Biochem. J. (1968) **108**, 861
Printed in Great Britain

The Location of the Active-Site Histidine Residue in the Primary Sequence of Papain

BY S. S. HUSAIN AND G. LOWE
The Dyson Perrins Laboratory, University of Oxford

(*Received* 23 *February* 1968)

Papain that had been irreversibly inhibited with 1,3-dibromo[2-^{14}C]acetone was reduced with sodium borohydride and carboxymethylated with iodoacetic acid. After digestion with trypsin and α-chymotrypsin the radioactive peptides were purified chromatographically. Their amino acid composition indicated that cysteine-25 and histidine-106 were cross-linked. Since cysteine-25 is known to be the active-site cysteine residue, histidine-106 must be the active-site histidine residue.

In the preceding paper (Husain & Lowe, 1968*b*) the bifunctional reagent 1,3-dibromoacetone was reported to irreversibly inhibit the proteolytic enzyme papain. The reagent alkylated the thiol group of a cysteine residue and N-1 of a histidine residue, thus generating a cross-link between these residues. Since papain contains only one thiol group, namely that of cysteine-25 (Light, Frater, Kimmel & Smith, 1964), it was assumed to be this residue that was alkylated. This has been confirmed and the identity of the histidine residue established.

A preliminary account of this work has been published (Husain & Lowe, 1968*a*).

MATERIALS AND METHODS

Papain. This was prepared and its activity determined as in the preceding paper (Husain & Lowe, 1968*b*).

Trypsin. Thrice-crystallized material obtained from Seravac Laboratories Ltd. (Maidenhead, Berks.) was used.

α-Chymotrypsin. Thrice-crystallized material obtained from Seravac Laboratories Ltd. was used.

Carboxypeptidase A. Di-isopropyl phosphorofluoridate-treated, dialysed and recrystallized material obtained from Sigma (London) Chemical Co., London, S.W. 6, was used.

Leucine aminopeptidase. This was obtained from Seravac Laboratories Ltd.

1,3-*Dibromo*[2-^{14}C]*acetone.* The preparation is described in the preceding paper (Husain & Lowe, 1968*b*).

Radioactivity. All radioactivity measurements were performed in solution in an Isotopes Development Ltd. scintillation counter type 6012. Radioactivity on paper chromatograms was located with Isotopes Development Ltd. chromatogram scanner 2029 in conjunction with scaler 1700 and a recorder.

Inhibition of papain with 1,3-*dibromo*[2-^{14}C]*acetone.* Twice-crystallized papain (1·26g.) was dissolved in 0·05M-sodium acetate buffer, pH 5·6 (100ml.), and activated by adding cysteine hydrochloride (0·21g.) and 0·1M-EDTA solution (0·5ml.). The activator was removed by passing (in two 50ml. batches) through a column (3cm. × 45cm.) of Sephadex G-25, equilibrated with 0·05M-sodium acetate buffer, pH 5·6, containing EDTA (0·5mM). The eluted enzyme was immediately inhibited with a solution of 1,3-dibromo[2-^{14}C]acetone (20mg., specific activity 0·40mc/m-mole) in 0·05M-sodium acetate buffer, pH 5·6. The solution was kept at 20° for 3hr.; no enzymic activity remained. The solution was concentrated (to 60ml.) and the excess of inhibitor removed by filtration through a column (3cm. × 45cm.) of Sephadex G-25.

Reduction of the inhibited papain with sodium borohydride and carboxymethylation. The inhibited enzyme solution was evaporated to dryness in a rotary evaporator at 25° and the residue dissolved in an aqueous solution of guanidine hydrochloride (43g. in 50ml.) through which N_2 had been bubbled. The solution was made 1mM with respect to EDTA and adjusted to pH 7·5. Fresh aqueous $NaBH_4$ solution (25mg./ml., 25ml.) was added and the solution maintained under N_2 at 40° for 1hr. A few drops of octan-1-ol were added to minimize frothing. The final pH of the solution was 9·8.

Iodoacetic acid (0·8g.) in aqueous solution was added and the solution adjusted to pH 8·6. After 20min. excess of borohydride was decomposed by lowering the pH to 6·0. The solution was concentrated (to 60ml.) in a rotary evaporator at 25° and the solution passed through a column (3cm. × 45cm.) of Sephadex G-25 equilibrated with 50% (v/v) acetic acid. The solvent was removed in a rotary evaporator and protein suspended in water (100ml.)

