Extensive Sequencing of Tryptic Peptides of a Rabbit Reticulocyte 66-kDa Protein That Promotes Recycling of Hsp 70

HOMOLOGY TO STRESS-RELATED PROTEINS*

(Received for publication, January 19, 1996, and in revised form, March 28, 1996)

Martin Gross†, Suzanne Hessefort, Annette Olin, and Giri Reddy§

From the Department of Pathology, The University of Chicago, Chicago, Illinois 60637

Trypsinization and sequence analysis of the 66-kDa rabbit reticulocyte protein (RF-hsp 70), shown in the preceding article to function as a recycling protein for hsp 70, demonstrates striking similarity to the transformation-sensitive human protein IEF SSP 3521 (Honoré, B., Leffers, H., Madsen, P., Rasmussen, H. H., Vandekerckhove, J., and Celis, J. E. (1992) J. Biol. Chem. 267, 8485–8491) and mouse extendin (Blatch, G. L., Lassle, M., Takatori, T., Gandhi, T., Kundra, V., and Zetter, B. R. (1995) Proc. Am. Assoc. Cancer Res. 36, 68). The human and mouse proteins share 97% sequence identity, and sequencing of 20 polypeptides (225 residues) from RF-hsp 70 reveals only 10 differences between the rabbit and human proteins and 13 differences between the rabbit and mouse proteins (96 and 94% identity, respectively). In addition, all three proteins are of similar size, and each contains 11 cysteines. These findings strongly suggest that these three proteins are homologs of the same activity. All differences (but one) between the human and mouse proteins occur within the amino-terminal half of the protein, and there is only one difference among 121 sequenced residues between RF-hsp 70 and the human or mouse protein which occurs within the carboxyl-terminal 70% of the molecule. In addition, where partial sequences of RF-hsp 70 and p60, a chick oviduct protein that shows 70% identity to the human protein (Smith, D. F., Sullivan, W. P., Marion, T. N., Zaitsu, K., Madden, B., McCormick, D. J., and Toft, D. O. (1993) Mol. Cell. Biol. 13, 869–876), overlap (a total of 54 residues), RF-hsp 70 and chick p60 show 78% sequence identity. Studies of the initial digestion of RF-hsp 70 by trypsin indicate that it is first converted to 58- and 54-kDa components, each of which is then converted to a 43-kDa polypeptide. This 43-kDa component is located in the human and mouse proteins at position 124 to about 470. It is converted subsequently to a 31-kDa polypeptide by trypsin hydrolysis at position 207. This 31-kDa component is finally split into 17- and 14-kDa polypeptides that are located at positions 208 to approximately 351 and 352 to approximately 470, respectively. The 14-kDa polypeptide is relatively resistant to further digestion with trypsin, and seven tryp tic peptides from other parts of RF-hsp 70 contain internal lysine and/or arginine residues (as do several tryptic peptides produced from IEF SSP 3521 and chick p60). Both features may be due to interference with trypsin action by secondary structure in the protein, since trypsinization of reduced and carboxymethylated RF-hsp 70 results in hydrolysis of the 14-kDa polypeptide and reduces the level of peptides that contain internal lysine and/or arginine, although it does not eliminate them.

In the preceding article (1) we demonstrated that a 66-kDa protein, purified from rabbit reticulocyte lysate and termed RF-hsp 70,1 promotes the recycling of hsp 70 by increasing the rate of dissociation of hsp 70-ADP in the presence of ATP, thus enhancing the folding activity of hsp 70 as measured by the reactivation of heat-denatured luciferase. We suggested that the recycling role of RF-hsp 70 may be similar to that envisioned for the activities termed DnaJ and GrpE in bacteria (2) and YDJ 1 in yeast (3). In an effort to identify RF-hsp 70 and to compare it more precisely with proteins described previously, we subjected RF-hsp 70 to trypsin digestion and the purified peptides to sequence analysis. This has permitted us to demonstrate a very high degree of sequence identity between RF-hsp 70 and several proteins described recently. One is the human protein IEF SSP 3521 (4), a component that is increased 2-fold in SV40-transformed MRC-5 fibroblasts and has 42% amino acid sequence identity to STI1, a stress-inducible mediator of the heat shock response in yeast (5). A second is mouse extendin, whose cDNA was isolated from a M27 mouse lung carcinoma cDNA expression library by screening with an antibody to an actin-associated protein from Dictyos
tium (6). Purified recombinant extendin was found to associate with actin, hsp 70, and hsp 90; to localize, in part, to extending cellular pseudopodia; and to have 97% sequence identity to IEF SSP 3521 (6).

