Amino Acid Sequences around the Cysteine Residues of Rabbit Muscle Triose Phosphate Isomerase

BY JANET C. MILLER AND S. G. WALEY

Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, U.K.

(Received 21 December 1970)

1. The nature of the subunits in rabbit muscle triose phosphate isomerase has been investigated. 2. Amino acid analyses show that there are five cysteine residues and two methionine residues/subunit. 3. The amino acid sequences around the cysteine residues have been determined; these account for about 75 residues. 4. Cleavage at the methionine residues with cyanogen bromide gave three fragments. 5. These results show that the subunits correspond to polypeptide chains, containing about 230 amino acid residues. The chains in triose phosphate isomerase seem to be shorter than those of other glycolytic enzymes.

Crystallographic evidence showed that the molecule of triose phosphate isomerase (EC 5.3.1.1) from rabbit muscle was a dimer in which the two subunits were related by a twofold rotation axis (Johnson & Waley, 1967). Likely types of molecules are thus α_2 or $\alpha_2\beta_2$, where α and β represent polypeptide chains. The subunit would be α (one polypeptide chain) in the first case or $\alpha\beta$ (two different polypeptide chains) in the second case. The subunit cannot be α_2 since there are approximately ten thiol groups/molecule, and five unique residues of cysteine (Burton & Waley, 1968). The subunits thus contain an odd number of cysteine residues, and so they cannot consist of a pair of identical chains.

Hence one aim in the work to be described was to see whether the subunits did represent polypeptide chains. Another aim was to make a start on finding the amino acid sequence, knowledge of which would be necessary to interpret a high-resolution X-ray image. Further, sequence information is necessary to assess homologies between enzymes from different sources: Perham (1969) raises the question whether the high degree of sequence homology found in glyceraldehyde 3-phosphate dehydrogenases is peculiar to that enzyme or general to glycolytic enzymes. Finally, structural information may eventually shed some light on relationships between glycolytic enzymes (Waley, 1969).

MATERIALS AND METHODS

Enzymes. Rabbit muscle triose phosphate isomerase was obtained from Boehringer Corp. (London) Ltd., London W.5, U.K. Trypsin (twice-crystallized, salt-free), crystalline bovine chymotrypsin, porcine pepsin and carboxypeptidase A treated with di-isopropyl phosphorofluoridate, were from Worthington Biochemical Corp.,

Freehold, N.J., U.S.A. Subtilisin was from Novo Terapeutisk Laboratorium, Copenhagen, Denmark, or from BDH Chemicals Ltd., Poole, Dorset, U.K. Aminopeptidase M (Rohm and Haas Co.) was from The Hexoran Co. Ltd., Belper, Derby, U.K. Thermolysin was from Serva, Heidelberg, Germany.

Enzymic digestion. Peptides were treated with trypsin, chymotrypsin, subtilisin or carboxypeptidase A [dissolved by the method of Smillie & Hartley (1966)], in 1% (w/v) NH_4HCO_3 , with pepsin in 5% (v/v) formic acid, and with thermolysin in ammonium acetate (0.2M with respect to acetate), pH 8.5, containing 5mM- CaCl_2 . The substrate/enzyme ratio was 50:1 or 100:1 (w/w). Digestions with aminopeptidase M were carried out in 33mM-sodium phosphate buffer, pH 7.4.

Purification of peptides. The tryptic digest of carboxymethylated triose phosphate isomerase was fractionated by ion-exchange chromatography on a column (80cm high $\times$ 0.9cm diam.) of sulphonated polystyrene essentially as described by Burton & Waley (1968), except that two linear gradients were used: the first was from 0.2M-pyridine acetate buffer, pH 3.1, to 0.8M-pyridine acetate buffer, pH 4.1; 500ml of each buffer was used. The flow rate was 6ml/h. The second gradient was from the pH 4.1 buffer to 2M-pyridine acetate buffer, pH 5; 200ml of each buffer was used. The molarity of the buffers refers to the pyridine (which was redistilled after addition of 0.1% of bromine); the pH was measured at room temperature, and the elution was carried out at 40°C. Chromatography on columns (140cm long $\times$ 0.9cm diam.) of Sephadex G-25 or G-50 was carried out in 20mM- NH_3 solution saturated with chloroform, or in 0.5% formic acid; peptides were detected by their extinction at 230nm or 235nm.

Paper chromatography of peptides was carried out in solvent BAWP,* and paper electrophoresis at pH 1.9, 3.5 or 6.5 in the buffers described by Ambler (1963), and at

* Abbreviations: BAWP, butan-1-ol-acetic acid-water-pyridine (15:3:12:10, by vol.) (Waley & Watson, 1953); Cmc, *S*-carboxymethylcysteine.

pH 4 in the buffer described by Grassman, Hannig & Plöckl (1955). Mobilities at pH 6.5, relative to aspartic acid -1, were used to calculate charges (Offord, 1966). Peptides were detected with the ninhydrin-cadmium reagent (Heilmann, Barollier & Watzke, 1957); the Ehrlich reagent for tryptophan was as described by Smith (1953), and the Pauly reagent for histidine or tyrosine utilized diazotized sulphanilic acid prepared from equimolar amounts of sulphanilic acid and NaNO_2 .

For peptide 'maps' the first step was high-voltage electrophoresis at pH 6.5; the basic and acidic peptides were then run in solvent BAWP, and the neutral peptides further separated by electrophoresis at pH 3.5 or 1.9 before chromatography in solvent BAWP. Peptides were sometimes detected with 0.02% ninhydrin in acetone after electrophoresis and chromatography; the proportion of *N*-terminal amino acid destroyed seems to be negligible (Beale, 1967).

