Dresden, M. H., and Gross, J. (1970), *Develop. Biol. 22*, 129.
Dunker, A. K., and Rueckert, R. R. (1969), *J. Biol. Chem. 244*, 5074.
Eisen, A. Z., Bauer, E. A., and Jeffrey, J. J. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 248.
Eisen, A. Z., Bloch, K. J., and Sakai, T. (1970), *J. Lab. Clin. Med. 75*, 258.
Eisen, A. Z., and Gross, J. (1965), *Develop. Biol. 12*, 408.
Eisen, A. Z., Jeffrey, J. J., and Gross, J. (1968), *Biochim. Biophys. Acta 151*, 637.
Evanson, J. M., Jeffrey, J. J., and Krane, S. M. (1968), *J. Clin. Invest. 47*, 2639.
Fantes, K. H., and Furminger, I. G. S. (1967), *Nature (London) 215*, 750.
Fullmer, H. M., and Gibson, W. (1966), *Nature (London) 209*, 728.
Fullmer, H. M., and Lazarus, G. S. (1969), *J. Histochem. Cytochem. 17*, 793.
Grillo, H. C., and Gross, J. (1967), *Develop. Biol. 15*, 300.
Grillo, H. C., and Lapiere, C. M., Dresden, H. M., and Gross, J. (1968), *Develop. Biol. 17*, 571.
Gross, J., and Lapiere, C. M. (1962), *Proc. Nat. Acad. Sci. U. S. 48*, 1014.
Gross, J., and Nagai, Y. (1965), *Proc. Nat. Acad. Sci. U. S. 54*, 1197.
Harper, E. (1970), *in* The Chemistry and Molecular Biology of the Intercellular Matrix, North Atlantic Treaty Organization Advanced Study Institute, Santa Margherita, Italy, 1969, Vol. 3, Balazs, E. A., Ed., London, Academic Press, p 1653.
Harper, E., Bloch, K. J., and Gross, J. (1970), *Amer. Rheum. Assn. Meeting, No. 28*, 33.
Harper, E., and Gross, J. (1970), *Biochim. Biophys. Acta 198*, 286.
Hartley, B. S. (1964), *Nature (London) 201*, 1284.
Jeffrey, J. J., and Gross, J. (1967), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 26*, 670.
Jeffrey, J. J., and Gross, J. (1970), *Biochemistry 9*, 268.
Kabat, E. A., and Mayer, M. M. (1961), Experimental Immunochemistry, Springfield, Ill., C. C. Thomas, p 153.
Lapiere, C. M., and Gross, J. (1963), *in* Mechanisms of Hard Tissue Destruction, Sognnaes, R. F., Ed., New York, N. Y., American Association for the Advancement of Science, Publication No. 75.
Nagai, Y., Gross, J., and Piez, K. A. (1964), *Ann. N. Y. Acad. Sci. 121*, 494.
Nagai, Y., Lapiere, C. M., and Gross, J. (1966), *Biochemistry 5*, 3123.
Ornstein, L. (1964), *Ann. N. Y. Acad. Sci. 121*, 321.
Raymond, S. (1962), *Clin. Chem. 8*, 455.
Sakai, T., and Gross, J. (1967), *Biochemistry 6*, 518.
Schwabe, C. (1969), *Biochemistry 8*, 783.
Shimizu, M., Glimcher, M. J., Travis, D., and Goldhaber, P. (1969), *Proc. Soc. Exp. Biol. Med. 130*, 1175.
Slansky, H. H., Gnadinger, M. C., Itoi, M., and Dohlman, C. H. (1969), *Arch. Ophthalmol. 82*, 108.

Amino- and Carboxyl-Terminal Amino Acid Sequences of the *Peptostreptococcus elsdenii* and *Clostridium pasteurianum* Flavodoxins*

Masaru Tanaka, Mitsuru Haniu, Gary Matsueda, Kerry T. Yasunobu,‡ Steven Mayhew, and Vincent Massey

