Analysis of Intracellular Substrates and Products of Thimet Oligopeptidase in Human Embryonic Kidney 293 Cells

Denise A. Berti§, Cain Morano§, Lilian C. Russo§, Leandro M. Castro§, Fernanda M. Cunha§, Xin Zhang§, Juan Sironi§, Clélio F. Kiltzke**, Emer S. Ferro§, and Lloyd D. Fricker§

From the Departments of §Cell Biology and Development and §Pharmacology, Biomedical Science Institute, University of São Paulo, São Paulo SP 05508-900, Brazil, the Department of Biochemistry, Federal University of São Paulo, São Paulo SP 04044-020, Brazil, the **Center for Applied Toxicology, CAT/CEPID, Butantan Institute, São Paulo 05503-900 SP, Brazil, and the §Department of Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, New York 10461

Thimet oligopeptidase (EC 3.4.24.15; EP24.15) is an intracellular enzyme that has been proposed to metabolize peptides within cells, thereby affecting antigen presentation and G protein-coupled receptor signal transduction. However, only a small number of intracellular substrates of EP24.15 have been reported previously. Here we have identified over 100 peptides in human embryonic kidney 293 (HEK293) cells that are derived from intracellular proteins; many but not all of these peptides are substrates or products of EP24.15. First, cellular peptides were extracted from HEK293 cells and incubated in vitro with purified EP24.15. Then the peptides were labeled with isotopic tags and analyzed by mass spectrometry to obtain quantitative data on the extent of cleavage. A related series of experiments tested the effect of overexpression of EP24.15 on the cellular levels of peptides in HEK293 cells. Finally, synthetic peptides that corresponded to 10 of the cellular peptides were incubated with purified EP24.15 in vitro, and the cleavage was monitored by high pressure liquid chromatography and mass spectrometry. Many of the EP24.15 substrates identified by these approaches are 9–11 amino acids in length, supporting the proposal that EP24.15 can function in the degradation of peptides that could be used for antigen presentation. However, EP24.15 also converts some peptides into products that are 8–10 amino acids, thus contributing to the formation of peptides for antigen presentation. In addition, the intracellular peptides described here are potential candidates to regulate protein interactions within cells.

Intracellular protein turnover is a crucial step for cell functioning, and if this process is impaired, the elevated levels of aged proteins usually lead to the formation of intracellular insoluble aggregates that can cause severe pathologies (1). In mammalian cells, most proteins destined for degradation are initially tagged with a polyubiquitin chain in an energy-dependent process and then digested to small peptides by the 26S proteasome, a large proteolytic complex involved in the regulation of cell division, gene expression, and other key processes (2, 3). In eukaryotes, 30–90% of newly synthesized proteins may be degraded by proteasomes within minutes of synthesis (3, 4). In addition to proteasomes, other extralysosomal proteolytic systems have been reported (5, 6). The proteasome cleaves proteins into peptides that are typically 2–20 amino acids in length (7). In most cases, these peptides are thought to be rapidly hydrolyzed into amino acids by aminopeptidases (8–10). However, some intracellular peptides escape complete degradation and are imported into the endoplasmic reticulum where they associate with major histocompatibility complex class I (MHC-I) molecules and traffic to the cell surface for presentation to the immune system (10–12). Additionally, based on the fact that free peptides added to the intracellular milieu can regulate cellular functions mediated by protein interactions such as gene regulation, metabolism, cell signaling, and protein targeting (13, 14), intracellular peptides generated by proteasomes that escape degradation have been suggested to play a role in regulating protein interactions (15). Indeed, oligopeptides isolated from rat brain tissue using the catalytically inactive EP24.15 (EC 3.4.24.15) were introduced into Chinese hamster ovarian-S and HEK293 cells and were found capable of altering G protein-coupled receptor signal transduction (16). Moreover, EP24.15 overexpression itself changed both angiotensin II and isoproterenol signal transduction, suggesting a physiological function for its intracellular substrates/products (16).

EP24.15 is a zinc-dependent peptidase of the metallopeptidase M3 family that contains the HEXXH motif (17). This enzyme was first described as a neuropeptide-degrading enzyme was first described as a neuropeptide-degrading enzyme.
Peptidomic Analysis for EP24.15

enzyme present in the soluble fraction of brain homogenates (18). Whereas EP24.15 can be secreted (19, 20), its predominant location in the cytosol and nucleus suggests that the primary function of this enzyme is not the extracellular degradation of neuropeptides and hormones (21, 22). EP24.15 was shown in vivo to participate in antigen presentation through MHC-I (23–25) and in vitro to bind (26) or degrade (27) some MHC-I associated peptides. EP24.15 has also been shown in vitro to degrade peptides containing 5–17 amino acids produced after proteasome digestion of β-casein (28). EP24.15 shows substrate size restriction to peptides containing from 5 to 17 amino acids because of its catalytic center that is located in a deep channel (29). Despite the size restriction, EP24.15 has a broad substrate specificity (30), probably because a significant portion of the enzyme-binding site is lined with potentially flexible loops that allow reorganization of the active site following substrate binding (29). Recently, it has also been suggested that certain substrates may be cleaved by an open form of EP24.15 (31). This characteristic is supported by the ability of EP24.15 to accommodate different amino acid residues at subsites S4 to S3’, which even includes the uncommon post-proline cleavage (30). Such biochemical and structural features make EP24.15 a versatile enzyme to degrade structurally unrelated oligopeptides.

Previously, brain peptides that bound to catalytically inactive EP24.15 were isolated and identified using mass spectrometry (22). The majority of peptides captured by the inactive enzyme were intracellular protein fragments that efficiently interacted with EP24.15. The smallest peptide isolated in these assays contained 5 and the largest 17 amino acids (15, 16, 22, 32), which is within the size range previously reported for natural and synthetic substrates of EP24.15 (18, 30, 33, 34). Interestingly, the peptides released by the proteasome are in the same size range of EP24.15 competitive inhibitors/substrates (7, 35, 36). Taken altogether, these data suggest that in the intracellular environment EP24.15 could further cleave proteasome-generated peptides unrelated to MHC-I antigen presentation (15).