Sequence methods. *N*-Terminal analyses were done by the 'dansyl technique' (Gray, 1967a; Hartley, 1970), and the DNS-amino acids were identified by t.l.c. on polyamide sheets (from Cheng-Chin Trading Co. Ltd., Taipei, Taiwan) with the solvents of Woods & Wang (1967); the plates were sometimes redeveloped in the second dimension with ethyl acetate-triethylamine acetate, pH 3 (40:1, v/v) (Waley, 1965), or with ethyl acetate-methanol-acetic acid (20:1:1, by vol.) (Crowshaw, Jessup & Ramwell, 1967). The dansyl-Edman technique (Gray, 1967b) was used for sequential analysis; destruction of DNS-serine was often unexpectedly extensive.

Amino acid analysis. Protein was dialysed, freeze-dried and hydrolysed with approx. 6M-HCl (BDH Aristar grade, diluted with 1 vol. of water) in evacuated tubes at 108°C for 24 or 96 h (two samples at each time); two further hydrolyses were carried out on carboxymethylated triose phosphate isomerase for 24 h. Destruction of serine and threonine was assumed to be first order (Hirs, 1954), and valine and isoleucine were assumed to have been completely released after hydrolysis for 96 h. Peptides were hydrolysed in evacuated tubes for 24 h unless otherwise stated; the addition of mercaptoethanol (0.1%) (Keutmann & Potts, 1969) diminished losses of tyrosine and carboxymethylcysteine. Analyses were carried out on a Beckman model 120C or a Locarte amino acid analyser.

Determination of extinction coefficient. A portion of dialysed triose phosphate isomerase was used for measurement of E_{280} ; dilutions were made by weight, and allowance was made for scattered light. The E_{280} of the undiluted solution was calculated to be 2.45. A portion of this solution was then treated with an equal volume of 20% (w/v) trichloroacetic acid, and the precipitate washed with 10% (w/v) trichloroacetic acid and dried to constant weight at 110°C (Hoch & Vallee, 1953). The concentration was 2.06 mg/ml and thus $E_{280}^{1\%} = 1.21$. Another portion of triose phosphate isomerase was dialysed for 4 days against NaHCO_3 (10 mg/l) and then against water; a portion was dried, first at 95°C and then at 105°C for 16 h; the concentration determined from the dry weight was 2.48 mg/ml and E_{280} was 3.01; these values also give $E_{280}^{1\%} = 1.21$.

Determination of cysteine. The protein was alkylated with sodium iodo[2- ^{14}C]acetate, and both the radioactivity of the protein and the amount of *S*-carboxy-

Table 1. *Determination of cysteine by reaction with sodium iodo[2- ^{14}C]acetate*

Alkylation was carried out as described in the text; the extent of alkylation is based on the amount of *S*-carboxymethylcysteine liberated by acid hydrolysis (second column), or the radioactivity of the protein (third column); the amount of protein was determined by amino acid analysis.

Time of reaction (h)	Number of SH groups/mol of mol.wt. 50000	
	Amino acid analysis	Extent of radioactivity
1	9.2	9.9
2	9.2	9.9

methylcysteine were determined (Table 1). A portion of the stock suspension of triose phosphate isomerase was dialysed against 0.2M-sodium borate, pH 8.5, containing 0.1M-EDTA and 25 μM -dithiothreitol. Sodium iodo[2- ^{14}C]acetate (0.2M, 0.2 mCi/ml) was added in eightfold excess over thiol groups, and the solution then made 4M in recrystallized guanidine hydrochloride. After being kept under N_2 for (usually) 2 h in the dark, mercaptoethanol (20-fold excess over the sodium iodoacetate) was added, and the reagents were removed by dialysis against 20 mM- NH_3 solution saturated with chloroform.

Determination of tryptophan. This was determined from the u.v. absorption (Beaven & Holiday, 1952) and by titration with *N*-bromosuccinimide in 8M-urea (Witkop, 1961), and by reaction with 2-hydroxy-5-nitrobenzyl bromide (Barman & Koshland, 1967).

Peptide nomenclature. The primary products are tryptic peptides. Further digestion of these peptides with chymotrypsin, pepsin, subtilisin or thermolysin gave peptides denoted by C, P, S or Th respectively. A refers to partial acid hydrolysis. Thus 4P1 refers to a peptic peptide obtained from (tryptic) peptide 4, and 4P1Th2 to a product obtained by further digestion of peptide 4P1 with thermolysin.

Alkylation and digestion with trypsin. The protein (100–260 mg) was alkylated as described under determination of cysteine and digested in 1% (w/v) NH_4HCO_3 with one-hundredth of its weight of trypsin for 4 h at 37°C.

Cleavage with cyanogen bromide. After carboxymethylation of 60 mg of triose phosphate isomerase, the solution was dialysed against 3 mM-dithiothreitol and freeze-dried. The residue was dissolved in 4 ml of 70% (v/v) formic acid that had been saturated with N_2 , and treated with 20 mg of CNBr at room temperature. The use of dithiothreitol should help to prevent oxidation of methionine residues (Drapeau & Yanofsky, 1967); after 8 h, a further 30 mg of CNBr was added. The next day, the solution was freeze-dried and the residue extracted with 2 ml of pH 3.1 buffer (300 ml of acetic acid and 16 ml of pyridine/l). The gelatinous residue was spun down and similarly extracted twice more. The residue (fraction R) was not purified further; the supernatant was fractionated by chromatography on Sephadex G-75 with the pH 3.1 buffer.

