2.4 RP chromatography

The HFMD-treated serum was fractionated using a protein fractionation system, Proteome Lab PF2D (Beckman Coulter). The separation was performed, at 0.75 mL/min, using a Beckman ODS column. The column was equilibrated with solvent A (0.1% TFA in water). All of the HFMD-treated serum sample (~250 μg) was injected into the column and eluted with solvent B (0.08% TFA in 100% ACN) for 49 min using the following linear gradient conditions: 0–40% B over 42 min, to 100% B at 49 min, and holding at 100% B for 5 min. Fractions were pooled (7 fractions/pool), lyophilized and resuspended in 100 μL 20 mM ammonium bicarbonate buffer (pH 8.0) prior to trypsin digestion. Aliquots of each fraction (5 μL) were collected, and the protein concentrations each fraction and the total protein were measured by BCA protein assay.

2.5 Trypsin digestion

The sample was reduced in 20 mM ammonium bicarbonate buffer (pH 8.0) containing 5 mM DTT at 80°C for 20 min. After cooling to room temperature, the sample was alkylated using 10 mM iodoacetamide at 37°C for 30 min in the same buffer. Trypsin was added to the reduced and alkylated sample at a protein:enzyme ratio of 50:1, followed by incubation at 37°C for 2–3 h. The same amount of trypsin was added again, and the mixture was incubated at 37°C overnight. The digestion was terminated by adding formic acid to a final concentration of 0.1%.

2.6 2-D LC

The digested peptides were fractionated by strong cation exchange (SCX) chromatography followed by RP chromatography. An online nano-LC system (Dina system, KYA technologies, Japan) was used for high-efficiency 2-D-LC experiments, with a 35 mm × 0.32 mm id SCX column, an 1 mm × 0.5 mm id RP trap column, and an 50 mm × 0.15 mm id RP separation column. The tryptic peptides were then diluted by adding 400 μL 0.1% formic acid and injected into the SCX column and washed with solvent A (0.1% formic acid). The peptides were separated into four fractions; one unbound fraction and three eluted fractions with 50 mM, 100 mM and 500 mM ammonium acetate. Peptides in each fraction were separately trapped and desalted into the online-connected RP trap column with solvent A, and eluted from the trap column and fractionated using the RP separation column under solvent B (70% ACN) with the following gradient conditions: 100% A for 10 min, 0–40% B over 140 min, 40–100% B for 20 min, holding 100% B for 10 min, and re-equilibration by solvent A for 10 min at a constant flow rate of 200 nL/min.

2.7 ESI-MS/MS

The peptides eluted from the RP column were analyzed using an online coupled tandem mass spectrometer. An online nanospray emitter (New Objective, MA, USA) was connected to the outlet of the RP separation column for ESI. A quadrupole-TOF mass spectrometer (Q-TOF Ultima, Waters-Micromass, UK) was used for MS/MS experiments with a heated capillary temperature of 150°C and an ESI voltage of 2.0 kV. The operation protocols were in accordance with the manufacturer's specifications.

2.8 Data analysis

Peak lists from acquired product ion spectra were generated using Masslynx software (Waters-Micromass). The peak lists were then searched against the Swiss-Prot human protein database containing 13 433 entries of human proteins, using MASCOT software (Matrix Science, MA, USA) [31]. The peptide tolerance and MS/MS tolerance were set to 1.0 Da and 0.8 Da, respectively. One missed cleavage per peptide was allowed, and oxidation of methionine and carbamidomethyl modification of cysteine were considered. For detection of proteins, we employed the following restrictions. The proteins with resulting MASCOT scores $\geq$50 were selected. Then, the assigned peptides that have minimum peptide scores of $\geq$10 were listed, and the real positive rates of these peptides were estimated as described in previous reports [32–34]. The reliabilities of the proteins were, then, calculated from the real positive rates of the assigned peptides. Here we discuss the proteins that had a reliability of >95% (p <0.05) and >80% (p <0.2).

3 Results and discussion

3.1 Design of the HFMD

The HFMD, designed originally for this research, is a fully automated system that can separate and concentrate LMW serum proteins from high-abundance HMW proteins within 1 h. The device consists of a protein separation unit and a concentration unit (Fig. 1a). The protein separation unit can discriminate LMW proteins in the serum. The HFM, made of polysulfone, is an asymmetric membrane that possesses a progressive change in pore size across its cross-section between the nanoporous inner surface and the coarse-pored outer surface (Fig. 1b). The highly asymmetric structure enables LMW serum proteins to pass through the membrane with a low filtration resistance. HFMs with molecular weight cut-off values of 13 000 and 1400 were designed for protein separation and concentration, respectively (Fig. 1c). The device also employs a multistage filtration and cascaded cross-flow process for separating LMW proteins from the serum (Fig. 1a). The multistage filtration process using the high-performance asymmetric membrane enables a highly sharp separation in comparison with a single-stage process. In the cross-flow process, the fluid flowing in the direction parallel to the membrane serves to clean the filter's inner surface and the prevent deposition or clogging of HMW proteins.

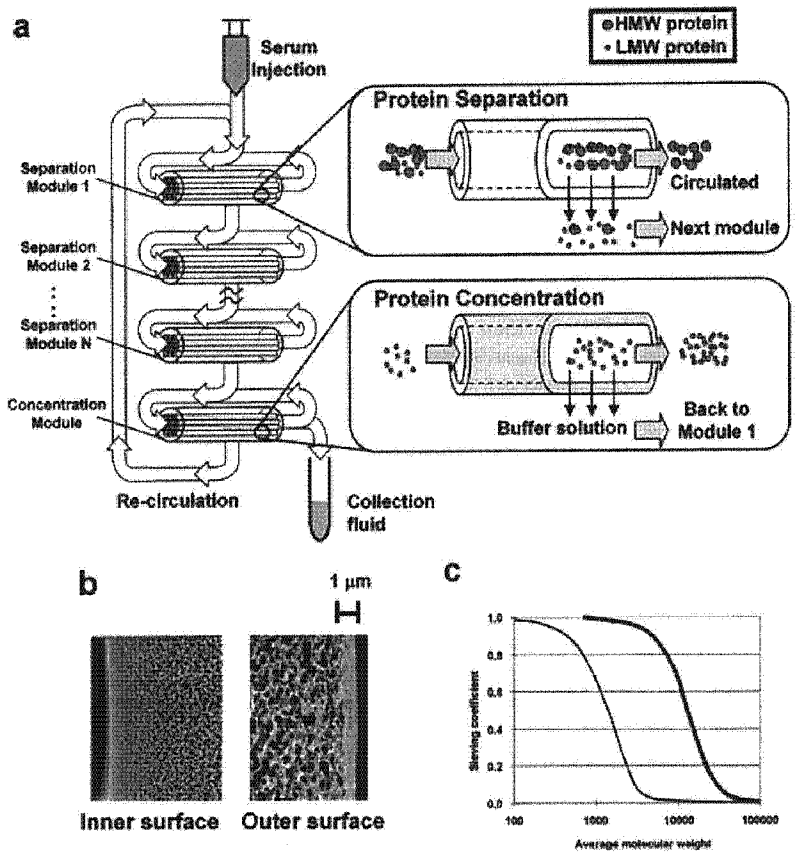

Figure 1. Principle of separation and concentration of LMW proteins using the HFMD. (a) Schematic representation of the HFMD. HFMD employs a combination of a multistage and cascaded cross-flow process for separation and a sequential concentration process. During separation, diluted serum circulates in the first separation module, which is fitted with the HFMs for protein separation. Only LMW serum proteins can pass through the HFM, and flow into the next module. Trace amounts of the HMW proteins might be able to leak from the HFM, but the multistage filtration process depletes them. The LMW proteins are finally processed in the concentration module, where only buffer solution and small molecules pass through the membranes. The solution passing through the concentration module is re-circulated to the first separation module. (b) Scanning electron microscopic image of cross-section of polysulfone HFM. (c) Dextran sieving profile of the membrane of the HFMD. Thick and thin curves indicate the sieving efficiencies of molecular weight cut-off values 13 000 and 1400 membranes, for protein separation and concentration, respectively.

3.2 Depletion of HMW serum proteins

A 1-mL aliquot of the serum sample was diluted and injected into the HFMD. After 1 h of operation, the solution of LMW serum proteins were collected and lyophilized. An aliquot of the resuspended sample was analyzed by SDS-PAGE on a 4–12% gel (Fig. 2a). HMW proteins were efficiently depleted from the serum by the treatment with this device (Fig. 2a, lane 4), while only HMW proteins could be recognized in crude serum (Fig. 2a, lane 2). When the amount of HSA, the most abundant human serum protein, was quantified by ELISA, the HSA concentration in the HFMD-treated sample was decreased to less than 0.001% of that in the non-treated sample. Although the affinity-based method also efficiently depletes HMW proteins including HSA from the human serum, the protein profile after the treatment is completely different from that of the HFMD-treated sample. As shown in Fig. 2a, lane 3, many of the HMW proteins still remained after the treatment, since the affinity-based method selectively depletes only 6 HMW proteins from the serum.

The treatment with the HFMD recovered about 200 μg proteins from 1 mL serum, which originally contained 50 mg total proteins. We believe that this recovery rate is reasonable as 22 abundant proteins comprise more than 99% of serum proteins (http://www.plasmaproteome.org and [23]), indicating that the amount of serum LMW proteins we are targeting here is less than 500 μg.

3.3 Increase in mass ratio of LMW proteins to total proteins

The densitometric measurement revealed that the serum protein fraction treated with the HFMD contained 85.3% LMW proteins (>47 kDa), while that treated with the affinity column contained 23.3% LMW proteins (Fig. 2b). The ELISA of β_2-microglobulin, one of well-known LMW proteins (11.5 kDa), revealed that the amount ratio of β_2-microglobulin against the total protein increases by 150–200-fold after HFMD treatment.