Digestion of the inhibited and carboxymethylated papain. The protein suspension was digested with trypsin (25mg.) at 25°, the pH being maintained at 8·0 on a pH-stat with 0·5N-NaOH. After 4hr., α-chymotrypsin (25mg.) was added and the digestion continued for a further 5hr. To stop the digestion the pH was lowered to 3·0 by adding dilute HCl. A small amount of precipitate, which was centrifuged off, contained an insignificant amount of radioactivity. The solution was concentrated (to 15ml.) in a rotary evaporator.

Sephadex chromatography of the digest. The digest (15ml.) was applied to a column (160cm. × 2cm.) of Sephadex G-25 (fine grade) equilibrated with 20% (v/v) acetic acid containing thiodiglycol (0·2%) and eluted with the same solvent at

15ml./hr. Fractions of volume 4ml. were collected (see Fig. 1).

Phosphocellulose chromatography of the radioactive fractions from the Sephadex column. Cellulose phosphate (Whatman P11, batch 242, medium fibrous powder, nominal total capacity 7·4m-equiv./g.) was washed with N-NaOH, water, N-HCl, water and finally 0·05M-pyridine–acetate buffer, pH3·91.

Fractions 65–82 (zone 2 in Fig. 1) from the Sephadex chromatogram were combined, the solvent was removed in a rotary evaporator, and the residue was dissolved in 10% (v/v) acetic acid (5ml.) and applied to a phosphocellulose column (35cm. × 1cm.). The peptides were eluted with an eight-chamber gradient as described by Canfield & Anfinsen (1963) [all buffers contained thiodiglycol (0·2%)]. Fractions of volume 2·2ml. were collected at 15ml./hr. (see Fig. 2).

Fractions 50–64 (zone 1 in Fig. 1) from the Sephadex chromatogram were similarly chromatographed on a phosphocellulose column (35cm. × 1cm.) but with a four-chamber gradient: chamber 1, 0·05M-pyridine–acetate buffer, pH3·91 (120ml.); chamber 2, 0·10M-pyridine–acetate buffer, pH4·02 (120ml.); chamber 3, 0·15M-pyridine–acetate buffer, pH4·21 (120ml.); chamber 4, 0·20M-pyridine–acetate buffer, pH4·42 (120ml.). Fractions of volume 2·2ml. were collected at 15ml./hr. After 150 fractions the column was eluted with 0·5M-pyridine–acetate buffer, pH5·40 (see Fig. 3).

Paper chromatography of the radioactive peaks from the phosphocellulose columns. The following fractions from the phosphocellulose columns were combined: peptides I and IV (referred to below as peptide IV), fractions 35–38 (Fig. 2) plus fractions 36–37 (Fig. 3); peptide II, fractions 60–68 (Fig. 2); peptide III, fractions 70–77 (Fig. 2); peptide V, fractions 73–79 (Fig. 3); peptide VI, fractions 184–188 (Fig. 3). Each of these peptides was purified on Whatman no. 3MM paper with butan-1-ol–acetic acid–water (40:6:15, by vol.) by descending chromatography for 15hr. Guide strips cut from the chromatograms showed that (with the exception of peptide II) the radioactive peptides were well separated from other peptides, and gave positive ninhydrin and Ehrlich reactions. The radioactive zones were cut out and eluted off the paper with 10% (v/v) acetic acid. Each radioactive peptide was passed through a column (140cm. × 1cm.) of Sephadex G-25 with 50% (v/v) acetic acid containing thiodiglycol (0·2%) to remove contamination arising from the paper.

Amino acid analysis of the radioactive peptides. Each peptide (0·1–0·5μmole) was hydrolysed with 6N-HCl (2ml.) at 110° for 24hr. in an evacuated tube. The hydrolysate was evaporated to dryness under reduced pressure and the residue taken up in 0·1N-HCl. The amino acid analyses were performed on a Technicon AutoAnalyzer with the buffer system of Thomson & Miles (1964). The analyses are shown in Table 1.

Carboxypeptidase A digestion of the radioactive peptides. Carboxypeptidase A (0·01ml. containing 20mg./ml.) was dissolved in 0·5% Na_2CO_3 solution (0·1ml.) and a sample (0·05ml.) added to 0·1μmole of the peptide in 0·1M-sodium phosphate buffer, pH7·6. The molar ratio of the peptide to enzyme was 100:3. Solutions were incubated for various times at 25°. The digestion was stopped by acidification with HCl to pH2·5 and the digest applied to the amino acid analyser. The analyses are shown in Table 2.