Although the mutated inactive enzyme “capture” assay was successful in identifying several cellular protein fragments that were substrates for EP24.15, it also found some interacting peptides that were not substrates. In this study, we used several approaches to directly screen for cellular peptides that were cleaved by EP24.15. The first approach involved the extraction of cellular peptides from the HEK293 cell line, incubation in vitro with purified EP24.15, labeling with isotopic tags, and analysis by mass spectrometry to obtain quantitative data on the extent of cleavage. The second approach examined the effect of EP24.15 overexpression on the cellular levels of peptides in the HEK293 cell line. The third set of experiments tested synthetic peptides with purified EP24.15 in vitro, and examined cleavage by high pressure liquid chromatography and mass spectrometry. Collectively, these studies have identified a large number of intracellular peptides, including those that likely represent the endogenous substrates and products of EP24.15, and this original information contributes to a better understanding of the function of this enzyme in vivo.

EXPERIMENTAL PROCEDURES

Reagents—Dimethyl sulfoxide (Me 2SO), α-cyano-4-hydroxycinnamic acid, 4-aminobutyric acid, N-hydroxysuccinimide, 1,3-dicyclohexylcarbodiimide, potassium hydrogen carbonate (KHCO 3), anhydrous methanol, acetone, tetrahydrofuran, chloroform, formic acid, iodomethane (CH 3I), and iodomethane-d 3 (CD 3I) were obtained from Aldrich. Acetonitrile and trifluoroacetic acid were purchased from Fisher. Hydrochloridric acid and hydroxylamine were supplied by Merck. Glycine and sodium hydroxide (NaOH) were obtained from Sigma. Sodium phosphate, dibasic anhydrous (Na 2HPO 4), and sodium phosphate monobasic anhydrous (NaH 2PO 4) were from Amresco (Solon, OH). The 4-trimethylammoniumbutyryl (TMAB) stable isotopic labeling reagents, containing either 9 atoms of deuterium (D9-TMAB) or no deuteriums (D0-TMAB) were synthesized as described (37–39). Rabbit anti-EP24.15 was from Proteimax Biotechnology (Cotia, São Paulo, Brazil), and goat anti-rabbit IgG conjugated to peroxidase was from Amersham Biosciences. Fluorescamine was from Invitrogen.

Peptide Extraction—Crude peptide extracts from HEK293 cells were prepared from 8 × 10 7 cells that were first washed three times with phosphate-buffered saline by centrifugation at 800 × g for 5 min. The pellet was resuspended in 10 ml of H 2O and was heated in a water bath for 20 min at 80 °C to inactivate proteases. After cooling in ice, 20 μl of 5 N HCl was added to give a final concentration of 10 mM. The cells were sonicated three times with 20 pulses (4 Hz). The homogenates were centrifuged at 1,500 × g for 40 min at 4 °C. The supernatants were collected in plastic ultracentrifuge tubes and centrifuged at 100,000 × g for 30 min at 4 °C. The supernatant was again collected and filtered through a Millipore centrifugal filter unit with a molecular mass cutoff of 5,000 Da. Peptides contained in the samples were further purified and concentrated with C18-like Oasis columns (Waters) and dried in a vacuum centrifuge. Peptide concentration in these samples was determined at pH 6.8 using fluorescamine as described previously (28, 40). A peptide mixture of known composition and concentration was used as the standard reference for determining the peptide concentration in HEK293 extracts (data not shown).

EP24.15 Expression—Rat testis recombinant EP24.15 was bacterially expressed as a glutathione S-transferase fusion protein as described previously (41). Briefly, protein purification was done by affinity chromatography using a glutathione-Sepharose column (Amersham Biosciences) and concentrated by centrifugation onto Millipore membranes with nominal molecular exclusion limits of 50,000 Da. The concentration of recombinant protein was determined using the Bradford assay (42). The protein homogeneity was analyzed by Coomassie Brilliant Blue staining after electrophoresis in 8% polyacrylamide gels containing SDS-PAGE. After confirming the homogeneity greater than 95% (data not shown), the protein preparation was stored at −80 °C.

Digestion of Cellular Peptides with EP24.15—The peptide extracts (160 μg) from HEK293 cells were incubated with either 260 or 2600 ng of recombinant EP24.15 in a final volume of 200 μl (final concentrations of EP24.15 were 16 and 160 nM, respec-
tively) in TBS containing 1 mM β-mercaptoethanol. The mixture was incubated for 20 min at room temperature, and the reaction was interrupted by heating at 100 °C for 4 min. The control reaction was prepared similarly except that the EP24.15 used was previously heated at 100 °C for 4 min before adding to the HEK293 peptide extract. After the reaction, the extracts were labeled with TMAB isotopic tags, as described below.

**HEK293 Cell Culture and Transfections**—Human embryonic kidney cells 293T (HEK293; Invitrogen) were cultured in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum, penicillin (100 units/ml), and streptomycin (100 µg/ml; Invitrogen) and incubated at 37 °C in a humidified atmosphere of 5% CO₂ and 95% air. Rat testis EP24.15 full-length cDNA was subcloned into the a modified pShooter vector (Invitrogen) in-frame with the c-Myc epitope for C-terminal fusion to the recombinant proteins, as described previously (19). DNA sequencing confirmed the correct frame orientation for all constructs (data not shown). Vector DNA was purified using a Wizard Plus kit (Promega, Madison, WI). Transient HEK293 cell transfections were obtained using the cell-specific liposome protocols. EP24.15 protein overexpression was analyzed by both enzyme activity and Western blotting using mock-transfected cells (transfected with the modified pShooter vector only) as a control. To check transfection efficiency, a green fluorescent protein coding plasmid (pEGFP-N1, Clontech) was also used in some assays. To assay the effect of EP24.15 overexpression on the level of cellular peptides, HEK293 cells were transfected with EP24.15 or plasmid alone, and after 2 days the cells were removed from the dish, washed with phosphate-buffered saline, and heated in an 80 °C water bath for 20 min to inactivate cellular proteases. The peptides were extracted, labeled with isotopic tags, and analyzed by mass spectrometry as described below.