ABSTRACT: The NH_2-terminal sequences of 2 flavodoxins were examined in the protein sequencer. The first 51 residues of the *Clostridium pasteurianum* and the first 41 residues of the *Peptostreptococcus elsdenii* flavodoxin were determined. The carboxyl-terminal sequences of the two flavodoxins were found to be Leu-Val-COOH and Lys-Ala-COOH for the *C. pasteurianum* and *P. elsdenii* proteins, respectively. The cysteine residue which is involved in binding the single FMN is not in the portion of the molecule which was sequenced. A number of constant residues were observed for the flavodoxins but the most interesting portion was a 12-residue section from residues 6 to 17 in the *P. elsdenii* protein and from residues 7 to 18 in the *C. pasteurianum* protein which probably has an important function. The flavodoxins have a relatively low molecular weight of 14,000 and are ideal proteins for studying the structure–function relationships and should serve to act as a model for other flavoproteins whose structures are as yet unknown. The fact that the X-ray structural studies of crystalline flavodoxin are in progress simply added further impetus for investigating the primary structure of flavodoxin. In addition, the capabilities of the sequencer are of general interest to biochemists at this time and the sequencer results obtained in our laboratory with different proteins are presented.

Flavodoxin, a FMN protein, was initially isolated in crystalline form from *Clostridium pasteurianum* (Knight *et al.*, 1966; Knight and Hardy, 1966). These investigators demonstrated that flavodoxin was produced by *C. pasteurianum* when the iron content in the growth medium was low (Knight *et al.*, 1966; Knight and Hardy, 1966, 1967). Furthermore,

* From the Department of Biochemistry-Biophysics, University of Hawaii, Honolulu, Hawaii (M. T., M. H., G. M., and K. T. Y.), and the Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan (S. M. and V. M.). *Received March 24, 1971.*

† Supported in part by grants from the National Institutes of Health No. GM 16784-01, GM 16228-02, and GM 11106, and the National Science Foundation No. GB 6608.

‡ To whom to address correspondence.

it was shown that flavodoxin would replace ferredoxin in certain reactions normally involving ferredoxin. Since the original isolation of flavodoxin from *C. pasteurianum*, the protein has been isolated in pure form from *Desulfovibrio gigas* (Dubourdieu and Le Gall, 1970), *Peptostreptococcus elsdenii* (Mayhew *et al.*, 1969; Mayhew and Massey, 1969), and *Desulfovibrio vulgaris* (Dubourdieu and Le Gall, 1970). In the above-mentioned reports, some of the physicochemical properties of the proteins were presented (Knight and Hardy, 1966). Also, X-ray structural studies of flavodoxin are in progress (Ludwig *et al.*, 1969).

Methods and Materials

Growth of Bacteria and Preparation of Flavodoxin. The conditions used for the growth and the preparation of crude extracts of *P. elsdenii* LC 1 (Mayhew *et al.*, 1969) and *C. pasteurianum* have been described previously (Knight and Hardy, 1966). The method for purifying the flavodoxin from the crude extracts has been published (Knight and Hardy, 1966; Mayhew *et al.*, 1969). The crystalline flavodoxins isolated met the usual criteria of purity (Knight and Hardy, 1966).

Preparation of the Derivative. The flavodoxin was converted into the carboxymethylcysteine derivative by the procedure of Crestfield *et al.* (1963) after first removing the FMN from the protein by either the trichloroacetic acid method or the KBr method of Mayhew *et al.* (1969).

NH_2-Terminal Sequence Determination. The NH_2-terminal sequence of the protein was determined in the Beckman protein sequencer which is a commercial adaptation of the instrument designed by Edman and Begg (1967). The procedure for converting the thiazolinone derivatives of the amino acid to the PTH[1]-amino acid has been published by Edman and Begg (1967). The procedure for operating the Beckman protein sequencer is given in the instrument manual.

The PTH-amino acids soluble in ethyl acetate were identified by gas–liquid chromatography in the Beckman Model GC-45 gas chromatograph. The two columns used and the experimental conditions necessary to separate the PTH-amino acids or the silyl derivatives of the PTH-amino acids have been described by Pisano and Bronzert (1969, 1970). The 10% DC-560 on Chromosorb W separated the more volatile group I PTH-amino acids and the silylated derivatives of PTH-glutamic acid and PTH-aspartic acid. The $^2/_3$ CFC on Chromosorb W separated the group II PTH-amino acids. PTH-arginine and PTH-histidine were not extracted from the aqueous phase by ethyl acetate unless the pH was raised above 8 by the addition of 0.2 ml of 1 M Na_2HPO_4 (Edman and Begg, 1967). The yield of these PTH-amino acids was quantitated after acid hydrolysis of the extracted sample by 6 N HCl for 18 hr at 150° (Van Orden and Carpenter, 1964). The sum of the ornithine and arginine content was used to calculate the total arginine content.