3.4 Key features of the HFMD

The key features of this unique device are as follows: (1) It can deplete almost all serum HMW proteins by a fully automated operation within 1 h. (2) It can handle up to 4 mL of serum sample at one time, while most commercially available affinity removal kits can handle only 10–40 μL. The use of 1–4 mL serum sample enables the detection of extremely

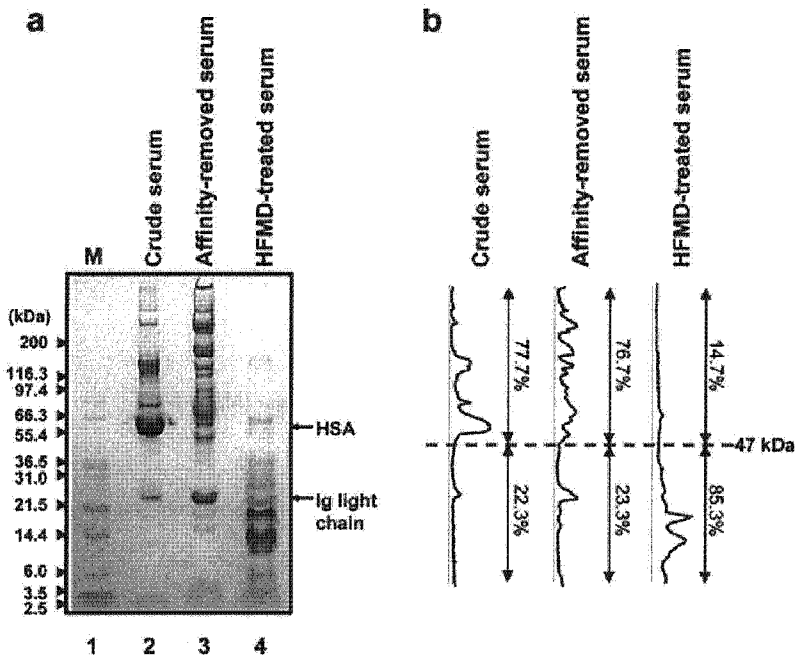

Figure 2. Efficient depletion of HMW serum proteins using HFMD. (a) The serum samples were analyzed by SDS-PAGE. Lane 1, molecular mass markers (M); lane 2, crude serum (20 μg protein); lane 3, serum passed through antibody-immobilized affinity column (Agilent Technologies) (20 μg proteins); lane 4, HFMD-treated serum (50 μg protein). (b) Densitometric analysis of gel shown in (a).

low abundance proteins. (3) The device is equipped with a disposable cartridge unit that can separate serum LMW proteins from HMW proteins, and concentrate the resulting solution simultaneously. (4) The device can be operated under general buffer conditions, thus the concentrated proteins can easily be applied to chromatographic columns (RP column, ion exchange column, and affinity removal column, for example), and/or be analyzed by 2-D (1-D)-PAGE and MS. HFMD, however, might not be suitable for searching low-abundant but HMW biomarker proteins, because they will be removed with other high-abundant HMW proteins, although their fragments could be collected by this device.

The high efficiency of this device for the depletion of serum HMW proteins may arise from large sample-membrane contact area and mild pressure conditions for the protein separation. Since this device employs HFMs and a cross-flow molecular sieving mechanism, a surface area and mild operation pressure are achieved during protein separation in contrast to that employing flat membrane. This is also supported by the fact that a filtration device employing a mild centrifugal pressure enabled better separation of LMW proteins from HMW proteins [22, 23].

3.5 Mass measurement of HFMD-treated sample

The crude and the HFMD-treated serum samples were analyzed by 2-D-LC-MS/MS after enzymatic digestion with trypsin. The MS/MS spectra of the samples were searched against the Swiss-Prot human proteomic database using MASCOT software [31], and further data mining was performed (see Materials and methods). The proteins hit by the data mining are listed in Tables 1 and 2. While only 24 proteins were detected from the crude serum sample, more than four times of this number, 104 proteins were detected from the HFMD-treated serum sample. As shown in Table 1, 20 of the 24 proteins detected from the crude serum sample were high-abundance proteins in human serum, while only 7 of the 104 proteins detected in the HFMD-treated sample were abundant serum proteins (Table 2) ([23], and http://www.plasmaproteome.org). The presence of these 7 abundant proteins in the HFMD-treated sample is reasonable, because these proteins are LMW proteins or cleaved during their maturation processes and thus are expected to pass through the HFMs and be detected by MS/MS analysis. These results indicate that almost 24 times more low-abundance serum proteins were newly detected following the HFMD treatment.

3.6 3-D-LC-MS/MS measurement

Even though the HFMD treatment followed by 2-D-LC/MS/MS enables the detection of more than 100 low-abundance serum proteins, it is generally estimated that up to 10 000 proteins commonly exist in human serum, most of which are present in relatively very low abundance [11]. In addition, we realized that we still miss many peptides when we measured the serum samples by 2-D-LC-MS/MS. The MS profiles for 2-D-LC-MS/MS seemed to be overloaded when performing the MS/MS scanning for all observed peptides (data not shown).

We, therefore, employed an additional fractionation method, RP chromatography of the HFMD-treated sample to reduce the complexity, prior to 2-D-LC-MS/MS (3-D-LC-MS/

Table 1. Proteins detected from the crude serum by 2-D-LC-MS/MS (bold: high-abundance serum proteins; http://www.plasmaproteome.org and [23])

Rank	Accession no.	MASCOT score	Molecular mass	Protein
1	**P02768**	**2534**	**69 321**	**Serum albumin precursor**
2	**P02787**	**1267**	**77 000**	**Serotransferrin precursor (Transferrin)**
3	**P02647**	**589**	**30 759**	**Apolipoprotein A-I precursor (Apo-AI)**
4	**P01857**	**453**	**36 083**	**Ig gamma-1 chain C region**
5	**P00738**	**445**	**45 177**	**Haptoglobin precursor**
6	**P01009**	**388**	**46 707**	**Alpha-1-antitrypsin precursor (Alpha-1 protease inhibitor)**
7	**P01834**	**260**	**11 602**	**Ig kappa chain C region**
8	**P01024**	**229**	**187 046**	**Complement C3 precursor**
9	**P01861**	**222**	**35 918**	**Ig gamma-4 chain C region**
10	**P01842**	**207**	**11 230**	**Ig lambda chain C regions**
11	**P01023**	**200**	**163 175**	**Alpha-2-macroglobulin precursor (Alpha-2-M)**
12	**P01876**	**199**	**37 631**	**Ig alpha-1 chain C region**
13	**P01859**	**194**	**35 862**	**Ig gamma-2 chain C region**
14	**P01877**	**180**	**36 485**	**Ig alpha-2 chain C region**
15	**P01860**	**162**	**32 310**	**Ig gamma-3 chain C region (Heavy chain disease protein)**
16	P02652	146	11 168	Apolipoprotein A-II precursor (Apo-AII)
17	P04264	130	65 847	Keratin, type II cytoskeletal 1 (Cytokeratin 1)
18	P02765	108	39 300	Alpha-2-HS-glycoprotein precursor (Fetuin-A)
19	Q14624	74	103 294	Inter-alpha-trypsin inhibitor heavy chain H4 precursor
20	**P02763**	**74**	**23 497**	**Alpha-1-acid glycoprotein 1 precursor (AGP 1)**
21	**P08603**	**64**	**139 034**	**Complement factor H precursor (H factor 1)**
22	**P01620**	**62**	**11 768**	**Ig kappa chain V-III region SIE**
23	**P00751**	**58**	**85 479**	**Complement factor B precursor (C3/C5 convertase)**
24	**P00450**	**57**	**122 128**	**Ceruloplasmin precursor (Ferroxidase)**

MS, Fig. 3a). As shown in Fig. 3b, the HFMD-treated serum proteins were injected into a C18 silica-based column, and fractionated every 7 minutes to give 7 fractions. The proteins in each fraction were lyophilized and resuspended in 100 μL ammonium bicarbonate buffer (pH 8.0). Each fraction contained 10–15 μg proteins, except fractions 6 and 7 (5.6 and 6.2 μg proteins, respectively), and the total amount of the proteins recovered from the RP chromatography was calculated as 71 μg.

The proteins in each fraction were digested with trypsin, and subsequently analyzed by 2-D-LC-MS/MS separately (Fig. 3c). After the seven sets of 2-D-LC/MS/MS, which took 3.7 days, the resulting MS/MS spectrum data for all the seven fractions were then merged and searched against the human proteomic database. As a result, 12 083 MS/MS spectra were obtained. On analyzing these spectra, 1822 proteins (>95% reliability) and 2294 proteins (>80% reliability) were detected (see Materials and Methods for data mining protocol). Table 3 shows the number of detected proteins in each MASCOT score. The top 100 proteins with high scores are listed in Table 4. In this analysis, HSA was found at the 70th position after the HFMD treatment, and, as in the case of 2-D-LC-MS/MS, several major serum proteins were detected in 3-D-LC-MS/MS, as described above. 3-D-LC-MS/MS allowed us to detect more than 15 times and 75 times of the serum proteins compared with 2-D-LC/MS/MS with and without the HFMD treatment, respectively. We performed this approach (HFMD treatment and 3-D-LC-MS/MS measurement) on three more independent serum samples. The amounts of the proteins recovered from the RP chromatography were 65, 77, 62 μg, and 1831, 1873, 1847 protein hits (>95% reliability) were obtained, respectively. Comparing the four sets of measurements, 852 proteins were found in all four datasets (Supplementary material), and 1502 proteins appeared more than three times. In case of >80% reliability, these numbers of protein hits were 2315, 2372, 2325, respectively, and 1033 proteins were found in all four datasets and 1804 protein appeared more than three times (data not shown). This protocol developed here could routinely analyze HFMD-treated serum samples every 3.7 days, and the number of detected proteins by this protocol is reproducible. Although the fractionation of proteins/peptides from crude samples by various column chromatographic approaches (RP, ion exchange, hydrophobic, etc.) is widely employed, serum samples are not commonly applied directly to those types of column because of its high sample complexity [35–39]. The removal of most serum HMW proteins, which are mainly abundant proteins, by HFMD markedly reduces sample complexity, and further fractionation of the HFMD-treated sample by RP chromatography is highly effective. Recently, about 1000 proteins in serum have been detected in some studies by multidimensional separation methods. The detection of around 2000 proteins by our procedure is similar to the level of the previous studies, and opens the possibility of discovering more protein biomarkers.