RESULTS

Amino acid composition

Published estimates of the molecular weight of rabbit muscle triose phosphate isomerase (Meir & Cotton, 1966; Burton & Waley, 1966; Johnson & Waley, 1967; Krietsch, Pentchev, Klingenburg, Hofstatter & Bucher, 1970) vary somewhat, but are distributed around a value of 50 000. Hence the values in the second column of Table 2 have been calculated on this basis; our values agree reasonably well with those given by Krietsch *et al.* (1970), and their values for methionine and tyrosine are probably more accurate. There is considerable variation in the analytical values for tryptophan

(Table 3), and there must be systematic errors in several of the methods. The high value obtained by the method of Barman & Koshland (1967) might suggest reaction with thiol groups, but *S*-(2-hydroxy-5-nitrobenzyl)cysteine is eluted with histidine on amino acid analysis (Horton & Koshland, 1967) and there was the usual proportion of histidine in the hydrolysate of treated protein. Perhaps one (or more) residues of tryptophan were disubstituted, as can happen with *N*-acetyltryptophan methyl ester (Spande, Wilchek & Witkop, 1968). A value of eight residues of tryptophan best accounts for the absorption of the protein, a check recommended by Wetlaufer (1962): the calculated molar extinction coefficient is then $(8 \times 5550) + (8 \times 1340) = 55\,120$, and the observed

Table 2. *Amino acid composition of triose phosphate isomerase*

Column (1) gives the results obtained in this work, and column (2) gives the results of Krietsch *et al.* (1970), recalculated for a molecular weight of 50 000.

Amino acid	This work	Krietsch <i>et al.</i>	Residues/226 residues (nearest integer)
	(1)	(1970) (2)	
Lys	40.5	39.4	20
His	7.5	7.7	4
Arg	13.8	15.0	7
Asp	38.2	37.6	19
Thr	29.1	27.0	14
Ser	26.6	22.0	13
Glu	50.2	50.4	25
Pro	17.8	18.6	9
Gly	44.8	44.5	22
Ala	48.5	49.0	24
Cys	9.9*	9.2	5
Val	44.2	45.9	22
Met	3.1	3.9	2
Ile	25.5	27.9	13
Leu	28.0	27.8	14
Tyr	6.5†	7.6	3-4
Phe	14.2	14.7	7
Trp	8‡	8.4	4

* Determined as *S*-carboxymethylcysteine (see the text).

† Found by the method of Beaven & Holiday (1952): 8.4 residues.

‡ See the text and Table 3.

Table 3. *Estimations of tryptophan in rabbit muscle triose phosphate isomerase*

Method	Reference to method	Residues/mol of mol.wt. 50 000	Reference
<i>p</i> -Dimethylaminobenzaldehyde Ultraviolet absorption	Spies & Chambers (1949)	4.6	Burton & Waley (1966)
	Beaven & Holiday (1952)	7.8	This work
	Edelhoc (1967)	8.6	Krietsch <i>et al.</i> (1970)
<i>N</i> -Bromosuccinimide	Witkop (1961)	8	This work
2-Hydroxy-5-nitrobenzyl bromide	Barman & Koshland (1967)	12	This work
Formylation	Previero, Colletti-Previero & Cavadore (1967)	9.4	Krietsch <i>et al.</i> (1970)

Table 4. *Amino acid composition of tryptic peptides containing carboxymethylcysteine*

The amino acid compositions are given as molar ratios, derived from hydrolysis for 24 and 72 h (peptides 1 and 2), or two hydrolyses for 24 h (peptides 3 and 4). Analytical values for peptide 5 are incomplete (see the text). The methods of purification are abbreviated thus: 1.9, 3.5, electrophoresis at pH 1.9 or 3.5 respectively; BAWP, paper chromatography in solvent BAWP; G-25, G-50, chromatography on Sephadex G-25 or G-50, in 20 mM-NH₃ or 0.5% formic acid. The first step in the purification was ion-exchange chromatography. The yields are derived from the radioactivity of the peaks after ion-exchange chromatography. The electrophoretic mobilities (*m*) refer to pH 6.5, and are relative to aspartic acid -1.00.

Amino acid	Peptide	...	...	1	2	3	4	5
Cmc				0.92	0.69	0.58	0.76	0.78
Asp				2.22		1.04		1.02*
Thr				1.96	2.14			0.87
Ser					1.03		1.15	1.06
Glu				1.19		1.1	2.35	1.40
Pro				2.98				
Gly					3.26		2.92	2.27
Ala				3.92	1.17	3.06	3.12	1.25
Val				3.02*	1.12	1.08	1.61*	1.35
Ile				1.09	1.81*	1.12*	1.78	
Leu							2.13	0.99
Tyr				0.81	0.58	0.35		
Phe				1.05				
Trp								1.00†
Lys					1.17	1.03	1.08	(1)
His							1.23	
Arg				0.79				(1)
Total no. of residues				20	13	10	18	14
Yield (%)				60	66	65	61	47
<i>m</i>				-0.4	0	0	-0.18	-0.23
Method of purification					1.9, BAWP; or BAWP, 3.5	1.9, BAWP, G-25; or BAWP, 3.5	G-50	G-25

* *N*-Terminal residue.

† Estimated from the E_{280} .

is $1.21 \times 50\,000 = 60\,500$; the ratio of the observed to the calculated (1.1) is within the usual range (1.0 to about 1.15).

