The Comparative Structure of Mammalian Glyceraldehyde 3-Phosphate Dehydrogenases

By R. N. PERHAM

Department of Biochemistry, University of Cambridge

(Received 25 July 1968)

1. The amino acid sequences around the thiol groups of glyceraldehyde 3-phosphate dehydrogenase from badger and monkey skeletal muscle were compared with the sequences around the thiol groups in the enzyme isolated from other organisms. 2. Preliminary evidence of the existence of isoenzymes in the badger was obtained. Only the major form, however, could be purified completely. 3. The monkey enzyme contains only three cysteine residues per polypeptide chain compared with the four found in all the other mammalian enzymes so far examined, including that of badger, and the two in yeast. The missing thiol group in monkey was identified as residue 281 in the corresponding sequence of the pig enzyme. 4. These experiments rule out any essential role for cysteine-281 in the function of the mammalian enzymes. 5. Further evidence of the remarkable conservation of amino acid sequence in this enzyme during evolution is presented and discussed.

Thiol groups of GPDH* have, for many years, been implicated in the catalytic activity of the enzyme (Dixon & Lutwak-Mann, 1937; Velick, 1953; Velick & Furfine, 1963). More recently it has proved possible to identify chemically those thiol groups that react selectively in the native enzyme with either a substrate, *p*-nitrophenyl acetate, or an inhibitor, iodoacetic acid (Harris, Meriwether & Park, 1963; Perham & Harris, 1963). The same reactive thiol group was found to mediate the reaction in both cases. Other evidence has shown that the enzyme from a number of different organisms is composed of four identical polypeptide chains, each of molecular weight approx. 36 000 and each containing one reactive thiol group (Harris & Perham, 1963, 1965; Harrington & Karr, 1965). The molecule of active enzyme accordingly has a molecular weight of about 145 000 and contains four catalytic sites. The complete amino acid sequences of both lobster (Davidson, Sajgò, Noller & Harris, 1967) and pig (Harris & Perham, 1968) GPDH are now available.

One fruitful approach to an analysis of the mechanism of enzymic action has been a comparative study of the structure of the enzyme from a variety of organisms (Perham & Harris, 1963; Allison & Harris, 1965; Perham, 1966; Harris & Perham, 1968). This has revealed that the amino acid sequence of the polypeptide chains of the

enzyme has been remarkably conserved during evolution. For example, the enzymes from such widely separated sources as pig and lobster have 72% of their amino acid residues in identical sequence. Of particular interest is the observation that each protein chain from any of the mammalian enzymes yet examined contains a total of four thiol groups, whereas the yeast enzyme contains but two (Perham, 1964; Harris & Perham, 1965). If one assumes that the enzymes from disparate sources all share a basic common reaction mechanism, and the indications are that this is so, then it becomes clear that two of the four thiol groups in each polypeptide chain of the mammalian enzymes cannot play a fundamental part in the catalysis effected by the enzyme. This is not to deny those thiol groups some secondary role, perhaps facilitation of the operation of the enzyme in a mammalian environment: in fact, this might be a possible role, in view of the conservation of their presence during evolution of the mammalian enzymes.

Reference to Table 1 shows that only limited amino acid-sequence alterations have been found in the vicinity of the reactive cysteine residue of enzymes from a large number of sources. To extend those observations to cover other species closely related in evolution to the ones previously examined, preparations of the enzyme from monkey and badger have now been investigated (Perham, 1968). These studies have revealed that the monkey and badger enzymes, unlike that of man, share the identical heptadecapeptide sequence

* Abbreviation: GPDH, glyceraldehyde 3-phosphate dehydrogenase [D-glyceraldehyde 3-phosphate-NAD oxidoreductase (phosphorylating), EC 1.2.1.12].

Table 1. *Amino acid sequence around the reactive cysteine residue of GPDH from a number of sources*

The residues in *italics* are the only ones in which variations have been found in the proteins examined. The reactive cysteine residue is indicated in **bold type**. References: ^a Harris *et al.* (1963); ^b Perham & Harris (1963); ^c Perham (1966); ^d Allison & Harris (1965).