Leucine aminopeptidase digestion of the radioactive peptide

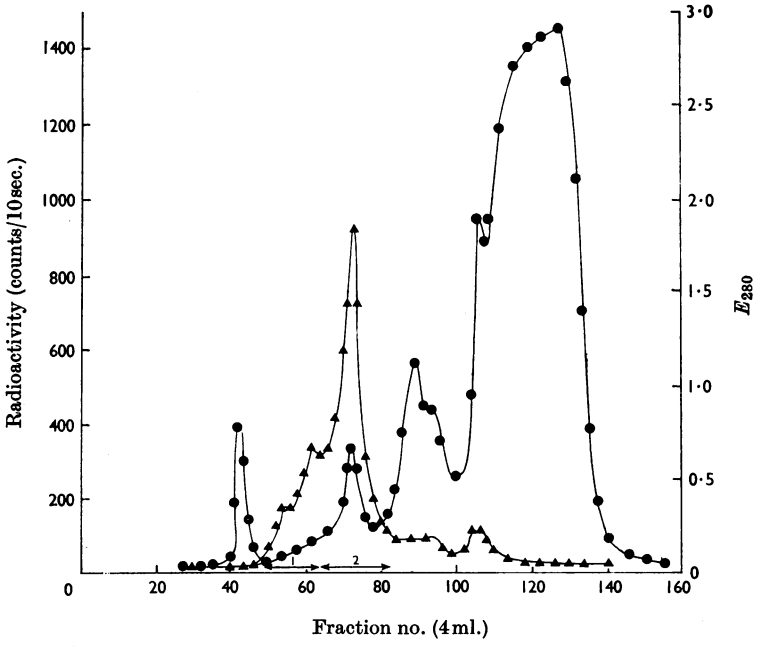


Fig. 1. Gel filtration on Sephadex G-25 of the tryptic and α-chymotryptic digest of 1,3-dibromo[2-^{14}C]acetone-inhibited and carboxymethylated papain, showing the radioactivity (▲) and E_{280} (●).

(*III*). Leucine aminopeptidase (0·5mg.) was dissolved in 0·05M-tris–HCl buffer, pH8·55 (0·1ml.), containing $MgCl_2$ (25mM). A sample (0·03ml.) was added to 0·1μmole of peptide (III) dissolved in 0·05M-tris–HCl buffer, pH8·55 (1·0ml.), containing $MgCl_2$ (5mM). The solution was incubated for 24hr. at 25°. The pH was adjusted to 2·5 and the solution applied to the amino acid analyser directly. The analysis is shown in Table 3.

Edman degradation of the radioactive peptide (III). The radioactive peptide (III) (1μmole) was subjected to three cycles of Edman degradation by the method of Konigsberg & Hill (1962). The analysis of the residual peptide after each cycle is shown in Table 4.

RESULTS AND DISCUSSION

The rapid and complete irreversible inhibition of papain was accomplished with only a slightly greater than equimolar quantity of 1,3-dibromo-[2-^{14}C]acetone. The inhibited enzyme was treated with sodium borohydride to reduce the ketone, which might otherwise have caused problems during the digestion and peptide separation. Reduction of the disulphide bridges was also accomplished and the cysteine residues were carboxymethylated with iodoacetic acid.

Digestion of the inhibited papain with trypsin and α-chymotrypsin gave a mixture of peptides that were fractionated on Sephadex G-25 (see Fig. 1). Fractions 50–82 contained almost all of the radioactivity. Fractions 50–64 (zone 1 in Fig. 1) and fractions 65–82 (zone 2 in Fig. 1) were chromatographed separately on phosphocellulose, but with different gradient elution buffers. Six radioactive peaks were collected (indicated by roman numerals in Figs. 2 and 3), but peaks I and IV were identical (referred to below as peptide IV) and were combined. The radioactive peptides, with the exception of peptide II, were finally purified by paper chromatography. The four peptides (III, IV, V and VI) after elution from the paper were passed through a Sephadex G-25 column to remove contaminants arising from the paper. The purified