The results from amino acid analyses can be used to determine the size of the subunit; the total number of amino acid residues is varied, and the value chosen that minimizes the sum of the squares of the deviations from integral values (Katz, 1968; Delaage, 1968). This gives a total of 226 residues, and the last column of Table 2 gives the calculated number of residues/subunit. The molecular weight of the subunit estimated on this basis is about 24 400; experimental values are about 26 000 (Johnson & Waley, 1967; Coulson, Knowles & Offord, 1970).

Sequences of cysteine peptides

The protein was alkylated with sodium iodo-[2-¹⁴C]acetate and digested with trypsin. The digest was first fractionated by ion-exchange chromatography and the five radioactive peptides, num-

bered according to the order in which they were eluted, were then further purified (Table 4). The work on each peptide is now described.

Peptide 1. This peptide, which is virtually ninhydrin-negative, is the same as peptide 8 of Burton & Waley (1968). It contains 20 residues (Table 4). The peptides obtained by digestion with chymotrypsin, and pepsin, are described in Table 5. Treatment of peptide 1C1 with carboxypeptidase gave alanine and tyrosine, and the specificity of chymotrypsin suggests that the *C*-terminal sequence is alanyltyrosine. Peptide 1C2 was uncharged at pH 4, so that aspartic acid, rather than the amide, was probably present.

Peptide 1P3 was the same as peptide 1C2, and its sequence was determined by the dansyl-Edman method as:

Ile-Asp-Phe-Ala-Arg

Although peptide 1 is ninhydrin-negative its amino group is not blocked. Similarly the peptic peptide representing its *N*-terminal sequence was

missed when the ninhydrin reaction was used for detection. This peptide was therefore detected and isolated as its dansyl derivative; peptide 1 was treated with DNS-chloride, desalted on Sephadex G-25 and digested with pepsin. The results (Table 5) appear to indicate regeneration of valine from the DNS-derivative after hydrolysis (incompletely dansylated peptide should have been separated by electrophoresis).

Peptide DNS-1P1, which contains the first six residues of peptide 1, was further digested with subtilisin and the two products were separated by electrophoresis at pH 6.5. The fluorescent peptide DNS-1P1S1 had the composition: proline (1.1), alanine (0.9), valine (0.5); the other product [aspartic acid (0.7), threonine (0.9), glutamic acid (1.3)] had a mobility of 0.95 at pH 6.5, and this

mobility shows that both aspartic acid and glutamic acid were present as such rather than as their amides. Limited digestion of DNS-1P1 with carboxypeptidase gave glutamic acid and threonine (ratio 1.0:0.8), so that we tentatively infer that the glutamic acid residue is *C*-terminal in peptide DNS-1P1.

The *N*-terminal sequence of peptide 1 was found to be Val-Pro-Ala-, and so these results gave the sequence of the first six residues in peptide 1 (Fig. 1). The sequence of the next nine residues was derived from further study of peptide 1P2. The first question was the number of valine residues: acid hydrolysis (for about 20h) gave approx. 1.5 residues. Digestion of 1P2 with aminopeptidase M at 37°C for 10h released 2mol. prop. of valine (and 1 of carboxymethylcysteine). During determination of the *N*-terminal residue as valine there was another fluorescent spot, which is now believed to be DNS-Val-Val; hydrolysis of the Val-Val peptide bond is notoriously slow. The presence of this bond was established during determination of the sequence of peptide 1P2 by the dansyl-Edman method, which led to the sequence:

Val-Val-Cmc-Ala-Pro-Pro-Thr-Ala-Tyr

This sequence was confirmed by amino acid analysis of the peptide after one, three and four degradations.

No useful results were obtained from digestion of peptide 1P2 with subtilisin or elastase, and so partial acid hydrolysis was carried out; peptides containing threonine are preferentially hydrolysed (Hill, 1965). The *C*-terminal peptide, containing threonine, was isolated after hydrolysis for 24h at 37°C in conc. hydrochloric acid. The peptide giving a positive Pauly reaction was isolated after electrophoresis at pH 1.9 and chromatography in solvent BAWP; this peptide 1P2A1 [threonine (0.8), alanine (1.2), tyrosine (1.1)] had *N*-terminal threonine. These results confirm the *C*-terminal region of peptide 1P2.

Peptide 2. This peptide is the same as peptide 17C1 of Burton & Waley (1968) and contains 13

Table 5. *Peptides obtained by digestion of peptide 1 with chymotrypsin and with pepsin*

Peptide 1 (0.5 μ mol) was treated with chymotrypsin for 16h at room temperature and the digest fractionated on Sephadex G-25 in 20 mM-ammonia; peptide 1C3 was further purified by electrophoresis at pH 5; its composition was found from paper chromatography. The digestion of peptide 1 with pepsin was carried out for 6h at 37°C, and the digest fractionated by electrophoresis at pH 1.9. The peptic digest from dansylated peptide 1 (DNS-1) was fractionated at pH 6.5. The amino acid compositions are expressed as molar ratios, except for peptide 1C1 where the basis is that the first five amino acids account for 10 residues. The *N*-terminal amino acid is shown in *italics*.