**Isotopic Labeling**—HEK293 peptide extracts were combined with 200 µl of 0.4 M phosphate buffer, pH 9.5. The pH was adjusted to 9.5 with 1 M NaOH. For each sample, 7.0/4H11003 reagent. After 40 min at room temperature, the hydrogen-and was added to the reaction to quench any remaining labeling was incubated for 30 min. After incubation, 30 µl of 2.5 M glycine was added to the reaction to quench any remaining labeling reagent. After 40 min at room temperature, the hydrogen- and deuterium-labeled samples were combined and centrifuged at 800 × g for 5 min at 4 °C. The pH was adjusted to 9.0–9.5, and 10 µl of 2.0 M hydroxylamine was added to remove TMAB labels from Tyr residues. The addition of hydroxylamine was repeated twice more over 30 min. The samples were desalted with a C18 column (Oasis, Millipore). The peptides were eluted from the C18 column with 1 ml of 100% methanol and 0.1% trifluoroacetic acid. The eluate was dried in a vacuum centrifuge and resuspended in 20 µl of water. Aliquots of 4.5 µl of sample were analyzed as described below.

**Peptidomic Analysis for EP24.15**

**Liquid Chromatography and Tandem Mass Spectrometry (LC-MS/MS) Analysis**—LC-MS/MS experiments were carried out on a Q-Tof-Ultima mass spectrometer (Micromas, Manchester, UK) or an API Q-Star Pulsar-i™ quadrupole time-of-flight mass spectrometer (Applied Biosystems/MDS Sciex, Foster City, CA). For the analysis performed on the Q-Tof, the peptide mixture was desalted on line for 15 min using a Symmetry C18 trapping column (5-µm particles, 180 µm inner diameter × 20 mm, Waters). The mixture of trapped peptides was then separated by elution with a water/acetonitrile, 0.1% formic acid gradient through a BEH 130-C18 column (1.7-µm particles, 100 µm inner diameter × 100 mm, Waters). Data were acquired in data-dependent mode, and multiply charged protonated peptide ions generated by electrospray ionization (ESI) were automatically mass selected and dissociated in MS/MS by 10–30-eV collisions with argon. Typical LC and ESI conditions were a flow rate of 600 nl/min, nanoflow capillary voltage of 3.5 kV, block temperature of 100 °C, and cone voltage of 100 V. For the analysis performed on the Q-Star, the peptide mixture was loaded on a PepMap™ C18 trapping column (5 µm, 100 Å, 300-µm inner diameter × 5 mm, LC Packaging, Marlton, NJ) and desalted for 30 min with 5% acetonitrile in 0.1% formic acid. Then the peptides were separated on a Vydac MS C18 capillary column (3 µm, 300 Å, 300 µm inner diameter × 250 mm, Vydac, Hesperia, CA) by gradient elution using solvents A and B at a flow rate of 0.25 µl/min. Solvent A was 2% acetonitrile, 0.1% formic acid in water. Solvent B was 98% acetonitrile, 0.1% formic acid in water. The gradient was 5%–45% solvent B over 50 min. The most intense ions in MS scans were fragmented to generate MS/MS spectra. The dynamic exclusion time for fragmented ions was 240 s. The collision energy was in the 20–45-eV range and dynamically changed based on the m/z value and charge state of the ion.

**MS Data Analysis**—To identify peptides, the raw data files were converted to a peak list format (mgf) by the software Mascot Distiller version 2.1.1 (Matrix Science Ltd., London, UK) and analyzed using the search engine MASCOT version 2.2 (Matrix Science Ltd.). The variable modifications were selected as TMAB (GIST-Quat (K) and GIST-Quat (N-term) for the labels with nine hydrogen and GIST-Quat:2H (9) (K) and GIST-Quat:2H (9) (N-term) for the labels with nine deuterium at N-terminal and lysine residues, respectively. The searches used the NCBI (data base from 01/18/2008, 5,303,346 sequences) with the taxonomy *Homo sapiens* (human), “no enzyme,” and 0.1 Da of mass tolerance for MS and MS/MS spectra. Mascot searches were followed by manual interpretation to eliminate false positives. Several criteria were used to accept or decline the peptides that were identified by Mascot as follows: (i) the majority (>80%) of the major MS/MS fragment ions matched predicted a, b, or y ions or parent ions with loss of trimethylamine; (ii) a minimum of five fragment ions matched b or y ions; and (iii) the number of tags incorporated into the peptide matched the number of free amines (N terminus and side chains of Lys). In addition, the Mascot score was either the top score of all potential peptides, or the other peptides with comparable scores could be excluded by the above criteria, leaving only one peptide that matched all criteria. For example, the peptide Ac-ASKRALVIL met the above criteria and had a
Peptidomic Analysis for EP24.15

Mascot score of 28 (see supplemental material). Other peptides had slightly lower scores. But most of these other peptides contain both a free N-terminal amine and an internal Lys; these would therefore have incorporated two isotopic tags. Because the observed peptide incorporated just one isotopic tag (based on the mass difference between the heavy and light peaks), these other peptides are not as likely to be correct. Furthermore, another MS/MS spectrum was identified by Mascot as Ac-ASKRALVILA, with the score of 29 (see supplemental material). Although this score was only marginally higher than the next peptide on the list, this other peptide could be excluded because it should have incorporated two tags and the detected peptide only incorporated one tag. Finally, the MS/MS spectra for the peptides identified as Ac-ASKRALVILA and Ac-ASKRALVILA by Mascot showed many of the same b ions, and therefore are likely to be related. This “manual validation” approach was applied to all of the peptides reported in this study. Quantification was performed by measuring the ratio of peak intensity for the D0- and D9-TMAB-labeled peptides pairs in the MS spectra. For this analysis, the monoisotopic peak and the peaks containing one and two atoms of 13C were used. Multiple scans of the MS spectra were combined prior to quantitation.