Table 2. Proteins detected from the HFMD-treated serum by 2-D-LC-MS/MS (104 proteins; top 30 are indicated; bold: high-abundance serum proteins; http://www.plasmaproteome.org and [23])

Rank	Accession no.	MASCOT score	Molecular mass	Protein
1	**P02671**	**567**	**94 914**	**Fibrinogen alpha/alpha-E chain precursor**
2	P02775	472	13 885	Platelet basic protein precursor (PBP) (Small inducible cytokine B7)
3	P00734	471	69 992	Prothrombin precursor (Coagulation factor II)
4	**P01028**	**401**	**192 650**	**Complement C4 precursor**
5	P02753	366	23 029	Plasma retinol-binding protein precursor (PRBP) (RBP) (PRO2222)
6	**P00746**	**245**	**26 987**	**Complement factor D precursor (C3 convertase activator)**
7	P61769	191	13 706	Beta-2-microglobulin precursor (HDCMA22P)
8	O05682	175	93 194	Caldesmon (CDM)
9	O15942	169	61 238	Zyxin (Zyxin 2)
10	P01253	142	4 918	Thymosin beta-4 (FX)
11	O96PK2	122	669 735	Microtubule-actin cross-linking factor 1, isoform 4
12	**P01024**	**120**	**187 046**	**Complement C3 precursor**
13	P07998	118	17 633	Ribonuclease pancreatic precursor (RNase 1) (RNase A)
14	P00915	117	28 721	Carbonic anhydrase I (Carbonate dehydratase I) (CA-I)
15	P00488	112	83 084	Coagulation factor XIII A chain precursor
16	Q9H299	110	10 431	SH3 domain-binding glutamic acid-rich-like protein 3
17	**P02675**	**106**	**55 892**	**Fibrinogen beta chain precursor**
18	**P02768**	**103**	**69 321**	**Serum albumin precursor**
19	P04264	96	65 847	Keratin, type II cytoskeletal 1 (Cytokeratin 1) (K1) (CK 1)
20	P60709	94	41 710	Actin, cytoplasmic 1 (Beta-actin)
21	P35579	91	226 392	Myosin heavy chain, nonmuscle type A
22	P49792	91	357 993	Ran-binding protein 2 (RanBP2)
23	Q9UPN3	87	619 975	Microtubule-actin cross-linking factor 1, isoforms 1/2/3
24	Q7Z333	82	302 761	Probable helicase senataxin (SEN1 homolog)
25	O00763	80	279 516	Acetyl-CoA carboxylase 2 (ACC-beta)
26	O14686	80	563 831	Myeloid/lymphoid or mixed-lineage leukemia protein 2
27	P50552	78	39 674	Vasodilator-stimulated phosphoprotein (VASP)
28	**P01834**	**76**	**11 602**	**Ig kappa chain C region**
29	O14185	76	215 239	Dedicator of cytokinesis protein 1 (180 kDa protein downstream of CRK)
30	P35556	75	314 131	Fibrillin 2 precursor

Table 3. The number of detected proteins in each MASCOT score

MASCOT score	Detected protein (95% reliability)	Detected protein (80% reliability)
>100	479	482
>90	638	647
>80	880	904
>70	1142	1217
>60	1463	1682
>50	1822	2294

3.7 Sensitivity of the HFMD treatment and 3-D-LC-MS/MS system

To address the detection limit of our approach, the lists of protein detected in each of the four independent analyses were compared with proteins already known as low-abundance proteins in human serum (http://www.plasmaproteome.org). The lowest abundant protein we could find in our list was α-fetoprotein, which is normally present at 1–10 ng/mL in the serum of healthy people. We successfully detected this protein with scores of 57, 54, 60, and 81 in each experiment. The ELISA tests for α-fetoprotein in each serum sample analyzed showed that the actual concentrations of this protein were in the range 2–3 ng/mL (data not shown). Therefore, our unique approach could detect the low-abundant serum proteins present in serum at only a few nanograms per milliliter. The most abundant serum protein (HSA) has a concentration of 50–60 mg/mL and the dynamic range with serum is known to exceed 10^{10} [4, 40]. Current technologies to detect low-abundant serum proteins that combine independent pre-fractionation methods (principally chromatography, immunoaffinity or affinity dye-based depletion, organic solvent precipitation and ultrafiltration, etc.) with MS/MS measurements have typical dynamic ranges of ~10^5 (~0.5 μg/mL protein) [4, 40]. Combining the HFMD with 3-D-LC-MS/MS methods, we confirmed that the dynamic range to detect low-abundant serum proteins could be extended to 10^7.

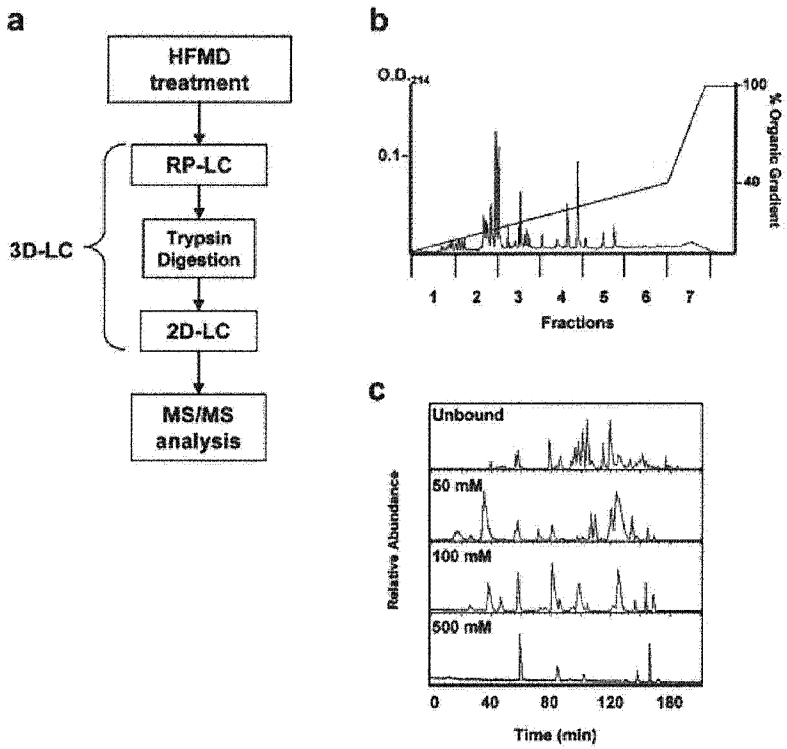

Figure 3. 3-D-LC-MS/MS system for reducing sample complexity. (a) Schematic presentation of protocol for combination of HFMD treatment and 3-D-LC-MS/MS. 3-D LC entails the RP separation of proteins (RP-LC), and the SCX and RP fractionation of tryptic-digested peptides (2-D LC). (b) RP chromatography of HFMD-treated serum sample (approximately 200 μg protein). (c) Base peak chromatograms of 2-D-LC-MS/MS analysis of the fraction 3 shown in (b). The tryptic peptides were separated by SCX chromatography into four fractions; one unbound fraction and three fractions eluted with 50 mM, 100 mM and 500 mM ammonium acetate. Peptides in each fraction were fractionated using the RP chromatography.

Table 4. Human proteins ($n = 1822$) detected in the HFMD-treated serum by 3-D-LC-MS/MS (top 100 proteins; bold: high-abundance serum proteins; http://www.plasmaproteome.org and [23])

Rank	Accession no.	Mascot score	Molecular mass	Protein
1	P02775	468	13 885	Platelet basic protein precursor (PBP) (Small inducible cytokine B7)
2	**P02671**	**386**	**94 914**	**Fibrinogen alpha/alpha-E chain precursor**
3	**P01028**	**373**	**192 650**	**Complement C4 precursor**
4	O05682	371	93 194	Caldesmon (CDM)
5	O96PK2	343	669 735	Microtubule-actin cross-linking factor 1, isoform 4
6	O9UPN3	342	619 975	Microtubule-actin cross-linking factor 1, isoforms 1/2/3
7	O8WXH0	341	795 885	Nesprin 2 (Nuclear envelope spectrin repeat protein 2)
8	O8NF91	320	1 010 412	Nesprin 1 (Nuclear envelope spectrin repeat protein 1)
9	O15149	313	531 412	Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1)
10	O9NU22	309	632 420	Midasin (MIDAS-containing protein)
11	P02753	300	23 029	Plasma retinol-binding protein precursor (PRBP) (RBP) (PRO2222)
12	P20929	291	772 727	Nebulin
13	P00734	286	69 992	Prothrombin precursor (Coagulation factor II)
14	O99996	279	453 387	A-kinase anchor protein 9 (Protein kinase A anchoring protein 9)
15	O9NZR2	256	515 171	Low-density lipoprotein receptor-related protein 1B precursor
16	O7Z7M0	242	254 405	Multiple EGF-like-domain protein 4
17	**P01024**	**241**	**187 046**	**Complement C3 precursor**
18	O9NQ38	241	120 681	Serine protease inhibitor Kazal-type 5 precursor
19	**P01834**	**240**	**11 602**	**Ig kappa chain C region**
20	O15413	240	551 578	Ryanodine receptor 3 (Brain-type ryanodine receptor)
21	P98164	238	521 589	Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
22	O92736	238	564 137	Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor)
23	P46013	234	358 526	Antigen KI-67
24	O14204	233	526 953	Dynein heavy chain, cytosolic (DYHC)

