Structure of *Neurospora* Malate Dehydrogenase. II. Isolation and Partial Characterization of Polypeptide Subunits*

Kenneth D. Munkres

ABSTRACT: Heterologous polypeptide subunits of *Neurospora* malate dehydrogenase are separated by gel filtration after denaturation of the enzyme with acid. Analyses of gel filtration properties, immunodiffusion–precipitin reactions, molecular weights and shapes, total amino acid compositions, sequentially terminal amino acids, and tryptic peptides of the polypeptide subunits and the undissociated protein indicate that two polypeptides, of equal molecular weight (13,500) and of nonidentical amino acid composition and sequence, associate in the complete enzyme protein to yield a tetrameric structure of the form $\alpha\alpha\alpha\beta$.

Results of previous investigations of the structure and function of this enzyme by genetic mutation and by reversible denaturation with acid and urea not only confirm the structural formula but also elucidate genetical and physiological relationships.

Previous investigations of the structure and function of *Neurospora* malate dehydrogenase (E.C. 1.1.1.37) by genetic mutation and by reversible denaturation from acid and urea indicated that this enzyme, of compact and spherical shape and mol wt 54,000, is composed of polypeptides with different amino acid sequence in the asymmetric tetramer of the form $\alpha\alpha\alpha\beta$ (Munkres and Richards, 1965a).[1] The present paper reports the isolation of the subunits and their amino acid compositions, molecular weights, tryptic peptides, and several other physical and chemical properties. In some experiments properties of the undissociated protein and the subunits are compared.

Materials and Methods

Chemicals. The following special commercial chemicals were used: oxaloacetate, iodoacetate, iodoacetamide, and PTCK-trypsin[2] (California Corp. for Biochemical Research); NADH, 98%, Grade III (Sigma Chemical Co.); Sephadex G-75 Medium (Pharmacia Fine Chemicals); difluoroisopropylcarboxypeptidase A, three-times-crystallized, Lot No. CoA-DFP-6129 (Worthington Biochemicals Corp.); trypsin, two-times-crystallized, Lot No. H 2176 (Mann Research Lab.); subtilopeptidase A (E.C. 3.4.4.16), crystalline, Batch No. CEL 1369 (Nagarse and Co., Osaka; distributed by Enzyme Development Corp., New York); NEM (Nutritional Biochemical Corp). All other commercial chemicals were analytical reagent grade. Poly-L-lysine was a gift from E. Katchalsky. Bovine pancreatic ribonuclease A was kindly donated by F. M. Richards.

Enzyme Purification and Assay. Malate dehydrogenase is assayed and purified as previously described (Munkres and Richards, 1965b).

Purification of Polypeptide Subunits. Pure, lyophilized, salt-free malate dehydrogenase (MDH) is dissociated at pH 2.8 and 20° for 1–2 hr as previously described.[1] Sephadex G-75, medium grade, is suspended in 0.1 M NaCl, allowed to swell at room temperature for at least 2 days, and washed on a Büchner funnel with a solution containing: 10 mM sodium citrate, 1 mM EDTA, 1 mM mercaptoethanol, 50 mM NaCl, and 57 mM citric acid at a final pH of 2.8. This solution is called buffer B. Buffer A at pH 6.8 contains no citric acid. Fine particles are removed from Sephadex suspended in buffer B by flotation and the suspension is deaerated in a vacuum flask with an aspirator. Columns are poured at room temperature and equilibrated at 4° by washing with buffer B for 12–18 hr at an average flow rate of 10 ml/hr. Samples of acidified MDH in buffer B at 4° are applied to Sephadex G-75 columns in volumes less than or equal to one-third of the total gel bed volume, washed into the column with one sample volume of buffer B, and eluted with buffer B at 4° at an average flow rate of 10 ml/hr. The effluent is collected at constant time intervals in fraction volumes one-tenth to one-twentieth of the gel bed volume. Protein is estimated by absorbancy measurements at 280 mμ or by a microbiuret method (Munkres and Richards, 1965b). Protein fractions are concentrated by one of the following three methods. Method 1: for concentration at constant pH and ionic strength, solutions are extracted with dry

* From the Department of Biological Sciences, Stanford University, Stanford, Calif. *Received May 5, 1965.* This investigation was supported by research contracts from the National Institutes of Health and the National Science Foundation administered by F. M. Richards, Yale University, and D. O. Woodward, Stanford University, and by postdoctoral fellowships (GM-9487 and GM-158-07) from the National Institutes of Health to the author. A preliminary report of this investigation was previously given (Munkres, 1965a).

[1] K. D. Munkres, submitted for publication.

[2] Abbreviations: PTCK-trypsin, L-1-tosyl-2-phenethyl chloromethyl ketone derivative of trypsin; NADH, reduced nicotinamide–adenine dinucleotide; MDH, malate dehydrogenase; NEM, *N*-ethylmaleimide.

Sephadex G-25 (coarse grade) and vacuum filtration at 4° by the method of Flodin *et al.* (1960). This method yields low recoveries of the B fraction since this protein adsorbs to Sephadex under these conditions. Method 2: protein solutions are dialyzed against distilled water in $^{46}/_{60}$ cm acetylated dialysis tubing for 24–48 hr at 4°. (Dialysis tubing is acetylated for 24 hr at room temperature in pyridine–acetic anhydride (90:10, v/v) and thoroughly washed with water.) Dialysis in unacetylated tubing leads to disastrous losses of protein, particularly with the B protein fraction, apparently because of its low molecular weight. Dialyzed solutions are lyophilized and stored *in vacuo* over $CaCl_2$. Method 3: protein solutions are lyophilized directly from buffer B, taken up in a minimum volume of distilled water, and dialyzed against distilled water for 24–48 hr at 4° in acetylated dialysis tubing.

Immunochemical Analysis. The preparation and purification of rabbit anti-MDH globulin was previously described.[1] Double diffusion–precipitin tests are performed by the method of Ouchterlony (1949).

Ultracentrifugation. Sedimentation velocity measurements are made with a Spinco Model E ultracentrifuge equipped with a phase contrast schlieren optical system and a temperature control system. In some instances measurements of the sedimentation constants near 1.8 were made at 4° with a synthetic boundary cell at 59,780 rpm.

The procedure for sedimentation equilibrium measurements of molecular weights with interference optics was previously described (Munkres and Richards, 1965b). Photographs are measured with a Mann two-dimensional microcomparator.

Total Amino Acid Analyses. Protein samples of 1 mg are hydrolyzed with 1 ml of constant-boiling HCl at 110° *in vacuo*. Amino acid analyses are made by the method of Spackman *et al.* (1958) with a Beckman amino acid analyzer. Calculations of amino acid residues/mole of protein are based upon the total anhydrous weight of amino acids recovered. Cysteic acid and methionine sulfone are determined after performic acid oxidation by the method of Moore (1963). Tryptophan is estimated by the methods of Goodwin and Morton (1946) and Spies and Chambers (1949).