EXPERIMENTAL PROCEDURES

Isolation of Tryptic Peptides of RF-hsp 70 and Microsequence Analysis—RF-hsp 70, purified as described (1) and at a concentration of 0.5 mg/ml in 0.02 M Tris-HCl, pH 7.5, 0.05 M KCl, 1 mM dithiothreitol, and 0.10 mM EDTA, was digested with 0.05 mg/ml trypsin for 18 h at 34 °C. EDTA and trifluoroacetic acid were added to a final concentration of 2.5 mM and 0.05% (v/v) respectively, and the sample was subjected to reverse-phase chromatography on an Aquapore RP-300 column (4.6 × 250 mm, Brownlee Laboratories) using a 0–60% linear acetonitrile gradient in 0.05% (v/v) trifluoroacetic acid in water at a flow rate of 1.0 ml/min over a 110-min period and employing a Varian model 5020 chromatograph equipped with a UV 100 with the absorbance set at 215 nm. Separated, tryptic peptides were subjected to sequence analysis in an Applied Biosystems 473A Sequencer using the pulsed liquid technology and equipped with an on-line ATZ to phenylthiohydantoin autoconverter, a gradient-programmable microbor chromatography system, and a Macintosh-based 610A data system to identify and

---

1 The abbreviations used are: hsp 70, a member of the 70-kDa heat shock protein family; RF-hsp 70, a 66-kDa rabbit reticulocyte protein that recycles hsp 70; hsp 90, a member of the 90-kDa heat shock protein family; RCM-RF-hsp 70, reduced and carboxymethylated RF-hsp 70.
quantitate the amino acid residues.

Amino Acid Compositional Analysis—Aliquots of purified RF-hsp 70 plus a known concentration of l-α-amino-β-guanidinopropionic acid (an internal standard) were hydrolyzed in 300 μl of 6 N constant boiling HCl (Sigma) at 120°C for 24 h. After being taken to dryness, each sample was dissolved in dilution buffer and applied onto a Beckman Model 121 µ microfuge, ion exchange amino acid analyzer, using both 440 and 570 nm, postcolumn ninhydrin/dimethyl sulfoxide reagent for detection and quantitation.

Reduction and Carboxymethylation of RF-hsp 70—Purified RF-hsp 70 (75 µg) was incubated for 4 h at 37°C in 25 mM Tris-HCl, pH 7.5, 8 µM urea, 0.4 mM NaCl, and 10 mM dithiothreitol in a final volume of 0.75 ml. lodoacetic acid was then added at a final concentration of 25 µg/ml, and the sample was incubated in the dark at 25°C for 20 min. It then received a final concentration of 100 mM β-mercaptoethanol and was filtered through a 0.7-µm filter. The sample was then electrophoresed on 7% polyacrylamide-SDS slab gels at 36 V (2.4 V/cm) for 16 h (see Figs. 3 and 5) or at 44 V for 19 h (see Fig. 4) followed by silver staining as described previously (7, 8).

Materials—Three times crystallized trypsin and chymotrypsin were obtained from Worthington, and sequencing grade endoproteinase Glu-C and endoproteinase Lys-C were from Boehringer Mannheim. Gel electrophoresis reagents were from Bio-Rad, and gel electrophoresis equipment was from Hoefer.

RESULTS

A comparison of the amino acid sequence of multiple tryptic peptides derived from RF-hsp 70 with reported sequence data indicates a striking similarity between RF-hsp 70 and the transformation-sensitive human protein IEF SSP 3521, identified by Honoré et al. (4), and mouse extendin, identified by Blatch et al. (6). Fig. 1 shows the complete translated amino acid sequences deduced from the cDNAs of the human (4) and mouse (6) proteins with aligned, partial sequences of trypsin peptides derived from RF-hsp 70 (see “Experimental Procedures”). Fig. 1 also includes partial sequences derived from gel-purified IEF SSP 3521 (4) and from gel-purified chick ovul
duct p60 (9). The latter protein immunopurifies with hsp 90 and hsp 70, found in certain assembly complexes of chick progesterone receptor, and shows 70% sequence identity to IEF SSP 3521 (9-11). The data in Fig. 1 demonstrate that the amino acid sequence of 20 tryptic peptides (225 residues) from RF-hsp 70 differs from that of the translated cDNA of human IEF SSP 3521 in only 10 positions and from that of mouse extendin at only 13 sites, corresponding to a sequence identity (to 41% of the total protein) of 96 and 94%, respectively. This strongly suggests that RF-hsp 70 is the homolog of the human and mouse proteins in rabbit reticulocytes. Interestingly, all differences (but one at position 209) between RF-hsp 70 and the human and mouse proteins occur within the amino-terminal 30% of the protein so that there is 100% identity (99% to mouse) between 121 residues from position 162 to the carboxyl terminus, position 543. In addition, all differences (but one) between the human and mouse proteins, which show 97% sequence identity to each other, occur within the amino-termi
tal half of the protein. When the degree of identity to the human protein of RF-hsp 70 and chick p60 (9) partial sequences is expressed as a function of the distance from the amino terminus (Fig. 2), both proteins show the greatest difference with the human protein within the amino-terminal portion of the protein and the greatest similarity within the carboxyl-terminal portion. A similar result is obtained when the sequence of mouse extendin is used for comparison. One exception to this is the fact that all four proteins are identical to each other at positions 84-90, and RF-hsp 70 and the human and mouse proteins are identical at positions 78-118 except for differences at positions 104 and 106. It is also of interest that where p60 sequences overlap the differences between RF-hsp 70 and the human protein (positions 13, 17, 24, 25, and 145), the rabbit and chick proteins are the same (Fig. 1). Overall, RF-hsp 70 and the chick p60 partial sequences overlap a total of 54 residues that show 78% identity (Fig. 1). Thus, from the data available, RF-hsp 70 shows much greater sequence identity to human IEF SSP 3521 and mouse extendin than does chick p60, whereas chick p60 shows about the same degree of sequence identity to RF-hsp 70 (78%) as it does to IEF SSP 3521 and mouse extendin (70%).