Further confirmation of the PTH-amino acid was obtained by thin-layer chromatography in which the solvent systems developed by Edman and Begg (1967) were used. The PTH-amino acids were visible because of the fluorescent indicator added to the silica plates (Camag, Inc., Milwaukee, Wis.). PTH-amino acids stable to acid hydrolysis (PTH-serine, PTH-threonine, PTH-tryptophan, and PTH-carboxymethylcysteine were unstable) were hydrolyzed in 6 N HCl at 150° for 18 hr (Van Orden and Carpenter, 1964) and the free amino acids were identified by the Beckman Model 120C automatic amino acid analyzer (Spackman *et al.*, 1967).

COOH-Terminal Amino Acid Analyses. The COOH-terminal amino acid of flavodoxin was identified by hydrazinolysis (Bradbury, 1958) as well as reaction with carboxypeptidase A and then carboxypeptidase B (Ambler, 1967). For the carboxypeptidase experiments, the carboxymethylcysteine derivatives of the flavodoxins in separate experiments were incubated with the peptidase (40:1, w/w) at pH 8.0 for the various time intervals at 40°. In order to terminate the reaction, the equilibrating buffer for the amino acid analyzer was added and the sample was then applied directly on the automatic amino acid analyzer.

Results

NH_2-Terminal Sequence. Table I summarizes the various PTH-amino acids which were released during the various steps of the Edman degradation of the *P. elsdenii* flavodoxin. It was possible to identify the first 41 sequences from the NH_2-terminal end. The PTH-amino acids were confirmed not only by gas chromatography but also by thin-layer chromatography and by hydrolysis of the PTH-amino acids to the free amino acids at each step of Edman degradation. PTH-serine, PTH-threonine, PTH-tryptophan, and PTH-carboxymethylcysteine did not survive the acid treatment at 150° and therefore, could not be identified by amino acid analyses of the hydrolysate of the PTH-amino acids. However, the major component could be identified through step 41 since the main contaminant was the PTH-amino acid from the previous step. From steps 42 to 45, it was possible to identify the major PTH-amino acid but we felt it would be advisable to confirm the sequence on the peptide fragment from this portion of the molecule. Furthermore, the sequence determination was repeated three times on the particular protein to obtain verification of the data.

Table II summarizes the results of the analysis of the *C. pasteurianum* carboxymethylcysteine derivative of flavodoxin in the protein sequencer. It was possible in this case to determine the sequence of 51 residues from the NH_2-terminal end of the molecule. This particular flavodoxin appears to be especially suitable for the instrument and should be considered as an alternative to myoglobin for making test runs on the instrument. The sequence analysis was repeated three times in order to ascertain that the proposed sequence is correct. The presence of PTH-serine, PTH-threonine, and PTH-carboxymethylcysteine which are always partially destroyed, and the fact that multiple products appeared in low yields at the later stages of Edman degradation made it possible to continue the sequence analysis. Noteworthy, again was the fact that there was a breakdown in the chemistry of Edman degradation and that at some of the steps, the PTH-amino acid from the previous step appeared. The gas chromatograms obtained from the 41st step of sequence analysis of the *P. elsdenii* and the 51st step of sequence analysis of the *C. pasteurianum* flavodoxins are shown in Figure 1 to illustrate the type of result obtained during gas chromatography of the PTH-amino acids.

COOH-Terminal Amino Acids Analysis. Hydrazinolysis of the deflavinized carboxymethylcysteine derivative of *P. elsdenii* yielded alanine in about 50% yield. Similar studies with the *C. pasteurianum* carboxymethylcysteine derivative yielded only valine in a low yield. In order to obtain additional data on the carboxyl-terminal se-

[1] Abbreviation used is: PTH, phenylhydantion.