Peptide	Amino acid composition
1C1	Asp (1.2), Thr (2.0), Glu (1.1), Pro (2.7), Ala (3.1), Cmc (0.4), Val (2.4), Tyr (0.7)
1C2	Asp (1.1), Ala (1.0), <i>Ile</i> (1.0), Phe (0.9), Arg (+)
1C3	<i>Ala</i> (+), Arg (+)
DNS-1P1	Asp (1.0), Thr (0.8), Glu (1.1), Pro (1.3), Ala (1.0), <i>Val</i> (0.8)
1P2	Cmc (1.1), Thr (1.1), Pro (2.2), Ala (2.3), <i>Val</i> (1.3), Tyr (0.4)
1P3	Asp (0.9), Ala (1.0), <i>Ile</i> (1.1), Phe (1.0) Arg (+)
1P4	<i>Ala</i> (+), Arg (+)

Fig. 1. Sequence of peptide 1. Sequences obtained by the dansyl-Edman method are indicated thus: $\rightarrow$ etc.; amino acids released by carboxypeptidase A are indicated thus: $\leftarrow$. Chymotryptic peptides are denoted by C, peptic by P, subtilisin peptides by S and peptides from partial acid hydrolysis by A.

residues. The sequence was studied by the dansyl-Edman method (Fig. 2), and by digestion with thermolysin (Table 6).

Carboxypeptidase A released serine from peptide 2Th1, and a little lysine together with traces of threonine and carboxymethylcysteine from peptide 2 itself. The sequence of the first four residues of peptide 2Th2 were obtained by the dansyl-Edman method. These results led to the sequence given in Fig. 2. The partial sequence given earlier (Burton & Waley, 1968) missed one of the isoleucine residues (hydrolysis of the Ile-Ile bond is slow), and misplaced the residue of serine.

Peptide 3. This, the smallest radioactive peptide, contained ten residues; the mobility at pH 6.5 was zero, so that both aspartic acid and glutamic acid

are present as the amides. The sequence was obtained by the dansyl-Edman method (Fig. 3) and confirmed by study of the peptides obtained by digestion with thermolysin and subtilisin (Table 7). This peptide is probably the same as the peptide 17B, which was not obtained pure by Burton & Waley (1968).

Peptide 4. Peptide 4 is the radioactive peptide that stays at the origin in paper chromatography in solvent BAWP; it corresponds with peptide 20A (Burton & Waley, 1968), from which the peptide Cmc-(Ile,Gly,Glu)Lys had been obtained by digestion with subtilisin. Peptide 4 contains 18 residues, and the initial sequence is: Val-Ala. Useful fragments were obtained by the action of thermolysin and pepsin (Table 8).

The sequence of the four peptides obtained by digestion with thermolysin were determined by the dansyl-Edman method; the two *C*-terminal ones (peptides 4Th3 and 4Th4) could then be ordered from the previously isolated subtilisin peptide already mentioned. There are four residues missing from the thermolysin peptides, and these can be assigned to the *N*-terminal region from the (partial) sequence of peptide 4P1 (Fig. 4). Further ordering of fragments was made more difficult by the fact that peptides 4Th1 and 4Th2 have the same *N*- and *C*-termini. Hence the action of thermolysin on the peptic peptide 4P1 was studied. The resulting peptides enabled the fragments to be ordered (Fig. 4). The structure of peptide 4P1Th2 would entail thermolysin's cleaving a terminal bond, if peptide 4P1 represents residues 1-9 of peptide 4. This would be unexpected (Ambler & Meadway, 1968), and it is more likely that peptide 4P1 is a mixture of residues 1-9 and residues 1-11, as is also suggested by the analytical values for glycine and valine.

Peptide 5. The fifth radioactive peptide had been previously obtained as an acidic hexapeptide containing tryptophan but lacking arginine or lysine, and we had supposed that this peptide (peptide 9) was formed by a chymotrypsin-like cleavage, as this is not uncommon (Burton & Waley, 1968). The

Table 6. *Peptides obtained from digestion of peptide 2 with thermolysin and subtilisin*

Peptide 2 (0.45 μmol) was digested with 5 μg of thermolysin for 3 h at 37°C, and the digest fractionated by chromatography in solvent BAWP. The digest of peptide 2 (0.45 μmol) with subtilisin (10 μg) for 17 h at room temperature was fractionated by electrophoresis at pH 6.5 and at 1.9. The *N*-terminal amino acid is shown in *italics*.

Peptide	Composition (molar ratios)
2Th1	Ser (1.0), Gly (2.2), <i>Ile</i> (1.8), Tyr (0.7)
2Th2	Cmc (0.5), Thr (1.9), Gly (1.3), Ala (1.3), Val (0.9), Lys (1.0)
2S1	Cmc (0.5), Thr (1.0), Gly (0.9), Ala (1.0), Lys (1.0)
2S2	Thr (1.0), Ser (1.0), Val (1.0)

Fig. 2. Sequence of peptide 2. The symbols are as used in Fig. 1. Th, Thermolysin; S, subtilisin.

Table 7. *Peptides obtained from digestion of peptide 3 with thermolysin and subtilisin*

The conditions of the digestion were the same as described for peptide 2 in Table 6. The digests were fractionated by electrophoresis at pH 6.5 and at 1.9. The compositions shown as (+) were determined by paper chromatography. Electrophoretic mobilities are relative to aspartic acid -1.00.

