Vol. 181, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 27, 1991 Pages 226-231

Comparison of the Amino Acid Sequences of Tissue-Specific Parvalbumins from Chicken Muscle and Thymus and Possible Evolutionary Significance

John M. Brewer[1,*], Jonathan Arnold[2], G.G. Beach[3], W.L. Ragland[3] and John K. Wunderlich[4]

Departments of [1]Biochemistry, [2]Genetics and [3]Avian Medicine and [4]Molecular Genetics Instrumentation Facility, University of Georgia, Athens, GA 30602

Received October 2, 1991

Chicken leg muscle parvalbumin was digested with cyanogen bromide or trypsin or trypsin after citraconylation. Peptides isolated by reverse phase HPLC at pH 7.0 were subjected to acid hydrolysis and amino acid analysis and, in some cases, sequencing. The chicken muscle parvalbumin amino acid sequence has ca. 80% sequence identity with α-type parvalbumins from mammalian (rabbit, human and rat) muscle. By contrast, the chicken thymus parvalbumin ("avian thymic hormone") sequence is very similar to reptile (turtle, salamander and frog) muscle β-type parvalbumins. We hypothesize that the evolutionary appearance of the warm-blooded reptiles was accompanied by recruitment of the β parvalbumin isozyme for promotion of lymphocyte maturation.

The discovery that "Avian Thymic Hormone" (ATH), a protein that promotes lymphocyte maturation (1), is a parvalbumin (2,3) focused attention on the parvalbumin isolated from chicken leg muscle (4). The latter protein exhibits a very different amino acid composition (2,4,5). Since both parvalbumins are tissue-specific (4,6,7) and the Avian Thymic Hormone circulates in the bloodstream (6,7), we sequenced the chicken muscle parvalbumin as part of an ongoing attempt to see where the recognition site(s) for the Avian Thymic Hormone might be located.

Materials and Methods

Chicken leg muscle parvalbumin was isolated essentially as described by Strehler et al (4). It appeared as a separate peak (monitored at 220 nm) in the wash from the DEAE column (10 mM sodium acetate, pH 5.7) (4), eluting shortly before the salt gradient was applied. Identification of the protein fraction as chicken muscle parvalbumin presented a problem: since it does not cross-react with monoclonal antibody to Avian Thymic Hormone, we were obliged

*To whom reprint requests should be addressed.

0006-291X/91 $1.50

to use the amino acid composition and calcium content as criteria for identification. We obtained an amino acid composition which was very similar to the published composition of chicken muscle parvalbumin (4) (not shown). In addition, our protein fraction had 2.3 moles of Ca^{2+} per mole of protein, assuming 108 residues per mole. The sequencing data confirmed that we were dealing with a parvalbumin and a reasonably homogeneous protein as well.

Amino acid analysis data indicated our protein had no cys or trp, one tyr and his, three met and two or three arg per mole. Consequently, once the protein fraction was identified as chicken muscle parvalbumin, we subjected it to digestion by trypsin (8), by trypsin after treatment of the protein with citraconic anhydride (9) and using cyanogen bromide (10). Separation of the resulting peptides by reverse-phase HPLC (Applied Biosystems 130A Separation System) could not be performed in the usual acid pH solvent system (11) because the peptides precipitate on the column support at low pH. Instead, we used a solvent system buffered at pH 7.0 with 0.05M triethylammonium acetate on a 2.1 x 30 mm Aqua Pore RP 300 (C8) column.

Subsequent acid hydrolysis and analysis was performed on an Applied Biosystems Model 420A Amino Acid Analysis System. Sequence analysis was done using Applied Biosystems Model 477A or 470A Sequencers equipped with a Spectra Physics data station and WINer/286 software. Metal content of protein fractions was determined by plasma emission spectroscopy and was done by the Chemical Analysis Laboratory at the University of Georgia.

Parvalbumin sequences were obtained from ref. 12, Swiss Protein Data Bank (version 18) and Protein Identification Resource (PIR) (version 28) and analyzed using PHYLIP programs (version 3.2) (1989) (provided by Felsenstein, J., University of Washington, Seattle, Washington 98195) with the Biological Sequence/Structure Computational Facility at the University of Georgia.

Formic acid (98%) was from Fluka A.G. Trypsin (TPCK treated) was from Worthington. Citraconic anhydride came from Sigma. Other chemicals used were reagent grade or better. The water was distilled in a glass still from deionized water (Continental Water Systems Corp.).