TABLE I: NH_2-Terminal Sequence of the *P. elsdenii* Flavodoxin.[a]

Step No.	Major PTH-Amino Acid	Yield (%)	Minor PTH-Amino Acid[b] (% Yield)
1	Met	61[c]	
2	Val	88[c]	
3	Glu	81[c]	
4	Ile	70[c]	
5	Val	88[c]	
6	Tyr	57[c]	
7	Trp	70[d]	
8	Ser	60[d]	
9	Gly	63[c]	
10	Thr	55[d]	
11	Gly	60[c]	
12	Asn	52[c]	
13	Thr	40[d]	Asn (11)[c]
14	Glu	45[c]	
15	Ala	45[c]	Glu (13)[c]
16	Met	33[c]	Ala (11)[c]
17	Ala	45[c]	Met (8)[c]
18	Asn	44[c]	Ala (9)[c]
19	Glu	42[c]	Asn (12)[c]
20	Ile	29[c]	Glu (10)[c]
21	Glu	35[c]	Ile (7)[c]
22	Ala	33[c]	Glu (11)[c]
23	Ala	35[c]	
24	Val	26[c]	Ala (9)[c]
25	Lys	19[c]	Val (13)[c]
26	Ala	32[a,c]	Lys (8)[c]
27	Ala	38[a,c]	
28	Gly	27[a,c]	Ala (9)[c]
29	Ala	29[a,c]	Gly (12)[c]
30	Asp	23[a,c]	Ala (8)[c]
31	Val	21[a,c]	Asp (9)[c]
32	Glu	17[a,c]	Val (8)[c]
33	Ser	10[a,d]	Glu (6)[c]
34	Val	13[a,c]	
35	Arg	1[a,e]	Val (10)[c]
36	Phe	6[a,c]	Arg (1)[e]
37	Glu	6[a,c]	Phe (4)[c]
38	Asp	6[a,c]	Glu (4)[c]
39	Thr	4[a,d]	Asp (3)[c]
40	Asn	7[a,c]	
41	Val	3[a,c]	Asn (2)[c]

[a] Sequence of residues 26–41 have been confirmed on peptide fragments and will be discussed elsewhere. [b] Minor components present in less than 20% of the major component are not reported in the table. [c] Yields quantitated by gas chromatography and identification confirmed by thin-layer chromatography and by hydrolysis to the free amino acid. [d] Quantitated by gas chromatography of the PTH-amino acid and identification confirmed by thin-layer chromatography. [e] Determined by hydrolysis to the free amino acid only.

TABLE II: NH_2-Terminal Sequence of the *C. pasteurianum* Flavodoxin.

Step No.	Major PTH-Amino Acid	Yield (%)	Minor PTH-Amino Acid[a] (% Yield)
1	Met	50[b]	
2	Lys	27[b]	
3	Val	70[b]	
4	Asn	63[b]	Val (18)[b]
5	Ile	50[b]	Asn (15)[b]
6	Ile	63[b]	
7	Tyr	37[b]	Ile (13)[b]
8	Trp	50[c]	Tyr (12)[b]
9	Ser	40[c]	Trp (10)[c]
10	Gly	36[b]	
11	Thr	30[c]	Gly (12)[b]
12	Gly	26[b]	
13	Asn	28[b]	Gly (13)[b]
14	Thr	25[c]	Asn (8)[b]
15	Glu	28[b]	
16	Ala	25[b]	Glu (10)[b]
17	Met	18[b]	Ala (7)[b]
18	Ala	19[b]	Met (5)[b]
19	Lys	8[b]	Ala (3)[b]
20	Leu	14[b]	Lys (3)[b]
21	Ile	14[b]	Leu (4)[b]
22	Ala	18[b]	Ile (6)[b]
23	Glu	18[b]	Ala (6)[b]
24	Gly	19[b]	Glu (6)[b]
25	Ala	20[b]	Gly (7)[b]
26	Gln	17[b]	Ala (6)[b]
27	Glu	20[b]	
28	Lys	12[b]	Glu (7)[b]
29	Gly	19[b]	Lys (4)[b]
30	Ala	13[b]	Gly (6)[b]
31	Glu	14[b]	Ala (5)[b]
32	Val	11[b]	Glu (5)[b]
33	Lys	9[b]	Val (4)[b]
34	Leu	7[b]	Lys (3)[b]
35	Leu	8[b]	
36	Asn	10[b]	Leu (3)[b]
37	Val	10[b]	Asn (4)[b]
38	Ser	8[c]	Val (4)[b]
39	Asp	7[b]	
40	Ala	5[b]	Asp (3)[b]
41	Lys	4[b]	Ala (3)[b]
42	Glu	5[b]	Lys (1)[b]
43	Asp	4[b]	Glu (2)[b]
44	Asp	4[b]	
45	Val	3[b]	Asp (2)[b]
46	Lys	1[b]	Val (2)[b]
47	Glu	2[b]	
48	Ala	2[b]	Glu (1)[b]
49	Asp	2[b]	Ala (1)[b]
50	Val	2[b]	Asp (1)[b]
51	Val	2[b]	