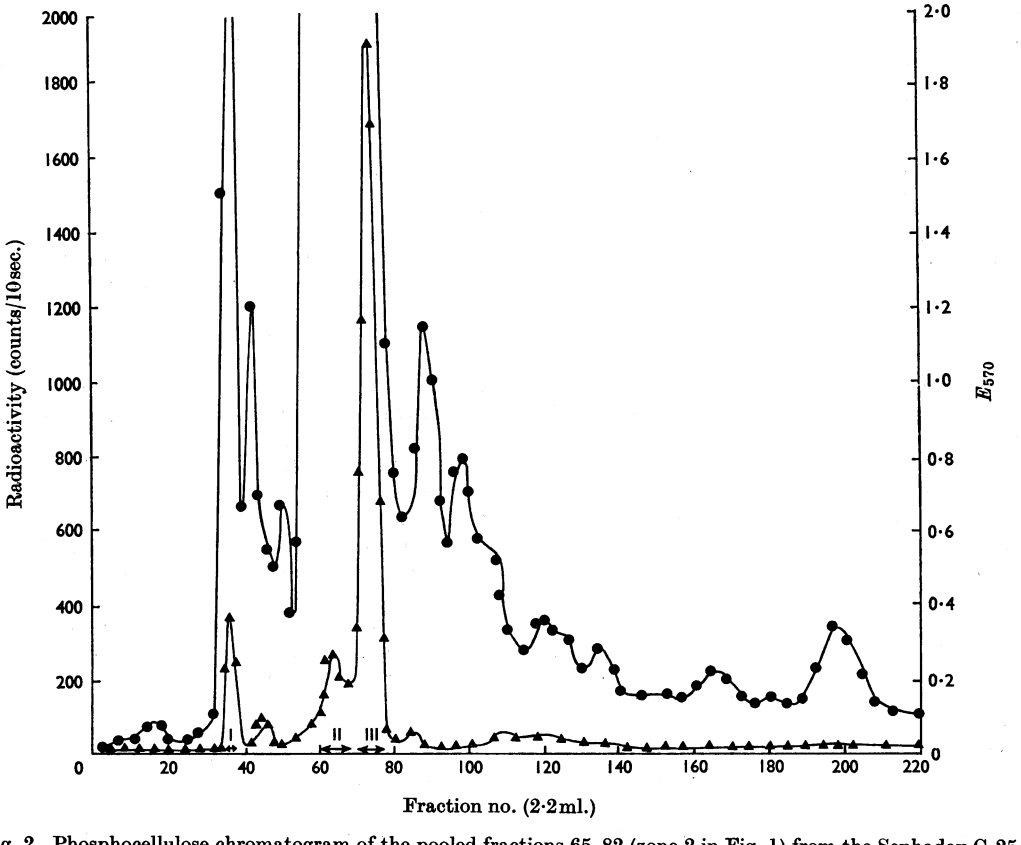


Fig. 2. Phosphocellulose chromatogram of the pooled fractions 65–82 (zone 2 in Fig. 1) from the Sephadex G-25 chromatogram showing the radioactivity (▲) and E_{570} (●) after reaction with ninhydrin in the analytical system of the Technicon AutoAnalyzer.

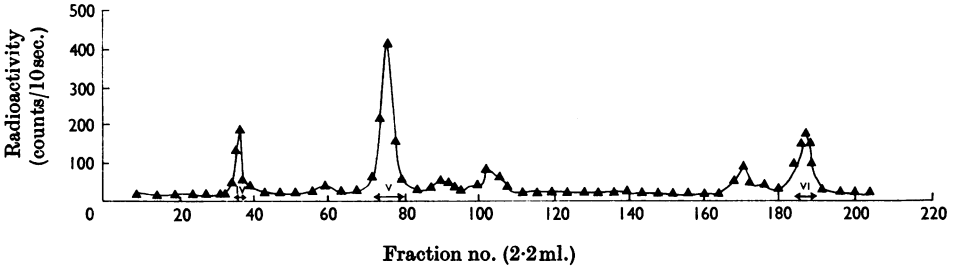


Fig. 3. Phosphocellulose radiochromatogram of the pooled fractions 50–64 (zone 1 in Fig. 1) from the Sephadex G-25 chromatogram.

Table 1. *Amino acid analyses of the acid hydrolysates of peptides III, IV, V and VI*

Tryptophan is not reported because of the variable results; each peptide, however, gave a positive reaction with Ehrlich reagent. Peptides III, IV and VI were assumed to contain 2 valine residues and peptide V 3 valine residues. CM-Cys, *S*-carboxymethylcysteine.