Results and Discussion

Based on amino acid analyses of peptide fractions obtained (Figure 1), amino acid sequence analysis was done on selected fractions. Examination of the amino acid compositions of the tryptic digest peptide fractions showed there were three with integral amounts of methionine: fractions T1, T2 and T3. Fraction T1 gave no sequence data until digested with cyanogen bromide; after digestion, we obtained the sequence represented by residues 3-12 (Figure 2). Peptide T1 is clearly the amino terminal fragment, which is evidently blocked as it is in most other parvalbumins examined (12). Comparison of its sequence with its composition suggested the amino terminus is ala-met-. Fraction T2 contained no lys or arg, so was the carboxyl terminal peptide. Fraction T3 bridged the cyanogen bromide fragments 3-32 and 33-105. Sequence analysis thus located all the methionines at residues 2, 32, and 105.

The early eluting cyanogen bromide fraction CB1 contained significant amounts of glx, ser, ala, val and homoserine upon hydrolysis, but insignificant amounts of free amino acids. These appeared to be the amino terminal and carboxyl terminal fragments. Sequence analysis gave val-ala-glu-ser; part of CB1 was heated to 110° for 20 minutes in 1N HCl (15), dried and sequenced. A marked increase in ala was obtained in the first cycle, also

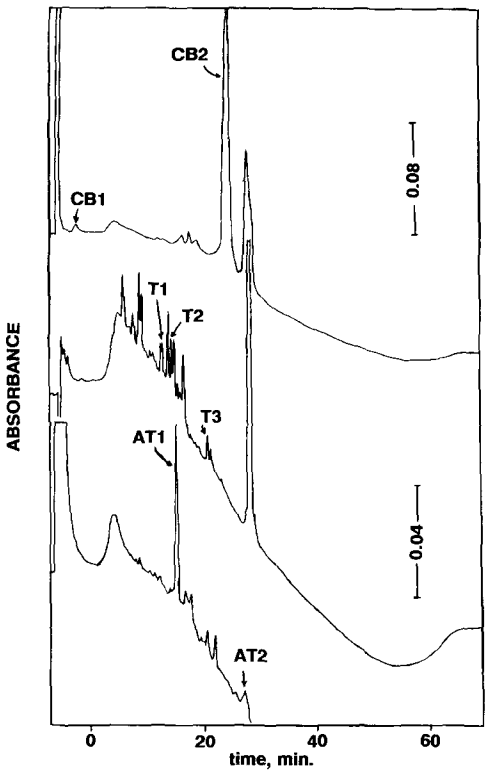
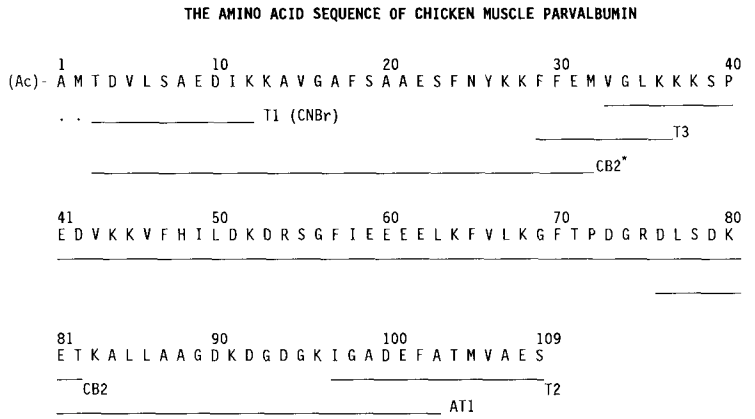


Figure 1. HPLC elution profiles of cyanogen bromide (upper), trypsin (middle) and trypsin after citraconylation (bottom) digests of chicken muscle parvalbumin. No peptides were observed after elution of AT2 (not shown). The quantities chromatographed were of the order of 300 pmol-3 nmol. The fractions from which sequence information was obtained are identified by numbers. These were collected by hand. The "time" is the time of application of the gradient: at 30 minutes the acetonitrile concentration is 40%. The absorbance is recorded at 225 nm (upper) or at 230 nm because of the absorbance of the buffer.

suggesting the amino terminal residue is ala. The blocking group is not known, but is generally acetyl.

Fraction CB2 had an amino acid composition similar to the original material, except for its lack of methionine; the later eluting peak had a similar composition but gave no amino terminus. The major sequence of CB2 obtained is shown as residues 33-. There was however a minor sequence (about 20% of the major) which appeared to be residues 3-32. By accident, the cyanogen bromide digestion was done in 98% formic acid rather than our customary 70%. The amino terminus of fragment 3-32 is thr-asp. It appears that the amino terminus becomes largely blocked in strong acid. As a consequence of these fortuitous circumstances, we were easily able to determine the major sequence (residues 33-) and at the same time obtain a large part of the minor (residues 3-32). The obtaining of sequence information from mixtures of peptides is an accepted practice (16).