Species	Sequence																	
Rabbit ^a	}	<i>Ile - Val-Ser-Asn-Ala-Ser-Cys-Thr-Thr-Asn-Cys-Leu-Ala-Pro-Leu-Ala-Lys</i>																
Pig ^b		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Yeast ^b																		
Ox ^c																		
Chicken ^d																		
Ostrich ^d	}	<i>Ile - Ile - Ser-Asn-Ala-Ser-Cys-Thr-Thr-Asn-Cys-Leu-Ala-Pro-Leu-Ala-Lys</i>																
Sturgeon ^d																		
Human ^d																		
Halibut ^d		<i>Val - Val-Ser-Asn-Ala-Ser-Cys-Thr-Thr-Asn-Cys-Leu-Ala-Pro-Leu-Ala-Lys</i>																
Lobster ^d		<i>Asp-Met-Thr-Val-Val-Ser-Asn-Ala-Ser-Cys-Thr-Thr-Asn-Cys-Leu-Ala-Pro-Val-Ala-Lys</i>																

around the reactive cysteine residues found in all the other mammalian proteins, but that the monkey enzyme has a total of only three thiol groups per polypeptide chain in contrast with the four regularly found in other mammalian enzymes. The missing thiol group has been identified and this, together with some other amino acid-sequence changes in these enzymes, is also described.

MATERIALS AND METHODS

Enzymes. The monkey GPDH used in these experiments was an acetone-dried powder prepared from rhesus monkeys (*Macaca mulata*), generously given by Mrs E. Szorenyi.

The badger enzyme was prepared from skeletal muscle taken from a single badger by a slight modification of the method of Elödi & Szorenyi (1956). Minced badger skeletal muscle was extracted once with 1 vol. of cold glass-distilled water containing 2 mM-EDTA and squeezed through two layers of muslin at 2°. Solid (NH₄)₂SO₄ was added to the filtrate to a final concentration of 70% of saturation, and the precipitate was removed by filtration on Whatman no. 3 paper. Assay by the method of Allison & Kaplan (1964) showed that about 30% of the total enzyme activity was precipitated with 70%-saturated (NH₄)₂SO₄; the remainder was not precipitated until 80% saturation was reached. The enzyme precipitated at 80% saturation was freed from contaminating haem-proteins by gel filtration through Sephadex G-100 at 2° in 5 mM-tris-HCl buffer, pH 8.5, containing 2 mM-EDTA and 10 mM-2-mercaptoethanol, and recrystallized from unbuffered 80%-saturated (NH₄)₂SO₄ containing 2 mM-EDTA in the cold. The yield was approx. 500 mg. from 2 kg. of muscle. The enzyme from the precipitate with 70%-saturated (NH₄)₂SO₄ proved impossible to purify completely from the mass of contaminating protein. All attempts to purify the enzyme from both the 70%- and 80%-saturation precipitates were hindered by an unfortunate spontaneous loss in activity after chromatography and crystallization. In this respect the badger enzyme resembles the yeast enzyme, which also rapidly loses activity during preparation and storage (Taylor, Meriwether & Park, 1963; Harris & Perham, 1965).

Trypsin (twice recrystallized) was purchased from

Worthington Biochemical Corp., Freehold, N.J., U.S.A. Subtilisin was obtained from British Drug Houses Ltd., Poole, Dorset.

Paper electrophoresis and chromatography. Details of the conditions used for paper electrophoresis and chromatography were given by Perham (1967). The electrophoretic mobility of peptides at pH 6.5 was used to assign amide groups to aspartic acid and glutamic acid residues.

Amino acid analysis. Peptides and proteins were hydrolysed and analysed for amino acids as described by Perham (1967). Performic acid oxidation was carried out by the method of Hirs (1956).

Reaction with iodo[2-¹⁴C]acetic acid. Samples of GPDH were treated with iodo[2-¹⁴C]acetic acid in 8 M-urea as described by Harris & Perham (1965).

Peptide 'maps' and radioautographs. Proteins were digested with 1% (w/w) trypsin in 0.5% (w/v) NH₄HCO₃, pH 8.0, at 37° for 4 hr., and the buffer was removed by freeze-drying. Peptide 'maps' and radioautographs were then prepared as described by Harris & Perham (1965). Peptides (0.2 μmole) were digested with subtilisin (50 μg.) in a similar fashion.

N-Terminal residues of peptides. N-Terminal residues of peptides were identified by the 'dansyl' technique (Gray & Hartley, 1963a). Amino acid sequences were established by means of the 'dansyl'-Edman degradation (Gray & Hartley, 1963b).

