Direct Sequence Analysis of a Cyanogen Bromide-Generated Peptide Corresponding to the Region of Subfragment-1 of Adult Chicken Gizzard Myosin Heavy Chain Predicted to Contain a Seven-Amino Acid Insertion

Masaharu Sakurai1), Bunji Watanabe2), Shuichi Nagata1)

1) Department of Hospital Pharmacy, Nagasaki University School of Medicine, 1-7-1 Sakamoto, Nagasaki 852, Japan
2) The School of Allied Medical Sciences, Nagasaki University, 1-7-1 Sakamoto, Nagasaki 852, Japan

The amino acid sequence of a cyanogen bromide-generated peptide corresponding to the region of subfragment-1 of chicken gizzard myosin, predicted from the cDNA sequence to contain a seven-amino acid insertion, was determined directly. Subfragment-1 was prepared from adult chicken gizzard myosin by limited digestion with papain, and a 137-residue peptide was obtained by cleaving subfragment-1 with cyanogen bromide. Amino acid composition analysis and sequence analysis of the fragments generated from this peptide by formic acid or α-chymotrypsin revealed that the peptide corresponded to leucine-165 to methionine-301 of chicken gizzard myosin heavy chain previously deduced from the cDNA sequence by Yanagisawa, M. et al.. The predicted seven-amino acid insertion (Gln-Gly-Pro-Ser-Phe-Ser-Tyr), absent from other smooth muscle myosins, was confirmed to be present in chicken gizzard myosin. The amino acid sequence surrounding the insertion site of chicken gizzard myosin shows greater homology to the corresponding sequence of other smooth muscle myosins than to those of chicken skeletal and cardiac muscle myosins. The insertion sequence may confer gizzard-specific functions to the gizzard myosin heavy chain, whereas the amino acid sequence surrounding the insertion site may be specific to smooth muscle.

Key words: subfragment-1, adult chicken gizzard myosin, amino acid sequence, isoform, smooth muscle

Introduction

The contractile protein myosin consists of two heavy chains and two pairs of light chains. Myosin heavy chain (MHC) comprises a globular head (subfragment-1, S-1) and an α-helical, coiled-coil rod, which can be separated from each other by limited proteolysis. S-1 possesses ATPase activity and binds actin5). The skeletal and cardiac muscle MHCs of vertebrates are encoded by a family of at least 10 genes whose expression is regulated developmentally and hormonally in a tissue-specific manner5)-9). Three MHC isoforms, SM1, SM2 and SMemb, have been identified in smooth muscle. The structural difference between SM1 and SM2 lies in the COOH-terminal portion of light meromyosin15); both isoforms (SM1, SM2) are generated by alternative RNA splicing and are also regulated developmentally10,11). Although the expression of SMemb is predominant in embryonic and fetal smooth muscle12), 13), the protein is reexpressed in proliferating smooth muscle cells such as those present in arteriosclerotic intimal tissue14). On the basis of its cDNA sequence, chicken gizzard S-1 was recently predicted to contain an insertion of seven amino acids compared with the corresponding region of S-1 from uterus15) and aorta smooth muscle myosins. Chicken gizzard myosin containing this insertion exhibits a higher ATPase activity than aorta smooth muscle myosin without it16). However, because differences between amino acid sequences determined directly and those predicted from cDNA have been detected17), it is important to verify the seven-amino acid insertion in chicken gizzard S-1 at the protein level. We have now determined the primary structure of the ATP-binding region of chicken gizzard S-1 thought to contain the seven-amino acid insertion and compared the sequence with those of other myosins.

Materials and Methods

Preparation of S-1 and isolation of CNBr peptides. Myosin was prepared from adult chicken gizzard muscle essentially as described by Ikebe et al.18). S-1 was obtained by digesting myosin with papain in 50 mM MgCl2. The purity of myosin and S-1 was examined by SDS-polyacrylamide gel electrophoresis. S-1 was further purified by gel filtration on a column (2.4 × 132 cm) of Sephacyr S-300 (Pharmacia) after subjecting to S-
carboxymethylation.

