AMINO ACID SEQUENCE HOMOLOGY IN THE ACTIVE SITE OF RABBIT AND STURGEON MUSCLE ALDOLASES

I. GIBBONS, P.J. ANDERSON and R.N. PERHAM

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, England

Received 21 July 1970

1. Introduction

In a previous study [1], we showed that the primary structure of fructose-1,6-diphosphate aldolase (EC 4.1.2.13) from a variety of mammalian muscles was unusually highly conserved, the homology being comparable to that found for another glycolytic enzyme, glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) [2]. We have now shown by means of amino acid sequence analysis around the substrate-binding lysine residue in the active site that the aldolase from muscle from a single sturgeon is also highly homologous with the aldolases from rabbit muscle and rabbit liver, at least in this region. The extent of the homology is again remarkably high for a protein isolated from such phylogenetically distinct sources, suggesting that such homology may be a common feature of enzymes of the glycolytic pathway.

During the course of this work and also of a study of the reactivity of the thiol groups of rabbit muscle aldolase [3], it became apparent that there were slight discrepancies between our results and the amino acid sequence around the substrate-binding lysine residue of the rabbit muscle enzyme previously established elsewhere [4,5]. We have therefore reinvestigated that sequence and present it here in comparison with the sturgeon enzyme.

2. Materials and methods

Aldolase from rabbit and a single sturgeon muscle was prepared as described earlier [6] and *S*-carboxymethylated with 2-^{14}C-iodo acetic acid in 8 M urea [1]. Enzymic digestion and cyanogen bromide cleavage, amino acid analysis, and *N*-terminal analysis by the dansyl technique were carried out as previously described [1]. Separation of peptides by paper electrophoresis and chromatography was performed according to Perham and Jones [7] and dansyl-Edman degradation of peptides was carried out essentially as described by Gray and Hartley [8]. Amide assignments in peptides were made from the electrophoretic mobility (m) at pH 6.5 [9], defining the mobility of aspartic acid as −1.00.

The peptides produced by cyanogen bromide cleavage of the *S*-carboxymethylated enzymes were, at the outset of this investigation, fractionated by gel filtration on Sephadex G-75 in pyridine-acetic acid buffers [1,10]. More recently it has been found convenient to render the peptides more soluble before fractionation by reversible reaction of the amino groups with citraconic anhydride [6,11]. The citraconylated peptides were then separated by gel filtration on Sephadex G-75 (105 cm × 2 cm) in 0.5% NH_4HCO_3, pH 8.0, recovered by freeze-drying and the citraconyl groups removed by incubating the peptides in 5% formic acid at 20° for 10 hr. Additional gel filtration on Sephadex G-50 or G-100 was required to purify the larger citraconylated peptides from sturgeon muscle aldolase (I. Gibbons, unpublished work).

Acrylamide gel electrophoresis in the presence of sodium dodecyl sulphate was carried out as described by Shapiro and Maizel [12].

3. Results

The elution profile on Sephadex G-75 of the cyanogen bromide cleavage products of *S*-carboxymethylated rabbit and sturgeon muscle aldolases are substantially different, although both digests contain a peptide (X4) of 18 residues eluting in identical positions [1]. The amino acid sequence of this peptide from rabbit has already been reported [5] and shown to derive from close to the substrate-binding lysine residue in the active site. The peptide SX4, from sturgeon had the composition: Lys_2, Cmc_1*, Thr_3, Ser_1, Glu_3, Gly_1, Ala_3, Val_1, Ile_1, Tyr_1, Hsr_1 and dansylation established the *N*-terminal residue as valine. Tryptic digestion of peptide SX4 gave the following fragments: SX4T1, (m=−0.27), with the composition Thr_1, Ser_1, Glu_2, Ala_1, Ile_1, Tyr_1, Hsr_1; SX4T2b, (m=0), with the composition Lys_1, Cmc_1, Thr_2, Glu_1, Gly_1, Ala_2, Val_1; SX4T2c, (m=0), with the composition Lys_1, Thr_1, Ser_1, Glu_2, Ala_1, Ile_1, Tyr_1, Hsr_1; and SX4T3, (m=0.91), which was free lysine. After chymotryptic digestion, a peptide SX4C1 with the composition Thr_1, Ser_1, Glu_2, Ala_1, Ile_1, Hsr_1 was isolated. The sequences of these peptides were established by dansyl-Edman degradation and carboxypeptidase digestion and can be fitted together to give a unique sequence for peptide SX4 as shown in fig. 1. Since peptide SX4T2b is neutral at pH 6.5, the glutamic acid residue therein must be present as glutamine. Carboxypeptidase digestion of peptide SX4C1 gave a peptide SX4C1K1, (m=−0.45), having the composition Thr_1, Ser_1, Glu_2 the electrophoretic mobility of which is consistent with only one of the glutamyl residues being amidated. Examination of the electrophoretic mobility of the peptide after successive rounds of Edman degradation showed the sequence of this peptide to Thr-Ser-Gln-Glu. The amide distribution in peptide SX4 follows.