RESULTS

Samples of the monkey and badger enzymes were allowed to react with iodo[2-¹⁴C]acetic acid in the presence of 8 M-urea. The S-carboxymethylated proteins were digested with trypsin, and peptide 'maps' and radioautographs were then prepared (Fig. 1). The badger enzyme preparation was found to have a pattern of radioactive tryptic peptides identical with that described for the pig and other mammalian enzymes (Harris & Perham, 1965; Perham, 1964), indicating an identical distribution of four enzyme thiol groups. The monkey enzyme, however, has one less radioactive peptide, the missing peptide corresponding to

Fig. 1. Radioactive peptides in the tryptic digest of *S*-carboxy[¹⁴C]methylated GPDH. Electrophoresis was at pH 6.5 and then at pH 3.5 at 60v/cm. for 1 hr. ↑, Origin. (a) GPDH from badger muscle; (b) GPDH from monkey muscle.

peptide T2a derived from cysteine-281 in the pig enzyme (Harris & Perham, 1968). Hence only three unique thiol groups seem to be present in the monkey-enzyme polypeptide chain. This result was confirmed by amino acid analysis of the performic acid-oxidized proteins. Table 2 lists the amino acid composition of the monkey and badger enzymes, with that of the pig enzyme shown alongside for comparison. All analyses are given for molecular weight 36 000. The major significant difference in the analyses is in the cysteic acid content. This close similarity in amino acid composition of mammalian GPDH enzymes has been observed before and is reflected in the near-identity of the peptide 'maps' of the pig, rabbit and ox enzymes (Perham, 1964; Harris & Perham, 1965).

Those peptides containing *S*-carboxymethyl-cysteine in the tryptic digest of the monkey enzyme were isolated by gel filtration on Sephadex G-25 and high-voltage paper electrophoresis by the methods

Table 2. Amino acid composition of performic acid-oxidized GPDH from badger and monkey skeletal muscle

The values given are the means of two analyses and are expressed for molecular weight 36 000. The protein samples were hydrolysed with 6*N*-HCl for 24 hr. at 105° in sealed evacuated tubes. The composition of the pig enzyme is taken from the sequence given by Harris & Perham (1968). CySO₃H, Cysteic acid.

Amino acid	Species ...	Amino acid composition (residues/mol. of mol.wt. 36 000)		
		Monkey	Badger	Pig
CySO ₃ H		2.9	3.7	4
Asp		42.6	40.3	38
Thr*		20.0	19.3	22
Ser*		20.1	19.9	19
Glu		20.5	20.3	18
Pro		14.0	11.2	12
Gly		34.3	32.4	32
Ala		32.2	32.5	32
Val†		28.1	28.7	34
Met		8.7	8.4	9
Ile†		17.6	17.3	21
Leu		20.1	18.0	18
Tyr		8.4	8.4	9
Phe		14.3	14.0	14
Lys		26.6	27.0	26
His		10.5	9.1	11
Arg		10.0	10.0	10
Trp‡				3

* Corrected for destruction during acid hydrolysis.
† Incompletely released in 24 hr. hydrolysis.
‡ Destroyed during performic acid oxidation.

described for the pig enzyme (Harris & Perham, 1965; Perham, 1964). The amino acid composition of peptide T4a was found to be identical with that of peptide T4a from pig GPDH (Perham & Harris, 1963). Peptide T4a was digested with subtilisin and the resulting peptide fragments were isolated by successive electrophoreses at pH 3.5 and pH 2.0. They were characterized as described by Perham & Harris (1963) and found to be identical with those derived from the subtilisin digest of peptide T4a from pig GPDH. This showed that peptides T4a from monkey and pig GPDH share an identical amino acid sequence. The amino acid composition of peptide T4b from monkey GPDH was found to be: cysteic acid (1.04), Asp (1.94), Thr (1.95), Ser (0.94), Pro (1.04), Ala (0.95), Val (3.80), Leu (1.01), Arg (0.99). This is identical with that of peptide T4b from the pig enzyme with the exception of the appearance of one residue of alanine and the loss of one residue of proline, suggesting a proline-alanine replacement. This was confirmed by showing, by four steps of the 'dansyl'-Edman