Table 4. Continued

Rank	Accession no.	Mascot score	Molecular mass	Protein
25	P12882	233	222 976	Myosin heavy chain, skeletal muscle, adult 1 (Myosin heavy chain IIx/d)
26	O9NYO7	227	358 033	Cadherin EGF LAG seven-pass G-type receptor 3 precursor
27	O9UKX2	226	222 906	Myosin heavy chain, skeletal muscle, adult 2 (Myosin heavy chain IIa)
28	O60673	224	352 566	DNA polymerase zeta catalytic subunit (EC 2.7.7.7) (hREV3)
29	O9Y623	223	222 874	Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain IIb)
30	P35555	221	312 098	Fibrillin 1 precursor
31	P46100	220	282 391	Transcriptional regulator ATRX (X-linked helicase II)
32	O96SN8	220	214 965	CDK5 regulatory subunit associated protein 2
33	O14789	219	375 848	Golgi autoantigen, golgin subfamily B member 1 (Giantin) (Macrogolgin)
34	P78527	218	468 788	DNA-dependent protein kinase catalytic subunit
35	P04275	216	309 092	Von Willebrand factor precursor (vWF)
36	O9NR09	215	527 276	Baculoviral IAP repeat-containing protein 6
37	O7Z7G8	213	448 449	Cohen syndrome protein 1
38	O13459	212	243 406	Myosin IXb (Unconventional myosin-9b)
39	O9Y6V0	212	566 309	Piccolo protein (Aczonin)
40	O01484	211	430 079	Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid)
41	O9UKX3	211	223 540	Myosin heavy chain, skeletal muscle, extraocular (MyHC-eo)
42	P49792	211	357 993	Ran-binding protein 2 (RanBP2) (Nuclear pore complex protein Nup358)
43	P21359	208	319 168	Neurofibromin (Neurofibromatosis-related protein NF-1)
44	P35579	204	226 392	Myosin heavy chain, nonmuscle type A
45	P11532	203	426 411	Dystrophin
46	O12986	201	123 068	Transcriptional repressor NF-X1
47	O15942	199	61 238	Zyxin (Zyxin 2)
48	P21817	197	564 815	Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor)
49	O13315	196	350 419	Serine-protein kinase ATM (Ataxia telangiectasia mutated)
50	O8NFP9	195	327 604	Neurobeachin protein (Lysosomal trafficking regulator 2)
51	O9Y618	193	273 866	Nuclear receptor corepressor 2 (N-CoR2)
52	O7Z333	193	302 761	Probable helicase senataxin (SEN1 homolog)
53	O07954	191	504 245	Low-density lipoprotein receptor-related protein 1 precursor (LRP)
54	P13535	191	222 625	Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal)
55	P12111	190	343 340	Collagen alpha 3(VI) chain precursor
56	O92608	190	211 812	Dedicator of cytokinesis protein 2
57	O12955	189	480 107	Ankyrin 3 (ANK-3) (Ankyrin G)
58	O9P2N4	187	216 417	ADAMTS-9 precursor
59	O14517	187	505 963	Cadherin-related tumor suppressor homolog precursor
60	O8IUG5	187	285 008	Myosin XVIIIB (Myosin 18B)
61	O02224	186	311 897	Centromeric protein E (CENP-E protein)
62	P49454	185	367 367	CENP-F kinetochore protein (Centromere protein F) (Mitosin)
63	O96T58	185	402 004	Msx2-interacting protein (SMART/HDAC1 associated repressor protein)
64	P25391	184	336 941	Laminin alpha-1 chain precursor (Laminin A chain)
65	P57071	184	169 092	PR-domain zinc finger protein 15 (Zinc finger protein 298)
66	O03164	182	431 622	Zinc finger protein HRX (ALL-1) (Trithorax-like protein)
67	O92896	181	134 505	Golgi apparatus protein 1 precursor (Golgi sialoglycoprotein MG-160)
68	O95613	181	376 130	Pericentrin 2 (Kendrin)
69	O9HC84	180	590 122	Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial)
70	**P02768**	**179**	**69 321**	**Serum albumin precursor**
71	O94833	178	590 528	Bullous pemphigoid antigen 1, isoforms 6/9/10 (Trabeculin-beta)
72	O02388	178	295 041	Collagen alpha 1(VII) chain precursor (Long-chain collagen)
73	O14839	178	217 853	Chromodomain helicase-DNA-binding protein 4 (CHD-4)
74	O8NEZ4	178	540 965	Myeloid/lymphoid or mixed-lineage leukemia protein 3 homolog
75	O43497	177	262 304	Voltage-dependent T-type calcium channel alpha-1G subunit
76	O8TD26	176	304 961	Chromodomain-helicase-DNA-binding protein 5 (CHD-5)
77	O14185	175	215 239	Dedicator of cytokinesis protein 1
78	O99698	175	428 852	Lysosomal trafficking regulator (Beige homolog)
79	P11055	175	223 898	Myosin heavy chain, fast skeletal muscle, embryonic
80	P11277	175	246 170	Spectrin beta chain, erythrocyte (Beta-I spectrin)

Table 4. Continued

Rank	Accession no.	Mascot score	Molecular mass	Protein
81	P61769	174	137 06	Beta-2-microglobulin precursor (HDCMA22P)
82	Q9UKN7	173	394 926	Myosin XV (Unconventional myosin-15)
83	Q8IWT3	173	281 049	p53-associated parkin-like cytoplasmic protein
84	P11137	172	199 489	Microtubule-associated protein 2 (MAP 2) (MAP-2)
85	O14524	170	227 015	Sodium channel protein type V alpha subunit
86	P15924	170	331 571	Desmoplakin (DP) (250/210 kDa paraneoplastic pemphigus antigen)
87	Q9Y6X9	170	110 655	Zinc finger CW-type coiled-coil domain protein 1
88	O15911	169	404 217	Alpha-fetoprotein enhancer binding protein (AT motif-binding factor)
89	O13439	168	260 980	Golgi autoantigen, golgin subfamily A member 4 (Trans-Golgi p230)
90	Q96L73	168	296 464	Nuclear receptor binding SET domain containing protein 1
91	Q9NSI6	168	257 062	WD-repeat protein 9
92	Q9NZJ4	167	436 697	Sacsin
93	Q9NYQ8	166	479 093	Protocadherin Fat 2 precursor (hFat2)
94	P12883	166	222 975	Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta)
95	Q96662	166	438 825	Genome polyprotein
96	O14683	166	143 144	Structural maintenance of chromosome 1-like 1 protein
97	P60709	165	41 710	Actin, cytoplasmic 1 (Beta-actin)
98	O7Z407	165	401 440	CUB and sushi multiple domains protein 3 precursor
99	P55268	165	195 954	Laminin beta-2 chain precursor (S-laminin) (Laminin B1s chain)
100	P17252	165	767 14	Protein kinase C, alpha type (EC 2.7.1.37) (PKC-alpha) (PKC-A)

3.8 Molecular mass distribution

The molecular mass distribution of proteins that were detected by proteomic analysis combined with HFMD treatment and 3-D-LC-MS/MS were examined. When the molecular mass of the proteins were defined from the MASCOT search against the Swiss-Prot database, 79% of the detected proteins had molecular masses much greater than 60 kDa, and 21% of the proteins had molecular masses lower than 60 kDa. Since we analyzed tryptic-digested peptides to detect intact-protein information, this result suggests that many proteolytic fragments are present in the HFMD-treated serum sample (*i.e.*, LMW serum fraction). This result matches well with a previous report, which showed the serum LMW proteins are enriched by peptide fragments of HMW proteins [23]. Proteolytic fragmentation of HMW proteins may occur during the HMFD treatment. To confirm whether proteolytic digestion during the separation is indeed the cause of HMW protein fragmentation, we performed HFMD treatment at 4°C in the presence of a protease inhibitor. The resulting mass distribution pattern did not change (data not shown), revealing that fragments of HMW proteins are commonly present in human serum.

4 Concluding remarks

Here, we described a novel protocol that enables the detection of around 2000 low-abundance serum proteins. This protocol consists of an originally developed serum pretreatment device, HFMD, and a 3-D-LC-MS/MS detection system. The combination of the reduction in sample complexity using HFMD and multidimensional fractionations could be used for analyzing low-abundance proteins including physiological biomarkers. Furthermore, many applicable combinations of HFMD with other already developed serum pretreatment devices (affinity columns, for example) can be considered for proteomic research. HFMD would be diversely applicable as a rapid and robust sample pretreatment tool that can solve the bottleneck in proteomic research.

5 References

[1] Grossklaus, D. J., Smith, J. A., Jr, Shappell, S. B., Coffey, C. S. *et al.*, *Urol. Oncol.* 2002, *7*, 195–198.

[2] Whitehouse, C., Solomon, E., *Gynecol. Oncol.* 2003, *88*, S152–157.

[3] Zhang, Z., Bast, R. C. Jr, Yu, Y., Li, J. *et al.*, *Cancer Res.* 2004, *64*, 5882–5890.

[4] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[5] Lathrop, J. T., Anderson, N. L., Anderson, N. G., Hammond, D. J., *Curr. Opin. Mol. Ther.* 2003, *5*, 250–257.

[6] Rosenblatt, K. P., Bryant-Greenwood, P., Killian, J. K., Mehta, A. *et al.*, *Annu. Rev. Med.* 2004, *55*, 97–112.

[7] Jacobs, I. J., Menon, U., *Mol. Cell. Proteomics* 2004, *3*, 355–366.

[8] Ardekani, A. M., Liotta, L. A., Petricoin, E. F. 3rd., *Expert Rev. Mol. Diagn.* 2002, *2*, 312–320.

[9] Sasaki, K., Sato, K., Akiyama, Y., Yanagihara, K. *et al.*, *Cancer Res.* 2002, *62*, 4894–4898.

[10] Kennedy, S., *Biomarkers* 2002, *7*, 269–290.

[11] Wrotnowski, C., *Genet. Eng. News* 1998, *18*, 14.

[12] Travis, J., Pannell, R., *Clin. Chim. Acta* 1973, *49*, 49–52.

[13] Gianazza, E., Arnaud, P., *Biochem. J.* 1982, *201*, 129–36.