[a] See footnote *b* in Table I. [b] See footnote *c* in Table I. [c] See footnote *d* in Table I.

quences of the flavodoxins, additional experiments with carboxypeptidase were carried out. Carboxypeptidase A digestion of the carboxymethylated *P. elsdenii* protein released only one mole of alanine per mole of protein but on the further addition of carboxypeptidase B, additional amino acids were liberated which were chiefly lysine and alanine (Figure 2). On the other hand, carboxypeptidase A released only valine and leucine from the carboxymethyl-

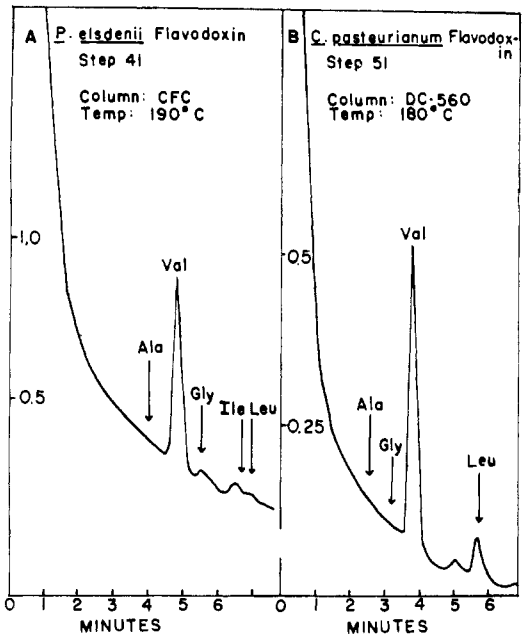

FIGURE 1: Gas chromatograms of the last step of Edman degradation of the flavodoxins from (A) *P. elsdenii* and (B) *C. pasturianum*. Detector response is plotted against retention time.

cysteine derivative of the *C. pasteurianum* flavodoxin. From the kinetics of the reaction (Figure 3) the carboxyl-terminal sequence was -Leu-Val-COOH.

Discussion

The amino acid composition of the *P. elsdenii* (Mayhew *et al.*, 1969) and *C. pasteurianum* (Knight and Hardy, 1967) flavodoxins have already been determined and are summarized in Table III along with the data on the *D. gigas* and *D. vulgaris* (Dubourdiew and Le Gall, 1970) proteins. Although the flavodoxins have amino acid compositions which are quite similar to one another, there are some marked differences. For example, the lysine content varies from 4 to 10. The *D. vulgaris* flavodoxin is the only flavodoxin isolated thus far that contains histidine. In addition, the cysteine content varies from 1 to 5 and it appears that at least

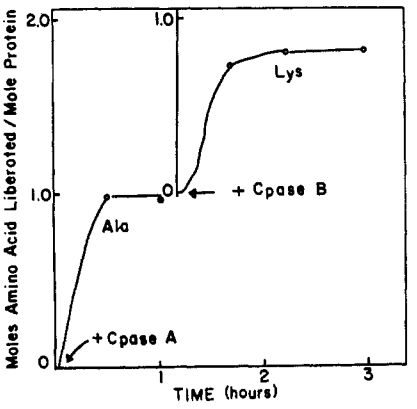

FIGURE 2: Kinetics of the release of amino acids from the *P. elsdenii* flavodoxin. The reaction was started by the addition of carboxypeptidase A and at 1 hr, carboxypeptidase B was added. The enzyme: substrate ratio and other experimental conditions are described in the Experimental Section.