Table 3. Amino acid sequences of the tryptic peptides containing the three cysteine residues of monkey GPDH, together with the corresponding peptides from pig and badger GPDH

CMC, S-Carboxymethylcysteine.		
Peptide	Species	Sequence
T4a	Monkey	Ile-Val-Ser-Asn-Ala-Ser-CMC-Thr-Thr-Asn-CMC-Leu-Ala-Pro-Leu-Ala-Lys
	Pig	
	Badger	
T4b	Pig	Val-Pro-Thr-Pro-Asn-Val-Ser-Val-Val-Asp-Leu-Thr-CMC-Arg
	Badger	
	Monkey	Val-Pro-Thr-Ala-Asn-Val-Ser-Val-Val-Asp-Leu-Thr-CMC-Arg

degradation, that the *N*-terminal sequence of peptide T4b from the monkey enzyme is Val-Pro-Thr-Ala-Asn-, whereas that from pig is Val-Pro-Thr-Pro-Asn-. Further evidence was obtained from a comparison of the subtilisin-digestion products of peptides T4b from monkey and pig GPDH (Perham, 1964). These were identical with one exception, the peptide Val-Pro-Thr-Pro-Asn from pig peptide T4b being replaced by the peptide Val-Pro-Thr-Ala and free asparagine from monkey peptide T4b. The sequences deduced for peptides T4a and T4b from the monkey enzyme are shown in Table 3 together with the corresponding peptides from the pig enzyme. Similar experiments were also undertaken with peptides T4a and T4b from badger GPDH. No differences could be found between the badger and the pig peptides; these results are also given in Table 3.

DISCUSSION

The appearance of two peaks of enzyme activity during ammonium sulphate fractionation could indicate the presence of isoenzymes of GPDH in the badger and, indeed, preliminary experiments with starch-gel electrophoresis did indicate differing mobilities for the two peaks of activity when examined by the technique of Fine & Costello (1963). However, it would be premature to draw a firm conclusion until both fractions have been completely purified and other badgers examined in detail. Rigorous evidence for the existence of isoenzymes in yeast GPDH has, however, recently been obtained (W. S. Allison, personal communication), and slight differences in the amino acid sequence of the polypeptide chains of the yeast enzyme have also been found (Jones & Harris, 1968). In addition, electrophoretically distinct forms of GPDH have recently been detected in a variety of other organisms (Lebherz & Rutter, 1967).

The replacement of proline by alanine (a single base change in the *Escherichia coli* code; Crick, 1966) at position 4 in peptide T4b of the monkey enzyme is the first sequence change recorded in this

region of the molecule among the mammalian GPDH enzymes examined (pig, ox, rabbit, monkey and badger). It seems likely that the very high degree of homology in amino acid sequence of these enzymes would militate against any resultant gross change of three-dimensional structure. Further progress in the crystallographic analysis of the enzyme structure (Watson & Banaszak, 1964) is desirable for more definitive interpretation of chemical analyses such as these. Such knowledge would also help to determine the significance of the replacement (or deletion) of a thiol group (cysteine-281) in the monkey enzyme. Intramolecular acyl-transfer reactions between the reactive cysteine-149 (in peptide T4a) and cysteine-281 (in peptide T2a) have been detected in the rabbit enzyme (Park, Meriwether & Halcomb, 1967), suggesting that those segments of the polypeptide chain are ordinarily in relatively close proximity at the catalytic site in the mammalian enzymes (although not necessarily contributed by the one sub-unit polypeptide chain; Perham, 1966). If this were so, an apparently redundant thiol group could perhaps play some useful minor role in the catalytic activity of the enzyme.

These experiments reinforce the conclusions of extensive previous work on the primary structure of various GPDH enzymes, that the sequence homology in these proteins is truly remarkable. One is led to enquire why this should be. Is it a reflection of the central role of this enzyme in glycolysis and its high specificity for substrates and coenzyme, or is it a general property of glycolytic enzymes? Similar experiments on other enzymes of the glycolytic pathway would be of great value in answering this question.