The rabbit peptide, RX4, had the following composition: Lys_1, His_2, Cmc_1, Thr_2, Ser_1, Glu_3, Pro_1, Gly_1, Ala_2, Val_1, Ile_1, Tyr_1, Hsr_1 and dansylation gave the *N*-terminal residue as valine. After tryptic digestion, the following fragments were isolated: RX4T1, (m=−0.27), composition His_1, Ser_1, Glu_2, Ala_1, Ile_1, Tyr_1, Hsr_1 and RX4T2, (m=+0.28), composition Lys_1, His_1, Cmc_1, Thr_2, Glu_1, Pro_1, Gly_1, Ala_1, Val_1. Chymotryptic digestion of peptide RX4 gave a peptide RX4C1, (m=−0.30), composition His_1, Ser_1, Glu_2, Ile_1, Ala_1, Hsr_1. Again, dansyl-Edman degradation and carboxypeptidase digestion gave sequences that together formed a unique sequence for peptide RX4 (fig. 1). The glutamyl residue in peptide RX4T2 must be amidated since the peptide is basic at pH 6.5. Carboxypeptidase digestion of peptide RX4C1 gave a peptide RX4C1K1, (m=−0.39), composition His_1, Ser_1, Glu_2. Since peptide RX4C1K1 is acidic, both glutamyl residues cannot be amidated. Examination of the mobility of the peptide RX4C1K1 after successive rounds of Edman degradation showed that its sequence must be Ser-His-Glu-Glu. The amide assignments in peptide RX4 are therefore as shown.

Comparison of the amino acid sequence of the rabbit peptide RX4 with the amino acid sequence around the substrate-binding lysine residue of rabbit muscle aldolase [4] has shown that peptide RX4 occurs in the primary structure on the *C*-terminal side of and close to that lysine residue [5] (see fig. 3). Since the rabbit and sturgeon peptides RX4 and SX4 are clearly closely homologous (fig. 1), we thought it reasonable to suppose that the rabbit and sturgeon muscle aldolases would also be homologous around the reactive lysine residue. The tryptic peptides of the separated cyanogen bromide fragments of the *S*-carboxymethylated sturgeon enzyme were therefore examined to see whether a peptide containing the substrate-binding lysine residue could be isolated, assuming the homology with the rabbit enzyme sequence would be sufficient to identify it. As a check, the corresponding tryptic peptide from the rabbit enzyme was also isolated for comparison, a simple matter since its amino acid sequence was known [4] and it had been shown to be part of the rabbit cyanogen bromide fragment X3 [13]. The initial search was made easier by the presence of two residues of histidine in the rabbit peptide that acted as a 'handle' on the peptide during purification and which, happily, by their presence in the corresponding sturgeon peptide also helped identify it. The amino acid composition of the appropriate rabbit peptide, RX3T5, was found to be Lys_1, His_2, Asp_2, Thr_1, Ser_1, Glu_1, Pro_1, Gly_1, Ala_1, Ile_1, Leu_4, Tyr_1, Hsr_1 and the *N*-terminal residue shown to be alanine by the dansyl procedure.

* *Abbreviations:*
Cmc: *S*-carboxymethylcysteine;
Hsr: Homoserine.