Proteolysis. The time course of proteolysis is followed with a recording automatic titrator (Radiometer, Type TTT 1a) at pH 8.0 and 30° with 0.01 N NaOH. Equivalents of NaOH in titrations are taken as 90% of the peptide bond equivalents (Richards, 1955). A slight correction for CO_2 absorption of 45 mμ equiv hr^{-1} is employed in experiments requiring maximum precision.

Trypsin used in initial experiments is treated with acid to minimize chymotrypticlike activity (Northrup and Kunitz, 1936). PTCK-trypsin (Kosta and Carpenter, 1964) is used in subsequent experiments after it became commercially available. Preliminary experiments on the tryptic hydrolysis of urea-denatured ribonuclease A with acid-treated trypsin in the titrator are performed to establish the appropriate conditions to obtain limit proteolysis.

Carboxyl-Terminal Amino Acid Analyses. MDH for these experiments is dialyzed against distilled water at 4° for 48 hr in acetylated tubing and lyophilized. MDH is dissolved in 6 ml of 0.1 M KCl in the reaction cup of the titrator, equilibrated at 30°, and adjusted to pH 8.0 with NaOH. A commercial mixture of difluoroisopropylcarboxypeptidase A and B[3] in 5% $LiCl_2$ is added to the reaction mixture in the proportion 50 moles of substrate to 1 mole of enzyme. While the course of the reaction is being monitored by the titrator, 1-ml aliquots are removed at given intervals and adjusted to pH 3 with HCl. Samples of the hydrolysate are prepared for automatic amino acid analysis by batch purification with Dowex 50-X2 as described by Davie *et al.* (1959). Control samples of MDH incubated without carboxypeptidase in the titrator yield no evidence of endogenous proteolysis. Analyses for endogenous free amino acid in the system at "zero time" are made with samples with carboxypeptidase added to MDH solutions at pH 3 that are subsequently purified with Dowex 50 as described above.

Peptide Analyses. Protein is denatured prior to hydrolysis with trypsin or subtilopeptidase A by incubation with 8 M urea, pH 8.2, for 1–3 hr at 37 or 50°. The urea concentration was diluted to 0.05 M prior to proteolysis and the urea was separated from peptides prior to paper chromatography with a column of Dowex 50-X2 as described by Helinski and Yanofsky (1962). In some instances, the protein is reduced with mercaptoethanol in 9.4 M urea at pH 8.2 and protein sulfhydryl

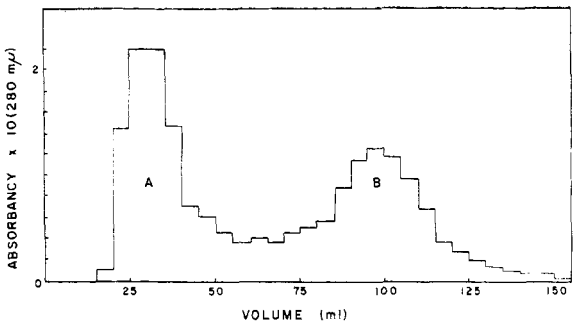

FIGURE 1: Separation of polypeptides of malate dehydrogenase by gel filtration. MDH (12 mg in buffer A, pH 6.8, is acidified to pH 2.8 with citric acid and incubated 2 hr at 25°. The sample (15 ml) is applied to a 20 × 2.2 cm of Sephadex G-75 equilibrated with buffer B, pH 2.8, at 4°. The column is eluted with buffer B at an average flow rate of 10 ml/hr at 4° and the effluent is collected in 5-ml fractions. See Materials for compositions of buffer A and B. The total gel bed volume is 76 ml. Void and internal volumes are 22 and 47 ml, respectively.

[3] Although the preparation was labeled difluoroisopropylcarboxypeptidase A, the presence of carboxypeptidase B was indicated by the hydrolysis of poly-L-lysine (0.6 mole of lysine released/mole) and by the hydrolysis of a limit tryptic digest of ribonuclease A to yield 7.9 moles of lysine/mole.

TABLE I: Weight Fractions and Distribution Coefficients of Polypeptides of Malate Dehydrogenase Obtained after Acid Dissociation and Gel Filtration.

Fraction[a]	Weight of Protein (mg) Experiment No. 1	2	3	4	5	6	Total	Distribution[b] Coefficient (K_D) n	$\bar{x}$	s
A	21.3	8.6	16.4	86	95	48	275	6	0.16	0.04
B	8.2	2.4	7.8	26.2	31.4	14	90	11	1.81	0.34
A/B	2.6	3.6	2.1	3.3	3.0	3.4	3.06			
Yield(%)[c]	..	..	..	100	93	98				

[a] See Figure 1 and text for definition of fractions. [b] $K_D = (V_e - V_0)/V_i$ where V_e is the elution volume, V_0 is the void volume, and V_i is the internal gel volume (Flodin, 1962). n is the number of experiments, $\bar{x}$ is the mean K_D, and s is the standard deviation. [c] Yield is expressed as the percentage of the initial total protein recovered in the sum of fractions A and B.

groups are blocked with iodoacetic acid or iodoacetamide (Anfinsen and Haber, 1961) or with *N*-ethylmaleimide.[1] The latter procedures allow the removal of urea by dialysis prior to proteolysis and chromatography, prevent the reassociation of the polypeptide subunits accompanying dilution from urea,[1] and yield protein that may be completely hydrolyzed by trypsin.

Trypsin and subtilopeptidase A are dissolved in water to 10 mg/ml and added to reaction mixtures at 0.01 to 0.02 part of enzyme to 1 part of substrate (w/w). All urea solutions were freshly prepared, passed through a column of mixed bed resin (Amberlite MB-1), and adjusted to pH 8.2 with triethylamine to minimize cyanate concentration (Stark *et al.*, 1960).

Results

Separation of Polypeptides. Gel filtration of acidified MDH at pH 2.8 characteristically yields two proteins in weight fractions of 3:1 and in near 100% yield (Figure 1, Table I). Separation of the two proteins is facilitated not only because of their molecular weight difference but also because the B fraction adsorbs strongly to Sephadex with an average distribution coefficient of 1.8. Nevertheless, the protein concentration never has been observed to go to zero between the two fractions even with columns up to 70 cm in length. The lowest protein concentration between the two peaks is taken as the midpoint in the calculation of weight fractions (Table I) and for pooling the fractions for subsequent analysis. The efficiency of separation is a function of the ratio of total protein to the total gel bed volume. Under nearly optimal conditions, such as in Figure 1, when the two fractions are separately pooled, concentrated, and rerun again on the column, the major fraction is obtained in homogeneous ($\geq$90%) form with respect to the other fraction.