[14] Lollo, B. A., Harvey, S., Liao, J., Stevens, A. C. *et al.*, *Electrophoresis* 1999, *20*, 854–859.

[15] Sato, A. K., Sexton, D. J., Morganelli, L. A., Cohen, E. H. *et al.*, *Biotechnol. Prog.* 2002, *18*, 182–192.

[16] Adkins, J. N., Varnum, S. M., Auberry, K. J., Moore, R. J. *et al.*, *Mol. Cell. Proteomics* 2002, *1*, 947–955.

[17] Pieper, R., Gatlin, C. L., Makusky, A. J., Russo, P. S. *et al.*, *Proteomics* 2003, *3*, 1345–1364.

[18] Steel, L. F., Trotter, M. G., Nakajima, P. B., Mattu, T. S. *et al.*, *Mol. Cell. Proteomics* 2003, *2*, 262–270.

[19] Greenough, C., Jenkins, R. E., Kitteringham, N. R., Pirmohamed, M. *et al.*, *Proteomics* 2004, *4*, 3107–3111.

[20] Bjorhall, K., Miliotis, T., Davidsson, P., *Proteomics* 2005, *5*, 307–317.

[21] Baussant, T., Bougueleret, L., Johnson, A., Rogers, J. *et al.*, *Proteomics* 2005, *5*, 973–977.

[22] Georgiou, H. M., Rice, G. E., Baker, M. S., *Proteomics* 2001, *1*, 1503–1506.

[23] Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J. *et al.*, *Mol. Cell. Proteomics* 2003, *2*, 1096–1103.

[24] Harper, R. G., Workman, S. R., Schuetzner, S., Timperman, A. T. *et al.*, *Electrophoresis* 2004, *25*, 1299–1306.

[25] Chernokalskaya, E., Gutierrez, S., Pitt, A. M., Leonard, J. T., *Electrophoresis* 2004, *25*, 2461–2418.

[26] Hood, B. L., Lucas, D. A., Kim, G., Chan, K. C. *et al.*, *J. Am. Soc. Mass. Spectrom.* 2005, *16*, 1221–1230.

[27] Chertov, O., Simpson, J. T., Biragyn, A., Conrads, T. P. *et al.*, *J. Expert Rev. Proteomics* 2005, *2*, 139–145.

[28] Rubin, R. B., Merchant, M., *Am. Clin. Lab.* 2000, *19*, 28–29.

[29] Basso, D., Valerio, A., Seraglia, R., Mazza, S. *et al.*, *Pancreas* 2002, *24*, 8–14.

[30] Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J. *et al.*, *Lancet* 2002, *359*, 572–577.

[31] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, *20*, 3551–3567.

[32] Kollmann, K., Mutenda, K. E., Balleininger, M., Eckermann, E. *et al.*, *Proteomics* 2005, *5*, 3966–3978.

[33] Weatherly, D. B., Astwood, J. A. 3rd, Minning, T. A., Cavola, C. *et al.*, *Mol. Cell. Proteomics* 2005, *4*, 762–772.

[34] Rudnick, P. A., Wang, Y., Evans, E., Lee, C. S. *et al.*, *J. Proteome Res.* 2005, *4*, 1353–1360.

[35] Raida, M., Schulz-Knappe, P., Heine, G., Forssmann, W. G., *J. Am. Soc. Mass. Spectrom.* 1999, *10*, 45–54.

[36] Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. *et al.*, *J. Proteome. Res.* 2003, *2*, 43–50.

[37] Washburn, M. P., Wolters, D., Yates, J. R. 3rd, *Nat. Biotechnol.* 2001, *19*, 242–247.

[38] Shen, Y., Zhang, R., Moore, R. J., Kim, J. *et al.*, *Anal. Chem.* 2005, *77*, 3090–3100.

[39] Jin, W. H., Dai, J., Li, S. J., Xia, Q. C. *et al.*, *J. Proteome Res.* 2005, *4*, 613–619.

[40] Jacobs, J. M., Adkins, J. N., Qian, W. J., Liu, T. *et al.*, *J. Proteome Res.* 2005, *4*, 1073–1085.

the FEBS Journal

Identification of yeast aspartyl aminopeptidase gene by purifying and characterizing its product from yeast cells

Ryo Yokoyama*, Hiroshi Kawasaki and Hisashi Hirano

Supramolecular Biology, International Graduate School of Arts and Sciences, Yokohama City University, Japan

Keywords
aspartyl aminopeptidase; evolution; M18 family of metalloproteases; matrix-assisted laser desorption ionization time-of-flight mass spectrometry

Correspondence
Dr. Hiroshi Kawasaki, Yokohama City University, Maioko-cho 641-12, Totsuka-ku, Yokohama, Kanagawa, 244-0813, Japan
E-mail: kawasaki@yokohama-cu.ac.jp

***Present address**
Laboratory for Immunogenomics, Research Center for Allergy and Immunology, Institute of Physical and Chemical Research, 1-7-22 Suehiro, Tsurumi, Yokohama 230-0045, Japan

(Received 8 September 2005, revised 6 November 2005, accepted 9 November 2005)

doi:10.1111/j.1742-4658.2005.05057.x

Aspartyl aminopeptidase (EC 3.4.11.21) cleaves only unblocked N-terminal acidic amino-acid residues. To date, it has been found only in mammals. We report here that aspartyl aminopeptidase activity is present in yeast. Yeast aminopeptidase is encoded by an uncharacterized gene in chromosome VIII (YHR113W, Saccharomyces Genome Database). Yeast aspartyl aminopeptidase preferentially cleaved the unblocked N-terminal acidic amino-acid residue of peptides; the optimum pH for this activity was within the neutral range. The metalloproteases inhibitors EDTA and 1.10-phenanthroline both inhibited the activity of the enzyme, whereas bestatin, an inhibitor of most aminopeptidases, did not affect enzyme activity. Gel filtration chromatography revealed that the molecular mass of the native form of yeast aspartyl aminopeptidase is $\approx$680 000. SDS/PAGE of purified yeast aspartyl aminopeptidase produced a single 56-kDa band, indicating that this enzyme comprises 12 identical subunits.

Although sequencing of the yeast genome was completed several years ago, protein databases still contain numerous hypothetical proteins that have yet to be identified in yeast cells. While analyzing the higher-molecular-mass fraction of yeast proteins, we found a homo-multimeric complex, the subunit of which was encoded by the uncharacterized gene, *yhr113w*. Sequence analysis of the Yhr113w protein suggested that it was an aminopeptidase.

Aminopeptidases remove amino acids sequentially from the unblocked N-termini of peptides and proteins [1]. Various aminopeptidases with different substrate specificities are distributed widely in prokaryotes and eukaryotes [2]. These aminopeptidases are classified into 19 groups, based on substrate specificity. Most aminopeptidases are metalloproteases [2], although a few have been reported to act at serine sites.

The amino-acid sequence of Yhr113wp is similar to that of human aspartyl aminopeptidase (EC 3.4.11.21) and yeast vacuole aminopeptidase I (EC 3.4.11.22). These enzymes belong to the M18 family of metalloproteases, each member of which comprises 8–12 identical subunits that contain zinc ions. However, there is no general metalloprotease motif for the sequences of M18 family proteins, such as the HEXXH + E motif for zinc binding in other aminopeptidases. Little is known about the active sites of M18 family proteins.

We purified the protein complex of Yhr113wp from yeast cells and identified it as yeast aspartyl aminopeptidase on the basis of its enzymatic activity. Aspartyl aminopeptidase has been reported to be present in mammals, but the enzyme has not previously been purified from other organisms such as plants, fungi, and bacteria, although it has been suggested, on the basis of sequence similarity, that aspartyl aminopeptidases are widely distributed. Here, we confirm that the enzyme with aspartyl aminopeptidase activity is widely distributed in eukaryotes.

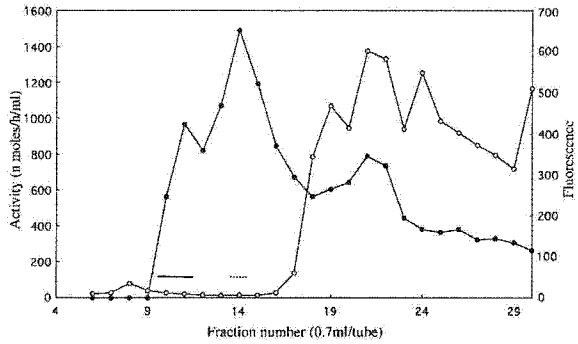

Fig. 1. Purification of yeast aspartyl aminopeptidase by Mono Q anion-exchange chromatography. Three peaks were associated with aspartyl aminopeptidase activity. The third peak exhibited leucine aminopeptidase activity. Fractions I (solid line; Fr. 10–11) and II (dotted line; Fr. 14) were pooled before being applied to a Superose 6 column. Leucine aminopeptidase activity (O) was measured using a fluorogenic substrate. Aspartyl aminopeptidase activity (●) was measured using MALDI-TOF MS.