Enzyme Assays and Analysis of Synthetic Peptides—The enzymatic activity of EP24.15 was determined in triplicate using a continuous assay with a quenched fluorescent substrate (7-methoxycoumarin-4-acetyl-PLGPdK-(2,4-dinitrophenyl), as described previously (43)). To discern peptidolytic activity (7-methoxycoumarin-4-acetyl-PLGPdK-(2,4-dinitrophenyl), using a continuous assay with a quenched fluorescent substrate enzymatic activity of EP24.15 was determined in triplicate quantitation.

JOURNAL OF BIOLOGICAL CHEMISTRY

14108

FIGURE 1. Diagram of HEK293 peptide isotopic labeling. Experiment 1, HEK293 cellular peptide extract (160 µg) was incubated with active EP24.15 (260 or 2600 ng). As a control, the same amount of heat-inactivated EP24.15 was incubated for 20 min with 160 µg of HEK293 cellular peptide extract. After extract, but before purification of the peptides and subsequent analysis, the peptides were labeled with one of two isotopic forms of the TMAB label, either containing 9 atoms of deuterium (D9-TMAB) or only the hydrogenated form (D0-TMAB). The TMAB reagent labels the free amine group present on the N terminus of most peptides and also on the side chain of Lys residues (unless the amine group is acetylated or otherwise modified). For each concentration of enzyme, duplicate incubations were performed, and the duplicate digestion products were labeled with D0- or D9-TMAB in the opposite orientation from each other (runs 1 and 3 versus runs 2 and 4). Experiment 2, HEK293 cells were transfected with EP24.15-expressing cDNA or with the control cDNA plasmid. Extracts of EP2415-expressing and control cells were labeled with D0- or D9-TMAB, as indicated, for a total of four independent samples and LC/MS runs.

RESULTS

In this study, several approaches were used to identify and characterize endogenous substrates of EP24.15. The first approach tested HEK293 cell extracts with purified enzyme, incubated in vitro, followed by a quantitative peptidomics analysis using differential isotopic tags and mass spectrometry (Fig. 1). The second approach tested the effect of overexpression of EP24.15 on the levels of peptides in the HEK293 cells (Fig. 1). Finally, the third approach examined the cleavage of synthetic peptides with purified EP24.15.

In this study, the first quantitative peptidomics approach resulted in the identification of 93 peptides from MS/MS anal-
Analysis using the Mascot search program followed by manual interpretation (supplemental Table 1 and Fig. 2). Mascot search results are included as supplemental material, and a representative MS/MS spectrum is indicated in Fig. 2. Manual interpretation of the Mascot search results was necessary because the program does not consider the loss of trimethylamine from the parent ion, and these ions are often the major ions in the spectra (i.e. the 555.936$^{3+}$ ion in Fig. 2). Thus, the Mascot scores were lower than typically considered for positive identification, and so manual assignment of the major peaks was necessary to confirm the identity of each peptide.

The size of identified peptides ranged from 5 to 22 amino acids in length, with most peptides in the range of 7–14 amino acids (supplemental Table 1); this is consistent with the average size of proteasome-mediated protein degradation products (7, 45). Moreover, hydrophobic, basic, and acidic amino acid residues were present on the C terminus of >90% of the peptides, whereas relatively common amino acids like Gly, Pro, Ser, Thr, Gln, and Asn were rarely found on the C terminus (supplemental Table 1). These data are also consistent with the specificity of proteasome-mediated protein degradation for hydrophobic, basic, and acidic residues (46).

Some peptides decreased upon incubation of the cell extract with active EP24.15, and likely represent enzyme substrates (supplemental Table 1 and Fig. 3, A and B; note the change in peak intensity when the labeling scheme was reversed). Other peptides showed an increase upon incubation with active EP24.15 (supplemental Table 1 and Fig. 3, C and D), likely representing products of enzyme cleavage. Other peptides were not affected by the incubation with active EP24.15 (supplemental Table 1 and Fig. 3, E and F). These peptides are not likely to be substrates or products of the enzyme.

Upon incubation with 2600 ng of active EP24.15, the majority of peptides identified in the HEK293 extracts were decreased 30% or more (active/inactive enzyme ratio is less than 0.70; supplemental Table 1). Some of the peptides were not detectable after incubation with active EP24.15, and these are listed in supplemental Table 1 with an active/inactive enzyme ratio
When incubated with the high concentration of EP24.15, 10 peptides increased more than 70%, and three peptides showed a smaller increase of 30–70% (supplemental Table 1). These peptides represent products of EP24.15. With the lower concentration of EP24.15, only five peptides were elevated more than 70%, and two showed intermediate levels of elevation (supplemental Table 1). As expected, all five of the peptides substantially elevated by the low concentration of EP24.15 were also elevated by the high concentration of enzyme. One of the peptides partially elevated by the low concentration showed a larger increase with the higher concentration of enzyme. However, the other peptide partially elevated by the low concentration of enzyme (Splicing factor, arginine/serine-rich 1-derived peptide AIRDIDIL) showed a large decrease with the higher concentration of enzyme (supplemental Table 1). A possible interpretation of this result is that AIRDIDIL is both a product of EP24.15 that results from cleavage of a larger peptide and also a substrate of EP24.15 that is cleaved by the higher concentration of the enzyme.

A total of 26 peptides did not change more than 30% when incubated with either concentration of EP24.15; these are neither substrates nor products of the enzyme. For all of the analyses, the forward and reverse labeling generally produced similar results, and the variation among the duplicate determinations was typically less than 10%.

When considering the length of the peptides, the majority of the EP24.15 substrates (i.e. active/inactive enzyme ratios of 0.70 or less) are 9–11 amino acids long for either the low or high concentration of enzyme (Fig. 4). Furthermore, the four best substrates were 9–11 amino acids in length (supplemental Table 1). The products of EP24.15 (i.e. active/inactive enzyme ratios of 1.30 or more) range from 6 to 12 amino acids for the low concentration of enzyme and 6 to 13 amino acids for the high concentration of enzyme (Fig. 4). Although more of the products are in the lower part of these ranges, both concentrations of enzyme resulted in products that were 10–11 amino acids in length.