The behavior of fraction A in a second gel filtration is a function of the method of concentration of the effluent from the first filtration. (See Methods for details of the concentration methods.) If fraction A is either allowed to stand for a day or so at 4° in the acid buffer after separation and concentrated by method 1 at constant pH and ionic strength or is concentrated by method 3 immediately after separation by lyophilization from the acid buffer, then two additional protein peaks are obtained in gel filtration, in addition to the original A peak, that are intermediate in distribution between the original A and B peaks and exhibit distribution coefficients of 0.39 and 0.68. In contrast, if the A fraction is immediately concentrated (by method 1) and placed on the column, or if the fraction is concentrated, after dialysis, from water by lyophilization (method 2), then the elution pattern is essentially unchanged. None of the concentration methods alter the elution pattern of the B fraction.

The foregoing results of acid dissociation and gel filtration experiments are interpreted, in light of the results of experiments reported below, in terms of the following reaction mechanisms.

$$\underset{\text{mol wt } 54{,}000}{\alpha\alpha\alpha\beta} \underset{\longleftarrow}{\overset{\text{pH } 2.8}{\longrightarrow}} \underset{\text{mol wt } 40{,}500\text{ fraction A}}{\alpha\alpha\alpha} + \underset{\text{mol wt } 13{,}500\text{ fraction B}}{\beta} \quad (1)$$

$$\alpha\alpha\alpha \xrightarrow{\text{pH } 2.8} \alpha\alpha + \beta \quad (2)$$

Dissociation of the native enzyme at pH 2.8 and 20° proceeds rapidly; the half-time of the reaction is 4–6 min.[1] Since the α trimer is included in the internal gel volume during filtration, its molecular weight is apparently near 40,000–50,000, the exclusion limit of Sephadex G-75 (Flodin, 1962), a conclusion that is supported by the results of other kinds of experiments (see below).

The reason the B fraction (β) is so greatly retarded on the Sephadex column is not clear. It has been proposed that proteins that adsorb to Sephadex contain relatively high proportions of heterocyclic and aromatic amino acids (Gelotte, 1960; Porath, 1960); however, the follow-

TABLE II: Physical Properties of *Neurospora* Malate Dehydrogenase and Polypeptide Subunits.

Preparation	Buffer[b]	pH	$s_{20,w}$ ($\times 10^{13}$)[c]	Mol wt ($\times 10^{-4}$)	Partial[g] Sp. Vol (cc/gm)	Fractional Coefficient (f/f_0)	Axial[h] Ratio (a/b)
Native MDH[a]	P_i	7.0	4.77[d]	5.40[e,f]	0.739	1.03	1.8
	P_i	7.0	4.69				
	Citrate	6.8	4.57				
Fraction A	Citrate	2.8	3.65	4.05[f]			
Fraction B	Citrate	2.8	1.80	1.35[e]	0.728	1.28	5.5
CAM-*A*[i]	P_i	7.0	1.75	1.34[f]	0.725	1.18	4.0
CAM-*B*[i]	P_i	7.0	1.70	1.36[f]			

[a] Munkres and Richards (1965b). [b] P_i, phosphate; citrate, buffer A or B (see Materials). [c] At 3–5 mg of protein/ml, except as noted. [d] Extrapolated to infinite dilution. [e] From sedimentation equilibrium. [f] From amino acid composition and other methods (see text). [g] Calculated from amino acid composition (McMeekin *et al.*, 1949). [h] As unhydrated prolate ellipsoid. [i] Carboxamidomethyl derivatives of fraction A and B.

ing compositional analyses of the α and β proteins do not support this hypothesis.

The second reaction, the dissociation of the α trimer, is apparently much slower but may be accelerated by lyophilization from the acid buffer at pH 2.8. If the molecular weights of the three components in the A fraction are taken as the series 40,500, 27,000, and 13,500, and the respective distribution coefficients, 0.16, 0.39, and 0.68, are expressed as elution volumes for a Sephadex column of given dimensions, the logarithm of the molecular weight is linearly related to the elution volume. Attempts to relate these results to the elution volumes of proteins of known molecular weight have not been made since there are few available proteins that may not undergo dissociation or unfolding at pH 2.8 under the conditions of these experiments.

Immunochemical Analysis. One of the most convenient tests for homogeneity of the polypeptide preparations is the double diffusion-precipitin test with anti-MDH globulin. A tracing of the results of a typical test is given in Figure 2. Fractions A and B of prototroph MDH yield single precipitin lines that are not homologous either in position or in quality of the precipitate. Precipitin lines from the two fractions are readily distinguishable in synthetic mixtures. Since it is possible to detect at least 10% inhomogeneity in these tests, it is apparent that the two fractions are at least 90% pure with respect to the other fraction. In this experiment, the proteins from MDH of a malate mutant (M-24) are also examined. In previous experiments (Munkres and Richards, 1965a), this mutant protein exhibited three less arginine and three extra tryptophan residues per molecule; hence, the amino acid replacement is in the α-polypeptide. The results of the precipitin test support this hypothesis. No detectable difference in the position or quality of the precipitin lines of the β subunits of M-24 and the prototroph are found, but the α-polypeptides differ in number and position of the precipitin lines. The nature of the multiple forms of the precipitin lines of fraction A from M-24 is not certain.

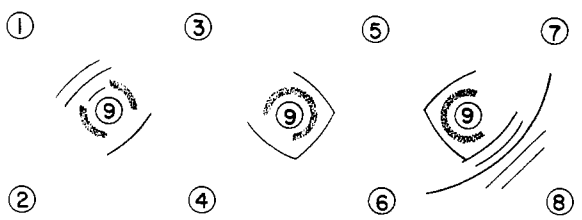

FIGURE 2: Tracing of an Ouchterlony plate and the precipitin lines formed by fractions A and B against rabbit anti-MDH globulin. The wells contain: (1) fraction A, mutant M-24; (2) fraction B, mutant M-24; (3) fraction B, prototroph; (4) fraction A, prototroph; (5) and (6) a 3:1 (w/w) mixture of fractions A and B, prototroph; (7) buffer; (8) prototroph MDH; and (9) rabbit anti-prototroph MDH globulin. Lyophilized protein of fractions A and B (method 2, Methods) and MDH are in 50 mM sodium phosphate buffer, pH 7.0, at 10 mg/ml. The reactions are carried out with 50 mM sodium phosphate buffer, pH 7.0, for 2 days at 25°.

Conceivably they represent dissociation of the α trimer in analogy to the results obtained in gel filtration of fraction A of the prototroph. This postulate is consistent with previous comparative observations of the molecular weight of this mutant enzyme that indicate it dissociates more readily at neutral pH than does the prototroph enzyme (Munkres and Richards, 1965a).