S-Carboxymethylated S-1 (CM-S-1) in 70% (v/v) formic acid was treated with a 300-fold excess of CNBr per methionine residue under nitrogen for 24 hr at 25°C. The CNBr-generated peptides were applied to a column (1.8×132 cm) of Sephadex G-75 (Pharmacia) that had been equilibrated with 5% (v/v) acetic acid and were eluted with the same buffer. Peptides in each peak were further purified by reversed-phase high-performance liquid chromatography (RP-HPLC) at room temperature on a column (4.6×250 mm) of Wakosil iC8 (Wako Pure Chemical Industries, Ltd.) that had been equilibrated with 0.1% trifluoroacetic acid; peptides were eluted with a linear gradient of acetonitrile. A large CNBr peptide was fragmented with 0.4 M formic acid at 108°C for 2 hr in an evacuated sealed tube, or further digested with a-chymotrypsin type VI; SIGMA) at 37°C for 2 hr in 1% NH₄HCO₃ (pH 8.0).

Amino acid analysis and sequence determination

Amino acid analysis of peptides was performed with a JLC-300 automatic amino acid analyzer (JEOL Co.) after hydrolysis with 5.7 M HCl at 110°C for 22 hr in evacuated sealed tubes by the method of Spackman et al. Amino acid sequences were determined with an Applied Biosystems model 476A protein sequencer equipped with a model 610A analysis program for the on-line detection of phenylthiohydantoin amino acids.

Results

S-1 was subjected to S-carboxymethylation and was purified by chromatography on a column of Sephacryl S-300 (Fig. 1).

CM-S-1 was cleaved with CNBr and the resulting peptides were separated by gel filtration on a column of Sephadex G-75 (Fig. 2). One large peptide (CN-1) was further purified by RP-HPLC on a column of Wakosil iC8 (Fig. 3). The amino acid composition of CN-1 is shown in Table 1. The NH₂-terminal 20 residues of CN-1 were sequenced by automated Edman degradation. CN-1 was thus shown to comprise 137 amino acid beginning with Leu-165 of the amino acid sequence of chicken gizzard MHC predicted from the cDNA sequence by Yanagisawa et al.

The primary structure of CN-1 surrounding the region of the seven-amino acid insertion was analyzed by first cleaving CN-1 with formic acid. The resulting peptides were separated by RP-HPLC on a column of Wakosil iC8 (Fig. 4). Three peptides (fa-1 to fa-3) were obtained and their amino acid sequences were determined either completely or partially.

CN-1 was also digested with a-chymotrypsin and the resulting peptides were separated by RP-HPLC on a
Table 1. Amino acid composition of peptide CN-1. Cysteine and methionine were determined as carboxymethyl cysteine and homoserine, respectively. Values in parentheses were obtained from the sequence data.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Number of residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cys</td>
<td>+ (1)</td>
</tr>
<tr>
<td>Asp</td>
<td>13.6 (13)</td>
</tr>
<tr>
<td>Thr</td>
<td>8.7 (9)</td>
</tr>
<tr>
<td>Ser</td>
<td>10.1 (11)</td>
</tr>
<tr>
<td>Glu</td>
<td>19.6 (18)</td>
</tr>
<tr>
<td>Pro</td>
<td>3.2 (2)</td>
</tr>
<tr>
<td>Gly</td>
<td>12.1 (11)</td>
</tr>
<tr>
<td>Ala</td>
<td>11.4 (11)</td>
</tr>
<tr>
<td>Val</td>
<td>6.0 (6)</td>
</tr>
<tr>
<td>Met</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Ile</td>
<td>10.3 (11)</td>
</tr>
<tr>
<td>Leu</td>
<td>10.2 (10)</td>
</tr>
<tr>
<td>Tyr</td>
<td>4.6 (6)</td>
</tr>
<tr>
<td>Phe</td>
<td>6.1 (7)</td>
</tr>
<tr>
<td>His</td>
<td>1.8 (2)</td>
</tr>
<tr>
<td>Lys</td>
<td>11.7 (12)</td>
</tr>
<tr>
<td>Arg</td>
<td>5.7 (6)</td>
</tr>
<tr>
<td>Total</td>
<td>137</td>
</tr>
</tbody>
</table>

column of Wakosil C8 (Fig. 5). Twelve peptides (Chy-1 to Chy-12) were obtained and sequenced either completely or partially. The sequences of both formic acid- and a-chymotrypsin-generated peptides are shown in Fig. 6.