We also subjected RF-hsp 70 to amino acid compositional analysis and compared the result with the amino acid composi
tion of IEF SSP 3521 and mouse extendin as deduced from the translation of each cDNA. The results (Table I) demonstrate that the three proteins have a very similar composition. In addition, many of the small differences reflect the known substitutions between RF-hsp 70 and the human and mouse proteins shown in Fig. 1. For example, these substitutions would result in two additional arginines and alanines, two fewer lysines, and three less aspartic acids plus asparagines in RF-hsp 70 when compared with IEF SSP 3521. They would also result in three additional arginines, two more valines, and two fewer leucines in RF-hsp 70 when compared with mouse extendin. In each instance, these substitutions correlate with differences in amino acid composition (Table I). It should also be noted that each protein has 11 cysteines, an amino acid that is not readily identified by sequence analysis. In addition, all cysteines are located identically in the human and mouse proteins, and at least three cysteines appear to be located identically in RF-hsp 70, since no amino acid was identified at three sites in RF-hsp 70 (presumptive cysteines) which would match cysteines at positions 62, 370, and 471 in IEF SSP 3521 and mouse extendin (Fig. 1).

Chromatographic separation of the products produced by trypsin digestion of RF-hsp 70 demonstrates a large peak eluting at 38% acetonitrile, considerably later than the other peptide peaks and close to the elution position (at 39.5% acetonitrile) of undigested RF-hsp 70 (data not shown). We sequenced this trypptic product through 35 cycles and found it to corre
spond to positions 352-386 in IEF SSP 3521 and mouse extendin (Fig. 1). Since no carboxyl terminus was identified, we analyzed this polypeptide by SDS-Polyacrylamide gel electro
phoresis (Fig. 3A, lane 4) and found it to migrate similar to α-lactalbumin (14.4 kDa) (Fig. 3A, lane 2), suggesting a molecular mass of about 14 kDa. This size suggests that the 14-kDa polypeptide extends from position 352 to 470 (arginine in IEF SSP 3521) or 119 amino acids in length. Consistent with this possibility, no other peptide from RF-hsp 70 was isolated from this region of the protein. Electrophoresis of RF-hsp 70, follow
ing incubation with the same concentration of trypsin (0.05 mg/ml) as used for the preparative isolation of peptides for sequence analysis, shows that within 1 h there is complete conversion of the 66-kDa protein to the 14-kDa polypeptide and smaller products (Fig. 3A, lane 5). The 14-kDa component is stable to further incubation with trypsin for 3 and 8.5 h (Fig. 3A, lanes 6 and 7). The 14-kDa polypeptide is not completely resistant to trypsin, since its stability to prolonged trypsinization varies somewhat with different RF-hsp 70 preparations (data not shown), and it is hydrolyzed slowly when RF-hsp 70 is incubated with more concentrated trypsin (Fig. 3B, lane 9).