The nature of the multiple precipitin lines observed with undissociated MDH is not clear. These results are somewhat analogous to the multiple electrophoretic forms described elsewhere that are shown to be genetically related to MDH and not protein impurities (Munkres and Richards, 1965a).

Molecular Weight and Shape. The sedimentation constants of fraction A (when judiciously concentrated as previously discussed) and fraction B at pH 2.8 are 3.65 and 1.81, whereas the constant of the undissociated

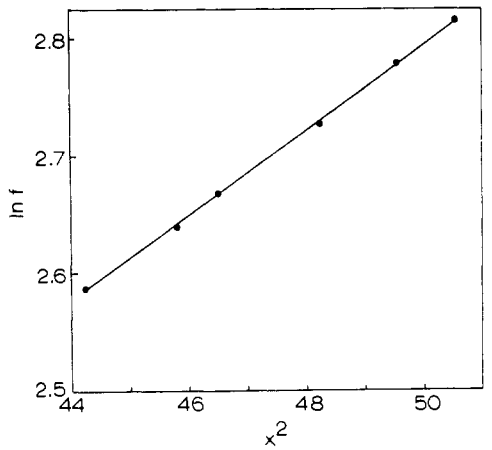

FIGURE 3: Sedimentation equilibrium plot of the β-polypeptide of *Neurospora* malate dehydrogenase. The natural logarithm of the fringe number is plotted *vs.* the square of the radial centrifugal distance (cm²). Centrifugation is at 20,410 rpm for 24 hr at 20° in citrate buffer, 0.125 M, pH 2.8, containing 1 mM EDTA, 50 mM NaCl, and protein, 2.8 mg/ml. Total fringe no., 14.6; bottom of cell, 51.1 cm²; meniscus at 41 cm²; fringe shift at 44.23 cm² is −1.3.

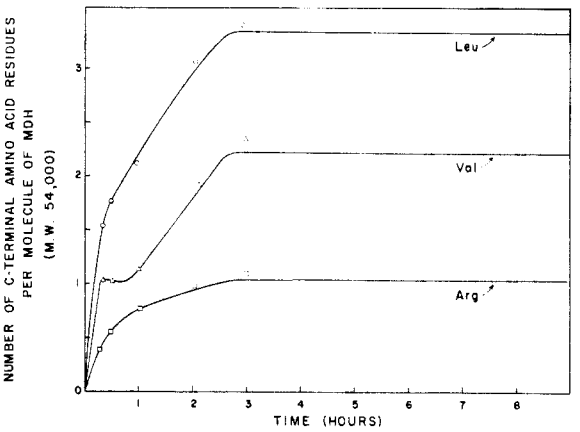

FIGURE 4: Hydrolysis of *Neurospora* malate dehydrogenase with a mixture of difluoroisopropylcarboxypeptidase A and B. The data from Table IV, for the amino acids indicated in the figure, are plotted to illustrate the three classes of stoichiometry: 3, 1, and 1 + 1 amino acid residues/mole.

protein, at pH 6.8 and comparable protein concentration, is 4.57 (Table II). Fractions A and B are denatured in 8 M urea, reduced with mercaptoethanol, and carboxamidomethylated with iodoacetamide by the method of Anfinsen and Haber (1961). After removal of the excess reagents by dialysis against neutral, dilute phosphate buffer, sedimentation constants of 1.75 and 1.70 S are obtained for A and B, respectively (Table II).

A molecular weight of 13,500 ± 500 is calculated for the β-polypeptide from sedimentation equilibrium measurements at pH 2.8 (Figure 3). The apparent molecular weight homogeneity is at least 95%. The minimal molecular weight estimated from the amino acid composition is 13,600 (Table III). By combining the sedimentation constant (1.80), the molecular weight (13,500), and the partial specific volume (0.728), the calculated frictional coefficient is 1.28. Assuming an unhydrated prolate ellipsoid, the apparent axial ratio is 5.5.

The minimal molecular weight of the α-polypeptide, estimated from the amino acid composition, is near 13,500 (Table III). Calculation of the shape factors from the sedimentation constant (1.75), the molecular weight (13,400), and the partial specific volume (0.725) yields a frictional coefficient of 1.18 and an axial ratio of 4.0 as an unhydrated prolate ellipsoid (Table II).

Although additional experiments are necessary to establish precisely the shapes of the proteins in solution, the available evidence indicates that acid denaturation of native MDH, which at neutral pH is compact and nearly spherical, is accompanied by extensive unfolding and elongation of the polypeptide chains.

The molecular weight and multimeric nature of the α-trimer is readily derived from the following experiments. First, from genetic amino acid replacement data and genetic theory, it was predicted that *Neurospora* MDH should have three polypeptides identical in amino acid sequence and at least one polypeptide of different sequence (Munkres and Richards, 1965b). Since the molecular weight of the enzyme is 54,000, the predicted molecular weight of the α-trimer is 40,500 with the assumption that the α- and β-polypeptides are of the same molecular weight. Also, evidence from the gel filtration experiments points to a molecular weight of 40,000–50,000 and the multimeric nature of the α-trimer. Third, the sedimentation constant of 3.65 is consonant with the predicted molecular weight, and the sedimentation constant of 1.75 after urea denaturation is consonant with the trimeric nature of α_3. Fourth, the carboxyl-terminal amino acid analyses given below indicate the existence of trimeric sequential homology in the carboxyl-terminal sequence of the whole enzyme. Moreover, the amino acid compositional data presented below allow the prediction of the minimal chemical molecular weight of the monomer, dimer, and trimer and indicate their compositional identity. Finally, the peptide analyses below confirm the minimal chemical molecular weight and indicate the identity of the amino acid sequence of the α-monomers. A summary diagram of the physical relationships of MDH and polypeptide subunits is given in Figure 6.

Amino Acid Composition. Analyses of the total amino acid composition of the α and β chains and the undissociated protein are summarized in Table III. α and β differ significantly in eight amino acids: His, Arg, Ser, Ala, half-Cys, Tyr, Phe, and Tryp. The near integral unity of four residues in the β chain and one residue in the α chain confirms the minimal chemical molecular weights and supports the physical estimates of molecular

TABLE III: Amino Acid Composition of *Neurospora* Malata Dehydrogenase and Polypeptide Subunits.

