Plant Physiol. (1992) 98, 1170–1174
0032-0889/92/98/1170/05/$01.00/0

Received for publication September 16, 1991
Accepted November 11, 1991

Posttranslational Modifications in the Amino-Terminal Region of the Large Subunit of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase from Several Plant Species[1]

Robert L. Houtz*, Loelle Poneleit, Samantha B. Jones[2], Malcolm Royer, and John T. Stults

Department of Horticulture and Landscape Architecture, Plant Physiology/Biochemistry/Molecular Biology Program, University of Kentucky, Lexington, Kentucky 40546 (R.L.H., L.P., S.B.J., M.R.); and Department of Protein Chemistry, Genentech, South San Francisco, California 94080 (J.T.S.)

ABSTRACT

A combination of limited tryptic proteolysis, reverse phase-high performance liquid chromatography, Edman degradative sequencing, amino acid analysis, and fast-atom bombardment mass-spectrometry was used to remove and identify the first 14 to 18 N-terminal amino acid residues of the large subunit of higher plant-type ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) from *Chlamydomonas reinhardtii*, *Marchantia polymorpha*, pea (*Pisum sativum*), tomato (*Lycopersicon esculentum*), potato (*Solanum tuberosum*), pepper (*Capsicum annuum*), soybean (*Glycine max*), petunia (*Petunia x hybrida*), cowpea (*Vigna sinensis*), and cucumber (*Cucumis sativus*) plants. The N-terminal tryptic peptide from acetylated Pro-3 to Lys-8 of the large subunit of Rubisco was identical in all species, but the amino acid sequence of the penultimate N-terminal tryptic peptide varied. Eight of the 10 species examined contained a trimethyllysyl residue at position 14 in the large subunit of Rubisco, whereas *Chlamydomonas* and *Marchantia* contained an unmodified lysyl residue at this position.

Rubisco (EC 4.1.1.39) is a large hexadecameric protein, with 8 large and 8 small subunits, that catalyzes the fixation of atmospheric CO_2 during photosynthesis (2). The large subunit is encoded by chloroplast DNA (8) and the small subunit by nuclear DNA (13). The small subunit is synthesized as a precursor with an N-terminal transit sequence that targets the polypeptide for import into the chloroplast. The transit sequence is proteolytically removed prior to assembly of small subunits with large subunits by chloroplast chaperonins (9). Synthesis of the large subunit in the chloroplast is also followed by posttranslational processing (16, 25). The N-terminal Met-1 and Ser-2 are removed and Pro-3 acetylated in Rubisco from spinach (*Spinacia oleracea* L.), wheat (*Triticum aestivum*), tobacco (*Nicotiana tabacum*), and muskmelon (*Cucumis melo*). Additionally, the tobacco and muskmelon large subunits contain a Me_3lys[3] residue at position 14, demonstrating additional posttranslational processing by a chloroplast localized *S*-adenosylmethionine:protein (lysine) $^{\epsilon}$*N*-methyltransferase (15).

In this report we describe the N-terminal sequence of the large subunit of Rubisco from several additional plant species. This structural information broadens the knowledge of both the species invariance of acetylated Pro-3, and the species variation of Me_3lys-14, as two posttranslational modifications that occur in the large subunit of higher plant-type Rubisco.

MATERIALS AND METHODS

Rubisco was purified from pea (*Pisum sativum*), tomato (*Lycopersicon esculentum*), potato (*Solanum tuberosum*), pepper (*Capsicum annuum*), soybean (*Glycine max*), petunia (*Petunia x hybrida*), cowpea (*Vigna sinensis*), cucumber (*Cucumis sativus*), *Marchantia polymorpha*, and *Chlamydomonas reinhardtii* by $(NH_4)_2SO_4$ precipitation followed by gel-permeation (Sepharose 4B) and anion-exchange (Whatman DE-52) chromatography (24). *C. reinhardtii* cells were ruptured prior to purification of Rubisco as described previously (14). The purified enzyme was activated with 20 mM $MgCl_2$ and 10 mM $NaHCO_3$ for 30 min and subsequently proteolyzed with 0.5% (w/w) trypsin for 2 to 3 h at 30°C. Rubisco activity was determined before and after proteolysis as previously described (16, 25). Peptides were purified (16, 25), and 1 to 10 nmol were sequenced with a gas-phase sequencer (Applied Biosystems 470A) or hydrolyzed with 6 $_N$ HCl for 22 h at 110°C before amino acid analysis with a Beckman system 6300 analyzer. In some cases, peptide identity was confirmed by FAB-MS and CAD using a JEOL HX110HF/HX110HF tandem mass spectrometer as previously described (16, 25). When possible, sequences were compared against known DNA sequences or previously published amino acid sequences. Protein was determined by a modified Lowry procedure (4). Discontinuous SDS-PAGE on 10 to 20% linear gradient gels was as described by Laemmli (20).