Quantitation of the 14-kDa polypeptide, using staining following gel electrophoresis, absorbance at 215 nm, and pmol of amino acids produced when subjected to sequence analysis,
indicates that its yield from RF-hsp 70 is approximately 50–80%. These findings indicate that RF-hsp 70 has a 14-kDa polypeptide core that is relatively resistant to a concentration of trypsin which hydrolyzes the remainder of the protein rapidly. This resistance is not due to a lack of tryptic hydrolytic sites in the 14-kDa core, since it has multiple lysine and arginine residues (Fig. 1). In addition, when RF-hsp 70 was digested for 17 h with 0.02 mg/ml trypsin, chymotrypsin, endoproteinase Glu-C, or endoproteinase Lys-C, the 14-kDa band was only seen following digestion with trypsin (Fig. 3), indicating that it had been hydrolyzed to smaller polypeptides by the other three proteases. Much briefer digestions have indicated that components of about 14 kDa are formed by incubating RF-hsp 70 with chymotrypsin or endoproteinase Lys-C, but they are only transiently stable (data not shown).

One possible explanation for the stability of the 14-kDa component to trypsin is that some lysine and arginine residues have undergone post-translational modification, rendering them resistant to cleavage. This explanation would require that such modifications be reversed by the conditions imposed by the amino acid sequence analysis. It also would require that the postulated modification(s) of lysine still permit recognition
and cleavage by endoproteinase Lys-C, since this protease hydrolyzes the 14-kDa component of RF-hsp 70 (Fig. 3B). An alternative explanation is that native RF-hsp 70 possesses sufficient secondary structure that its 14-kDa component is rendered relatively inaccessible to trypsin, although not to the other proteases examined. To test this possibility, we reduced and carboxymethylated a sample of RF-hsp 70 and compared its sensitivity to trypsin with that of native RF-hsp 70. We found that trypsin readily hydrolyzes the 14-kDa component in RCM-RF-hsp 70, but not in RF-hsp 70, when analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 3B, lanes 10 and 11) or reverse-phase chromatography (the large peak eluting at 38% acetonitrile is removed; data not shown). As noted previously (Table 1), RF-hsp 70, IEF SSP 3521, and mouse extendin have 11 cysteines each, and 7 of these in the human and mouse proteins occur within or near the region corresponding to the 14-kDa component in RF-hsp 70 (Fig. 1). This result suggests that the 14-kDa core in RF-hsp 70 is relatively resistant to trypsin because of protein folding that may be dependent upon the presence of one or more intramolecular disulfide bonds. Intermolecular disulfides are not involved, since complexes between RF-hsp 70 molecules are not observed when this protein is analyzed by density gradient centrifugation, as shown in the preceding article (1).

Characterization of the 14-kDa component prompted us to examine the cleavage of RF-hsp 70 by trypsin at earlier times of incubation to determine whether there is discrete processing of this protein. The results in Figs. 4 and 5 demonstrate that this does occur. When incubated with 0.005 mg/ml trypsin, the 66-kDa band of RF-hsp 70 (Fig. 4, lane 2) is reduced by about one-half after approximately 30 s and is almost completely removed after 2 min (Fig. 4, lanes 4–8). It is converted into 58- and 54-kDa components that are formed in approximately equal amounts. These two components are produced at the same time, becoming maximal after 1 min of digestion (Fig. 4, lane 6) and almost completely disappearing after 3.5–5 min (Fig. 4, lanes 9 and 10). The 58- and 54-kDa components are each converted to a 43-kDa product, since this 43-kDa component is almost the only band on the gel after 5 min of digestion (Fig. 4, lane 10), and its staining intensity at this time is close to that predicted if all 66-kDa protein is converted to the 43-kDa polypeptide. To localize the 43-kDa component within the intact 66-kDa protein, a scaled up trypsin digestion of RF-hsp 70, as shown in Fig. 4, was stopped after 5 min and fractionated by reverse-phase chromatography as described under “Experimental Procedures.” The largest peak, eluting at 37% acetonitrile, showed only the 43-kDa band when a small aliquot was subjected to SDS-polyacrylamide gel electrophoresis and silver staining (Fig. 4, lane 13). Amino-terminal sequence analysis of this polypeptide (through seven cycles) corresponds to positions 124–130 in IEF SSP 3521, but with valine substituting for phenylalanine at position 124, as noted in the sequencing of a peptide eluting at 19.5% acetonitrile which was isolated by reverse-phase chromatography of an 18-h trypsin digest of RF-hsp 70 (Fig. 1). Its estimated molecular size from SDS-gel electrophoresis suggests that the 43-kDa component extends from position 124 to about 470 in IEF SSP 3521 and mouse extendin, implying that the 14-kDa component characterized above would comprise the carboxyl-terminal portion of the 43-kDa polypeptide. This is confirmed by the additional digestion and sequencing that is indicated below. This result indicates that the 58- and 54-kDa components are formed from the 66-kDa protein by trypsin hydrolysis at positions 470 and 123, respectively, in IEF SSP 3521 and mouse extendin. The 58- and 54-kDa components would then each be converted to the 43-kDa polypeptide by hydrolysis at positions 123 and 470, respectively. If the 58- and 54-kDa components were formed by trypsin hydrolysis at about positions 73 or 77 and 123, respectively, another intermediate of about 47 kDa should occur, and this is not detected (Fig. 4). Formation of the 58- and 54-kDa components by hydrolysis at about positions 470 and 415 or 428, respectively, is precluded, since this is inconsistent with the formation of the 14-kDa component characterized above and would require that the 43-kDa polypeptide extend from position 124 to about 415 or 428, a length of only 291–304 residues.