	Number of Amino Acid Residues/Molecule of Protein						
					$\alpha\alpha\alpha\beta$		
	α		β				Calculated ($3\alpha + \beta$)
Amino Acid	Observed	Nearest Integer	Observed	Nearest Integer	Observed	Nearest Integer	Nearest Integer
Lys	8.92	9	10.7	11	39.8	40	38
His	1.67	2	3.04	3	9.07	9	9
Arg	2.78	3	4.86	5	14.9	15	14
Asp	13.4	13	12.3	12	50.4	50	51
Thr	10.2	10	8.97	9	38.8	39	39
Ser	9.45	9	12.9	13	36.9	37	40
Glu	13.1	13	15.4	15	46.6	47	54
Pro	6.85	7	7.6	8	28.6	29	29
Gly	13.8	14	12.9	13	54.4	54	55
Ala	13.8	14	10.1	10	52.3	52	52
Half-Cys[a]	1.78	2	1.21	1	6.50	7	7
Val	8.25	8	8.39	8	40.0	40	32
Met[b]	0.96	1	0.95	1	3.97	4	4
Ileu	6.40	6	6.00	6	23.6	24	24
Leu	9.30	9	7.75	8	38.0	38	35
Tyr	3.36	3	0.995	1	10.8	11	10
Phe	4.75	5	3.44	3	17.6	18	18
Tryp[c]	0	0	1.0	1	1.07	1	1
Total		128		128		515	512
Mol wt		13,413		13,613		54,223	53,852
No. of Analyses[d]							
22 hr	2		3		9		
62 hr	1		2		2		

[a] As half-Cys, cysteic, or half-cys + cysteic acid. [b] As Met, Met sulfone or Met sulfone plus Met. [c] Determined spectrophotometrically and colorimetrically (see Methods). [d] Time of acid hydrolysis at 110°. Data are corrected for hydrolytic losses.

weights. Moreover, the amino acid composition of the undissociated protein is precisely predicted from the sum of the compositions of the monomers (Table III, columns 7 and 8). These results indicate that the proposed structure is compositionally sufficient; no other different polypeptides or nonprotein moieties are indicated or necessary.

Sequentially Terminal Amino Acids. Analyses of the carboxyl-terminal amino acids of undissociated MDH are performed by hydrolysis with a mixture of difluoroisopropylcarboxypeptidase A and B (Table IV; Figure 4). The hydrolysis is characterized by three types of stoichiometry; 3, 1, and 1 + 1 moles of amino acid/mole of protein. The inferred carboxyl-terminal sequence of the α-polypeptide is *-Tyr-Lys-Phe-Leu*. The trimeric proportion of α in the whole protein is illustrated by the 3:1 molar ratio of each of these amino acids, with the exception of tyrosine. Tyrosine is the one exception to the general stoichiometric pattern, being of the 1 + 2 type. This observation is interpreted as indicating the early release of *Tyr* from β during hydrolysis followed by incomplete hydrolysis of α chains at the fourth residue, *Tyr*, an interpretation that is supported by the observation that β has only 1 *Tyr* residue/molecule (Table III).

The 1 + 1 pattern apparently reflects the sequential release of the same amino acid at different points in the C-terminal sequence of β. The first four residues in the β chain are (tentatively): *-Ala-Ileu*(Tyr, Val).

The C-terminal analyses with whole MDH were performed before the isolated chains were available with the object of determining if the enzyme is composed of nonidentical polypeptides. Additional C-terminal analysis of the isolated chains is required to precisely establish their sequences.

In contrast to the foregoing results, Grimm and Doherty (1962) observed that the C-terminal amino acids of bovine and porcine mitochondrial MDH are:

TABLE IV: Hydrolysis of Malate Dehydrogenase with Difluoroisopropylcarboxypeptidase A and B.

Amino Acid[a]	Moles of Amino Acid/Mole of Protein[b] Time (min) 20	30	60	120	180	540	Nearest Integer[c]
Lys	0.68	1.01	1.50	1.96	3.52	3.04	3
Arg	0.40	0.55	0.80	0.98	1.10	0.99	1
His	0.13	0.30	0.26	0.50	0.32	0.60	1
Asp	0.09	0.18	0.26	0.36	0.33	0.66	1
Thr	0.34	0.51	0.55	0.85	1.02	1.09	1
Ser	0.05	0.40	0.47	0.76	0.62	1.29	1
Glu	0.31	0.30	0.38	0.71	0.97	1.37	1
Gly	0.04	0.14	0.01	0.02	0.02	0.41	0
Ala	0.48	0.67	0.89	1.54	1.80	2.01	1 + 1
Val	1.04	1.04	1.11	1.66	2.50	2.05	1 + 1
Met	0.00	0.40	0.40	0.43	0.46	0.63	1
Ileu	0.84	0.85	0.98	2.08	2.28	2.07	1 + 1
Leu	1.64	1.81	2.12	3.15	3.50	3.29	3
Tyr	0.85	1.18	1.07	1.78	2.90	2.47	1 + 2
Phe	1.15	1.22	1.61	2.35	3.94	3.00	3

[a] Not detected were: $AspNH_2$, $GluNH_2$, Pro, half-Cys, and Tryp. [b] MDH (0.93 μmole, 50 mg) was hydrolyzed with a mixture of difluoroisopropylcarboxypeptidase A and B (0.02 μmole, 0.7 mg) at pH 8.0 and 30° in a recording titrator as described in the text under Methods. Analyses are corrected for free amino acids in the system at time zero. [c] Nearest integer at limit proteolysis. 1 + 1 and 1 + 2 indicates biphasic kinetics (see Figure 5 and text for discussion).

Thr, *Ser*, and *Met*. This comparison emphasizes the previous conclusion, based on comparative amino acid analyses, that mammalian and *Neurospora* malate dehydrogenases are markedly different proteins (Munkres and Richards, 1965b).

Previous analyses of amino-terminal residues of *Neurospora* MDH yielded only 1 residue/molecule: alanine (Munkres and Richards, 1965b). Presumably this amino acid is derived from the β chain. The reason why other amino-terminal groups were not observed is not clear. Conceivably, the α chain possesses an acetylated amino-terminal residue. Such groups are not detected by the method employed, that of Sanger (Winstead and Wold, 1964). The natural occurrence of a number of proteins with acetylated amino-terminal residues is now well known (Winstead and Wold, 1964; Perlman and Block, 1963).

Proteolysis and Peptide Analyses. Before undertaking analyses of tryptic peptides of MDH and the polypeptide subunits, preliminary experiments were performed with the aid of an automatic recording titrator to establish the appropriate conditions to obtain complete hydrolysis, in an effort to avoid anticipated complications in peptide analysis associated with unhydrolyzed "core" protein. The theoretical and desired goal with MDH was the hydrolysis of 56 peptide bonds/molecule, the sum of the observed number of lysine and arginine residues (Table III).

"Native" MDH (freshly prepared and lyophilized) is extraordinarily resistant to proteolysis with either trypsin or subtilopeptidase A. The hydrolytic yield with trypsin is only 15–20 peptide bonds/mole. With subtilopeptidase A, of much broader specificity than trypsin, only 20–25 bonds are hydrolyzed. The experimentally determined hydrolytic limit with subtilopeptidase A against carboxymethyl-MDH (see below) is about 100–120 bonds/mole.