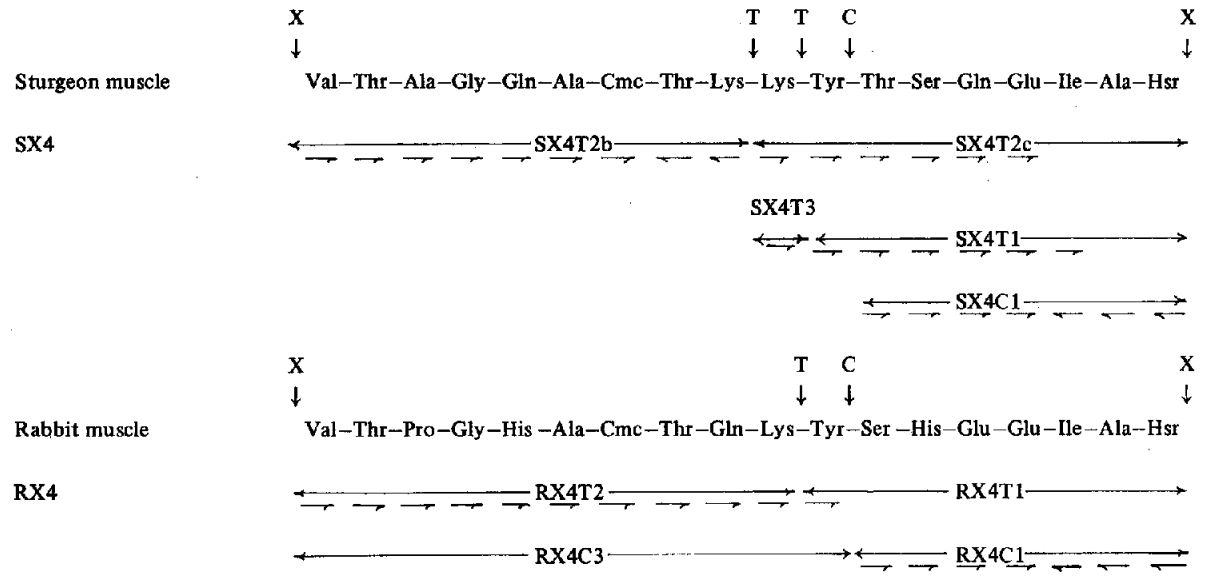

Fig. 1. The amino acid sequences of peptides SX4 and RX4 from sturgeon and rabbit muscle aldolases. – denotes a residue established by the dansyl-Edman degradation; – denotes a residue established by carboxypeptidase digestion; X↓, T↓, and C↓ indicate positions of cleavage by cyanogen bromide, trypsin and chymotrypsin respectively.

This composition is slightly different from that predicted from the published sequence evidence [4,5] in that it contains additional residues of aspartic acid, proline and homoserine. Since homoserine must be *C*-terminal in a peptide derived by cyanogen bromide cleavage, lysine cannot be the *C*-terminal residue in this tryptic peptide. After chymotryptic digestion of peptide RX3T5, the following peptides were separated by electrophoresis at pH 6.5 RX3T5C1, (m=−0.35) with the composition Thr_1, Glu_1, Gly_1, Leu_3; RX3T5C3, (m=+0.22), composition His_2, Asp_1, Ser_1, Ala_1, Ile_1, Leu_1, Tyr_1; RX3T5C4 (m= 0.38), composition Lys_1, Asp_1, Pro_1, Leu_1, Hsr_1; RX3T5C5, (m=+0.45), composition Lys_1, Asp_1, Pro_1, Hsr_1. The amino acid sequences of these peptides were established by means of the dansyl-Edman degradation and carboxypeptidase digestion, enabling a unique sequence for peptide RX3T5 to be formulated (fig. 2). The amide assignments are made in accord with the mobilities of the individual chymotryptic peptides at pH 6.5. In addition, treatment of peptide RX3T5C5 with carboxypeptidase left a peptide with the composition Lys_1, Asp_1, Pro_1 and mobility m=+0.53 at pH 6.5, confirming that the aspartic acid residue must be present as asparagine.

A peptide SXb1T3 with an electrophoretic mobility identical to that of the rabbit peptide RX3T5 was readily detected in the tryptic digest of the sturgeon cyanogen bromide fragment SXb1. This peptide had an identical composition to that of peptide RX3T5 with the exception of valine replacing isoleucine: the *N*-terminal residue was also alanine. The amino acid sequence was established in an identical fashion to that used for peptide RX3T5, as summarized in fig. 2. The sturgeon chymotryptic peptide SXb1T3C3 which extends from residues 1–8, on digestion with pronase gave a peptide SXb1T3C3R1 (m= 0.00) with the composition His_1, Asp_1, Ser_1, Ala_1, Leu_1. Thus the aspartyl residue at position 4 is assumed to be the free acid. The rabbit peptide RX3T5C3 corresponding to SXb1T3C3 had exactly the same electrophoretic mobility at pH 6.5 (+ 0.23) and is therefore also assumed not to have its aspartyl residue amidated.

In separate experiments, the *S*-carboxymethylated sturgeon enzyme was electrophoresed in the sodium dodecyl sulphate-acrylamide gel system of Shapiro and Maizel [12], using bovine serum albumin, rabbit glyceraldehyde-3-phosphate dehydrogenase and egg-white lysozyme as markers. The interpolated molecular weight was 42,000.