When incubated with 0.05 mg/ml trypsin, the 66-, 58-, and 54-kDa components are removed completely within the first 20 s of trypsinization, and the 43-kDa polypeptide is the major product (Fig. 5, lanes 2 and 4). The 43-kDa polypeptide is converted relatively quickly into a 31-kDa polypeptide that becomes maximal at about 2 min of trypsinization (Fig. 5, lanes 8 and 13). Finally, as the 31-kDa polypeptide disappears, it appears to be split into 17- and 14-kDa polypeptides (Fig. 5, lanes 13 and 14). The 17-kDa band is only transiently stable in this digestion (Fig. 5, lanes 13–16), but it is partially stable for a much longer time in digests of another preparation of RF-hsp 70 used for the preparative isolation of tryptic peptides for sequence analysis (see Fig. 3B, lanes 3, 10, and 14). Reverse-phase chromatography of such digests has revealed a large absorbance peak eluting at 34% acetonitrile which corresponds to the 17-kDa polypeptide, as identified by gel electrophoresis of this component (data not shown). In addition, reduction and carboxymethylation of RF-hsp 70 eliminates the relative resistance to trypsin hydrolysis of the 17-kDa polypeptide, as observed for the 14-kDa polypeptide (Fig. 3B, lanes 10 and 11). The 17-kDa polypeptide could only be sequenced through 10 cycles (because of a low yield in the sequenator), but the sequence corresponds exactly to positions 208–217 in IEF SSP.
Table I

| Amino acid | RF-hsp 70 | IEF SSP 3521 | Extendin | Amino acid | RF-hsp 70 | IEF SSP 3521 | Extendin |
|------------|-----------|--------------|-----------|------------|-----------|--------------|-----------|
| Asx        | 62        | 66           | 67        | Met        | 13        | 18           | 17        |
| Thr        | 18        | 20           | 17        | Ile        | 21        | 21           | 19        |
| Ser        | 17        | 18           | 18        | Leu        | 11        | 51           | 53        |
| Glx        | 87        | 83           | 85        | Tyr        | 23        | 29           | 28        |
| Pro        | 30        | 30           | 30        | Phe        | 11        | 11           | 11        |
| Gly        | 23        | 20           | 20        | His        | 11        | 9            | 9         |
| Ala        | 56        | 51           | 53        | Lys        | 61        | 63           | 62        |
| Val        | 17        | 14           | 14        | Arg        | 31        | 27           | 28        |
| Cys        | 11        | 11           | 11        | Trp        | 1         | 1            | 1         |
| (Total)    | 542       | 543          | 543       | (Total)    | 542       | 543          | 543       |

Fig. 3. Hydrolysis of RF-hsp 70 by trypsin produces a relatively stable 14-kDa product. RF-hsp 70 was incubated at 34°C at a concentration of 0.5 mg/ml with 0.05 mg/ml trypsin in 0.02 M Tris-HCl, pH 7.5, 0.05 M KCl, 1 mM dithiothreitol, and 0.10 mM EDTA. Aliquots (2.0 μl) were removed at 1, 3, and 8.25 h and analyzed by SDS-polyacrylamide gel electrophoresis followed by silver staining (panel A, lanes 5-7, respectively). Panel A also contains 0.25 μg each of phosphorylase b, bovine serum albumin, ovalbumin, and chymotrypsinogen (lane 1) and 0.35 μg of α-lactalbumin (lane 2) molecular mass standards, 0.20 μg of RF-hsp 70 (lane 3), and approximately 0.2 μg of the product of trypsin digestion of RF-hsp 70 eluting at 38% acetonitrile upon reverse-phase chromatography as indicated under “Experimental Procedures” (lane 4). Panel B shows the lower portion of the gel electrophoretogram of similar samples that were incubated at 34°C for 17 h and contained, in a final volume of 4.0 μl: 0.07 μg of trypsin alone (lanes 2 and 13) or plus 1.5 μg of RF-hsp 70 (lanes 3 and 14); 0.07 μg of endoproteinase Glu-C alone (lane 4) or plus 1.5 μg of RF-hsp 70 (lane 5); 0.07 μg of chymotrypsin alone (lane 6) or plus 1.5 μg of RF-hsp 70 (lane 7); 0.28 μg of trypsin alone (lane 8) or plus 1.5 μg of RF-hsp 70 (lane 9); and 0.07 μg of endoproteinase Lys-C alone (lane 15) or plus 1.5 μg of RF-hsp 70 (lane 16). Samples in lanes 10–12 contained, respectively: 1.5 μg of RF-hsp 70 plus 0.16 μg of trypsin; 1.5 μg of RCM-RF-hsp 70 plus 0.16 μg of trypsin; or 0.16 μg of trypsin alone (in a final volume of 13 μl). Lanes 1 and 17 contain 0.25 μg of chymotrypsinogen and 0.35 μg of α-lactalbumin. Lanes 1–12 and 13–17 represent separate gel analyses, and the samples in lanes 2–9, 10–12, and 13–16 represent separate incubations. Estimated molecular masses for the products of trypsin digestion of RF-hsp 70 are indicated with arrows on the right (panels A and B).