In an experiment on the effect of heat denaturation of MDH after treatment at 100° for 10 min in water, the unexpected observation was that no more peptide bonds were hydrolyzed with trypsin than with the native protein. In retrospect, these results are explicable in light of the interpretation of the reversibility of heat denaturation by urea and dialysis (Munkres, 1965b). The heat-denatured protein may be of compact structure and tightly folded due to illegitimate noncovalent bonds and, consequently, not readily susceptible to proteolytic attack.

The extent of tryptic hydrolysis was improved very little by treatment of native MDH with 8 M urea at pH 8.2 for 1 hr at either 37 or 50°. In these experiments, the urea concentration was diluted to 0.05 M before tryptic hydrolysis. Subsequent observations indicate that MDH tends to undergo renaturation and regain activity after dilution from 8 M urea (Munkres, 1965b).

Carboxymethyl-MDH is completely hydrolyzed by trypsin after denaturation in 9.4 M urea at pH 8.2 with strong mercaptoethanol, followed by treatment with iodoacetate, iodoacetamide, or *N*-ethylmaleimide and dialysis to remove excess reagents (Table V).

Tryptic hydrolysis of peptide bonds of carboxymethyl-MDH is correlated with the total number of

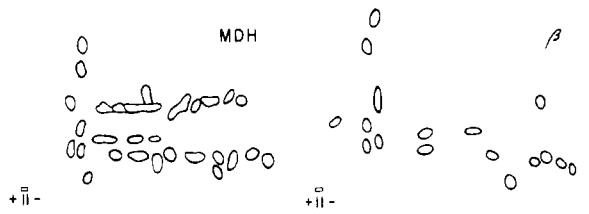

FIGURE 5: Tracing of peptide maps of tryptic peptides of *N*-ethylmaleimide derivatives of malate dehydrogenase and the β-polypeptide.

TABLE V: Number of Peptides Released during Tryptic Hydrolysis of Carboxymethyl-malate Dehydrogenase.

No. of Peptide Bonds Hydrolyzed (equiv/mole)[a]	Hydrolysis (%)[b]	No. of Peptides Obsd[c]	
		Strong Ninhydrin	Total
18	32	..	23
28	49	11	24
34	61	12	25
42	75	13	26
52	93	13	30

[a] Estimated with recording titrator. [b] No. of peptide bonds hydrolyzed/(56 total Lys + Arg residues) × 100. [c] By two-dimensional paper chromatography. "Strong ninhydrin" implies those peptides that developed color most rapidly and gave intense stain with ninhydrin.

different peptides formed (Table V). As complete tryptic hydrolysis is approached only 29–30 different peptides are found (Table V, Figure 5). Of these 30, 13 stain intensely with ninhydrin. Independent analysis of the tryptic peptides of the α chain yields 13 that are comparable in position to the intensely staining peptides observed with MDH. Thirteen is precisely the number expected on the basis of number of arginine and lysine residues in the α molecule (Table III). The relative intense staining of these 13 in the peptide maps of whole MDH apparently reflects the fact that each is present in a molar concentration threefold greater than that of each peptide from the β chain.

Seventeen tryptic peptides are derived from the β-polypeptide, in exact agreement with the number predicted on the basis of the number of arginine and lysine residues/β molecule (Table III, Figure 5).

The foregoing results lead to the conclusions: the sum of the different tryptic peptides from the α- and β-polypeptides is equal to the total number of different tryptic peptides yielded by ααaβ. Moreover, these results indicate that the two polypeptides have little, if any, homologous amino acid sequences in common. Finally, it appears that tryptic hydrolysis of denatured

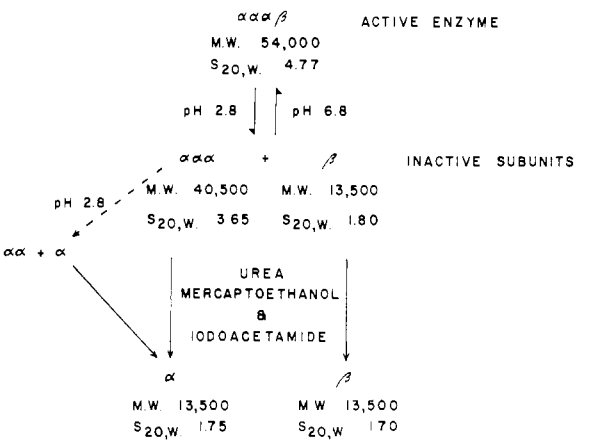

FIGURE 6: Association–dissociation of polypeptide subunits of *Neurospora* malate dehydrogenase.

MDH proceeds principally by the "one by one" mechanism (Tiselius and Eriksson-Quensel, 1939) rather than by the "zipper" mechanism (Linderstrøm-Lang, 1952), at least in the latter course of the reaction, since only seven additional peptides are released in going from 32 to 93% hydrolysis (Table V). [See Otteson and Széleky (1962) for a recent review of proteolytic mechanisms.]

Discussion

The pioneering studies of Beadle and Tatum (1941) with *Neurospora crassa* led to the concept that there is a 1:1 relationship between gene and enzyme. Subsequent studies on the structure of proteins (Sanger and Tuppy, 1951) and genetic material (Watson and Crick, 1953) led to a restatement of this relationship in molecular terms (Crick, 1958); the linear sequence of nucleotides in a gene specifies the linear sequence of amino acids in a polypeptide. Experimental evidence for this "collinearity" hypothesis has recently been reported (Yanofsky *et al.*, 1964; Sarabhai *et al.*, 1964). Meanwhile, prompted by the pioneering studies of Sanger and paralleling developments in molecular genetics, the existence of polypeptide subunits in proteins first demonstrated by Svedberg and Pedersen (1940) has been reaffirmed. Proteins may now be classified as *homopolymorphic* and *heteropolymorphic* with regard to homology of the amino acid sequences of their constituent polypeptides (Pontecorvo, 1963). Over 12 examples of heteropolymorphic proteins, most of which require the heteroassociation for biological activity, are now known (Reithel, 1963). In the genetic determination of the amino acid sequence of a protein, implicit in the collinearity relationship is the one gene (cistron):one polypeptide hypothesis (Benzer, 1957). Thus, it readily follows that the one gene:one enzyme relationship is a special case in which only *one* enzymatic activity is associated with *one* polypeptide or multiples thereof, and that the structure of heteropolymorphic proteins is determined by more than one gene. Examples of this relationship are

TABLE VI: Genetic Determination of the Structure of Heteropolymorphic Proteins.