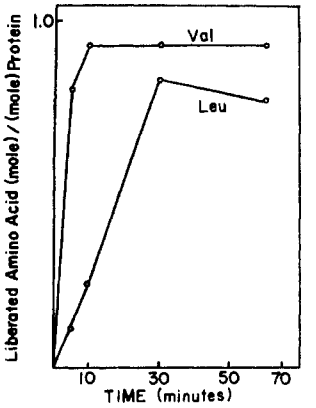

FIGURE 3: Kinetics of the release of amino acids from the *C. pasteurianum* carboxymethylcysteine derivative flavodoxin. The conditions were identical with those reported in the legend to Figure 2.

one of the cysteine residues is involved in the binding of the FMN to the protein (Mayhew *et al.*, 1969). In addition the methionine varies from 0 to 5 and therefore methionine does not appear to be important for activity. In general, the flavodoxins from the species isolated thus far show that there may be considerable differences in the primary structures of the flavodoxins from different microorganisms.

With the advent of the development of a commercial protein sequencer and the purchase of the instrument by our laboratory, a logical step in determining the primary structure was to investigate the NH_2-terminal sequence with this instrument. In the present investigation of the *C. pasteurianum* and *P. elsdenii* flavodoxins were run in the instrument. There is a big advantage in determining

TABLE III: Amino Acid Composition of Various Flavodoxins.[a]

Amino Acid	*D. vulgaris*[b]	*D. gigas*[b]	*C. pasteurianum*[c]	*P. elsdenii*[d]
Lysine	4	8	10	8
Histidine	1	0	0	0
Arginine	6	3	2	2
Aspartic acid	20	17	18	16
Threonine	6	9	4	8
Serine	7	8	14	7
Glutamic acid	17	18	19	18
Proline	4	6	4	5
Glycine	19	14–15	17	14
Alanine	18	15	14	18
Cysteine	5	5	1	2
Valine	10	16	15	13
Methionine	0	2	4	5
Isoleucine	9	5	5	5
Leucine	13	14	12	7
Tyrosine	5	5	2	2
Phenylalanine	6	3	3	4
Tryptophan	1–2	1	4	4
Total	151–152	149–150	148	138

[a] Expressed as moles of amino acid per mole of protein. [b] Dubourdieu and Le Gall (1970). [c] Knight and Hardy (1967). [d] Mayhew *et al.* (1969).

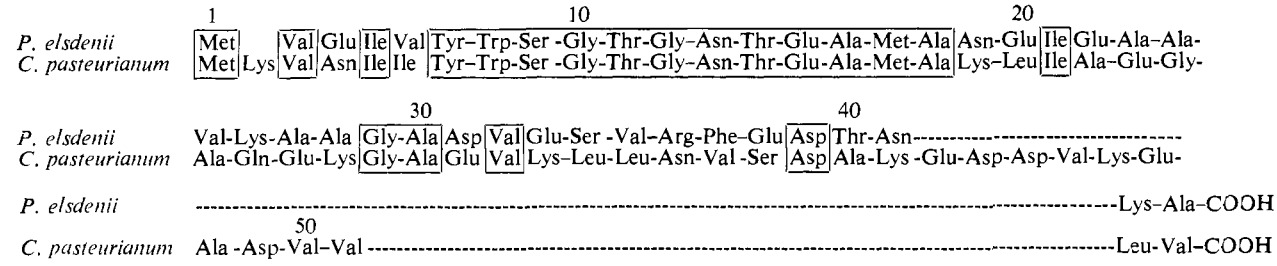

FIGURE 4: A comparison of the partial sequence of the *P. elsdenii* and *C. pasteurianum* flavodoxins. Identical residues are blocked off.

the *C. pasteurianum* flavodoxin sequence since it contains only one cysteine residue which is involved in binding the FMN. However, the organism produces much less flavodoxin than *P. elsdenii* and considerable time will be required to stockpile sufficient flavodoxin from this organism. We chose to investigate several flavodoxins for a number of reasons. For example, in order to gain confidence that there has been no error in the proposed sequence of flavodoxin, at least two flavodoxins from different organisms should be determined. Furthermore, our laboratory is investigating the sequences of not only flavodoxins from these anaerobic organisms but also the ferredoxins and rubredoxins which are also present in order to determine the relatedness of these bacteria from an evolutionary standpoint. Also, the constant residues will provide clues concerning the essential portions of the flavodoxins molecules.