[1] This research was supported by U.S. Department of Agriculture/Competitive Research Grants Office Grant 89-37262-4482 and Hatch Project KY 00586 to R.L.H. and is published as Kentucky Agricultural Experiment Station Article 91-10-14-2.

[2] Recipient of U.S. Department of Education summer fellowship to encourage minority participation in graduate education. Present address: 777 Gorgas Street, Mobile, AL 36605.

[3] Abbreviations: Me_3lys, *N*-trimethyllysine; Ribulose-P_2, D-ribulose-1,5-bisphosphate; FAB-MS, fast-atom bombardment mass spectrometry; CAD, collisionally activated dissociation; u, atomic mass unit(s).

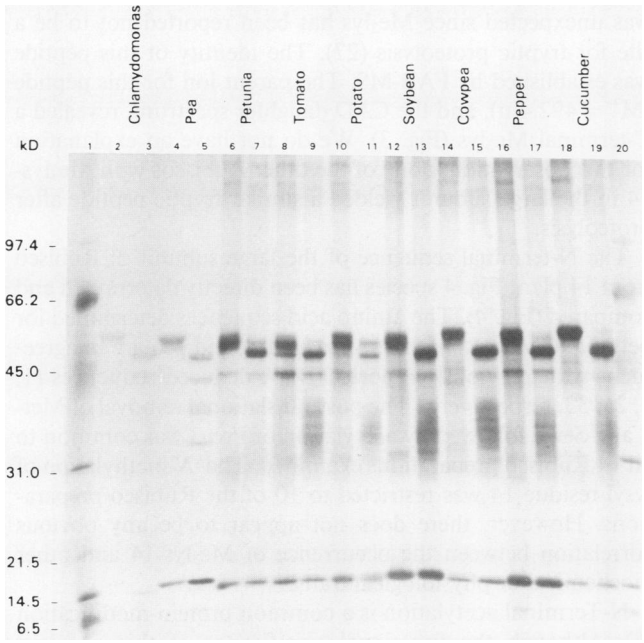

Figure 1. SDS-PAGE analysis of Rubisco from several plant species before and after tryptic proteolysis. Lanes 1 and 20 are molecular mass standards with kDa shown to the left. Lanes 2 to 19 contain 0.5 to 1.5 μg of Rubisco before (even lanes) and after (odd lanes) proteolysis with 0.5% trypsin. Proteins were visualized by silver-staining.

RESULTS AND DISCUSSION

Limited tryptic proteolysis decreased the molecular mass of the large subunit of Rubisco by approximately 1 to 2 kD as determined by SDS-PAGE (Fig. 1). With the conditions described, only the large subunit of Rubisco was proteolyzed, with the exception of Rubisco from *C. reinhardtii*, which also showed a molecular mass decrease in the small subunit as has been recently observed (5). Proteolysis also decreased the catalytic activity of Rubisco from all species by 60 to 70%, in agreement with reports that the penultimate N-terminal region of the large subunit of Rubisco is essential for maximum catalytic activity (12, 16, 18).

Table I. *Amino Acid Composition of the N-Terminal Tryptic Fragment from the Large Subunit of Cucumber Rubisco*

Approximately 3 nmol of HCl-hydrolyzed peptide were analyzed for amino acid composition. Molar ratio was calculated with proline equal to 1. All other residues were less than 0.2 nmol. Variation in the molar ratio of amino acid residues from other Rubisco preparations was ± 10%.

