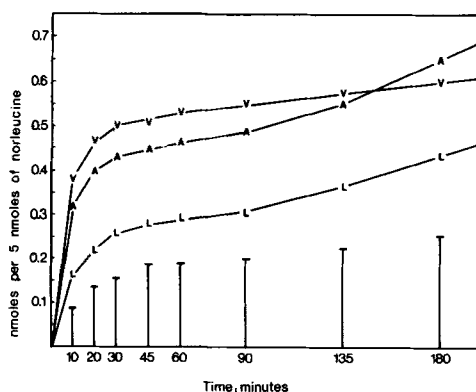


Figure 2. Release of amino acids from the C-terminus of RuBPCase large subunit digested with DFP-treated carboxypeptidase A as a function of time. V = valine, A = alanine, L = leucine. The bars indicate the sum of all other amino acids.

Figure 1. SDS-polyacrylamide gel electrophoresis of large subunit of RuBPCase from barley modified by S-carboxymethylation, citraconylation and decitraconylation. 50 μ g of protein was subjected to electrophoresis as previously described (12) on a 11–15% polyacrylamide gradient gel.

Table II

N-terminal amino acid sequences of ribulose-1,5-bisphosphate carboxylase large subunit (LS-1) and eight of its cyanogen bromide fragments. For further details see text.

	1	10
LS-1	Ala-Gly-Val-Lys-Asp-Tyr-Lys-Leu-Thr-Tyr-Tyr-Thr-Pro-Glu-Tyr-	
CBP-CM2-I	20	30
	Glu-Thr-Lys-Asp-Thr-Asp-Phe-Leu-Ala-Ala-Phe-Arg-Val-Ser-Pro-	
		40
	Gln-Pro-Gly-Val-Pro-Pro-Glu-Glu-Ala-Gly-Ala-Ala-Val-Ala-Ala-Glu-	
	1	10
CBP-CM1	Pro-Ala-Leu-Thr-Glu-Ile-Phe-Gly-Asp-Asp-Ser-Val-Leu-Gln-Phe-	
CBCam-IISP4	20	30
CBP-CM5X	Gly-Gly-Gly-Thr-Leu-Gly-His-Pro-Trp-Gly-Asn-Ala-Pro-Gly-Ala-	
		40
	Ala-Ala-Asn-Arg-Val-Ala-Leu-Glu-Ala-Cys-Val-Gln-Ala-Arg-Asn-	
		50
	Glu-Gly-Arg-Asp-Leu-Ala-Arg-Glu-	
	1	10
CBP-CM2-II	Ser-Gly-Gly-Asp-His-Ile-His-Ser-Gly-Thr-Val-Val-Gly-Lys-Leu-	
CBCam-IIISP2	Glu-Gly-Glu-Arg-	
	1	10
CBCam-IVSP2	Arg-Ile-Arg-Asp-Arg-Phe-Val-Phe-Cys-Ala-Glu-Ala-Ile-Tyr-Lys-	
CBCam-IVCM2	20	30
(Arg-peptide)	Ser-Gln-Ala-Glu-Thr-Gly-Glu-Ile-Lys-Gly-His-Tyr-Leu-Asn-Ala-	
	1	10
(His-peptide)	His-Asp-Tyr-Leu-Thr-Gly-Gly-Phe-Thr-Ala-Asn-Thr-Thr-Leu-Ala-	
		20
	His-Tyr-Cys-Arg-Asp-	
	1	10
CBP-CM8	Ile-Lys-Gly-Ala-Val-Phe-Ala-Arg-Gln-Leu-Gly-Val-Pro-Glu-	
	1	10
CBP-CM3	Pro-Gly-Val-Ile-Pro-Val-Ala-Ser-Gly-Gly-Ile-His-Val-Trp-Trp-	
	His-Met	
	1	
CBP-VI	Lys-Ala-Val-Ile-Asp-His-Arg-Gln-	

3.2. N-terminal sequences of the large subunit and of its cyanogen bromide fragments

The reduced and S-carboxymethylated large subunit (LS-1) was subjected to automated Edman degradation resulting in the identification of the first 20 amino acid residues. Further sequencing was not possible due to increasing back-ground of other amino acid derivatives.