The first 41 residues of the *P. elsdenii* flavodoxin were determined and the first 51 residues of the *C. pasteurianum* flavodoxin from the NH_2-terminal end of the protein were determined. These results are quite excellent as far as the capability of the instrument at the present time is concerned since we have been averaging about 30 residues for a variety of proteins. However, the results are not as spectacular as the report of Brewer and Ronan (1970) on the bovine parathyroid hormone in which case the first 66 residues from the NH_2 terminus were determined. In the case of the flavodoxin analyses in the sequenator, the reasons for the failure to go beyond the indicated residue was not due to the washing out of the protein from the cup by the organic solvents but was due to a failure in the chemistry of the reactions. The number of sequences of different proteins determined with our instrument are summarized in Table IV. At the present time, there is insufficient information to discuss intelligently the sequences which are difficult for the instrument to handle.

Hydrazinolysis of the *P. elsdenii* and the *C. pasteurianum* flavodoxins indicated that the COOH terminus was alanine and valine, respectively. The sequential hydrolysis of the *P. elsdenii* flavodoxin first by carboxypeptidase A and then by carboxypeptidase B indicated that lysine was the penultimate amino acid. Carboxypeptidase A alone liberated leucine and valine from the *C. pasteurianum* flavodoxin and in combination with the hydrazinolysis experiments, demonstrated that leucine was the penultimate amino acid.

The partial amino acid sequences determined thus far are compared in Figure 4. For the comparison it was assumed that there either was a deletion of a residue in position 2 of *P. elsdenii* flavodoxin or an insertion of a lysine residue in position 2 of the *C. pasteurianum* flavodoxin. Identical sequences were then observed at positions, 1, 3, 5, 7–18, 21, 29–30, 32, and 39. The constant section from residue 7–18 is interesting and may be an essential portion of the molecule. The cysteine residue which binds FMN (Mayhew *et al.*, 1969) was not present in the portion of the molecule sequenced thus far. The homology in the partial sequences of the two flavodoxins was 50% which is quite similar to the values of 54% observed for the ferredoxin from the two microorganisms.

In summary, the results of the sequence analyses of the flavodoxins from *P. elsdenii* and *C. pasteurianum* which can be determined on the intact protein are presented in this report. In order to complete the sequence, peptides will have to be isolated and the sequence determined by conventional methods. However, due to the great interest at the present time in the capabilities of the protein sequencer, we feel that the results presented here are of interest to protein sequencers.

TABLE IV: Number of Residues Determined in the Protein Sequencer in Our Laboratory.

Protein	Sequences Determined	References
1. *C. pasteurianum* flavodoxin (Cm derivative)	51	Present report
2. *P. elsdenii* flavodoxin (Cm derivative)	41	Present report
3. *Clostridium tartarivorum* ferredoxin (Cm derivative)	23	*a*
4. *P. elsdenii* ferredoxin (Cm derivative)	32	*b*
5. *Myxobacter* AL-1 protease (Cm derivative)	22	*c*
6. Chicken anti-DNP-γ-globulin		*d*
(a) Light chain (Cm derivative)	20	
(b) Heavy chain (Cm derivative)	25	

[a] M. Tanaka, M. Haniu, R. Himes, J. Akagi, L. Barnes, and K. T. Yasunobu, unpublished results. [b] P. Azari, J. Tsunoda, S. Mayhew, and K. T. Yasunobu, unpublished results. [c] G. Matsueda, R. Wolfe, and K. T. Yasunobu, unpublished results. [d] R. T. Kubo, Rosenbloom, and A. Benedict, unpublished results.

Acknowledgments

The authors are indebted to John J. Pisano for his invaluable assistance with the gas–liquid chromatography method.

ology and to Hugh Niall for his helpful discussions on the use of the protein sequencer.

Added in Proof

The partial amino acid sequence of the *C. pasteurianum* flavodoxin has been reported by Fox and Brown (1971). The placement of a lysine residue in position 46 is based on their report. In our studies, lysine and valine were observed in 1 and 2% yields, respectively, and it was not possible to decide which amino acid occupied the 46th position.