Residue	Amount	Molar Ratio
	nmol	
Glx	7.7	2.1
Lys	3.2	0.9
Thr	7.2	2.0
Pro	3.7	1.0

The N-terminal peptide from the large subunit of Rubisco in all species examined was blocked and could not be sequenced. An amino acid composition for the N-terminal peptide from cucumber Rubisco shown in Table I is essentially identical to the amino acid composition obtained for the large subunit N-terminal tryptic peptides from all other Rubisco preparations. The N-terminal tryptic peptides from the large subunit of *M. polymorpha*, *C. reinhardtii*, pepper, and pea Rubisco were also subjected to FAB-MS and CAD analysis, all of which yielded a Figure 2 molecular mass (M + H^+ = 745.2 u) and sequence (Fig. 2) identical with that previously reported for spinach, muskmelon, tobacco, and wheat (16). Based on HPLC retention times during purification, inability to sequence, and similar amino acid compositions, we conclude that the N-terminal sequence of the large subunit of Rubisco from the preparations examined in this study up to Lys-8 is acetyl-Pro-Gln-Thr-Glu-Thr-Lys-COOH.

The penultimate N-terminal tryptic peptides from the large subunit of Rubisco vary in length depending on the presence or absence of a Me_3lys residue at position 14 (16), which has been reported as not susceptible to tryptic proteolysis (16, 27). Thus, in species where the large subunit of Rubisco contains Me_3lys-14, the penultimate N-terminal tryptic peptides end with Lys-18 as the C-terminal residue.

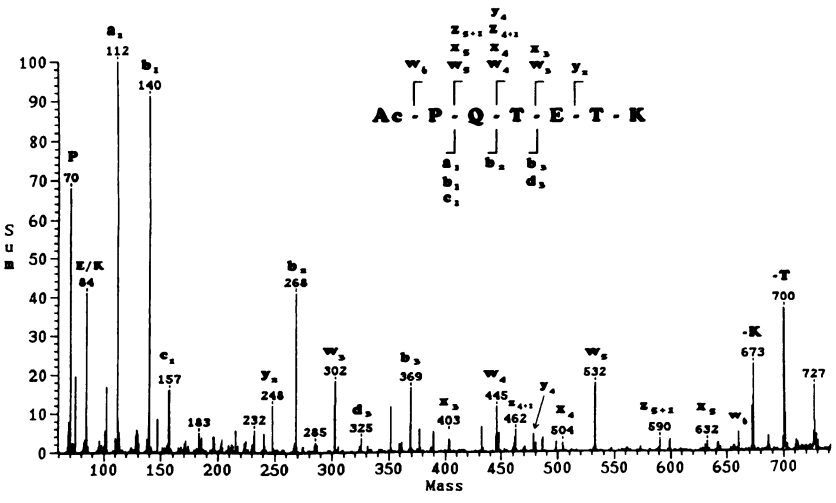

Figure 2. CAD daughter spectrum of the N-terminal tryptic peptide from the large subunit of pea Rubisco (M + H^+ = 745.2 u). Peaks are labeled according to Johnson *et al.* (17). The CAD daughter spectra were identical for the N-terminal tryptic peptides from the large subunit of *M. polymorpha* and *C. reinhardtii* Rubisco.

Table II. *Amino Acid Composition of the Penultimate N-Terminal Tryptic Fragment from the Large Subunit of Cucumber Rubisco*

Approximately 3 nmol of HCl-hydrolyzed peptide were analyzed by amino acid composition. Molar ratio was calculated with glycine equal to 2. All other residues were less than 0.2 nmol. Variation in the molar ratio of amino acid residues from other Rubisco preparations was ± 8%.

Residue	Amount	Molar Ratio
	nmol	
Ala	7.2	1.9
Gly	7.5	2.0
Lys	3.0	0.8
Me_3lys	2.8	0.8
Phe	4.0	1.1
Ser	3.9	1.0
Val	7.8	2.1

All of the penultimate N-terminal tryptic peptides from the large subunit of Rubisco could be sequenced. However, standard sequencing by Edman degradation does not result in an identifiable amino acid residue at position 14 if this residue is Me_3lys. Therefore, amino acid analysis following HCl hydrolysis was used to confirm the presence of Me_3lys-14.