Ten of the peptides identified as products were also found in larger forms that were identified as substrates (supplemental Table 1). In all cases, a single cleavage of the substrate would be
required to generate the observed product (supplemental Table 1). For example, the peptide GFGDLKSPAGLOQ was identified as a substrate, and both GFGDLK and SPAGLOQ were identified as products; therefore, it is highly likely that EP24.15 cleaved at the Lys-Ser bond. For the other sets of peptides identified as substrates/products, only the N-terminal product was observed and not the shorter C-terminal product; 2–4 amino acid-long peptides are difficult to identify by mass spectrometry. To test if the prediction was correct, peptides corresponding to three of the substrates were custom-synthesized, incubated with purified EP24.15, and then analyzed by LC/MS (Fig. 5). The peptide SAMTEEEAIVAIKAMAK was found to be cleaved at the predicted site, between the Lys and the Ala located four residues from the C terminus. The synthetic peptide VFDVELLKLLE was also cleaved four amino acids from the C terminus at the Leu-Leu bond (Fig. 5), as predicted from the analysis of the substrate/product pairs (Table 1). The peptide ELFADKVPKTAENFR was predicted to be cleaved at two sites as follows: one four residues from the end at the Ala-Glu bond and the other two residues from the end at the Asn-Phe bond (Table 1). Of these, the cleavage four residues from the end showed a larger increase in the peptidomics analysis of HEK293 cell extract (supplemental Table 1; ratio = 3.21) than the cleavage two residues from the end (ratio 1.74). Consistent with this observation, the synthetic peptide incubated with EP24.15 showed the major cleavage site to be the Ala-Glu bond four residues from the C terminus, and the other product was not detected (Table 1 and Fig. 5).

To evaluate the physiological involvement of EP24.15 in metabolism of peptides, HEK293 cells were transfected with either EP24.15 cDNA encoding plasmid or control modified pShooter vector only. Using the HEK293 cell-specific liposome HEKfectin™, transfection efficiency was ~95% as observed with a green fluorescent protein-expressing plasmid (data not shown). HEK293 cells transfected with EP24.15 cDNA encoding plasmid showed higher EP24.15 protein (Fig. 6A) and enzymatic activity (Fig. 6B) levels compared with mock-transfected cells. These data confirm the overexpression of EP24.15 protein and enzyme activity in our model. Peptidomics analysis of the EP24.15-transfected and mock-transfected cells identified 38 peptides. Of these 38 peptides, 5 were decreased more than 70% (average ratios smaller than 0.3) and another 12 were partially decreased (average ratios from 0.3 to 0.69) upon overexpression of EP24.15 (supplemental Table 2); these represent in vivo substrates of this enzyme. Three peptides were elevated (average ratios greater than 1.3) by the overexpression of EP24.15, two of these by 3-fold or more (supplemental Table 2). These elevated peptides represent in vivo products of EP24.15. Eighteen of the identified peptides were not altered (average ratios from 0.7 to 1.3) by the overexpression of EP24.15 and are neither substrates nor products of the enzyme (supplemental Table 2).

To determine whether the observed changes in peptide levels were statistically significant, additional controls were performed comparing untransfected HEK293 cells with mock-transfected cells (supplemental Table 2, column 9). These peptides were expected to show ratios close to 1.0; the actual values observed are a reflection of the natural variation of each peptide between replicate culture plates. A comparison of the ratio observed from these control experiments with that observed from the EP24.15-transfected cells (supplemental Table 2, column 8) showed that nine peptides were significantly altered by the EP24.15 expression (supplemental Table 2, column 10). In addition, several other peptides showed comparable changes as these nine peptides, but because these other peptides were not detected in the control LC/MS runs, it was not possible to validate the changes.

To further confirm that peptides altered by overexpression of EP24.15 in HEK293 cells are substrates or products of the enzyme, seven additional peptides were synthesized and tested with purified EP24.15, along with the three peptides shown in Fig. 5. All of the peptides predicted to be good substrates were efficiently cleaved by purified EP24.15 (Table 2). Importantly, the cleavage of these peptides performed under zero order
kinetics was completely prevented by the EP24.15 inhibitor N-(1-(RS)-carboxy-3-phenylpropyl)-Ala-Aib-Tyr-p-amino-benzoate (data not shown). The peptide AVA-9 was cleaved by purified EP24.15 with the highest efficiency of all peptides studied (Table 2); this peptide showed a very large decrease upon overexpression of EP24.15 in HEK293 cells (supplemental Table 2). The other peptides predicted to be substrates were also cleaved by purified EP24.15, but the efficiency of cleavage was lower for these peptides than for AVA-9 (Table 2). The predicted product of EP24.15 (SAM-12) and the peptide that

![FIGURE 5. Analysis of cleavage of synthetic peptides by purified EP24.15. Peptides were incubated with purified EP24.15 and analyzed on LC/MS as described under "Experimental Procedures." For each peptide, all MS spectra representing peptides were accumulated into a single spectrum as follows: A and C, the indicated spectrum represents all MS spectra from 20 to 30 min; B, indicated spectrum represents the sum of MS spectra from 25 to 31.5 min. A, most intense fragment from peptide SAMTEEAAVAIKAMAK was SAMTEEAAVAIK (1+ or 2+); B, most intense fragment from peptide VFDVELLKLLE was VFDVEL (1+); C, most intense fragment from peptide ELFADKVPKTAENFR was ELFADKVPKTA (1+ or 2+).