3521 and mouse extendin (Fig. 1). Its molecular size suggests that the 17-kDa component extends from position 208 to 351, an arginine in human and mouse proteins (although it could terminate at a trypsin site located slightly more amino-terminally), which would be continuous with the amino terminus of the 14-kDa polypeptide that extends from position 352 to, approximately, 470. Since the kinetics of trypsin digestion in Fig. 5 show that as the 31-kDa band disappears the 17- and 14-kDa bands emerge (lanes 13–15), the results strongly suggest that the 31-kDa band represents positions 208–470. Formation of the 31-kDa polypeptide from the 43-kDa polypeptide would then occur by cleavage at positions 206–207 (lysines in
Sequencing of an Hsp 70 Recycling Protein

16847

Fig. 4. Rate of hydrolysis of RF-hsp 70 by 0.005 mg/ml trypsin. RF-hsp 70 was incubated at 34 °C at a concentration of 0.5 mg/ml with 0.005 mg/ml trypsin under the conditions indicated in the legend to Fig. 3. Aliquots (1.5 μl) were removed at 0.33, 0.67, 1.0, 1.5, 2.0, 3.5, 5.0, and 10 min (lanes 4–11, respectively) and subjected to SDS-polyacrylamide gel electrophoresis and silver staining. Aliquots (1.5 μl) of similar samples, containing trypsin without RF-hsp 70, were removed at 2.0 and 10 min and analyzed similarly (lanes 3 and 12). Molecular mass standards as indicated in Fig. 3A were run in lanes 1 and 14, and 0.25 μg of RF-hsp 70 was run in lane 2. Lane 13 contains approximately 0.2 μg of the major product of a preparative trypsin digestion of RF-hsp 70, incubated under the same conditions as the sample in lane 10, which was purified by reverse-phase chromatography and eluted at 37% ace- tonitrile. Estimated molecular masses for the trypsin cleavage products of RF-hsp 70 are shown with arrows on the right. Lanes 1-12 and 13, 14 represent separate gel analyses.

the human and mouse proteins). The relative sensitivity to trypsin hydrolysis of this site in the 43-kDa polypeptide, when compared with the relative stability of the 17- and 14-kDa components derived from the 43-kDa polypeptide, suggests that the region adjacent to and including this site may be particularly exposed to proteolytic attack. This may be related to the run of four glutamic acids (at positions 192–195) and seven prolines (at positions 199–205) which immediately pre- cede lysines at 206 and 207 (Fig. 1). In contrast to the effect of trypsin, which also migrates as a 23-kDa band completely stable to incubation by itself for 17 h at 34°C (data not shown). In addition, none of these is of exceptional length, in contrast to the 14- and 17-kDa polypeptides, since six were sequenced through the carboxyl terminus (and are 9, 10, 10, 16, and 16 amino acids in length, respectively), whereas the seventh is, at most, 28 amino acids long. Furthermore, this phenomenon is not limited to RF-hsp 70 and our trypsinization conditions, since the partial amino acid se- quences of chick p60 obtained by Smith et al. (9) and of IEF SSP 3521 obtained by Honoré et al. (4) also show tryptic peptides that contain internal lysine or arginine (Fig. 1). We have sum- marized these data in Table II, excluding the 14- and 17-kDa components of RF-hsp 70. The results demonstrate that internal lysine and arginine residues occur in tryptic peptides from all three proteins and that these internal tryptic sites occur almost as frequently as carboxyl-terminal lysine and arginine residues. In the case of RF-hsp 70, lysine is more often internal than carboxyl-terminal, whereas arginine is more often carbox- yl-terminal than internal (Table II). In addition, as can be seen in Fig. 1, peptides with internal tryptic sites are derived from approximately all regions of the protein. Where there are over- lapping tryptic polypeptide sequences among the three pro- teins, there appears to be some correlation in the location of these internal sites. For example, there is an internal arginine at position 87 in RF-hsp 70 and p60, an internal lysine at position 210 in RF-hsp 70 and IEF SSP 3521, and an internal lysine at position 317 in RF-hsp 70 and p60, although arginine in position 153 is internal in p60 but carboxyl-terminal in RF-hsp 70 (Fig. 1). In addition, carboxyl-terminal sites in tryptic peptides also overlap: lysine at position 109 in RF-hsp 70 and IEF SSP 3521; arginine at position 160 in RF-hsp 70 and p60; and lysine at position 325 in RF-hsp 70 and p60 (Fig. 1). These findings suggest that there may be similarities in struc- tural organization among these three proteins.