The N-terminal sequence of the largest cyanogen bromide fragment (CBP-CM2-I) was found to be identical to the N-terminal sequence of the intact large subunit. From two runs on the sequenator the total of 46 of the 135 residues in this peptide were delineated (Table II).

The next largest peptide (CBP-CM1, CBCam-IISP4) was shown to contain 90 amino acid

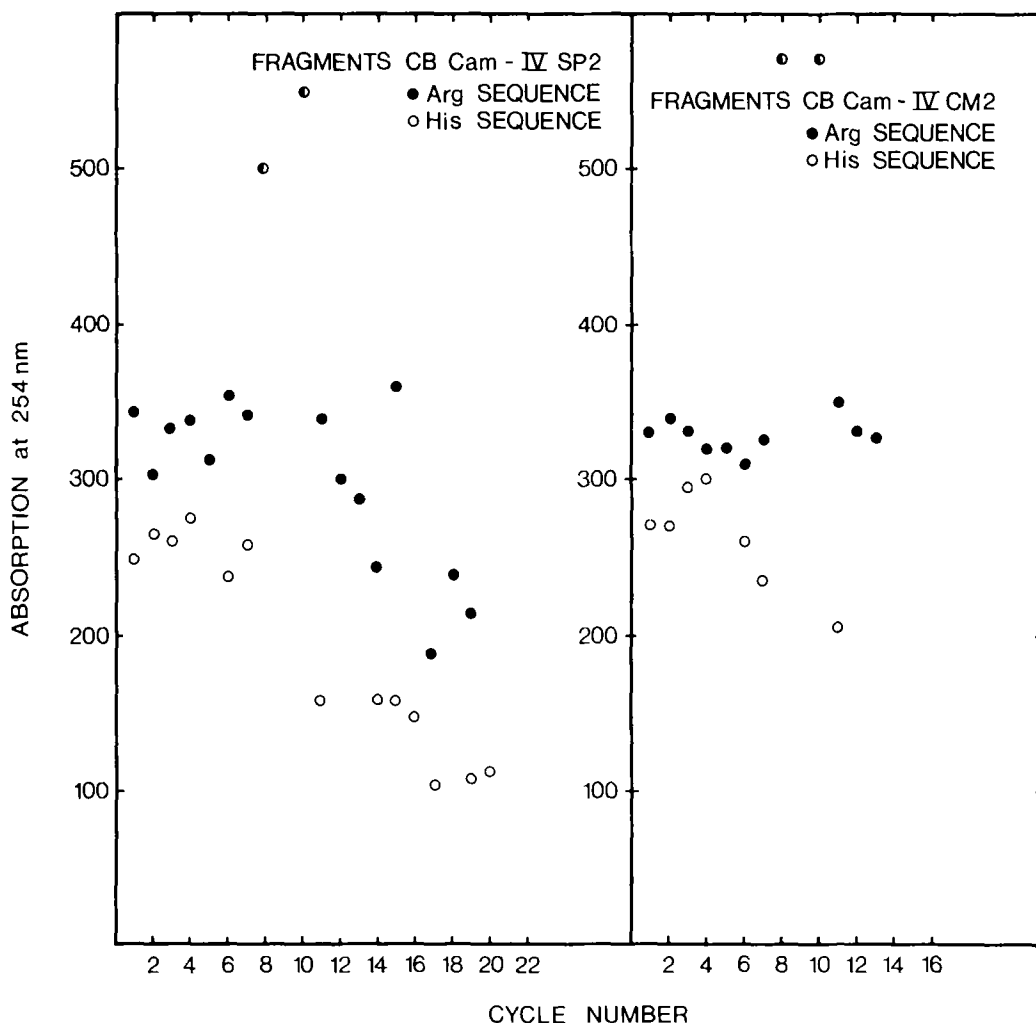


Figure 3. The sequential release of amino acids from the Arg- and His-CNBr-fragments of the RuBPCase large subunit. Left side: without polybrene. Arg-fragment corrected for overlaps. The half filled symbols indicate identical residues in the two sequences. Right side: with polybrene. Arg-fragment corrected for overlaps. For further details see text.