Protein	Organism	Structural Genes, No.	Different Polypeptides, No.	Functional Proteins, No.	Multimericity[a]	Ref.
Hemoglobin	Man Mouse	4	4	4	4	*b*
Lactate dehydrogenase	Man Mouse	2	2	5	4	*c*
Tryptophan synthetase	*E. coli* *Neurospora*	2	2	1–2	2–4	*d*
7S α-globulin	Man	2	2	1	4	*e*
Malate dehydrogenase	*Neurospora*	2	2	1	4	*f*

[a] Number of polypeptides per functional protein. [b] Ingram (1963). [c] Appella and Markert (1961); Cahn *et al.* (1962); Lindsay (1963); Markert (1963); Nance *et al.* (1963); Shaw and Barto (1963). [d] Yanofsky *et al.* (1964); Carsiotis *et al.* (1965). [e] Harboe *et al.* (1962). [f] Munkres *et al.* (1965); Munkres and Richards (1965a,b); Munkres[1] (1965a,b).

given in Table VI. In these systems there is an exact correspondence between the number of structural genes and the number of sequentially different polypeptides. However, the number of functional proteins varies and is dependent, in part, upon the number of different polypeptides and the multimericity of the functional protein.

Although precise correlation of gene–protein structures in these systems obtains, in general, there are certain experimental and theoretical limitations to this approach that may not be widely recognized. Strictly rigorous proof that a given gene is a *limited* and *continuous* segment of genetic material cannot be solely attained by classical analysis of recombination and mutation, particularly in instances where two (functionally defined) genes are contiguous segments. Although large numbers of mutants of a given gene may be isolated, allowing the recombination analysis to "run the map into the ground" (Benzer, 1959), the occurrence of "silent" genetic regions, either as a consequence of mutation–selection bias or as a phenomenon intrinsic to genetic material or gene product, cannot be rigorously excluded (Bonner, 1964). Conversely, if two nonlinked genes were found that apparently determined the structure of one protein, there are two alternatives to the inference that the protein is heteropolymorphic. First, there is the theoretical possibility that the two genes may have arisen recently in evolution by duplication followed by translocation (Lewis, 1950; Ingram, 1963). If no mutation occurred subsequent to duplication and translocation, then each gene would code for the amino acid sequence of the same polypeptide. No examples of such redundancy in a gene–protein system have yet been demonstrated. Another theoretical limitation to genetic inferences about protein structure lies in the initial definition of the wild-type or prototrophic state. Some apparently wild-type organisms may theoretically be "suppressed wild type." Suppressor genes may act at the level of translation in protein synthesis and lead to amino acid replacements such that the mutational effect on a protein, whose amino acid sequence is encoded by another genetic locus, is repaired or compensated (Garen and Siddiqi, 1962; Brody and Yanofsky, 1964). Thus, mutation in a "suppressed wild" of either the structural or suppressor gene could lead to structural alteration of the same protein. In the absence of information about amino acid sequences of the polypeptide subunits of the protein, such genetic results could lead to incorrect inferences about the heteropolymorphic nature of the protein.

Similarly, inferences concerning the number of genes in a system derived from studies of protein structure are also subject to experimental and theoretical limitations. The occurrence of a heteropolymorphic protein does not necessarily indicate that two or more genes carry the information for the amino acid sequence of the sequentially different polypeptides. Some heteropolymorphic proteins may originally be made as single polypeptides which, after folding, are converted to proteins with two or more sequentially nonidentical chains. The prime and sole example is the conversion of chymotrypsinogen to chymotrypsin. In this case, a one-chain protein is converted to a protein with three disulfide-linked chains by the proteolytic removal of two small peptides (Walsh and Neurath, 1964). Conversely, the isolation and characterization of a homopolymorphic protein does not ensure that heteroassociation does not occur. This situation is exemplified by the lactic dehydrogenase isoenzymes of animals in which both hetero- and homopolymorphic enzyme proteins may occur in equilibrium (Table VI).

In *Neurospora* two unlinked genes, probably on different chromosomes, have been designated *ma*-1 and *ma*-2 (Munkres *et al.*, 1965). Subject to the limitations of genetic analysis as discussed above, the two genes constitute two segments of limited continuous genetic

material. Mutations of either gene result in structural and functional alteration of malate dehydrogenase (Munkres and Richards, 1965a).[1] (In addition, another activity, aspartate aminotransferase, is associated with the same protein and is simultaneously altered.[1]) The observations, that mutations of either gene lead to dissociation of the protein and that in the case of a mutant of the *ma*-2 locus three arginine residues are replaced by three tryptophan residues/mole, indicate that the protein is heteropolymorphic and that the minimal structure is of the form $\alpha\alpha\alpha\beta$, where α and β are sequentially nonidentical polypeptides. This conclusion was of necessity still subject to the limitations discussed above. However, when the results of the present and preceding (Munkres, 1965b) investigations of the structure of malate dehydrogenase are analyzed together with the genetic data, then the reservations discussed are canceled in the solution of the gene–protein relationship.

The discovery of two unique, sequentially and compositionally different polypeptides of *Neurospora* malate dehydrogenase indicates that the theoretical reservations of the genetical analysis discussed above are not applicable to this gene–protein system. Each of the two genes apparently functions as one unit of genetic material (a cistron) by the specification of the amino acid sequences of the α and β polypeptides, respectively. Conversely, the occurrence of two mutationally, recombinationally, and functionally distinct genes in this system indicates that the two polypeptides are synthesized independently by the cell and are not derived from a single chain by limited proteolysis as in the chymotrypsinogen–chymotrypsin conversion. Moreover, the fact that the α and β subunits are found in a molar ratio of 3:1 with a number of preparations of the enzyme that are obtained in yields of 50 to 80% indicates that the principal form of the enzyme is $\alpha\alpha\alpha\beta$ and that an equilibrium, random mixture of hetero- and homopolymorphic associations of these subunits does not obtain as with animal lactate dehydrogenases.

Numerous and recent investigations of protein structure have supported the hypothesis that the folding process by which a protein arrives at its three-dimensional configuration is essentially a thermodynamic one and that no *genetic* information, other than that present in the amino acid sequence of the protein, is required (Reithel, 1963; Epstein *et al.*, 1963). As stated by Crick (1958), "folding is simply a function of the order of the amino acids." The results of previous investigations of the alteration of the structure and function of malate dehydrogenase by genetic mutation and reversible denaturation from acid and urea are fully compatible with this hypothesis (Munkres and Richards, 1965a; Munkres, 1965b). However, it is probable that the secondary and tertiary configurations of this enzyme are influenced by the coenzymes NAD and pyridoxal phosphate.[1] Nevertheless, these and other possible refinements (Epstein *et al.*, 1963) do not affect the validity of the assumption that no genetic information, other than the amino acid sequence, is required. The coenzymes (episemantic molecules) do not express genetic information contained in enzymes (tertiary semantides) but are a product of this information (Zuckerkandl and Pauling, 1965).