A representative amino acid composition of the penultimate N-terminal fragment derived from the large subunit of cucumber Rubisco is shown in Table II. All of the Rubisco preparations examined except *M. polymorpha* and *C. reinhardtii* contained a Me_3lys residue at position 14 in the large subunit. In the large subunit of pea Rubisco, position 10 is a lysyl residue (33). Therefore, the tryptic peptide bearing Me_3lys-14 contains Val-11 as the N-terminal residue. The dipeptide Ala-9 to Lys-10 was not recovered after limited tryptic proteolysis of pea Rubisco.

Although limited tryptic proteolysis of pea Rubisco yielded a large subunit penultimate N-terminal peptide from Val-11 to Lys-18 that contained a Me_3lys residue at position 14, another penultimate N-terminal peptide was recovered in equal proportions that corresponded to Val-11 to Lys-14, but with a C-terminal lysyl residue that was modified in a manner indicative of methylation based on amino acid analysis. This was unexpected since Me_3lys has been reported not to be a site for tryptic proteolysis (27). The identity of this peptide was established by FAB-MS. The parent ion for this peptide (M^+ = 492.4 u), and the CAD daughter spectrum, revealed a C-terminal Me_3lys (Fig. 3). We do not have an explanation for this observation; none of the other Rubiscos with Me_3lys-14 in the large subunit yielded a similar tryptic peptide after proteolysis.

The N-terminal sequence of the large subunit of Rubisco from 14 plant Fig. 4 species has been directly determined and compared (Fig. 4). The amino acid sequences determined for petunia, *Chlamydomonas*, *Marchantia*, and pea are in agreement with previously reported DNA-deduced sequences (1, 7, 26, 33, respectively). The posttranslational removal of Met-1 and Ser-2 followed by acetylation of Pro-3 was common to all 14 Rubisco preparations examined. The '*N*-methylation of lysyl residue 14 was restricted to 10 of the Rubisco preparations. However, there does not appear to be any obvious correlation between the occurrence of Me_3lys-14 and other biochemical or physiological traits.

N-Terminal acetylation is a common protein modification (6). Although the functional significance of this covalent modification is not firmly established, several lines of evidence suggest a possible role in protein stability (3). In pea chloroplasts, imino-peptidases have been reported that could possibly act on the large subunit of Rubisco were it not acetylated (21). In specific cases, such as for α-melanocyte-stimulating hormone (31) and β-endorphin (30), N-terminal acetylation has direct effects on biological activity. However, to date, there is no evidence for N-terminal acetylation regulating the activity or stability of Rubisco.

The role of '*N*-methylation of lysyl residues in proteins is also not clear (19, 22, 23). The methylation of Lys-115 in calmodulin has dramatic effects on its ability to activate NAD-kinase (28), but no effect on its ability to activate phosphodiesterase (29). Some studies suggest that the methylation of Lys-115 in calmodulin serves a protective role against ubiquitination and subsequent proteolytic degradation (10, 11). The occurrence of ubiquitin in *C. reinhardtii* chloroplasts has been demonstrated by immunoelectron microscopy (32). How the posttranslational acetylation and methylation of the

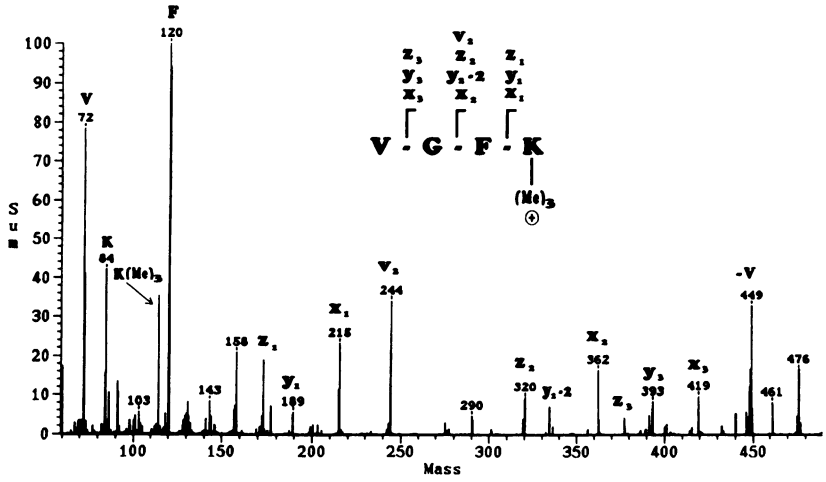

Figure 3. CAD daughter spectrum of the penultimate N-terminal tryptic peptide from the large subunit of pea Rubisco (M^+ = 492.4 u). The peaks are labeled according to Johnson *et al.* (17) except all peaks are 1 u lower in mass because the ions are not protonated but rather contain a fixed charge.