Results

Purification of yeast aspartyl aminopeptidase

A novel aspartyl aminopeptidase was purified from yeast cells by ultracentrifugation, ammonium sulfate fractionation, and chromatography. Mono Q chromatography was used to separate aspartyl aminopeptidase activity into three major peaks (Fig. 1). The first and second peaks were pooled as fraction I and II, respectively, and proteins within each fraction were separated by Superose 6 gel chromatography (Fig. 2A). Because the third peak exhibited leucyl aminopeptidase activity, proteins in this peak were not purified further. The aspartyl aminopeptidase activity was eluted from the Superose 6 column as a high-molecular-mass complex for both fractions I and II. SDS/PAGE revealed that the active fractions from fraction I contained a protein with a molecular mass of 56 kDa (Fig. 2B). This 56-kDa protein was identified as Yhr113wp by peptide mass fingerprinting using MALDI-TOF MS. Fractions from fraction II were pooled and separated further on a Superose 6 column. The active fraction from fraction II

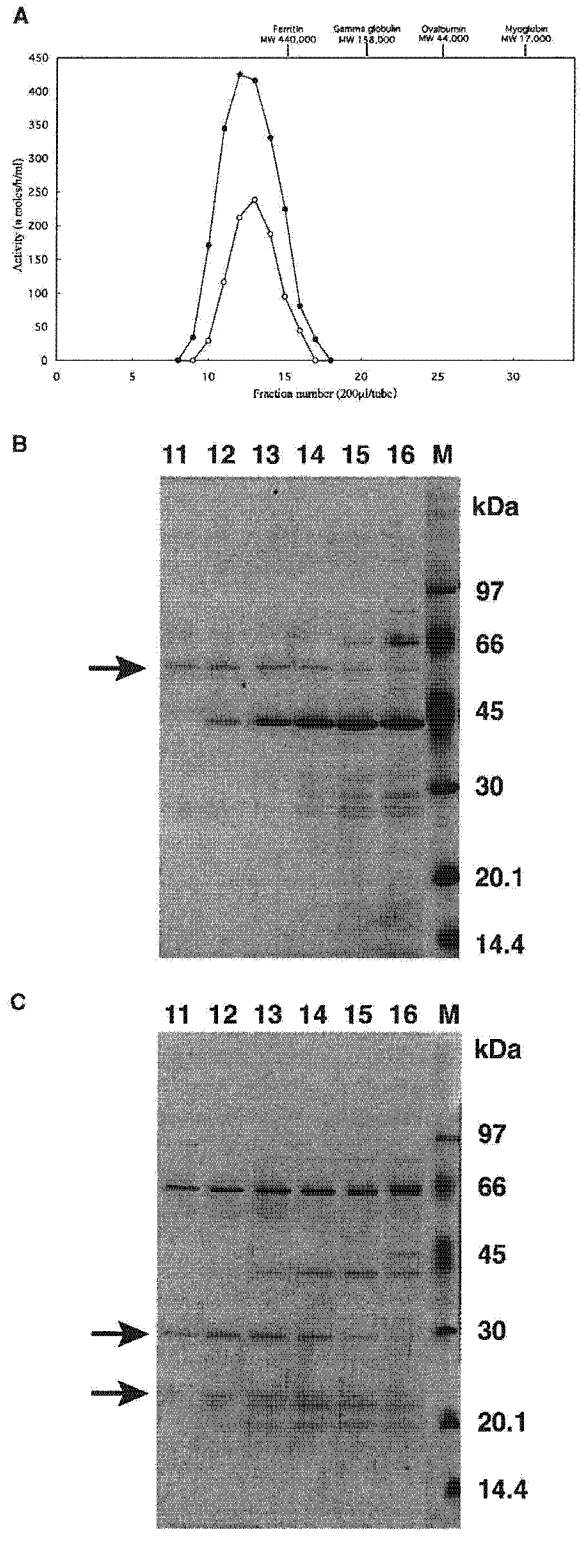

Fig. 2. Separation of yeast aspartyl aminopeptidase by chromatography. (A) Elution of the activity in Mono Q fractions I and II using a Superose 6 column. (O, ●) Activity of fractions I and II, respectively. (B, C) Results of SDS/PAGE analysis of fractions I and II. Lane numbers correspond to the chromatography fractions. Lane M contained a molecular mass marker. Arrows indicate the position of yeast aspartyl aminopeptidase.

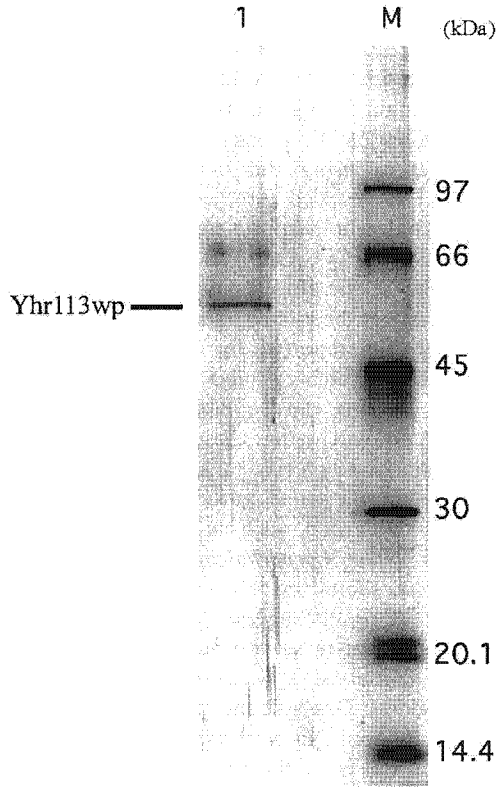

Fig. 3. Yeast aspartyl aminopeptidase (uncleaved form). The enzyme was purified from Mono Q fraction I. Lane 1 shows the purified enzyme stained with silver. Lane M contained a molecular mass marker.

was eluted at the same position as that from fraction I (Fig. 2A). The active fractions from fraction II, however, contained two polypeptides with molecular masses of 31 and 24 kDa (Fig. 2C). Both of these peptides were identified as Yhr113wp by peptide mass finger printing. These results indicate that the polypeptides in the Yhr113wp complex were cleaved by proteases without loss of aspartyl aminopeptidase activity.

We purified ≈400 ng yeast aspartyl aminopeptidase (uncleaved form) from fraction I (Fig. 3), which was used for the following experiments.

Yeast aspartyl aminopeptidase activity

We analyzed the initial degradation velocity of various concentrations of angiotensin I by yeast aspartyl aminopeptidase. The degradation reaction constant (K_m) was estimated to be 0.064 mM based on a Lineweaver–Burk plot.

Molecular mass of the native yeast aspartyl aminopeptidase complex

FPLC revealed that the molecular mass of the native yeast aspartyl aminopeptidase was 680 kDa. Because the molecular mass of Yhr113wp obtained using SDS/PAGE was 56 kDa, and the molecular mass of Yhr113wp calculated from the amino-acid sequence was 54.2 kDa, we deduced that the aspartyl aminopeptidase complex comprises 12 subunits.

Aminopeptidase digestion of peptide substrates

The purified yeast aspartyl aminopeptidase did not digest Leu-NH-Mec, Ala-NH-Mec, Met-NH-Mec, Phe-NH-Mec, or Lys-NH-Mec (data not shown). Yeast aspartyl aminopeptidase cleaved the unblocked N-terminal acidic amino-acid residue of several different peptides (Table 1), but failed to cleave N-acetylated Asp, Tyr, Ile, and Sar (sarcosine) residues at the N-terminus. The N-terminal Asn residue of peptide was cleaved after prolonged incubation. Although yeast aspartyl aminopeptidase cleaved the first N-terminal acidic amino-acid residue of peptides, it did not cleave the second N-terminal Gly or Arg residue. These results suggest that it only exhibits substrate specificity for unblocked N-terminal acidic amino-acid residues. This characteristic is shared by mammalian aspartyl aminopeptidases. Comparison of the digestion

Table 1. Activities of purified yeast aspartyl aminopeptidase against various peptide substrates. The values are mean ± SD from four independent experiments.

Substrate	Sequence	Activity (%)
Angiotensin I	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu	100
Angiotensin II	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe	144 ± 5.7
Angiotensinogen 1–14	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Leu-Val-Tyr-Ser	35 ± 6.4
Angiotensin II antipeptide	Glu-Gly-Val-Thr-Val-His-Pro-Val	116 ± 2.9
N-Acetyl-angiotensinogen 1–14	Ac-Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Leu-Val-Tyr-Ser	0
Tyr-bradykinin	Tyr-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg	0
Ile-Ser-bradykinin	Ile-Ser-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg	0
[Sar1,Ala8]Angiotensin II	Sar-Arg-Val-Tyr-Ile-His-Pro-Ala	0

Table 2. Effect of inhibitors on aspartyl aminopeptidase activity.

Inhibitor	Activity (%)
None	100
Bestatin (0.2 mM)	99.3 ± 16
EDTA (20 mM)	23.9 ± 4.2
1,10-Phenanthroline (5 mM)	66.5 ± 4.5

Enzyme activity was measured with the substrate angiotensin I. Each treatment was repeated three times.

of angiotensin I and II and angiotensinogen 1–14 by yeast aspartyl aminopeptidase revealed that the enzyme cleaved the shorter peptides at a higher rate.

Effect of inhibitors on aspartyl aminopeptidase activity

Table 2 shows the effects of various inhibitors on the activity of yeast aspartyl aminopeptidase. EDTA and 1,10-phenanthroline, which are metalloprotease inhibitors, caused inhibition. In contrast, bestatin, an inhibitor of most aminopeptidases, did not affect the activity of yeast aspartyl aminopeptidase.

Discussion

We purified a novel aspartyl aminopeptidase from yeast and identified it as Yhr113wp. It cleaved only the unblocked N-terminal acidic amino acid of peptides. The enzyme did not cleave N-terminal neutral and basic amino-acid residues. The substrate specificity and optimum pH (7.5–7.9; data not shown) were similar to those for mammalian aspartyl aminopeptidase extracted from the cytosol of rabbit brain cells [3]. However, there are several differences between the mammalian and yeast enzymes. First, the subunit composition differs. Mammalian aspartyl aminopeptidases comprise eight identical subunits, whereas the yeast enzyme contains 12 identical subunits. Second, the mammalian aspartyl aminopeptidase has been reported not to be affected by EDTA, whereas yeast aspartyl aminopeptidase was inhibited by EDTA in this study. This indicates that the surroundings of the zinc ion in yeast aspartyl aminopeptidase differ from those of the mammalian enzyme. The yeast enzyme was not affected by bestatin, an inhibitor of most aminopeptidases. Wilk *et al.* [3] reported that the mammalian enzyme was also unaffected by bestatin.