The problem of where and how the α- and β-monomers of *Neurospora* MDH are synthesized and assembled in the cell remains to be solved. Does the synthesis of the two extended polypeptide chains proceed concurrently on the same or different ribosomes? Are free α and β monomers assembled in a concerted reaction or does the assembly occur in two stages with the last step being the association of the α trimer and β monomer? The results of reassociation experiments *in vitro* favor the latter possibility (Munkres, 1965b). Finally, the intracellular site(s) of synthesis of the chains and the complete enzyme must be correlated with site(s) of intracellular localization of the functional enzyme.

References

Anfinsen, C. B., and Haber, E. (1961), *J. Biol. Chem. 236*, 1361.

Apella, E., and Markert, C. L. (1961), *Biochem. Biophys. Res. Commun. 6*, 171.

Beadle, G. W., and Tatum, E. L. (1941), *Proc. Natl. Acad. Sci. U. S. 27*, 499.

Benzer, S. (1957), *in* A Symposium on the Chemical Basis of Heredity, McElroy, W. D., and Glass, B., eds., Baltimore, The Johns Hopkins Press, p. 70.

Benzer, S. (1959), *Proc. Natl. Acad. Sci. U. S. 45*, 1607.

Bonner, D. M. (1964), *J. Exptl. Zool. 157*, 9.

Brody, S., and Yanofsky, C. (1964), *Proc. Natl. Acad. Sci. U. S. 50*, 9.

Cahn, R., Kaplan, N. O., Levine, L., and Zwilling, E. (1962), *Science 136*, 962.

Carsiotis, M., Apella, E., Provost, P., Germershausen, J., and Suskind, J. R. (1965), *Biochem. Biophys. Res. Commun. 18*, 877.

Crick, F. H. C. (1958), *Symp. Soc. Exptl. Biol. 12*, 138.

Davie, E. W., Newman, C. R., and Wilcox, P. E. (1959), *J. Biol. Chem. 234*, 2635.

Epstein, C. J., Goldberger, R. F., and Anfinsen, C. B. (1963), *Cold Spring Harbor Symp. Quant. Biol. 28*, 439.

Flodin, P. (1962), Dextran gels and Their Applications in Gel Filtration, Halstad, Sweden, Meijels Bokindustri.

Flodin, P., Gelotte, B., and Porath, J. (1960), *Nature 188*, 493.

Garen, A., and Siddiqi, O. (1962), *Proc. Natl. Acad. Sci. U. S. 48*, 1121.

Gelotte, B. (1960), *J. Chromatog. 3*, 330.

Grimm, F. C., and Doherty, D. G. (1962), *Biochim. Biophys. Acta 57*, 381.

Harboe, M., Osterland, O. K., and Kunkel, H. G. (1962), *Science 136*, 979.

Helinski, D. R., and Yanofsky, C. (1962), *Biochim. Biophys. Acta 63*, 10.

Ingram, V. M. (1963), The Hemoglobins in Genetics and Development, New York, Columbia Univ. Press.

Kosta, V. A., and Carpenter, F. H., (1964), *J. Biol. Chem. 239*, 1799.

Lewis, E. B. (1950), *Advan. Genetics 3*, 78.

Linderstrøm-Lang, K. (1952), Proteins and Enzymes,

Lane Medical Lectures, Vol. VI, Stanford, Stanford Univ. Press.
Lindsay, D. T. (1963), *J. Exptl. Zool. 152*, 75.
Markert, C. L. (1963), *Science 140*, 1329
McMeekin, T. L., Groves, M. L., and Hipp, N. J. (1949), *J. Am. Chem. Soc. 71*, 3298.
Moore, S. (1963), *J. Biol. Chem. 238*, 235.
Munkres, K. D. (1965a), *Federation Proc. 24*, 469.
Munkres, K. D. (1965b), *Biochemistry 4*, 2180 (this issue; preceding paper).
Munkres, K. D., and Richards, F. M. (1965a), *Arch. Biochem. Biophys. 109*, 457.
Munkres, K. D., and Richards, F. M., (1965b), *Arch. Biochem. Biophys 109*, 465.
Munkres, K. D., Giles, N. H., and Case, M. E. (1965), *Arch. Biochem. Biophys. 109*, 397.
Nance, W. E., Calflin, A., and Smithies, O. (1963), *Science 142*, 1075.
Northrup, J. H., and Kunitz, M. (1936), *in* Handbuch der biologischen Arbeitsmethoden, Vol. IV, pt. 2, Abderhalden, E., ed., Berlin, Urban and Schwarzenberg, p. 2213.
Ouchterlony, Ö. (1949), *Acta Pathol. Microbiol. Scand. 26*, 507.
Ottesen, M., and Székely, M. (1962), *Compt. Rend. Trav. Lab. Carlsberg, Sér. Chim. 32*, 319.
Perlman, R., and Block, K. (1963), *Proc. Natl. Acad. Sci. U. S. 50*, 533.
Pontecorvo, G. (1963), *Proc. Roy. Soc. (London) B158*, 1.
Porath, J. (1960), *Biochim. Biophys. Acta 39*, 193.
Reithel, F. J. (1963), *Advan. Protein Chem. 18*, 123.
Richards, F. M. (1955), *Compt. Rend. Trav. Lab. Carlsberg, Sér. Chim. 29*, 329.
Sanger, F., and Tuppy, H. (1951), *Biochem. J. 49*, 481.
Sarabhai, A. S., Stretton, A. O. W., Brenner, S., and Bolle, A. (1964), *Nature 201*, 13.
Shaw, C. R., and Barto, E. (1963), *Proc. Natl. Acad. Sci. U. S. 50*, 211.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.
Spies, J. R., and Chambers, D. C. (1949), *Anal. Chem. 21*, 1249.
Stark, G. M., Stein, W. H., and Moore, S. (1960), *J. Biol. Chem. 235*, 3177.
Svedberg, T., and Pedersen, K. O. (1940), The Ultracentrifuge, Oxford, Clarendon.
Tiselius, A., and Eriksson-Quensel, I. B. (1939), *Biochem. J. 33*, 1752.
Walsh, K. W., and Neurath, H. (1964), *Proc. Natl. Acad. Sci. U. S. 52*, 8884.
Watson, J. D., and Crick, F. H. C. (1953), *Cold Spring Harbor Symp. Quant. Biol. 18*, 123.
Winstead, J. A., and Wold, F. (1964), *Biochemistry 3*, 791.
Yanofsky, C., Carlton, B. C., Guest, J. R., Helinski, D. R., and Henning, U. (1964), *Proc. Natl. Acad. Sci. U. S. 51*, 266.
Zuckerkandl, E., and Pauling, L. (1965), *J. Theoret. Biol. 8*, 357.