![TABLE 1](image)

**Analysis of product/substrate relationships and prediction of likely cleavage sites**

Peptides that were found to be products of EP24.15 (as defined in Fig. 4 legend) and that corresponded to shorter forms of peptides identified as substrates (also as defined in Fig. 4) are indicated. As shown in supplemental Table 1, peptide sequences in parentheses were tentatively identified based on mass, charge, and number of TMAB tags incorporated; all other sequences were identified by these criteria and also by MS/MS sequence. The cleavage site within the substrate that would generate the observed product is indicated with a dash. Three of the substrate peptides were synthesized and tested with purified EP24.15, as shown in Fig. 5, and the results are summarized in this table.

| Protein name                        | Product          | Substrate (and likely cleavage site) | Comment                                      |
|--------------------------------------|------------------|--------------------------------------|----------------------------------------------|
| Cytochrome c oxidase subunit VIII    | IHSLPPE          | IHSLPPE-AMAK                          | Confirmed with synthetic peptide             |
| Eukaryotic translation elongation factor 1β2 | GFGDLK          | GFGDLK-SPAGLQV                        | Confirmed with synthetic peptide             |
| Eukaryotic translation elongation factor 1β2 | (SPAGLQV)       | GFGDLK-SPAGLQV                        | Confirmed with synthetic peptide             |
| Eukaryotic translation initiation factor 5A | SAMTEEAAVAIK    | SAMTEEAAVAIK-AMAK                     | Confirmed with synthetic peptide             |
| Eukaryotic translation initiation factor 5A | ELFADKVPKTAEN   | ELFADKVPKTAEN-FR                      | Not major cleavage site of synthetic peptide |
| FK506-binding protein                | DIEDVF-YKY      | DIEDVF-YKY                            |                                               |
| Peptidylprolyl isomerase A           | KDIEDVF         | KDIEDVF                               |                                               |
| Peptidylprolyl isomerase A           | (DIEDVF)        | DIEDVF                                |                                               |

The predicted product of EP24.15 (SAM-12) and the peptide that
was unaffected by overexpression of enzyme (VFD-7) were not cleaved by purified EP24.15 (Table 2).

The cleavage sites of these additional synthetic peptides were determined by LC/MS/MS (Table 2), as shown for three of the peptides (Fig. 5). In most cases, the cleavage site of the synthetic peptide matched the cleavage predicted from the analysis of the HEK293 peptidome (supplemental Table 1, Table 1, and supplemental Table 2). For example, the synthetic peptide AIR-10 was found to be cleaved into AIR-DIDL and KNR (Table 2). Similarly, the peptide AIR-DIDL-KNR was found in cells (supplemental Table 2) to be a substrate of EP24.15 and the peptide AIR-DIDL was found in vitro to be a product when the cell peptide extract was incubated with the high concentration of EP24.15 (supplemental Table 1).

**DISCUSSION**

In this study, three different approaches were used to identify natural substrates of EP24.15, and each approach has its advantages and disadvantages. The studies investigating the cleavage of synthetic peptides by purified enzyme are best for determining the relative hydrolysis ratio and cleavage site within a peptide, but because each synthetic peptide is examined individually, it is not a high throughput approach to screen for novel substrates. In contrast, the two peptidomics approaches, one using peptide extracts as substrates for purified enzyme and the other testing the effect of overexpression of the enzyme on endogenous peptides, measure dozens of peptides simultaneously. In addition, these studies are not limited to previously identified peptides. The peptidomics approaches more closely resemble the in vivo conditions where the protease is exposed to numerous substrates. However, these peptidomics approaches cannot conclusively identify the cleavage sites within a particular peptide. Thus, the combination of approaches used in this study provides a more thorough analysis than each individual approach alone.

In this study, we identified 116 peptide fragments of intracellular proteins in HEK293 cell extracts. Although some of the same peptides were found in all extracts of HEK293 cells, many peptides were found in just one of the experiments. One possible interpretation for these data is that some peptides are more stable than others and are regularly present in these cells, whereas other peptides are only transiently present within the cell. Previous studies on the mouse brain peptidome found that the relative levels of some peptide fragments of intracellular proteins varied from mouse to mouse more than the levels of neuropeptides, although other protein fragments were common to all brain extracts and showed low variability (47). Together, these results demonstrate that some peptide fragments of intracellular proteins are normally present in cells and can be consistently detected. This runs contrary to the current view that after proteasome-mediated cleavage, cellular peptidases rapidly convert all the peptides to amino acids (9, 25, 28, 35).

The quantitative peptidomics approach using stable isotopic labels and mass spectrometry provides quantitative information on the relative efficiency of the metabolism of these intra-
Peptidomic Analysis for EP24.15

Table 3: Comparative analysis between in vitro and in vivo substrates and products of EP24.15

| Protein Name                          | Sequence                        | HEK peptidome in vitro | Exp. in cells | Synthetic peptide in vitro |
|---------------------------------------|---------------------------------|------------------------|---------------|----------------------------|
| Activated RNA polymerase II transcript factor 4 | KEQISDIDDAVRKL                  | Not substrate          | Not substrate | Not substrate              |
| Cathepsin C                           | DFPNPFELTNH                     | Substrate              | Not substrate | Not substrate              |
| Cytochrome c oxidase subunit V1b      | AC-AEDMETKIKNY                  | Substrate              | Not substrate | Not substrate              |
| Eukaryotic translation elongation factor 1β2 | GFGDLKPSAGLQVL                | Substrate              | Not substrate | Not substrate              |
| Eukaryotic translation initiation factor 4A | VGNLIPNTV                     | Substrate              | Not substrate | Not substrate              |
| Eukaryotic translation initiation factor 5A | AVAIKAMAK                     | Substrate              | Not substrate | Not substrate              |
| Eukaryotic translation initiation factor 5A | SAMTEEAAVAIK                   | Product                | Product       | Product                    |
| Eukaryotic translation initiation factor 5A | MTEEAAVAIKAMAK                 | Substrate              | Not substrate | Not substrate              |
| Eukaryotic translation initiation factor 5A | SAMTEEAAVAIKAMAK                | Substrate              | Not substrate | Not substrate              |
| FK506-binding protein                 | VDFVELLKLE                      | Substrate              | Not substrate | Not substrate              |
| Peptidylprolyl isomerase A            | KHTPGILSM                       | Substrate              | Not substrate | Not substrate              |
| Peptidylprolyl isomerase A            | DIAVDPGDLPRV5SF                | Substrate              | Not substrate | Not substrate              |
| Peptidylprolyl isomerase A            | ELFADKVPKTAENFR                | Substrate              | Not substrate | Not substrate              |
| Ribosomal protein S21                 | KADGIVSKNF                      | Substrate              | Not substrate | Not substrate              |
| Ribosomal protein S21                 | AKADGIVSKNF                     | Substrate              | Not substrate | Not substrate              |
| S100 calcium-binding protein A11      | HDSFKLAVPSQKRT                  | Substrate              | Not substrate | Not substrate              |
| Splicing factor, arginine/serine-rich 1| AIRDIDLKNR                     | Substrate              | Not substrate | Not substrate              |
| Splicing factor, arginine/serine-rich 1| GAIRIDLKNR                     | Substrate              | Not substrate | Not substrate              |
| Splicing factor, arginine/serine-rich 1| GRRIPPAFVEF                    | Substrate              | Not substrate | Not substrate              |