COMPARISON OF MAP TURTLE, CHICKEN AND RABBIT MUSCLE AND CHICKEN THYMUS PARVALBUMINS

```
                1     5       10      15      20      25      30      35
TURTLE MUSCLE   A M T D I L S A K D I E A A L T S C Q A A D S F N Y K S F F S K V G L
CHICKEN THYMUS  A I T D I L S A K D I E S A L S S C Q A A D S F N Y K S F F S T V G L
CHICKEN MUSCLE  A M T D V L S A E D I K K A V G A F S A A E S F N Y K K F F E M V G L
RABBIT MUSCLE   A M T E L L N A E D I K K A I G A F A A A E S F D H K K F F Q M V G L

           36      40      45      50      55      60      65      70
TMP        K G K S T D Q V K K I F G I L D Q D K S G F I E E D E L Q L F L Q N F S S
CTP        S S K T P D Q I K K V F G I L D Q D K S G F I E E E E L Q L F L K N F S S
CMP        K K K S P E D V K K V F H I L D K D R S G F I E E E E L K F V L K G F T P
RMP        K K K S T E D V K K V F H I L D K D K S G F I E E E E L G F I L K G F S P

              75      80      85      90      95      100     105
TMP        T A R A L T A A E T K A F M A A G D T D G D G K I G V D E F Q A L V K A
CTP        S A R V L T S A E T K A F L A A G D T D G D G K I G V E E F Q S L V K A
CMP        D G R D L S D K E T K A L L A A G D K D G D G K I G A D E F A T M V A E S
RMP        D A R D L S V K E T K T L M A A G D K D G D G K I G A D E F S T L V S E S
```

Figure 2. The amino acid sequence of chicken muscle parvalbumin. This was derived from the sequences of the peptides shown (upper). CB2* was eluted using a 12-40% acetonitrile gradient over 50 minutes; see the text. Except for CB1, the peaks to the left of CB2 had low amino acid contents. The blocking group is probably acetyl-. See the text. The lower part of the Figure gives a comparison between the sequences of the Map turtle muscle parvalbumin (TMP) (uppermost) (13), the chicken thymic parvalbumin (CTP) (3), the chicken muscle parvalbumin (CMP) and the rabbit muscle parvalbumin (RMP) (bottom) (14).

The mixture of peptides prepared by HPLC after digestion with cyanogen bromide in 70% formic acid was chromatographed using a flatter gradient (12-40% acetonitrile over 50 minutes) (not shown). An accidented peak followed by two smaller peaks was obtained. The accidented peak resembled those normally obtained with intact parvalbumins (17). Evidently the two large cyanogen bromide peptides form a complex under the conditions we employed in the HPLC. However, we performed amino acid analysis on eight fractions comprising the accidented peak and also on the two smaller following peaks. The analyses suggested a partial resolution of the two peptides had occurred, since the first seven fractions from the accidented peak were enriched in tyr relative to his, arg and phe. These seven fractions were pooled and sequenced. The sequence data suggested that the pooled fractions contained about three times

as much of the shorter peptide (residues 3-32) as the longer (residues 33-105). Hence we were able to complete the sequencing of the chicken leg muscle parvalbumin.

Fraction AT1 appeared to be residues 76-109 by its amino acid composition, and its sequence - not completed as the overlap with the T2 sequence was judged sufficient - confirmed this. Fraction AT2 appeared to be residues 1-54. A relatively small amount of AT2 was sequenced after cyanogen bromide treatment, confirming that both sequences expected were present.

A partial sequence of a chicken muscle parvalbumin was determined from the cDNA sequence by Palmisano and Henzl (18). Their sequence corresponds to residues 57-101 of our sequence, except they have asp at residue 61 and val, glu and lys at residues 99-101. Our chemical sequence data are unequivocal and consistent and we are confident of our assignments.

After this work was completed, another group published an amino acid sequence of chicken leg muscle parvalbumin (19). Their peptides (from tryptic and V8 protease digestion) were analyzed by mass spectrometry and their sequence is identical to ours.

Figure 2 also shows the sequence data we have obtained, compared with the known sequence of the thymic parvalbumin from the same organism and the sequences of the two parvalbumins from other organisms that most closely resemble the two parvalbumins from chicken.