Although CN-1 was sequenced completely, four overlapping peptides were not obtained; the corresponding junction sites were aligned on the basis of the amino acid sequence of chicken gizzard MHC deduced from the cDNA.²⁰

![Fig. 4](image-url) Separation of cleavages of CN-1 with 0.4M formic acid by RP-HPLC. The cleavages of CN-1 with 0.4M formic acid were applied to a column of Wakosil C8 (4.6×250 mm). Conditions were the same as in Fig. 3 except for the CH3CN concentration gradient.

![Fig. 5](image-url) Separation of digests of CN-1 with a-chymotrypsin by RP-HPLC. The digests of CN-1 with a-chymotrypsin were applied to a column of Wakosil C8 (4.6 × 250 mm). Conditions were the same as in Fig. 3 except for the CH3CN concentration gradient.

![Fig. 6](image-url) Sequence determination of CN-1, a 137-residue peptide, containing the seven-amino acid insertion in S-1 from chicken gizzard muscle myosin.

The peptides obtained by fragmentation with 0.4 M formic acid (fa), a-chymotrypsin (chy) are shown. The arrows indicate the amino acids identified by automated Edman degradation. The double line and thick line indicate seven-amino acid insertion and the ATP binding site, respectively.
Discussion

A seven-amino acid insertion has been predicted near the ATP-binding site in chicken gizzard MHC and is associated with increased actin-activated Mg\(^{2+}\)-ATPase activity\(^{21}\). White et al.\(^{20}\) described a similar insertion in fetal rat visceral smooth muscle myosin that was not present in myosin from the adult rat stomach. Furthermore, a 10-amino acid insertion has been identified in chicken nonmuscle MHC at the same position as the seven-amino acid insertion in intestinal MHC\(^{22}\). It has been assumed that these various myosin isoforms result from alternative splicing of the primary transcript smooth muscle MHC gene.

However, these amino acid sequence were deduced from the corresponding cDNA sequences, we have now determined, by direct analysis, the amino acid sequence of a large CNBr fragment of chicken gizzard S-1 predicted to contain the seven-amino acid insertion. A seven-amino acid insertion with a sequence (Gln-Gly-Pro-Ser-Phe-Ser-Tyr) identical to the sequence deduced from the corresponding cDNA by Kelley et al.\(^{20}\) was indeed present in chicken gizzard S-1. Comparison of the sequence of the region surrounding the seven-amino acid insertion of chicken gizzard S-1 with that of other smooth muscle MHCs revealed a difference of only one residue with fetal rat visceral MHC and showed that this insertion is not present in rabbit uterus, chicken aorta, or rat stomach MHC (Fig. 7). Furthermore, the amino acid sequence of this region of chicken gizzard S-1 shows greater homology to the corresponding regions of smooth muscle myosins from other animals (Fig. 7) than to chicken skeletal\(^{20}\) or cardiac\(^{20}\) muscle MHC (Fig. 8). It is unclear how the seven-amino acid insertion in chicken gizzard myosin may affect the function of the protein, but its location near the ATP-binding site is consistent with a role in the increased ATPase activity of myosin from this tissue.

Acknowledgments

We thank Prof. Masataka Ichikawa, the Department of Hospital Pharmacy, Nagasaki University School of Medicine, for his useful advice and continuous encouragement in the present study.

References

2) Izumo, S., Nadal-Ginard, B., and Mahadavi, V.: All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. Science 231: 597-600, 1986.
5) Nguyen, H. T., Oubria, R. M., Wydro, R. M., and Nadal-Ginard, B.: Sarcomeric myosin heavy chain is coded by a highly conserved