Amino acid	Peptide III	Peptide IV	Peptide V	Peptide VI
CM-Cys	0·95	0·78	1·04	0·70
Asp	1·92	1·87	2·04	1·65
Ser	1·71	1·82	1·71	1·60
Glu	1·14	1·15	1·06	1·00
Gly	2·09	2·41	3·16	2·05
Ala	2·00	1·81	2·96	2·07
Val	2·00	2·00	3·00	2·00
^{14}C-labelled amino acid	+	+	+	+
Lys	−	−	−	0·98

peptides III and V contained 30% and 5% respectively of the radioactivity in the inhibited enzyme; peptides IV and VI contained about 2% each.

1,3-Dibromoacetone was known to react with a cysteine and histidine residue in papain (Husain & Lowe, 1968*b*). From the amino acid analyses of the acid hydrolysates of the four peptides (see Table 1) and inspection of the amino acid sequence of papain (Light *et al.* 1964), it was clear that cysteine-25 and histidine-106 were the alkylated residues. The structures of the peptides III and V are therefore as indicated in Fig. 4. Peptide IV, which is more acidic than peptide III, but which has the same amino acid analysis after acid hydrolysis, is probably an artifact arising from the spontaneous rearrangement of the *N*-terminal asparaginyl residue to an *N*-terminal β-aspartyl residue. There are now several reports of this type of rearrangement (Haley & Corcoran, 1967).

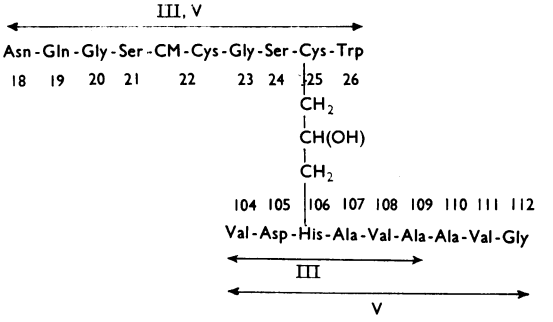


Fig. 4. Structures of peptides III and V. CM-Cys, *S*-carboxymethylcysteine.

Peptide VI contains a lysine residue, but otherwise has a similar amino acid composition to peptide III. It has not been established whether this residue arises from lysine-17 or lysine-103 in papain, because of the small amount of this peptide available. The cleavage at the *N*-terminus of this lysine residue may have arisen during the sodium borohydride reduction of the inhibited enzyme (cf. Crestfield, Moore & Stein, 1963). The *N*-terminal lysine residue so formed would not be susceptible to hydrolysis by trypsin.

The cleavage between alanine-109 and alanine-110 leading to peptide III has previously been observed during the sequence studies by Light & Smith (1962). Probably the very hydrophobic nature of neighbouring sequence is responsible. The cleavage between glycine-112 and tyrosine-113 leading to peptide V is less easily rationalized, but the yield of this peptide was much smaller.

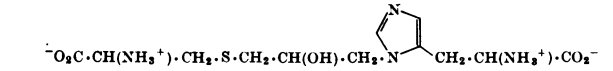
Each peptide gave a positive reaction (for tryptophan) with Ehrlich reagent and from their acid hydrolysates the same new amino acid emerged on the chromatogram just ahead of but well resolved from the ammonia peak. When the column eluate not used by the analytical system of the Auto-

Table 2. *Amino acid analyses of the carboxypeptidase A digests of peptides III, IV, V and VI*

The time of digestion is given for each analysis.

Amino acid	Peptide III		Peptide IV	Peptide V	Peptide VI
	30 min.	15 hr.	4 hr.	24 hr.	4 hr.
Gly	—	—	—	0·45	—
Ala	0·82	0·90	0·84	0·71	0·73
Val	0·82	0·96	0·88	0·72	0·74
Trp	0·51	1·09	0·80	1·08	0·68

Analyzer was collected and each fraction assayed radiochemically, all the radioactivity was found in the fractions associated with the new amino acid peak. It has already been established that the histidine residue is alkylated at N-1 and therefore the structure of the new amino acid must be:



The amino acid analyses of the carboxypeptidase A digest of the four peptides are shown in Table 2. It is evident that carboxypeptidase A rapidly hydrolyses the bond between alanine-109 and valine-108, but the bond between the alkylated histidine-106 and alanine-107 is not susceptible to hydrolysis. The more sluggish hydrolysis of the bond between valine-111 and glycine-112 in peptide V is in accord with the known specificity of this enzyme (Bailey, 1962). These results, together with the amino acid analysis of the leucine aminopeptidase digest (Table 3) and the Edman degradation (Table 4) of peptide III, confirm the amino acid sequence around residues cysteine-25 and histidine-106.