Comparison of the chicken parvalbumin sequences with the 23 known sequences of parvalbumins from other organisms shows clearly that the chicken thymic parvalbumin most closely resembles the β parvalbumins as a group, and specifically, those from Map turtle (83% sequence identity), salamander (76%) and frog (R. esculenta) muscle (69%). The chicken muscle parvalbumin is evidently an α parvalbumin which most closely resembles the rabbit (89%) (14), rat (76%) (20) and human (76%) muscle parvalbumins.

Both α and β parvalbumins are found in muscle tissue of frogs, fish (including the coelacanth) and the salamander, the "lower" (cold-blooded) vertebrates in general (13,21). By contrast, rabbit and chicken muscle contain only the α form (cf. for example ref. 5). Our working hypothesis, based inevitably on limited data, is that the appearance of the warm-blooded vertebrates was accompanied by recruitment of the β parvalbumin for expression in thymus only, in order to act on the immune system. Its exact function in this connection is the subject of intense investigation in our laboratories.

These data are the first to compare sequences of two parvalbumins with markedly different functions from the same species but different tissues. The much greater dissimilarity in sequences (58% sequence identity) between the two chicken parvalbumins explains how both proteins can exist in the same organism with one circulating in the bloodstream, functioning to promote

immune system maturation, without interference by the other. The putative receptor(s) for the thymic protein can readily discriminate between them.

Acknowledgment

Supported by grant 88-34116-3635 from the U.S. Department of Agriculture.

References

1. Murthy, K.K., Beach, F.G. and Ragland, W.L. (1984) In Thymic Hormones and Lymphokines (A.L. Goldstein, ed.), Plenum Press, N.Y. pp. 375-382.
2. Brewer, J.M., Wunderlich, J.K., Kim, D.-H., Carr, M.Y., Beach, G.G., and Ragland, W.L. (1989) Biochem. Biophys. Res. Comm. 160 1155-1161.
3. Brewer, J.M., Wunderlich, J.K. and Ragland, W.L. (1990) Biochimie 72 653-660.
4. Strehler, E.E., Eppenberger, H.M. and Heizmann, C.W. (1977) FEBS Lett. 78 127-133.
5. Blum, H.E., Lehky, P., Kohler, L., Stein, E.A. and Fischer, E.H. (1977) J. Biol. Chem. 252 2834-2838.
6. Murthy, K.K., Pace, J.L. Barger, B.O., Dawe, D.L. and Ragland, W.L. (1984) Thymus 6 43-56.
7. Hall, C.A., Beach, F.G. and Ragland, W.L. (1989) Immunobiology (suppl.) 4 143-144.
8. Smyth, D.G. (1967) Methods Enzymol. XI 214-231.
9. Atassi, M.Z. and Habeeb, A.F.S.A. (1972) Methods Enzymol. 25 546-553.
10. Gross, E. (1967) Methods Enzymol. XI 238-255.
11. Tempst, P., Hunkapiller, M.W. and Hood, L.E. (1984) Analyt. Biochem. 137 188-195.
12. Sasaki, T., Tanokura, M. and Asaoka, K. (1990) FEBS Lett. 268 249-251.
13. Maeda, W.M., Zhu, D. and Fitch, W.M. (1984) Mol. Biol. Evol. 1 473-488.
14. Enfield, D.L., Ericsson, L.H., Blum, H.E., Fischer, E.H. and Neurath, H. (1975) Proc. Nat. Acad. Sci. USA 72 1309-1313.
15. Chin, C.C.Q. and Wold, F. (1987) In: Methods in Protein Sequence Analysis (K.A. Walsh, ed.), Humana Press, Clifton, NJ. pp. 505-512.
16. Shieh, D.H. and Travis, J. (1987) J. Biol. Chem. 262 6055-6059.
17. Berchtold, M.W., Heizmann, C.W. and Wilson, K.J. (1983) Analyt. Biochem. 129 120-131.
18. Palmisano, W.A. and Henzl, M.T. (1991) Biochem. Biophys. Res. Comm. 176 328-334.
19. Kuster, T., Staudenmann, W., Hughes, G.J. and Heizmann, C.W. (1991) Biochemistry 30 8812-8816.
20. Berchtold, M.W., Heizmann, C.W. and Wilson, K.J. (1982) Eur. J. Biochem. 127 381-389.
21. Capony, J.-P., Demaille, J., Pina, C. and Pechere, J.-F. (1975) Eur. J. Biochem. 56 215-227.