Peptide	Amino acid composition (molar ratios)	Mobility at pH 6.5	N-Terminus
3Th1	Cmc (1.1), Asp (1.3), Glu (1.2), Ala (2.0), Val (0.7)	-0.36	Val
3Th2	Tyr (+), Lys (+)	+0.46	Tyr
3S1	Glu (1.4), Ala (2.6), Val (1.0), Ile (1.0)	-0	Ile
3S2	Asp (+), Glu (+), Cmc (+), Tyr (+)	-0.34	Glx
3S3	Tyr (+), Lys (+)	+0.46	Tyr

sequence of this hexapeptide (the original sample) has now been studied by the dansyl–Edman method, which gave the first four residues; carboxypeptidase A released tryptophan (when 1 μ g of enzyme was used), and tryptophan, threonine and alanine (when 10 μ g of enzyme was used). These results give the sequence for peptide 9:

Asp-Cmc-Gly-Ala-Thr-Trp

The analysis of peptide 5 (Table 4) is incomplete, as most of the sample from the first digest was lost; the sample from the second digest, immediately before chromatography on Sephadex G-25, contained a radioactive peptide of mobility –0.2 (relative to aspartic acid –1); but after storage this peptide was no longer present, and there were two acidic peptides and one basic one. The intact peptide 5 had apparently decomposed (perhaps owing to bacterial decomposition or an oxidative cleavage). The mixture of peptides was fractionated by electrophoresis at pH 6.5, to give 5A, 5B and 5C (Table 9 and Fig. 5); other peptides (which we did not succeed in purifying) were also present, so that

Fig. 3. Sequence of peptide 3. The symbols and abbreviations are as used in Fig. 2.

we cannot be certain that peptide 5C represents the C-terminal part of peptide 5. Peptide 5B is the same as peptide 9 referred to above.

Cleavage with cyanogen bromide

Isolation and composition of fragments. Part of the treated protein (see the Materials and Methods section) was insoluble in the pH3.1 buffer; the amino acid analysis (Table 10) of this fraction (R) suggests that it contains about 130 residues. These include two residues of carboxymethylcysteine (in peptides 2 and 4), and two residues of tyrosine, one in peptide 2 and one in the ‘active site’ peptide (Waley, Miller, Rose & O’Connell, 1970). The soluble material was fractionated on Sephadex G-75

Fig. 4. Sequence of peptide 4. The symbols and abbreviations are as used in Fig. 2, except for the dotted bar: peptide 4P1 was probably a mixture of residues 1–9 and 1–11 (see the text).

Table 8. Peptides obtained by digestion of peptide 4 with thermolysin and with pepsin

The digestion with thermolysin was carried out with 10 μ g of enzyme and 0.6 μ mol of peptide 4 at 37°C for 3.5 h; the digest was fractionated by chromatography in solvent BAWP, followed by electrophoresis at pH 3.5. Peptic digestion was carried out with 4 μ g of enzyme and 0.1 μ mol of peptide 4 at room temperature for 18 h; the digest was fractionated by electrophoresis at pH 6.5, and the neutral peptides were purified by electrophoresis at pH 1.9. The last three peptides were obtained from thermolysin on peptide 4P1; the digest was fractionated at pH 6.5. The N-terminal amino acid is shown in *italics*.

Peptide	Amino acid composition (molar ratios)	Mobility at pH 6.5
4Th1	Ser(1.2), Glu(1.1), Gly(1.0), <i>Leu</i> (0.7)	–0.37
4Th2	Gly(1.0), <i>Leu</i> (1.0)	0
4Th3	Cmc(0.9), Ala(1.4), <i>Val</i> (1.0), Ile(0.9)	–0.4
4Th4	Glu(1.0), Gly(1.1), <i>Ile</i> (0.9), Lys(1.0)	0
4P1	Ser(0.7), Glu(1.0), Gly(1.4), Ala(1.7), <i>Val</i> (1.5), Leu(1.8), His(1.0)	0
4P2	Cmc(0.6), Ala(1.3), <i>Ile</i> (1.1)	–0.43
4P3	Glu(1.1), Gly(0.9), <i>Ile</i> (1.1), Lys(0.9)	0
4P4	Cmc(0.7), <i>Gly</i> (0.9), Ala(1.3), Val(0.9), Ile(1.2)	–0.55
4P1Th1	Ala(2.1), <i>Val</i> (0.9), His(1.0)	+0.4
4P1Th2	Ser(1.1), Glu(1.2), Gly(1.2), <i>Leu</i> (0.6)	–0.48
4P1Th3	Ser(1.2), Glu(1.1), Gly(1.2), <i>Leu</i> (1.6)	–0.39

Table 9. *Peptides derived from peptide 5*

The origin of peptides 5A, 5B and 5C is described in the text. Peptides 5BS were obtained by digestion of peptide 5B with 10 μ g of subtilisin (15h at room temperature), and purified by electrophoresis at pH 6.5 and 1.9. Peptides 5CS were obtained similarly from peptide 5C. The *N*-terminal amino acid is shown in *italics*.

Peptide	Amino acid composition (molar ratios)	Mobility at pH 6.5
5A	Cmc(0.3), <i>Asp</i> (1.2), Thr (1.3), Gly (1.2), Ala (1.1)	-0.78
5B	Cmc(0.5), <i>Asp</i> (1.1), Thr (1.1), Gly (1.0), Ala (1.3), Trp (+)	-0.53
5C	Ser (0.9), Glu (1.0), Gly (1.0), <i>Val</i> (2.0), Leu (1.0), Lys (1.0), Arg (1.0)	+0.25
5BS1*	Gly (+), Ala (+), <i>Asp</i> (+), Cmc (+)	-0.92
5BS2	Gly (+), Ala (+), <i>Asp</i> (+), Cmc (+), Thr (+)	-0.81
5CS1	<i>Ser</i> (0.6), Glu (1.2), Arg (1.2)	0
5CS2†	Glu (+), Ser (+), <i>Gly</i> (+), Arg (+)	

* Carboxypeptidase A released alanine.