residues. Three sequence runs gave the sequence of the first 53 amino acid residues (Table II).

Two of the cyanogen bromide fragments isolated were of similar size (50–60 residues). One of these (CBP-CM6 + 7, CBCam-IIISP1) was found to be blocked to both manual N-terminal analysis (12) and automated Edman degradation. After treatment of CBCam-IIISP1 with pyroglutamate aminopeptidase, a valine residue was found by N-terminal analysis. However, due to

the small quantity of the deblocked peptide no further N-terminal sequence data was obtainable. The second peptide (CBCam-IIISP2, CBP-CM2-II) was composed of 55 amino acid residues and serine was found at the amino-terminal end. This fragment was not found in all cyanogen bromide cleavage reactions and was isolated in very low yields. Automated Edman degradation resulted in the sequence of the first 19 amino acid residues (Table II).

Table III

Comparison of three CNBr-fragments from barley D-ribulose-1,5 bisphosphate carboxylase large subunit to fragments obtained from the enzyme isolated from spinach (two CNBr-fragments and one tryptic fragment).

	1	10
Barley (CBP-CM8)	Ile-Lys-Gly-Ala-Val-Phe-Ala-Arg-Gln-Leu-Gly-Val-Pro-Glu-	
	1	
Spinach (CBP-VI)	Met-His-Arg-Ala-Val-Phe-Ala-Arg-Gln-	
	1	
Barley (CBP-VI)	Lys-Ala-Val-Ile-Asp-His-Arg-Gln-	
	1	
Spinach (CBP-VI)	Lys-Ala-Val-Ile-Asp-His-Arg-Gln-	
	1	10
Barley (CBP-CM2-II)	Met-Ser-Gly-Gly-Asp-His-Ile-His-Ser-Gly-Thr-Val-Val-Gly-Lys-	
	20	
	Leu-Glu-Gly-Glu-Arg-Glu-	
	1	10
Spinach (5)	Leu-Ser-Gly-Gly-Asp-His-Ile-His-Ser-Gly-Thr-Val-Val-Gly-Lys-	
	20	
	Leu-Glu-Gly-Glu-Arg-	

Automated Edman degradation on cyanogen bromide fragment CBP-CM8 was terminated after 23 cycles. This sequence run resulted in the definitive identification of 14 of the total 25 residues in this sequence (Table II).

The two cyanogen bromide fragments which were approximately 40 residues long (CBCam-IV) could not be separated from each other (12) by the chromatographic techniques employed. However, the mixture of the two fragments (CBCam-IVSP2, CBCam-IVCM2) was subjected to automated Edman degradation. In case of mixture CBCam-IVSP2 the two peptides were present in a ratio of 3:2 and could therefore be identified simultaneously (Figure 3). An additional aid to identification was the fact that the arginine-sequence ran with a step-wise overlap of approximately 30% from step 3 onward. Prior to running mixture CBCam-IVCM2, polybrene was added to the cup, in an attempt to increase the repetitive yields. Although in the initial step the two sequences were found in close to equimolar amounts the subsequent steps provided additional support for the sequence deduced from CBCam-IVSP2. The use of polybrene had a pronounced effect on the repetitive yield of the arginine sequence while

leaving the histidine sequence unaffected (Figure 3).