References

Ambler, R. B. (1967), *Methods Enzymol. 11*, 436.
Bradbury, J. H. (1958), *Biochem. J. 68*, 475.
Brewer, H., Jr., and Ronan, R. (1970), *Proc. Nat. Acad. Sci. U. S. 67*, 1862.
Crestfield, A. M., Moore, S., and Stein, W. H. (1963), *J. Biol. Chem. 238*, 662.
Dubourdieu, M., and Le Gall, J. (1970), *Biochem. Biophys. Res. Commun. 38*, 965.
Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80.
Fox, J. L., and Brown, J. R. (1971), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 30*, 1242.
Knight, E., Jr., D'Eustachio, A. J., and Hardy, R. W. F. (1966), *Biochim. Biophys. Acta 113*, 626.
Knight, E., Jr., and Hardy, R. W. F. (1966), *J. Biol. Chem. 241*, 2752.
Knight, E., Jr., and Hardy, R. W. F. (1967), *J. Biol. Chem. 242*, 1370.
Ludwig, M. L., Andersen, R. D., Mayhew, S. G., and Massey, V. (1969), *J. Biol. Chem. 244*, 6047.
Mayhew, S. G., Foust, G. D., and Massey, V. (1969), *J. Biol. Chem. 244*, 903.
Mayhew, S. G., and Massey, V. (1969), *J. Biol. Chem. 244*, 794.
Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244*, 5597.
Pisano, J. J., and Bronzert, T. J. (1970), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 29*, 916.
Spackman, D. H., Moore, S., and Stein, W. H. (1958), *Anal. Chem. 30*, 1190.
Van Orden, H., and Carpenter, F. (1964), *Biochem. Biophys. Res. Commun. 14*, 399.

Synthesis of Polypeptides and Oligopeptides with the Repeating Sequence L-Alanyl-L-prolylglycine*

G. P. Lorenzi,† B. B. Doyle,‡ and E. R. Blout§

ABSTRACT: The synthesis and characterization of the polytripeptide poly(Ala-Pro-Gly) and a series of oligopeptides (from the tripeptide to the octadecapeptide) with this sequence are reported. The polymer was synthesized by the active ester method, using the *N*-hydroxysuccinimide ester as the polymerizable tripeptide derivative in one preparation and the *p*-nitrophenyl ester in a second preparation. Good yields of relatively high average molecular weight polymer were obtained in both cases. The oligomers were prepared sequentially using the mixed-anhydride method, with isobutyl chloroformate as the mixed-anhydride-forming reagent. All the products were crystalline and gave only one spot on thin-layer chromatography after purification. Conformational studies on this polymer and the oligomers are reported in the accompanying paper.

One useful approach to the understanding of the physical and chemical properties of collagen is the study of model polypeptides having amino acid compositions and distributions resembling those of the natural protein. In particular, studying the conformational properties of such models in the solid state and in solution may help clarify the factors which stabilize the triple-helical structure of collagen. Thus in recent years there have been synthesized a number of polytripeptides similar to collagen in that every third residue is a glycine and at least one of the other two amino acids is a proline or hydroxyproline (Carver and Blout, 1967; Andreeva *et al.*, 1967).

To extend these studies, we have prepared the polytripeptide poly(Ala-Pro-Gly) and the first six members of the series of oligopeptides with the same sequence, having the general formula Boc-(Ala-Pro-Gly)$_n$-OMe.[1] We have also synthesized the corresponding trifluoroacetic acid salts, and some of the related Boc-protected peptide acids. The study of such oligopeptides of well-defined chemical structure and molecular weight provide a reliable complement to the study of the polydisperse fractions which were prepared from the polytripeptide. Experiments with the oligopeptides may help clarify relationships between structure and molecular weight

* From the Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115. *Received March 22, 1971.* We are pleased to acknowledge the support, in part, of this work by U. S. Public Health Grants AM-07300 and AM-10794. B. B. D. was a Public Health Predoctoral Fellow, No. 5-F1-GM-33,992-03.

† Present address: Department of Industrial and Engineering Chemistry, Swiss Federal Institute of Technology, Zurich, Switzerland.

‡ Present address: Laboratory of Biochemistry, National Institute of Dental Research, National Institutes of Health, Bethesda, Md.

§ To whom to address correspondence.

[1] The following abbreviations are used in this paper: Boc, *tert*-butyloxycarbonyl; Z, benzyloxycarbonyl; Pro, L-prolyl; Ala, L-Alanyl; Gly, glycyl; OMe, methyl ester; OBzl, benzyl ester; OSu, *N*-hydroxysuccinimide ester; ONp, *p*-nitrophenyl ester.