† Aminopeptidase M released: Ser (1.0), Gly (1.0), Glu (0.2).

Asp-Cmc-Gly-Ala-Thr-Trp,Val-Val-Leu-Gly-Ser-Glu-Arg-Lys

Fig. 5. Sequence of fragments derived from peptide 5. The symbols are as used in preceding figures.

with the pH 3.1 buffer. Fragment M was eluted at $V_e/V_0 = 1.48$ (where V_e is the volume required to elute the fragment, and V_0 is the void volume measured with Dextran Blue). This value of V_e/V_0 was used to estimate the size of fragment M as follows. Lai (1968) separated the fragments from aldolase by this method, and a plot of the elution volumes in his experiments against the logarithm of the number of residues in the fragments is linear, and was used to estimate the size of fragment M as about 85–95 residues. The method is an approximate one, but the solvent used (30% in acetic acid) is likely to prevent aggregation and to disrupt secondary and tertiary structure. The amino acid analysis (Table 10) is consistent with this estimate, and there are two residues of tyrosine, one each in the cysteine peptides 1 and 3. A low-molecular-weight fragment (L) was also obtained; on further purification by electrophoresis at pH 6.5 this separated into two peptides, La and Lb, which had mobilities (relative to aspartic acid -1) of +0.65 and +0.5. These mobilities are consistent with molecular weights of about 1800, and charges of +4 and +3 (Offord, 1966); the amino acid analyses (Table 10) suggest that these are variants of the same peptide, one with *C*-terminal homoserine lactone, and the other with *C*-terminal homoserine (Ambler, 1965).

Peptide 'maps'. After digestion with trypsin, peptide 'maps' were prepared from the two large fragments (M and R). These (Fig. 6) confirmed that the fragments represented different parts of the chain, and that the insoluble fraction (R) was reasonably free from uncleaved chains. Fraction M contained cysteine peptides 1 and 3, and fraction R contained peptides 2 and 4, but the location of peptide 5 has not been determined.

Terminal residues. Digestion with carboxypeptidase A released glutamine from the insoluble fraction R (2mg of peptide and 20mg of urea were treated with 60 μ l of 1% ammonium hydrogen carbonate containing 40 μ g of carboxypeptidase A, and samples taken after 0.5h and 5h at room temperature), identified by electrophoresis at pH 1.9 and chromatography in solvent BAWP. Since glutamine is at the *C*-terminal of triose phosphate isomerase (Burton & Waley, 1968), R is the *C*-terminal fragment. The other fragments, Lb and M, released homoserine. The *N*-terminal residue of La and Lb was alanine, but the *N*-terminal residues of M and R could not be identified. The *N*-terminal residue in triose phosphate isomerase is alanine (Burton & Waley, 1968) and possible alignments are L-M-R or M-L-R. The former is consistent with the evidence, but the latter is not excluded. However, since *N*-terminal alanine was found from the whole protein, it might be expected to be detectable, if present, in fragment M.

DISCUSSION

Are the subunits polypeptide chains? The work on the cysteine peptides described above confirms that there are five unique residues of cysteine/molecule. There are ten residues of cysteine/molecule and the yields of the peptides are consistent with the two subunit structure. There are also two residues of methionine/subunit, and cleav-

Table 10. Amino acid composition of fragments obtained by the use of cyanogen bromide

The amino acid compositions are given as residues/fragment, derived from hydrolyses for 24 and 48 h for the two larger peptides; homoserine and tryptophan were not estimated, and are omitted from the totals. The rounded-off values in the last column ('chain') are from Table 2.

Amino acid	Fragment	...	...	La	Lb	M	R	Chain
Cmc						2.3	2.0	5
Asp				1.0	1.1	11.4	9.9	19
Thr						6.8	7.0	14
Ser				1.2	1.0	1.9	7.0	13
Glu						6.3	16	25
Pro				0.7	0.9	5.8	5	9
Gly				2.0	2.2	6.4	13.7	22
Ala				0.9	0.6	12.0	13.3	24
Val				0.9	1.1	6.8	14	22
Met								2
Ile						5.4	8.1	13
Leu						5.4	8.9	14
Tyr						1.9	2.0	3-4
Phe				1.9	2.3	3.0	3.9	7
Trp				+	+	—	+	4
Lys				1.7	1.7	8.2	9.7	20
His							2.0	4
Arg				1.0	1.1	2.6	4.3	7
Total				12	12	86	127	226

Fig. 6. Tryptic peptide 'map' of *S*-carboxymethylated triose phosphate isomerase (upper part), and of the two larger cyanogen bromide fragments (lower parts). Radioactive peptides are shown as solid areas. The peptides were separated by electrophoresis at pH 6.5 (horizontal direction); the acidic and basic peptides were then separated by chromatography in solvent BAWP (vertical direction); the neutral peptides were fractionated by electrophoresis at pH 1.9 before chromatography.

age with cyanogen bromide gave three fragments. These results lead to the conclusion that the subunits represent polypeptide chains. There is, moreover, strong independent confirmation of this view. Electrophoresis on polyacrylamide gels in the presence of sodium dodecyl sulphate, which is one of the more reliable methods of estimating the molecular weights of polypeptide chains, gave a value of 26 500 (Coulson *et al.* 1970). This value for the molecular weight of the polypeptide chain is (within the limits of experimental error) identical with the value for the molecular weight of the subunits. As far as we know, no other glycolytic enzyme has as short a chain as triose phosphate isomerase.