Yeast aminopeptidase I and aspartyl aminopeptidase, both of which belong to the M18 family of metalloprotease, comprise 12 identical subunits, and the activity of these enzymes is inhibited by EDTA [4,5]. These two enzymes may be related evolutionarily. The substrate specificity of aminopeptidase I is different from that of yeast aspartyl aminopeptidase, aminopeptidase I preferentially cleaving hydrophobic N-terminal amino-acid residues. The M18 family has so far been found to contain 55 proteins (sequences) from bacteria and eukaryotes [6]. We selected 12 of these and aligned their sequences using the CLUSTAL_X program [7] (Table 3). This revealed three conserved His residues, two conserved Glu residues, and five conserved Asp residues. As His, Cys, Glu, and Asp are often ligands for zinc ions [8], the aforementioned conserved amino-acid residues may be related to active and zinc-binding enzyme sites, as reported by Wilk *et al.* [3,9]. The dendrogram of proteins similar to Yhr113wp suggests that these proteins can be divided into two groups (Fig. 4). Yeast and mammalian aspartyl aminopeptidase have the same specificity, and these aminopeptidases lie within the same group of the dendrogram. Proteins in this group, such as those from *Schizosaccharomyces pombe*, *Caenorhabditis elegans* and *Arabidopsis thaliana*, should exhibit aspartyl aminopeptidase activity. These results suggest that aspartyl aminopeptidases are present in several eukaryotes. Basic amino acids were conserved at several sites of the aspartyl aminopeptidases. One site, indicated by 'B' in Table 3, was conserved only in aspartyl aminopeptidases and not other members of the M18 family. This site may determine the substrate specificity of the enzyme.

As aspartyl aminopeptidases lack the N-terminal signal sequence present in the aminopeptidase I precursor, they are probably located in the cytosol. Therefore, aspartyl aminopeptidase activity may be related to the metabolism of cytosolic peptides, particularly during the final step in their degradation. Tamura *et al.* [10] reported that three aminopeptidases are involved in the final degradation of proteins in *Thermoplasma acidophilum*, one of which preferentially cleaves the N-terminal acidic amino-acid residues of short peptides. In the present study, yeast aspartyl aminopeptidase preferentially cleaved shorter peptides. Therefore, yeast aspartyl aminopeptidase may participate in the final step of protein degradation.

In conclusion, we have investigated the biochemical properties of yeast aspartyl aminopeptidase, which is the product of an uncharacterized gene, *yhr113w*. Our findings suggest that it may be involved in protein catabolism in the cytosol of yeast cells.

Experimental procedures

Materials

The yeast strain used was B-8032 (*MATa ura3-52 CYC1-963 cyc7-67 lys5-10*). Leupeptin, bestatin, and TPCK-tryp-