These results imply that N-1 of histidine-106 is within 5 Å of the sulphur atom of cysteine-25. Arguments have already been presented that suggest that N-3 of the histidine residue may be hydrogen-bonded to the sulphur atom of cysteine-25 (Husain & Lowe, 1968*b*).

Though the amino acid sequence around cysteine-25 is remarkably similar to the amino acid sequence around the active-site serine residue of trypsin (Lowe, 1966), it is now evident that the active-site histidine-106 of papain bears no recognizable homology with the active-site histidine of trypsin or α-chymotrypsin. However, the proximity of the imidazole group of histidine-106 to the thiol group of the active-site cysteine-25, together with the kinetic evidence (Lowe & Williams, 1965*a*,*b*; Wallenfels & Eisele, 1968), leaves little doubt that this residue also plays a key role in the mechanism of action of this enzyme.

Table 3. *Amino acid analysis of the leucine aminopeptidase digest of peptide III*

CM-Cys, *S*-carboxymethylcysteine.

Amino acid	No. of residues
CM-Cys	1·0
Asp	0·9
Ser, Asn, Gln	Unresolved
Gly	2·0
Ala	1·6
Val	2·0
Trp	1·1

Table 4. *Amino acid analyses of the acid hydrolysate of the residual peptide after each Edman degradative cycle of peptide III*

The amino acids lost in each cycle are indicated by the *italicized* values. CM-Cys, *S*-carboxymethylcysteine.

Amino acid	Peptide III	First cycle	Second cycle	Third cycle
CM-Cys	0·95	0·95	0·94	0·95
Asp	1·92	*1·12*	*0·20*	—
Ser	1·71	1·68	1·89	1·79
Glu	1·14	1·08	*0·35*	—
Gly	2·09	2·03	2·21	*1·33*
Ala	2·00	1·87	1·89	2·40
Val	2·00	*1·00*	1·00	1·39

Note added in proof. The structure of papain determined by X-ray crystallography (Drenth, Jansonius, Koekoek, Swen & Wolthers, 1968) has revealed that the tentative amino acid sequence determined by chemical methods (Light *et al.* 1964) requires revision. The active-site cysteine residue remains as residue 25, but histidine-106 in the chemically determined sequence becomes histidine-158. The imidazole group of this residue is 4 Å from the sulphur atom of cysteine-25 in the crystal structure (obtained from methanol–water, 2:1, v/v), which is in excellent agreement with the results reported in this and the preceding paper for papain in aqueous solution.

The authors gratefully acknowledge a grant from the Medical Research Council to support this work.

REFERENCES

Bailey, J. L. (1962). *Techniques in Protein Chemistry*, p. 203. Amsterdam: Elsevier Publishing Co.

Canfield, R. E. & Anfinsen, C. B. (1963). *J. biol. Chem.* **238**, 2684.

Crestfield, A. M., Moore, S. & Stein, W. H. (1963). *J. biol. Chem.* **238**, 622.
Drenth, J., Jansonius, J. N., Koekoek, R., Swen, H. M. & Wolthers, B. G. (1968). *Nature, Lond.*, **218**, 929.
Haley, E. E. & Corcoran, B. J. (1967). *Biochemistry*, **6**, 2668.
Husain, S. S. & Lowe, G. (1968*a*). *Chem. Commun.* p. 310.
Husain, S. S. & Lowe, G. (1968*b*). *Biochem. J.* **108**, 855.
Konigsberg, W. & Hill, R. J. (1962). *J. biol. Chem.* **237**, 2547.
Light, A., Frater, R., Kimmel, J. R. & Smith, E. L. (1964). *Proc. nat. Acad. Sci., Wash.*, **52**, 1276.
Light, A. & Smith, E. L. (1962). *J. biol. Chem.* **237**, 2537.
Lowe, G. (1966). *Nature, Lond.*, **212**, 1263.
Lowe, G. & Williams, A. (1965*a*). *Biochem. J.* **96**, 194.
Lowe, G. & Williams, A. (1965*b*). *Biochem. J.* **96**, 199.
Thomson, A. R. & Miles, B. J. (1964). *Nature, Lond.*, **203**, 483.
Wallenfels, K. & Eisele, B. (1968). *Europ. J. Biochem.* **3**, 267.