We tested whether trypsin-resistant lysine and arginine residues might occur at multiple sites in RF-hsp 70 because of protection afforded by secondary and possibly higher order structure which may be dependent upon intramolecular disul- fide bond formation. Therefore, equivalent amounts of RF-hsp 70 and RCM-RF-hsp 70 were digested with trypsin and ana- lyzed by reverse-phase chromatography. We found that where individual peptides could be quantitated readily from the chro- matogram (using the absorbance at 215 nm) and compared, reduction and carboxymethylation of RF-hsp 70 lowered the recovery of three polypeptides with internal lysine and/or arginine to 16, 17, and 23% of that seen with native RF-hsp 70, whereas the recovery of three polypeptides lacking internal lysine or arginine was only reduced to 59, 66, and 67% of that in the native protein (data not shown). In addition, we isolated these same peptides by reverse-phase chromatography and lyophilization, and then we reincubated one-half of each with 0.05 mg/ml trysin and the other one-half with buffer only for 7.5 h at 34 °C. Analysis of each by reverse-phase chromatogra- phy demonstrated 86, 95, and 100% trypsin hydrolysis of those peptides containing internal lysine and/or arginine and approxi- mately complete stability of those peptides lacking internal lysine and/or arginine (data not shown). These results strongly suggest that protein folding may protect RF-hsp 70 from tryp- sin hydrolysis at multiple sites in the protein. However, this may not be a complete explanation because tryptic peptides with internal lysine or arginine were obtained from chick p60 after it had been reduced and carboxymethylated (9), and we observe reduced recovery, but not elimination, of tryptic pep-
Extensive sequencing of tryptic peptides of RF-hsp 70, a 66-kDa rabbit reticulocyte protein shown in the previous article to function as a recycling protein for hsp 70 (1), has demonstrated that it is strikingly similar to a transformation-sensitive human protein (IEF SSP 3521), mouse extendin, and chick p60. Specifically, we have shown that RF-hsp 70 (20 polypeptides and 225 residues) has 96% sequence identity to IEF SSP 3521 and 94% identity to mouse extendin. Where partial sequences of RF-hsp 70 and p60 overlap (a total of 54 residues), they show 78% identity. All four proteins are of similar molecular size (about 63 kDa). In addition, IEF SSP 3521 and p60 show multiple isoforms due to charge differences when analyzed by isoelectric focusing (4, 9), and we have observed that RF-hsp 70, which is homogeneous when analyzed by SDS-polyacrylamide gel electrophoresis, shows microheterogeneity when chromatographed on reverse-phase or Mono Q columns (data not shown). Finally, we have noted (Table II) the presence of internal lysine and/or arginine residues in a number of tryptic peptides derived from RF-hsp 70, IEF SSP 3521 (4), and p60 (9). We have shown that this appears to be due to protection afforded by protein folding. We have also noted that the location of these internal sites is often the same in these three proteins (Fig. 1), suggesting that these proteins share similarities in structural organization. The sum of these observations strongly suggests that on a structural basis, RF-hsp 70, IEF SSP 3521, mouse extendin, and chick p60 are homologs of the same protein.

In previous studies of IEF SSP 3521 (4), mouse extendin (6), and chick p60 (9), purified cellular protein was not isolated, and the precise functional role of these proteins was not determined. In one study, Schumacher et al. (12) found that recombinantly expressed (in bacteria) IEF SSP 3521 failed to enhance the renaturation of heat-denatured luciferase by hsp 70 and/or hsp 90, in contrast to the activity of RF-hsp 70, purified from rabbit reticulocyte lysate (1). Schumacher et al. (12) did not comment upon this negative result. We believe that there are a number of possible explanations for their finding other than the implication that RF-hsp 70 and IEF SSP 3521 may be functionally different. For example, IEF SSP 3521 may not fold properly or may become unstable in bacteria, or it may fail to undergo a required covalent modification. In addition, we believe that the known functional characteristics of IEF SSP 3521, extendin, and p60 are consistent with the possibility that these proteins are functionally homologous to RF-hsp 70.