The two smallest cyanogen bromide fragments from the barley large subunit, having N-terminal proline and lysine (CBP-CM3, CBCam-VISP1 + 2, CBP-VI), have been analyzed by manual Edman degradation and peptide mapping (12). One fragment contained the sequenced 17 amino acids and the other a total of about 13 amino acids.

4. DISCUSSION

The partial amino acid sequences determined for the eight cyanogen bromide fragments and the carboxyl-terminal analysis account for 210 of the approximately 490 residues in the large subunit of RuBPCase. The N-terminal sequence of the undegraded polypeptide was found to be identical to that of the largest cyanogen bromide fragment. The sequence of the 46 established N-terminal amino acid residues will be useful in attempts to find by nucleotide sequencing the end of the chloroplast gene coding for the large subunit. The N-terminus of the 55.000 mol. wt. large subunit of the ribulosebisphosphate car-

boxylase of *Rhodospirillum rubrum* has been determined to be Met-Asx-Glx- by SCHLOSS et al. (13). The carboxylase of this photosynthetic bacterium has a quaternary structure very different from the higher plant enzymes. According to MCFADDEN and PUROHIT (7) it contains only two large subunits as compared to the octameric structure of the plant enzymes and furthermore it is devoid of a small subunit. It is therefore not surprising that the two N-termini bear no resemblance to one another. Attempts to determine the N-terminal residue of the large subunit in tobacco (G. C. GIBBONS, person. commun.) and *Oenothera* (B. MARTIN and A. HOLDER, unpublished) have so far not been successful because of blocked N-terminal residues.

Treatment of the barley large subunit with DFP-treated carboxypeptidase A released a number of amino acids from the carboxyl-terminal end. They are consistent with the sequence -Leu-Ala-Val-COOH. Valine is also the C-terminal amino acid of the large subunit of RuBPCase isolated from spinach (14).

Using bromodihydroxybutanone bisphosphate and N-bromoacetylanthranilic acid as substrate analogue affinity label HARTMAN et al. (5) have isolated and sequenced 3 labelled tryptic peptides from spinach with 20, 13 and 13 residues respectively. One of these is homologous to cyanogen bromide fragment CBP-CM2-II of barley in 19 of the 20 amino acid positions as shown in Table III. The single alteration in the two sequences is Met as the pre-N-terminal residue in barley and Leu in the corresponding position in the spinach peptide. As the tryptic peptide from spinach originates from a portion of the large subunit polypeptide which has chemically reacted with the substrate analogue this peptide is probably in the vicinity of the catalytic site. The homologous cyanogen bromide fragment CBP-CM2-II of barley can thus be expected to mark a portion of the polypeptide chain which plays a role in catalysis. The two other proposed active site tryptic peptides which have been sequenced from spinach large subunit have as yet no homologous counterpart in the sequenced sections of the barley large subunit.

Two cyanogen bromide fragments from spinach large subunit have been sequenced (12) and were found homologous to cyanogen bro-

mid fragment CBP-CM8 and the lysine CBP-VI fragment of barley (Table III). Complete homology is found for the 8 residues of the CBP-VI fragment, whereas in fragment CBP-CM8 six of the nine positions sequenced are homologous.

Comparing 37 amino acid residues of spinach and barley large subunit 4 differences were encountered. Three of these differences can be derived by single nucleotide substitutions, whereas the fourth (Lys/His) requires two nucleotide substitutions. This definitively demonstrates that the nucleotide sequence of the gene in the chloroplast coding for the large subunit of RuBPCase in barley differs from that in spinach.

ACKNOWLEDGEMENT

We would like to thank Professor M. OTTESEN for helpful advice and Ms. LONE SØRENSEN and Ms. BODIL CORNELIUSSEN for excellent technical assistance.