a Selected results from the experiment testing HEK293 cell peptidome extracts with 2600 ng of purified EP24.15 are shown (Experiment 1, runs 3 and 4).

b Selected results from the experiment testing overexpression of EP24.15 in HEK293 cells are shown (Experiment 2, runs 1–4).

c Results from the experiment testing synthetic peptides with purified EP24.15.

Cellular peptides by EP24.15. Previous studies have used mass spectrometry to evaluate the cleavage of peptides by proteases but not in a quantitative fashion (48, 49). Although EP24.15 is thought to be a key enzyme for the degradation of antigens produced by the proteasome (24, 25, 28, 50), it has not previously been investigated using a peptidomics approach.

The role of EP24.15 in producing antigenic peptides is controversial. EP24.15 was first suggested to protect antigenic peptides from further degradation and to increase hasp65 from Mycobacterium tuberculosis antigen presentation (23, 26). Later, it was proposed that EP24.15 degraded antigenic peptides and decreased antigen presentation (24, 25). Increasing intracellular EP24.15 enzymatic activity by 0.6 times was shown to stimulate MHC-I antigen presentation (23), although its overexpression by 15–16 times caused a reduction in MHC-I antigenic peptides (24, 25). Importantly, it has been observed that only the highest EP24.15-expressing cells decrease surface MHC-I expression independent of the antigen investigated (24). Because peptides of the size used for MHC-I antigen presentation (i.e. 8–10 residues) are both substrates and products of EP24.15 (supplemental Table 1, Table 1, supplemental Table 2, and Fig. 4), our results are consistent with a dual role for EP24.15 in the production of some MHC-I peptides and the degradation of others. However, the 6-fold increase of EP24.15 activity obtained in this study had no effect on approximately half of the observed peptides, and only 20% of the peptides were greatly affected by this large increase in enzyme levels. EP24.15 is known to be regulated by interferon-γ and other mechanisms, and based on the results of this study, relatively small changes in enzyme activity can alter levels of some cellular peptides (51, 52). Recently, a small 2–6 times increase in intracellular EP24.15 activity in either Chinese hamster ovary-S or HEK293 has been shown to significantly affect G protein-coupled receptor signal transduction, suggesting a physiological role for this enzyme in cell signaling (16).

The majority of HEK293 cell peptides identified in this study are likely to be generated by proteasomes as they are fragments of intracellular proteins that are 5–17 amino acids in length and have either a hydrophobic (~75%), basic (~12%), or acidic (~8%) amino acid at the C terminus. Only a very small percentage of the peptides have C-terminal Gly, Pro, Ser, Thr, Asn, Gln, or Cys residues, consistent with previous findings that proteosome-mediated cleavages rarely occur at these sites (46, 53). Acting after the proteasome, intracellular peptidases, such as prolyl oligopeptidase (54), insulin-degrading enzyme (55), tripeptidyl peptidase II (56), among others (57), could also be related to the generation of some of the oligopeptides identified in this study. In previous studies using the catalytically inactive EP24.15 in a "substrate capture" assay, the isolated peptides were in the same size range as the ones found here, and also had predominantly aliphatic/aromatic residues at the C terminus (15, 22, 32). Substrate capture assays previously performed used rat brain and mouse adipose tissue to isolate peptides, and in this study we used HEK293 cells. Therefore, this is likely to account for at least some of the differences in the peptide sequences identified herein and in previous studies. Moreover, the inactive EP24.15 capture assay selects specific peptides from a complex mixture based on their increased affinity for the enzyme, although the current analyses peptides are not previously selected, and their sequence identification relies on their abundance and ionization characteristics.

Most of the previous studies on the substrate specificity of EP24.15 examined either non-native peptides containing fluorescent and/or quenching groups on the N and C termini or used a relatively small number of neuropeptides (18, 22, 30, 41, 58–63). Thus, this study that used a large number of endogenous peptides provides much more information regarding the cleavage of peptides that EP24.15 will likely encounter within a cell. However, some peptides found to be cleaved in vitro do not appear to be substrates in cells overexpressing EP24.15 (Table 3), which could be due at least in part to distinctive compartmentalization. Altogether, these results suggest that in vivo, EP24.15 has a limited set of substrates despite the existence of a large number of intracellular oligopeptides.
Although the peptidomics analysis of EP24.15 substrates is an important contribution toward defining the function of this enzyme, knowledge of the substrate alone does not provide information regarding the cleavage specificity. However, a comparison of the substrates and products identified in this study allows for predictions regarding the cleavage site. For this analysis, peptide identified as products were compared with those identified as substrates. Of the 10 peptides found to be the strongest products of EP24.15 (ratio >1.5 when incubated with the high concentration of enzyme; see Table 1), all were also found in longer forms that would require a single cleavage to generate the product, and these longer forms were identified as substrates. In one case (the peptide GFGDLKSPAGLQV), both the N-terminal and C-terminal products were identified, strongly supporting the predicted cleavage site. The observed cleavage sites (Table 2), as well as the predicted ones that were not tested, fit the general consensus for cleavage of peptides by EP24.15 (18, 22, 30, 41, 58–63). These previous studies have established that EP24.15 has a broad substrate specificity for peptides 6–17 residues long, with a slight preference for hydrophobic residues (other than Ile) in the P1 and/or P3’ positions, with cleavage favored between 3 and 6 residues from the C terminus. A Pro residue was also reported to be preferred in the P2’ position. In this study, most of the cleavage sites appear to be 3–4 residues from the C terminus. Two exceptions are cleaved 7 residues from the C terminus, and another exception is cleaved 2 residues from the C terminus (Tables 1 and 2). Interestingly, many of the products result from cleavage after a Lys residue (Table 1). Therefore, it is tempting to speculate that EP24.15 has a preference for a Lys in the P1 position, along with the other preferences previously noted (i.e. Pro in the P2’ position, a hydrophobic group in the P1 and/or P3’ position, and cleavage close to the C terminus). If Lys is included as a preference in the P1 position and cleavages 2–7 residues from the C terminus are allowed, nearly all of the products observed in this study match at least 2 of the 4 preferences.