Minor components. Scopes (1968) detected several bands of enzyme activity on starch-gel electrophoresis of rabbit muscle extracts, and Krietsch *et al.* (1970) also found several bands in their isolated enzyme. Burton & Waley (1968) found one main band, representing about 85% of the protein, on starch-gel electrophoresis, and only one band on disc electrophoresis; similarly Coulson *et al.* (1970) found only one band on electrophoresis in the presence of sodium dodecyl sulphate. We have not detected variant structures in the cysteine peptides. It is possible that some of the minor components could represent variants in which an amide group has been converted into a carboxylic acid, as may happen with aldolase (Lai, Chen & Horecker, 1970).

This paper is a contribution from the Oxford Enzyme Group, which is supported by the Science Research Council. The support of the Medical Research Council is gratefully acknowledged. We are grateful to Dr R. E. Offord for the use of an amino acid analyser.

REFERENCES

- Ambler, R. P. (1963). *Biochem. J.* **89**, 349.
Ambler, R. P. (1965). *Biochem. J.* **96**, 32p.
Ambler, R. P. & Meadway, R. J. (1968). *Biochem. J.* **108**, 893.
Barman, T. E. & Koshland, D. E. (1967). *J. biol. Chem.* **242**, 5771.
Beale, D. (1967). *Biochem. J.* **103**, 129.
Beaven, G. H. & Holiday, E. R. (1952). *Adv. Protein Chem.* **7**, 320.
Burton, P. M. & Waley, S. G. (1966). *Biochem. J.* **100**, 702.
Burton, P. M. & Waley, S. G. (1968). *Biochem. J.* **107**, 737.
Coulson, A. F. W., Knowles, J. R. & Offord, R. E. (1970). *Chem. Commun.* p. 7.
Crowshaw, K., Jessup, S. J. & Ramwell, R. W. (1967). *Biochem. J.* **103**, 79.
Delaage, M. (1968). *Biochim. biophys. Acta*, **168**, 573.
Drapeau, G. R. & Yanofsky, C. (1967). *J. biol. Chem.* **244**, 5434.
Edelhoch, H. (1967). *Biochemistry, Easton*, **6**, 1948.
Grassman, W., Hannig, K. & Plöckl, M. (1955). *Hoppe-Seyler's Z. physiol. Chem.* **299**, 258.
Gray, W. R. (1967a). In *Methods in Enzymology*, vol. 11, p. 139. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
Gray, W. R. (1967b). In *Methods in Enzymology*, vol. 11, p. 469. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
Hartley, B. S. (1970). *Biochem. J.* **119**, 805.
Heilmann, J., Barollier, J. & Watzke, E. (1957). *Hoppe-Seyler's Z. physiol. Chem.* **309**, 219.
Hill, R. L. (1965). *Adv. Protein Chem.* **20**, 37.
Hirs, C. H. W. (1954). *J. biol. Chem.* **211**, 941.
Hoch, F. L. & Vallee, B. L. (1953). *Analyt. Chem.* **25**, 317.
Horton, H. R. & Koshland, D. E. (1967). In *Methods in Enzymology*, vol. 11, p. 566. Ed. by Hirs, C. H. W. New York: Academic Press Inc.
Johnson, L. N. & Waley, S. G. (1967). *J. molec. Biol.* **29**, 321.
Katz, E. P. (1968). *Analyt. Biochem.* **25**, 417.
Keutmann, T. & Potts, J. T. (1969). *Analyt. Biochem.* **29**, 175.
Krietsch, W. K. G., Pentchev, P. G., Klingenburg, H., Hofstatter, T. & Bucher, T. (1970). *Eur. J. Biochem.* **14**, 289.
Lai, C. Y. (1968). *Archs Biochem. Biophys.* **128**, 202.
Lai, C. Y., Chen, C. & Horecker, B. L. (1970). *Biochem. biophys. Res. Comm.* **40**, 461.
Meir, P. D. & Cotton, D. W. K. (1966). *Nature, Lond.*, **209**, 1022.
Offord, R. E. (1966). *Nature, Lond.*, **211**, 591.
Perham, R. N. (1969). *Biochem. J.* **111**, 17.
Previero, A., Colletti-Previero, M. A. & Cavadore, J. C. (1967). *Biochim. biophys. Acta*, **147**, 153.
Scopes, R. K. (1968). *Biochem. J.* **107**, 139.
Smillie, L. B. & Hartley, B. S. (1966). *Biochem. J.* **101**, 236.
Smith, I. (1953). *Nature, Lond.*, **171**, 43.
Spande, T. F., Wilchek, M. & Witkop, B. (1968). *J. Am. chem. Soc.* **90**, 3256.
Spies, J. R. & Chambers, D. C. (1949). *Analyt. Chem.* **21**, 1249.
Waley, S. G. (1965). *Biochem. J.* **96**, 722.
Waley, S. G. (1969). *Comp. Biochem. Physiol.* **30**, 1.
Waley, S. G., Miller, J. C., Rose, I. A. & O'Connell, E. L. (1970). *Nature, Lond.*, **227**, 601.
Waley, S. G., & Watson, J. (1953). *Biochem. J.* **55**, 328.
Wetlaufer, D. B. (1962). *Adv. Protein Chem.* **17**, 303.
Witkop, B. (1961). *Adv. Protein Chem.* **16**, 221.
Woods, K. R. & Wang, K. T. (1967). *Biochim. biophys. Acta*, **133**, 369.