One other problem remains unsolved and this is the nature of the selection pressure that causes stabilization of an altered sequence for the same enzyme in different organisms. It is known that *in vitro* the turnover numbers of the individual GPDH enzymes are more or less identical (Allison & Kaplan, 1964) and the assay *in vitro* therefore gives little clue to the functional significance of the

amino acid-sequence changes noted. Experience with haemoglobin has shown that many amino acid-sequence changes can be accommodated in the surface of the fundamental tetrameric structure of the molecule, without causing noticeable changes in the properties of the protein *in vitro* (Perutz & Lehmann, 1968). It is unclear what pressures select and maintain these altered structures: perhaps it is necessity to interact with other macromolecules in which there are compensatory structural changes. Why there are so few changes in the mammalian (and other) GPDH enzymes therefore becomes a question of some importance. At the moment it seems impossible to give a convincing answer in terms of protein structure.

I am indebted to Dr J. I. Harris for discussion and permission to quote unpublished results. I am grateful to Mr S. van Heyningen for much valuable help in preparing the badger enzyme and to Dr C. J. R. Thorne for the amino acid analyses. I thank Mrs E. Szorenyi for the generous gift of monkey GPDH.

REFERENCES

- Allison, W. S. & Harris, J. I. (1965). *Abstr. 2nd Meet. Europ. biochem. Soc., Vienna*, p. 140.
- Allison, W. S. & Kaplan, N. O. (1964). *J. biol. Chem.* **239**, 2140.
- Crick, F. H. C. (1966). *Cold Spr. Harb. Symp. quant. Biol.* **31**, 1.
- Davidson, B. E., Sajgó, M., Noller, H. F. & Harris, J. I. (1967). *Nature, Lond.*, **216**, 1181.
- Dixon, M. & Lutwak-Mann, C. (1937). *Biochem. J.* **31**, 1347.
- Elődi, P. & Szorenyi, E. (1956). *Acta physiol. Acad. sci. hung.* **9**, 339.
- Fine, I. H. & Costello, L. A. (1963). In *Methods in Enzymology*, vol. 6, p. 958. Ed. by Colowick, S. P. & Kaplan, N. O. New York: Academic Press Inc.
- Gray, W. R. & Hartley, B. S. (1963a). *Biochem. J.* **89**, 59 p.
- Gray, W. R. & Hartley, B. S. (1963b). *Biochem. J.* **89**, 379.
- Harrington, W. F. & Karr, G. M. (1965). *J. molec. Biol.* **13**, 885.
- Harris, I., Meriwether, B. P. & Park, J. H. (1963). *Nature, Lond.*, **197**, 154.
- Harris, J. I. & Perham, R. N. (1963). *Biochem. J.* **89**, 59 p.
- Harris, J. I. & Perham, R. N. (1965). *J. molec. Biol.* **13**, 876.
- Harris, J. I. & Perham, R. N. (1968). *Nature, Lond.*, **219**, 1025.
- Hirs, C. W. (1956). *J. biol. Chem.* **219**, 611.
- Jones, G. M. T. & Harris, J. I. (1968). *Abstr. 5th Meet. Europ. biochem. Soc., Prague*, p. 185.
- Lebherz, H. G. & Rutter, W. J. (1967). *Science*, **157**, 1198.
- Park, J. H., Meriwether, B. P. & Halcomb, S. (1967). *Abstr. 7th int. Congr. Biochem., Tokyo*, F-283.
- Perham, R. N. (1964). Ph.D. Thesis: University of Cambridge.
- Perham, R. N. (1966). *Biochem. J.* **99**, 14 c.
- Perham, R. N. (1967). *Biochem. J.* **105**, 1203.
- Perham, R. N. (1968). *Abstr. 5th Meet. Europ. biochem. Soc., Prague*, p. 194.
- Perham, R. N. & Harris, J. I. (1963). *J. molec. Biol.* **7**, 316.
- Perutz, M. F. & Lehmann, H. (1968). *Nature, Lond.*, **219**, 902.
- Taylor, E. L., Meriwether, B. P. & Park, J. H. (1963). *J. biol. Chem.* **238**, 734.
- Velick, S. F. (1953). *J. biol. Chem.* **203**, 563.
- Velick, S. F. & Furfine, C. (1963). In *The Enzymes*, vol. 7, p. 243. Ed. by Boyer, P. D., Lardy, H. & Myrback, K. New York: Academic Press Inc.
- Watson, H. C. & Banaszak, L. J. (1964). *Nature, Lond.*, **204**, 918.