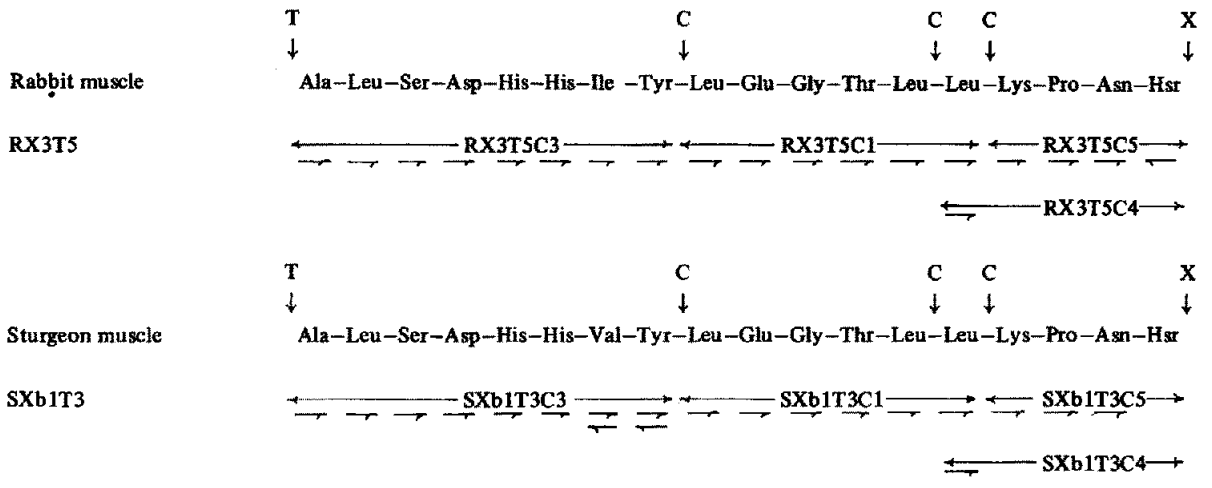

Fig. 2. The amino acid sequence of peptides RX3T5 and SXb1T3 from rabbit and sturgeon muscle aldolases.

4. Discussion

It has already been established [4,5] that in the rabbit muscle aldolase peptides RX3T5 and RX4 adjoin one another in the primary structure as shown in fig. 3. The extensive sequence homology demonstrated in this study strongly suggests that the same is true for the corresponding peptides derived from the sturgeon enzyme (fig. 3) and that the lysine residue at position 15 is involved in Schiff base formation with substrate. Structural studies on the sturgeon enzyme after inactivation by reduction with sodium borohydride in the presence of substrate confirm that this is so (P.J. Anderson, unpublished work). There are, however, several discrepancies between the amino acid sequence for the rabbit muscle enzyme derived here and that established elsewhere [4,5]. Thus the amide assignments as positions 4, 10, 32 and 33 are made differently and the sequence -Pro-Asn- at positions 16–17 is inverted. In fact, the sequence reported here would explain why trypsin fails to cleave at Lys-15 in peptide RX3, since it is well known that Lys-Pro bonds are not normally susceptible to tryptic attack. Moreover, it would also explain the presence of the peptide derived from positions 18–28 in tryptic digests of the intact *S*-carboxymethylated enzyme, an anomaly to which we have already drawn attention [3], on the reasonable supposition that the Asn-Met bond at position 17–18 is susceptible to chymotryptic-like cleavage during tryptic digestion.

The impressive sequence homology around the active lysine residue in the aldolases from rabbit muscle, sturgeon muscle, and rabbit liver [14] is clearly visible in fig. 3, providing further evidence for their derivation from a common ancestral gene [14]. (In view of the present evidence it may well be that in the liver enzyme positions 4 and 10 should be glutamic acid and the sequence at position 16–17 should be -Pro-Asn-.) All the variations observed can be accounted for by single base changes in the *E. coli* code [15]. This homology between proteins from such phylogenetically distinct sources is unusually high and recalls that previously found for glyceraldehyde 3-phosphate dehydrogenase [2,16], in accord with the suggestion that such homology may be a common feature of enzymes of the glycolytic pathway [1]. The similarities in primary structure would suggest that the aldolases have very similar tertiary structures. The results of ultracentrifugation [1,17,18] and the gel electrophoresis of the sturgeon enzyme indicate that these enzymes are all tetrameric with a subunit molecular weight of approximately 40,000. It is reasonable to suppose therefore that these aldolases will have very similar three dimensional structures.