Honé et al. (4) found that the level of IEF SSP 3521 is
increased about 2-fold with a temperature shift from 37 to 42°C, conditions that induce the heat-stress response and an increase in hsp 70. In addition, Blatch et al. (6) observed that purified recombinant extendin interacts with hsp 70 and hsp 90 and may participate in the formation of filamentous actin from its subunits and the extension of cellular pseudopodia. Finally, Smith et al. (9–11) have shown that p60 is associated with hsp 70 and hsp 90 and that p60 enhances the formation of progesterone receptor complexes mediated by these proteins. Smith et al. (9) also found that a monoclonal antibody, prepared against avian p60, reacted specifically with a similar sized protein in rabbit reticulocyte lysate and could be used to immunoprecipitate purify from reticulocyte lysate a complex containing hsp 90, hsp 70, and p60. Since we have shown RF-hsp 70 to be associated with and to bind to hsp 70 (1), these findings add strong support to our contention that RF-hsp 70 and p60 are homologous proteins, and they strongly suggest that IEF SSP 3521 and mouse extendin may also be functionally analogous to RF-hsp 70 activity in rabbit reticulocyte lysate.

One additional indication that these four proteins may have a similar functional role in the heat-stress response is that they all share considerable sequence identity with STI1, a heat- and stress-inducible protein in yeast (5), and GMSTI, a heat-inducible soybean protein (Glycine max) protein (13). Nicolet and Craig (5) observed a 10-fold increase in STI1 mRNA upon a temperature shift from 23 to 39°C or treatment with the amino acid analog canavanine. They also noted impaired growth at temperatures above 30°C in cells carrying a disruption mutation in the STI1 gene (5). Similarly, Torres et al. (13) found that the concentration of mRNA for the GMSTI protein increases sharply and markedly in soybean plants shifted from 25 to 42°C. They observed a 10-fold increase in STI1 mRNA upon a temperature shift from 37 to 42°C, conditions that induce the heat-stress response and an increase in hsp 70. In addition, Blatch et al. (4) that this portion of IEF SSP 3521 (positions 351–470) contains four of the 10 putative repeats identified in the complete protein which match the tetra-tripeptide repeat. This represents a 34-amino acid repetitive sequence motif that is characteristic of a family of proteins (14–16) including several that are involved in cell cycle control (14, 15, 17, 18) or RNA synthesis regulation (19, 20), an outer mitochondrial membrane protein (21), and the heat- and stress-inducible proteins STI1 in yeast (5) and GMSTI in soybean (13). Proteins possessing these repeats may be dependent upon them for their function. For example, as noted by Boguski et al. (16), a truncated form of the gene product which is involved in the import of proteins into mitochondria, which lacks the carboxy-terminal 203 residues including almost the entire tetra-tripeptide repeating domain, becomes normally attached to the outer mitochondrial membrane but cannot substitute for the function of the intact protein (22). In addition, Boguski et al. (16) reported that single amino acid substitutions in the cell division cycle protein CDC23, which result in a temperature-sensitive phenotype, cluster into two separate domains, one of which includes the carboxy-terminal, tetra-tripeptide repeating segment. The possible role of these repeating elements and of the 14-kDa component in the function of RF-hsp 70 and its human (IEF SSP 3521), mouse (extendin), and chicken (p60) counterparts remains to be determined.

Findings presented in this paper have also demonstrated that trypsin hydrolyzes RF-hsp 70 to 58- and 54-kDa polypeptides, each of which is then converted to a 43-kDa product and then a 31-kDa polypeptide, which is finally split into 17- and 14-kDa polypeptides. The 43-kDa component is located at position 124 to about 470 in IEF SSP 3521 and mouse extendin, and the 17- and 14-kDa components are located at approximately positions 208–351 and 352–470, respectively. The 17-kDa component is transiently stable to trypsin, whereas the 14-kDa component is stable to trypsin for relatively long periods of time. In contrast, both 17- and 14-kDa polypeptides are hydrolyzed by trypsin when RF-hsp 70 has been reduced and carboxymethylated. Our findings suggest that the relative resistance of the 14-kDa polypeptide to trypsin may be due to a high degree of protein folding in this region which may be dependent upon intramolecular disulfide bond formation. Consistent with this explanation is the presence of 7 cysteines in IEF SSP 3521 and mouse extendin located within or near the region corresponding to the 14-kDa component in RF-hsp 70.