Although this study was primarily focused on examination of the substrates and products of EP24.15, the finding that a large number of intracellular (cytosolic and nuclear) protein fragments were detectable in HEK293 cell extracts raises the possibility that some of these peptides have biological functions. It is interesting that of the 116 different peptides detected in this study, 64 of these are from precursors that have been found to produce detectable peptides in mouse brain proteomics studies (64–67 and data not shown). Furthermore, 19 of the peptides found in this study in HEK293 cells represent the identical fragment that was previously found in mouse brain. These peptides therefore represent major cellular peptides that are common to HEK293 cells and mouse brain (64–67).

Of the 32 precursor proteins identified in this study, nearly all were found in a previous proteomic analysis of HEK293 cells (68). However, only 5 of the 32 precursor proteins found in this study (triose-phosphate isomerase, profilin I, peptidyl-prolyl cis-trans isomerase A, vimentin, and 10-kDa heat shock protein) were among the most abundant 100 proteins in the previous proteomic study (68). Additionally, 16 of the 32 proteins identified herein were identified in a recent study that examined protein turnover in the human A549 adenocarcinoma cell line; the half-lives of these 16 proteins varied from ~131 h (proteasome subunit, β-type, 6) to 6.8 h (vimentin) (69). Therefore, the peptides identified in this study are not simply associated with the most abundant and/or least stable proteins. It is possible that interactions between the observed peptides and specific proteins prevent the degradation of these peptides (9, 16). Approximately half of the proteins from which these peptides are derived function in the binding of RNA or proteins. Often the RNA or protein binding domain is fairly small, and peptides that correspond to these domains could potentially disrupt the normal function of the protein. In fact, peptide aptamers rationally designed to mimic protein interaction motifs can disrupt specific protein interactions and that has been extensively shown to represent a useful method for rational prediction and manipulation of protein function in vivo and in silico (70–72). Other potential functions for intracellular peptides are possible, and peptides similar to those generated by proteasomes have been recently shown to interfere with G protein-coupled receptor cell signaling when added back into the cells (16). Thus, one possibility is that in vivo intracellular peptides such as the ones described herein could be important pathophysiological regulators of protein interactions involved in gene regulation, metabolism, cell signaling, and protein targeting among others (15, 16).

Acknowledgments—Mass spectrometry was performed through Rede de Proteoma do Estado de São Paulo in the Laboratório Nacional de Luz Sincrotron, Campinas, SP, Brazil and also in the Laboratory for Macromolecular Analysis of the Albert Einstein College of Medicine.

REFERENCES

1. Hershko, A., Ciechanover, A., and Varshavsky, A. (2000) Nat. Med. 6, 1073–1081
2. Ciechanover, A., and Schwartz, A. L. (2004) Biochim. Biophys. Acta 1695, 3–17
3. Lecker, S. H., and Goldberg, A. L. (2002) J. Physiol. (Lond.) 545, 729
4. Lippincott-Schwartz, J., Bonifacino, J. S., Yuan, L. C., and Klausner, R. D. (1988) Cell 54, 209–220
5. Geier, E., Pfeifer, G., Wilm, M., Lucchiari-Hartz, M., Baumeister, W., Eichmann, K., and Niedermann, G. (1999) Science 283, 978–981
6. Kessler, B. M., Glas, R., and Ploegh, H. L. (2002) Mol. Immunol. 39, 171–179
7. Wenzel, T., Eckerskorn, C., Lottspeich, F., and Baumeister, W. (1994) FEBS Lett. 349, 205–209
8. Falk, K., Rotzschke, O., and Ramn, H. G. (1990) Nature 348, 248–251
9. Reits, E., Griekspoor, A., Neijssen, J., Groothuis, T., Jalink, K., van Veelen, P., Janssen, H., Calafat, J., Drijfhout, J. W., and Neefjes, J. (2003) Immunity 18, 97–108
10. Goldberg, A. L. (2003) Nature 426, 895–899
11. Goldberg, A. L., Cascio, P., Saric, T., and Rock, K. L. (2002) Mol. Immunol. 39, 147–164
12. Ramn, H. G. (2002) Nature 419, 443–445
13. Dawson, T., and Scott, J. D. (1997) Science 278, 2075–2080
14. Dawson, T., and Nash, P. (2003) Science 300, 445–452
15. Ferro, E. S., Hyslop, S., and Camargo, A. C. (2004) J. Neurochem. 91, 769–777
16. Cunha, F. M., Berti, D. A., Ferreira, Z. S., Klitzke, C. F., Markus, R. P., and Ferro, E. S. (2008) J. Biol. Chem. 283, 24448–24459
17. Pierotti, A., Dong, K. W., Glucksman, M. J., Orlowski, M., and Roberts, J. L. (1990) Biochemistry 29, 10323–10329
18. Orlowski, M., Michaud, C., and Chu, T. G. (1983) Eur. J. Biochem. 14115

May 22, 2009 • Volume 284 • Number 21