REFERENCES

1. AMBLER, R. P.: Enzymatic hydrolysis with carboxypeptidases. In: *Methods in Enzymology*. C. H. W. Hirs and S. N. Timasheff, eds., Academic Press, New York and London, vol. XXV, 143-154 (1972).
2. ATASSI, M. Z. & A. F. S. A. HABEED: Reactions of protein with citraconic anhydride. In: *Methods in Enzymology*. C. H. W. Hirs and S. N. Timasheff, eds., vol. XXV, 546-553 (1972).
3. BEGG, G. S. & F. J. MORGAN: A non-volatile buffer with improved performance in automated sequencing. *FEBS Letters* 66, 243-245 (1976).
4. EDMAN, P. & A. HENSCHEN: Sequence determination. In: *Protein Sequence determination*, 2. ed., S. B. Needleman, ed., Springer-Verlag, Berlin-Heidelberg-New York, pp. 262 (1975).
5. HARTMAN, F. C., I. L. NORTON, C. D. STRINGER & J. V. SCHLOSS: Attempts to apply affinity labeling techniques to ribulose bisphosphate carboxylase/oxygenase. In: *Photosynthetic carbon assimilation*. H. W. Siegelman and G. Hind, eds., Plenum Press, New York-London, pp. 245-269 (1978).
6. KULBE, K. D.: Micropolyamide thinlayer chromatography of phenylthiohydantoin amino acids (PTH) at subnanomolar level. A rapid micro-technique for simultaneous multisample identifi-

- cation after automated Edman degradations. *Analytical Biochemistry* 59, 546–573 (1974).
7. MCFADDEN, B. A. & K. PUROHIT: Chemosynthetic, photosynthetic and cyanobacterial ribulose-bisphosphate carboxylase. In: *Photosynthetic carbon assimilation*. H. W. Siegelman and G. Hind, eds., Plenum Press, New York-London, pp. 179–207 (1978).
 8. MENDEZ, E. & C. Y. LAI: Regeneration of amino acids from thiazolinones formed in the Edman degradation. *Analytical Biochemistry* 68, 47–53 (1975).
 9. OVERBALLE-PETERSEN, C., B. MARTIN, V. HASEMANN, J. T. JOHANSEN, I. SVENDSEN & M. OTTESEN: The complete amino acid sequence of Cu, Zn-superoxide dismutase from *Saccharomyces cerevisiae*. *Carlsberg Res. Commun.* 44, (in preparation).
 10. PETERSEN, C., V. HASEMANN, B. MARTIN, J. T. JOHANSEN, I. SVENDSEN & M. OTTESEN: The amino terminal sequence of superoxide dismutase from *Saccharomyces cerevisiae*. *Carlsberg Res. Commun.* 42, 391–395 (1977).
 11. PODELL, D. N. & G. N. ABRAHAM: A technique for the removal of pyroglutamic acid from the amino terminus of proteins using calf liver pyroglutamate amino peptidase. *Biochem. Biophys. Res. Commun.* 81: 176–185 (1978).
 12. POULSEN, C.: Cyanogen bromide cleavage of the large subunit of ribulosebisphosphate carboxylase from barley. *Carlsberg Res. Commun.* 44, 163–189 (1979).
 13. SCHLOSS, S. V., E. F. PHARES, M. V. LONG, C. D. STRINGER & F. C. HARTMAN: Increased levels of ribulosebisphosphate carboxylase in *Rhodospirillum rubrum*. In: *Photosynthetic carbon assimilation*. H. W. Siegelman and G. Hind, eds., Plenum Press, New York-London, pp. 424–425 (1978).
 14. SYGIYAMA, T. & T. AKAZAWA: Subunit structure of spinach leaf ribulose-1,5-bisphosphate carboxylase. *Biochemistry* 9, 4499–4504 (1970).
 15. WETTSTEIN, D. von, C. POULSEN & A. A. HOLDER: Ribulose-1,5-bisphosphate carboxylase as a nuclear and chloroplast marker. *Theor. Appl. Genet.* 53, 193–197 (1978).