	3	8	14	18
Spinach	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-Lys			
Wheat	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Gly-Val-Gly-Phe-Lys			
Tobacco	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Muskmelon	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Cucumber	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Soybean	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Pea	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Lys-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Cowpea	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Tomato	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Potato	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Petunia	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Pepper	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Ser-Val-Gly-Phe-***Me$_3$lys***-Ala-Gly-Val-Lys			
Chlamydomonas	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Gly-Ala-Gly-Phe-Lys			
Marchantia	Acetyl-Pro-Gln-Thr-Glu-Thr-Lys-Ala-Gly-Val-Gly-Phe-Lys			

Figure 4. Summary of the amino acid sequences and posttranslational modifications in the N terminus of the large subunit of Rubisco from several plant species. Data for spinach, wheat, tobacco, and muskmelon are from Houtz *et al.* (16).

large subunit of higher plant-type Rubisco relates to functional aspects of the enzyme remains to be shown. However, those preparations of Rubisco where the large subunit is not modified at lysyl residue 14 are *in vitro* substrates for the recently isolated tobacco *S*-adenosylmethionine:Rubisco large subunit (lysine) '*N*-methyltransferase (15). Thus, *in vitro* methylation of Lys-14 in the large subunit of Rubisco and subsequent comparative enzymological studies may enable identification of the functional aspects of this covalent modification.

ACKNOWLEDGMENT

We thank Ms. Karen Goodlet for manuscript preparation.

LITERATURE CITED

1. **Aldrich J, Cherney B, Merlin E, Palmer J** (1986) Sequence of the *rbc* L gene for the large subunit of ribulose bisphosphate carboxylase-oxygenase from petunia. Nucleic Acids Res **14:** 9534
2. **Andrews JT, Lorimer GH** (1987) Rubisco: structure, mechanisms, and prospects for improvement. *In* MD Hatch, NK Boardman, eds, The Biochemistry of Plants, Vol 10. Academic Press, New York, pp 131–218
3. **Arfin SM, Bradshaw RA** (1988) Cotranslational processing and protein turnover in eukaryotic cells. Biochemistry **27:** 7979–7984
4. **Bensadoun A, Weinstein D** (1976) Assay of proteins in the presence of interfering material. Anal Biochem **70:** 241–250
5. **Chen Z, Spreitzer RJ** (1991) Proteolysis and transition-state-analogue binding of mutant forms of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Chlamydomonas reinhardtii*. Planta **183:** 597–603
6. **Driessen HPC, de Jong WW, Tesser GI, Bloemendal H** (1981) The mechanism of N-terminal acetylation of proteins. Crit Rev Biochem **18:** 281–325
7. **Dron M, Rahire M, Rochaix J-D** (1982) Sequence of the chloroplast DNA region of *Chlamydomonas reinhardii* containing the gene of the large subunit of ribulose bisphosphate carboxylase and parts of its flanking genes. J Mol Biol **162:** 775–793
8. **Ellis RJ** (1981) Chloroplast proteins: synthesis, transport, and assembly. Annu Rev Plant Physiol **32:** 111–137
9. **Goloubinoff P, Christeller JT, Gatenby AA, Lorimer GH** (1989) Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins ang Mg-ATP. Nature **342:** 884–889
10. **Gregori L, Marriott D, Putkey JA, Means AR, Chau V** (1987) Bacterially synthesized vertebrate calmodulin is a specific substrate for ubiquitination. J Biol Chem **262:** 2562–2567
11. **Gregori L, Marriott D, West CM, Chau V** (1985) Specific recognition of calmodulin from *Dictyostelium discoideum* by the ATP, ubiquitin-dependent degradation pathway. J Biol Chem **260:** 5232–5235
12. **Gutteridge S, Millard BN, Parry MAJ** (1986) Inactivation of ribulose-bisphosphate carboxylase by limited proteolysis. FEBS Lett **196:** 263–268
13. **Highfield PE, Ellis RJ** (1978) Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. Nature **271:** 420–424
14. **Houtz, RL, Ries SK, Tolbert NE** (1985) Effect of triacontanol on *Chlamydomonas*. II. Specific activity of ribulose-bisphosphate carboxylase/oxygenase, ribulose-bisphosphate concentration, and characteristics of photorespiration. Plant Physiol **79:** 365–370
15. **Houtz RL, Royer M, Salvucci ME** (1991) Partial purification and characterization of ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit '*N*-methyltransferase. Plant Physiol **97:** 913–920
16. **Houtz RL, Stults JT, Mulligan RM, Tolbert NE** (1989) Post-translational modifications in the large subunit of ribulose bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA **86:** 1855–1859
17. **Johnson RS, Martin SA, Biemann K, Stults JT, Watson JT** (1987) Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: Differentiation of leucine and isoleucine. Anal Chem **59:** 2621–2625
18. **Kettleborough CA, Parry MAJ, Burton S, Gutteridge S, Keys AJ, Phillips AL** (1987) The role of the N-terminus of the large subunit of ribulose-bisphosphate carboxylase investigated by