Although the specificity of the rabbit liver and muscle enzymes are somewhat different [19] and preliminary experiments show that the sturgeon muscle enzyme closely resembles the rabbit muscle enzyme in this respect (I. Gibbons, unpublished work), it is interesting to note that this difference is not ob-

	1	2	3	4	5	6	7	8	9	9a	10	11	12	13	14	15	16	17	18
Rabbit liver	Ala	Leu	Asn	Asn	His	His	Val	Tyr	Leu	Ser	Gln	Gly	Thr	Leu	Leu	LYS	Asn	Pro	Met
Rabbit muscle	Ala	Leu	Ser	Asp	His	His	Ile	Tyr	Leu	–	–	Glu	Gly	Thr	Leu	Leu	LYS	Pro	Asn	Met
Sturgeon muscle	Ala	Leu	Ser	Asp	His	His	Val	Tyr	Leu	–	–	Glu	Gly	Thr	Leu	Leu	LYS	Pro	Asn	Met

	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Rabbit liver	Val	Thr	Ala	Gly	His	Ala	Cmc	Thr	Lys									
Rabbit muscle	Val	Thr	Pro	Gly	His	Ala	Cmc	Thr	Gln	Lys	Tyr	Ser	His	Glu	Glu	Ile	Ala	Met
Sturgeon muscle	Val	Thr	Ala	Gly	Gln	Ala	Cmc	Thr	Lys	Lys	Tyr	Thr	Ser	Gln	Glu	Ile	Ala	Met

Fig. 3. The amino acid sequence around the substrate-binding lysine residue of aldolase from rabbit muscle, rabbit liver and sturgeon muscle. The rabbit liver sequence is that proposed by Morse and Horecker [14]. The substrate binding residue is at position 15.

viously reflected in the amino acid sequences around the active lysine residue. Further comparison of the primary structures of these enzymes will be of great interest in this respect. It is also important to note that the sturgeon aldolase used in these experiments was isolated from a single sturgeon fished in the Ottawa River, unlike that used in our earlier experiments [1], which was isolated from a single sturgeon fished in the North Sea. Work in progress on muscle aldolase from other sturgeons suggests that there are sequence differences in sturgeon aldolases (perhaps species differences) that may well correlate with the natural habitat of the sturgeon being salt water or fresh water.

Acknowledgements

We are glad to acknowledge the award of a Canadian Medical Research Council Fellowship to P.J.A. and of a Science Research Council Studentship to I.G.

References

[1] P.J. Anderson, I. Gibbons and R.N. Perham, European J. Biochem. 11 (1969) 503.
[2] R.N. Perham, Biochem. J. 111 (1969) 17.
[3] P.J. Anderson and R.N. Perham, Biochem. J. 117 (1970) 291.
[4] C.Y. Lai, P. Hoffee and B.L. Horecker, Arch. Biochem. Biophys. 112 (1965) 567.
[5] C.Y. Lai and C. Chen, Arch. Biochem. Biophys. 128 (1968) 212.
[6] I. Gibbons and R.N. Perham, Biochem. J. 117 (1970) 415.
[7] R.N. Perham and G.M.T. Jones, European J. Biochem. 2 (1967) 84.
[8] W. Gray and B.S. Hartley, Biochem. J. 89 (1963) 379.
[9] R.E. Offord, Nature 211 (1966) 591.
[10] C.Y. Lai, Arch. Biochem. Biophys. 128 (1966) 202.
[11] H.B.F. Dixon and R.N. Perham, Biochem. J. 109 (1968) 312.
[12] A.L. Shapiro and J.V. Maizel Jr., Anal. Biochem. 29 (1969) 505.
[13] W. Chan, D.E. Morse and B.L. Horecker, Proc. Natl. Acad. Sci. U.S. 57 (1967) 1013.
[14] D.E. Morse and B.L. Horecker, Arch. Biochem. Biophys. 125 (1968) 942.
[15] F.H.C. Crick, Cold Spring Harbor Symp. Quant. Biol. 31 (1966) 1.
[16] J.I. Harris, B.E. Davidson, M. Sajgo, H.F. Noller and R.N. Perham, FEBS Symposium 18 (1970) 1.
[17] K. Kawahara and C. Tanford, Biochemistry 5 (1966) 1578.
[18] R.W. Gracy, A.G. Lacko and B.L. Horecker, J. Biol. Chem. 244 (1969) 3913.
[19] D.E. Morse and B.L. Horecker, Advan. Enzymol. 31 (1968) 125.