construction and expression of chimaeric genes. Eur J Biochem **170:** 335–342

19. **Klee CG, Vanaman TC** (1982) Calmodulin. Adv Protein Chem **35:** 213–303
20. **Laemmli UK** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680–685
21. **Liu X, Jagendorf AT** (1986) Neutral peptidases in the stroma of pea chloroplasts. Plant Physiol **81:** 603–608
22. **Lukas TJ, Wiggins M, Watterson DM** (1985) Amino acid sequence of a novel calmodulin from the unicellular alga *Chlamydomonas.* Plant Physiol **78:** 447–483
23. **Marshak DR, Clarke M, Roberts DM, Watterson DM** (1984) Structural and functional properties of calmodulin from the eukaryotic microorganism *Dictyostelium discoideum.* Biochemistry **23:** 2891–2899
24. **McCurry SD, Gee R, Tolbert NE** (1982) Ribulose-1,5-bisphosphate carboxylase/oxygenase from spinach, tomato or tobacco leaves. Methods Enzymol **90:** 515–521
25. **Mulligan RM, Houtz RL, Tolbert NE** (1988) Reaction-intermediate analogue binding by ribulose bisphosphate carboxylase/oxygenase causes specific changes in proteolyic sensitivity: the amino-terminal residue of the large subunit is acetylated proline. Proc Natl Acad Sci USA **85:** 1513–1517
26. **Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S-I, Inokuchi H, Ozeki H** (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature **322:** 572–574
27. **Paik WK, Kim S** (1975) Protein methylation: chemical, enzymological and biological significance. Adv Enzymol Relat Areas Mol Biol **42:** 227–256
28. **Roberts DM, Rowe PM, Siegel FL, Lukas TJ, Watterson DM** (1986) Trimethyllsine and protein function. J Biol Chem **261:** 1491–1494
29. **Rowe PM, Wright LS, Siegel FL** (1986) Calmodulin *N*-methyltransferase. J Biol Chem **261:** 7060–7069
30. **Smyth DG, Massey DE, Zakarian S, Finnie MDA** (1979) Endorphins are stored in biologically active and inactive forms: isolation of an $^{\alpha}$*N*-acetylpeptide. Nature **279:** 252–254
31. **Waller JP, Dixon HBF** (1960) Selective acetylation of the terminal amino group of corticotrophin. Biochem J **75:** 320–328
32. **Wettern M, Parag HA, Pollmann L, Ohad I, Kulka RG** (1990) Ubiquitin in *Chlamydomonus reinhardtii.* Distribution in the cell and effect of heat shock and photoinhitibition on its conjugate pattern. Eur J Biochem **191:** 571–576
33. **Zurawski G, Whitfeld PR, Bottomley W** (1986) Sequence of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from pea chloroplasts. Nucleic Acids Res **14:** 3975–3992