Table 3. Alignment of amino-acid sequence of aminopeptidases in M18 family. The sequences, *Saccharomyces cerevisiae* (Yhr113wp) (P38821), *S. pombe* (O36014), *Homo sapiens* (Q9ULA0), *Mus musculs* (Q9Z2W0), *C. elegans* (Q19087), *A. thlaliana* (Q9LST0), *Pseudomonas aeruginosa* (Q9HYZ3), *Mycobacterium leprae* (Q50022), *Streptomyces coelicolor* (Q9XA76), *Borrelia burgdorferi* (O51572), *S. cerevisiae* (Lap4p) (P14904), and *Thermotoga neapolitana* (O86957) are aligned using CLUSTAL_X. Amino acids identical in at least 10 of 12 sequences are shaded, and histidines, glutamates, and aspartates conserved among all sequences are indicated with *. The sites of basic amino acids conserved among all of the aspartyl aminopeptidases are indicated with b or B. The site of B, which is near the putative active site of glutamate, may determine the substrate specificity of the enzyme.

```
Yhr113wp-S.cerevisiae   1     ----------------------------------MFRIQLRTMSSKTCKSDYPKEFVSFLNSSHSPYHTVHNIKKHLVSNG-FKELSERDSWAG-HVAQK
Ampep-S.pombe           1     ----------------------------------------MTATAKSCALD----FLDFVNASPTPYHAVQNLAEHYMSHG-FQYLSEKSDWQS-KIEPG
DNPEP-H.sapiens         1     -----------------------------------MQVAMNGKARKEAVQTAAKELLKFVNRSPSPFHAVAECRNRLLQAG-FSELKETEKWN---IKPE
DNPEP-M.musculs         1     -------------------------------------MAMNGRARKEAIQATARELLKFVNRSPSPFHVVAECRSRLLQAG-FRELKETEGWD---IVPE
Ampep-C.elegans         1     -------------------------------------MAAALKPSAPEIRKAAQEFINYLNKAVTPFHATQEVKDRLLQAG-FTELPESGHWD---IQPT
Ampep-A.thlaliana       1     ------------------------------------------MDKSSLVSD----FLSFLNASPTAFHAVDESKRRLLKAG-YEQISERDDWK---LEAG
Ampep-P.aeruginosa      1     -------------------------------------------MRAELNQG----LIDFLKASPTPFHATASLARRLEAAG-YRRLDERDAWH---TETG
Ampep-M.leprease        1     -----------------------------------------MPAS-------AADLCEFINASPSPFHVCATVAGRLLDAG-YAELSEVERWP----DHP
Ampep-S.coelicolor      1     -----------------------------------------MSAPSRFDRGHTDDLMTFLSASPTPYHAVASAAARLEKAG-FRQVAETDAWE----ATS
Ampep-B.burudorferi     1     --------------------------------------------MVXXKTLEPKFFQSLLDNSPTPYHLVNYIEEKLINYFNAQQLKLNEKWK----IKT
Lap4p-S.cerevisiae      1     MEEQREILEQLKKTLQMLTVEPSKNNQIANEEKEKKENENSWCILEHNYEDIAQEFIDFIYKNPTTYHVVSFFAELLDKHN-FKYLSEKSNWQDSIGEDG
Ampep-T.neapolitina     1     -----------------------------MKMERKNVWQHRNREE---IESFSKEYMDFMGRAKTERLAVREIRKFLEKEG-FVPIEDFAGDPM------

                                    b                         * *                                    b
Yhr113wp-S.cerevisiae   65    GKYFVTRNGSSIIAFAVGGKWEPGNPIAITGAHTDSPALRIKPI-SKRVSEKYLQVGVETYGGAIWHSWFDKDLGVAGRVFVKDAKTGKSIA--RLVDLN
Ampep-S.pombe           55    NSYFVTRNKSSIIAFSIGKKWKPGNGFSIIATHTDSPTLRLKPK-SQKSAYGYLQVGVEKYGGGIWHTWFDRDLSLAGRVMVEEED-GRVIQ--YNVHID
DNPEP-H.sapiens         62    SKYFMTRNSSTIIAFAVGGQYVPGNGFSLIGAHTDSPCLRVKRR-SRRSQVGFQQVGVETYGGGIWSTWFDRDLTLAGRVIVKCPTSGRLEQ--QLVHVE
DNPEP-M.musculs         60    NNYFLTRNSSSIIAFAVGGQYVPGNGFSLIGAHTDSPCLRVKRK-SRRSQVGYHQVGVETYGGGIWSTWFDRDLTLAGRVIIKCPTSGRLEQ--RLVHIE
Ampep-C.elegans         60    SKYFVTKNRSAILAFAVGGSYKPGSGFSIVVGHTDSPCLRVKPI-SHQKSDKFLQVGVSTYGGGIWRTWFDRDLSVAGLVIVKN--GEKLQH--KLIDVK
Ampep-A.thlaliana       51    KKYFFTRNYSTIVAFAIGHKYVAGNGFHIIGAHTDSPCLKLKPV-SKITKGGCLEVGVQTYGGGLWYTWFDRDLTVAGRVILKEEKAGSVSYSHRLVRIE
Ampep-P.aeruginosa      50    GRYYVTRNDSSLIAIRLGRRSPLESGFRLVGAHTDSPCLRVKPN-PEIARNGFLQLGVEVYGGALFAPWFDRDLSLAGRVTFRAN--GKLES--RLVDFR
Ampep-M.leprease        48    GRYFIVRAG-SLVAWSAGQGVKAHAPFRIVGAHTDSPNLRVKQH-PDLLVAGWRVVALQPYGGAWLNSWLDRDLGVSGRLSVRSAG-KGSEITDRLVRID
Ampep-S.coelicolor      55    GGKYVLRGG-AIVAWYVPEGAAAHTPFRIVGAHTDSPNLRVKPR-PDTGAHGWRQVAVEIYGGPLMNSWLDRDLGLAGRLSLRDG-------STRLVDVD
Ampep-B.burudorferi     53    GSYYIKKEGTSLIAFNIDVK-KKYEPFLIAAAHTDSPGLKLKID-ATEKVSGVFYNHIEVYGSPIISTWIDRDLSLAGIVYFKKNE----NIESKLINIE
Lap4p-S.cerevisiae      100   GKFYTIRNGTNLSAFILGKNWRAEKGVGVIGSHVDALTVKLKPVSFKDTAEGYGRIAVAPYGGTLNELWLDRDLGIGGRLLYKKKGTNEIKS-ALVDSTP
Ampep-T.neapolitina     62    DMAVYAVNRGKAIAAFRVVDD-LKKGLNLVVAHIDSPRLDFKPN-PLVEDEQIALFKTHYYGGIKKYHWFNIPLEIHGVLFRGDGEEIEIHVG---DKPE

                                         *
Yhr113wp-S.cerevisiae   162   RPLLKIPTLAIHLDRDVN-QKFEFNRETQLLPIGGLQEDKTEAKTEKEI---NNGEFTSIKTIVQRHHAELLGLIAKEL--AIDTIED---IEDFELILY
Ampep-S.pombe           151   RPLLRIPTLAIHLDPSAN-SSFSFNMETEFVPLIGLEN---ELAKEETS---DNGD--------KYHHPVLLSLLANEISKSLETTIDPSKIVDFELILG
DNPEP-H.sapiens         159   RPILRIPHLAIHLQRNIN-ENFGPNTEMHLVPILATAIQEELEKGTPEP---GPLNAVDER-----HHSVLMSLLCAHLG------LSPKDIVEMELCLA
DNPEP-M.musculs         157   RPILRIPHLAIHLQRNIN-ENFGPNTEIHLVPILATAVQEELEKGTPEP---GPLGATDER-----HHSVLMSLLCTHLG------LSPDSIMEMELCLA
Ampep-C.elegans         155   KPVLFIPNLAIHLETDR--TTFKPNTETELRPILETFAAAGIN-APQKP---ESTGFADPRNITNNHHPQFLGLIAKEAG------CQPEDIVDLDLYLY
Ampep-A.thlaliana       150   DPIMRIPTLAIHLDRNVNTEGFKPNTQTHLVPVLATAIKAELNKTPAESGEHDEGKKCAETSSKSKHHPLLMEIIANALG------CKPEEICDFELQAC
Ampep-P.aeruginosa      145   KAIAVIPNLAIHLNRAAN-EGWPINAQNELPPIIAQLAPGEAAD----------------------FRLLLDEQLLREHG------ITADVVLDYELSFY
Ampep-M.leprease        145   DPILRVPQLAIHLAE--DRKSLTLDPQRHVNAVWGVGD----------------------------KAGSLLEYVAERTG------VAVADVLAVDLMTH
Ampep-S.coelicolor      146   RPLLRVPQLAIHMDRNVSTEGLKLDKQRHLQPVWGLGDSV--------------------------RDGDLIAFLEDEAG------LARGEVTGWDLMTH
Ampep-B.burudorferi     147   N-IGIIPNLAIHLNRQIN-EGFKYNAHDNLTVISSTKK-------------------------------AIKDNILEQLG------IECENFLSCDLIFT
Lap4p-S.cerevisiae      199   LFVCRIPSLAPHFGKPAEGPFDKEDQTIPVIGFPTPDEEGNEPPTDDEK----------KSPLFGKHCIHLLRYVAKLAG------VEVSELIQMDLDLF
Ampep-T.neapolitina     157   DPVFTIPDLLPHLD---KEDAKISEKFKGENLMLIAGTIPLSGEEKEA------------------VKMNVLKILNEMYG------ISEEDFVSGEIEVV

                                                b *                            B     * **                   b
Yhr113wp-S.cerevisiae   253   DHNASTLGGFNDEFVFSGRLDNLTSCFTSMHGLTL-AADT-EIDRESGIRLMACFDHEEIGSSSAQGADSNFLPNILERLSILKGDGSDQTKPLFHSAIL
Ampep-S.pombe           236   DAEKARLGGIHEEFVFSPRLDNLGMTFCASQALTK-SLENNSLDNESCVRVVPSFDHEEIGSVSAQGAESTFLPAVLQRICELGKESS-----LFS----
DNPEP-H.sapiens         244   DTQPAVLGGAYDEFIFAPRLDNLHSCFCALQALIDSCAGPGSLATEPHVRMVTLYDNEEVGSESAQGAQSLLTELVLRRISASCQHPTA---------FE
DNPEP-M.musculs         242   DTQPAVLGGAYEEFIFAPRLDNLHSCFCALQALIDSCASPASLARDPHVRMVTLYDNEEVGSESAQGAQSLLTELILRRISASPQRLTA---------FE
Ampep-C.elegans         243   DTNKAAIVGMEDEFISGARLDNQVGTYTAISGLLESLTG-ESFKNDPQIRIAACFDNEEVGSDSAMGASSSFTEFVLRRLSAG-GSTTA---------FE
Ampep-A.thlaliana       244   DTQPSILAGAAKEFIFSGRLDNLCMSFCSLKALIDATSSGSDLEDESGIRMVALFDHEEVGSNSAQGAGSPVMIDAMSHITSCFSSDTK--------VLK
Ampep-P.aeruginosa      216   DTQSAAVVGLNDEFIAGARLDNLLSCHAGLEALLN--------AEGDENCILVCTDHEEVGSCSHCGADGPFLEQVLRRLLPEGDAFSR-----------
Ampep-M.leprease        209   DLVPSMVIGADANLLSAPRLDNQVSCYAGMEALLASVPH-------DCLPVLALFDHEEVGSTSDRGARSNLLSTVLERIVLAAGGGRD--------DYL
Ampep-S.coelicolor      214   SVEPPAYLGRDRELVAGPRMDNLLSVHAGTAALASVAASG---ADLPYIPVLAAFDHEETGSQSDTGADGPLLGGVLERSVFARGGSYE--------DRA
Ampep-B.burudorferi     208   ESQPSKIIGTEGEFLASKNLDNKSGCHAIMNSYVHTSND--------KNKIAVFFDNEEVGSLTSRGADSNFLSEVLERIDIALDLTREEH--------L
Lap4p-S.cerevisiae      283   DVQKGTIGGIGKHFLFAPRLDDRLCSFAAMIALICYAKDVN-TEESDLFSTVTLYDNEEIGSLTRQGAKGGLLESVVERSSSAFTKKPVD--------LH
Ampep-T.neapolitina     230   PAFPPREVGIDRSLIGAYGQDDRICAYTALRAF--LEANP------KKSIGVVFLDKEEIGSDGNTGAKARFYLRVLRQILKIQGGPEIQN-----SHSM

                                         *           b               b     b                b
Yhr113wp-S.cerevisiae   351   ETSAKSFFLSSDVAHAVHPNYANKYESQHKPLLGGGPVIKINANQR-----YMTNSPGLVLVKRLAEAAKVPLQLFVVANDSPCGSTIGPILASKTGIRT
Ampep-S.pombe           326   ISMVKSFLVSADMAHAMHPNYSSRYENSNTPFLNKGTVIKVNANQR-----YTTNSAGIVLLKKVAQLADVPIQSFVVRNDSPCGSTIGPKLAAMTGMRT
DNPEP-H.sapiens         335   EAIPKSFMISADMAHAVHPNYLDKHEENHRPLFHKGPVIKVNSKQR-----YASNAVSEALIREVANKVKVPLQDLMVRNDTPCGTTIGPILASRLGLRV
DNPEP-M.musculs         333   EAIPKSFMISADMAHAVHPNYSDKHEENHRPSFHKGPVIKVNSKQR-----YASNAVSESMIREVAGQVGVPLQDLMVRNDSPCGTTIGPILASRLGLRV
Ampep-C.elegans         332   EAIGKSMLISADQAHATHPNYSAKHEENHRPAFHGGVVVKVNVNQR-----YATTSTTHAALKQVAFEAQVPLQVVVVRNDSPCGSTVGPILATKLGLQT
Ampep-A.thlaliana       335   KAIQKSLLVSADMAHALHPNFMDKHEENHQPKMHGGLVIKHNANQR-----YATNAVTSFVFREIAEKHNLPVQDFVVRNDMGCGSTIGPILASSVGIRT
Ampep-P.aeruginosa      297   -AIQRSLLVSADNAHGVHPNYADRHDANHGPALNGGPVIKINSNQR-----YATNSETAGFFRHLCQDSEVPVQSFVTRSDMGCGSTIGPITASQVGVRT
Ampep-M.leprease        294   RRLPASLLVSADMAHATHPNYPECHEPSHLIEVNAGPVLKVHPNLR-----YATDGRTAAAFEVACQQAGVRLQRYEHRADRPCGSTIGPLASARTGIPT
Ampep-S.coelicolor      313   RAFAGTVCLSSDTGHAVHPNYAERHDPTHHPRINGGPILKVNVNNR-----YATDGSGRAVFAAACEKADIPFQTFVSNNSMPCGTTIGPITAARHGIST
Ampep-B.burudorferi     292   IKTNKSFNISIDSVHGIHPGYTSKHDPNYQANLGKGVVVKNSANFR-----YATTSTGFAKLKNLAIKNNIKIQEIIMKANVPSGTTIGPISNARTGIET
Lap4p-S.cerevisiae      374   TVWANSIILSADVNHLYNPNFPEVYLKNHFPVPNVGITLSLDPNGH-----MATDVVGTALVEELARRNGDKVQYFQIKNNSRSGGTIGPSLASQTGART
Ampep-T.neapolitina     317   RYWKKSAVISGDVCAAVNPPYKDVHDLHNAPRSGYGVALVKYTGARGKYSTNDAHAEFVARVRQVLNERNVIWQVATLGKVDQGGGGTYAKFFAERGADV

                               *        *  b
Yhr113wp-S.cerevisiae   446   LDLGNPVLSMHSIRETGGSADLEFQIKLFKEFFERYTSIESEIVV--
Ampep-S.pombe           421   LDLGNPMLSMHSCREMCGSKDFEYAVVLFSSFFQNFANLEEKIIIDE
DNPEP-H.sapiens         430   LDLGSPQLAMHSIREMACTTGVLQTLTLFKGFFELFPSLSHNLLVD-
DNPEP-M.musculs         428   LDLGSPQLAMHSIRETACTTGVLQTLTLFKGFFELFPSVSRNLLVD-
Ampep-C.elegans         427   VDVGCPQLAMHSIREFADTSSIYQATTLYSTFYERLSTVLSNMQ---
Ampep-A.thlaliana       431   VDVGAPQLSMHSIREMCAADDVKHSYEHFKAFFQEFTHLDAKLTIDV
Ampep-P.aeruginosa      391   VDIGLPTFAMHSIRELAGSHDLAHLVKVLGAFYASSELP--------
Ampep-M.leprease        389   VDVGAAQLAMHSARELMGAHDVAVYSAALQAFFSADLF---------
Ampep-S.coelicolor      398   VDIGVAILSMHSARELCGADDPHLLANALVAFLQP------------
Ampep-B.burudorferi     387   IDIGTPMWAMHSLRETVSIADHIEAIKLLRAFFEKGI----------
Lap4p-S.cerevisiae      469   IDLGIAQLSMHSIRAATGSKDVGLGVKFFNGFFKHWRSVYDEFGEL-
Ampep-T.neapolitina     417   YDMGPALLGMHSPFEISSKADLFETYRAYRALLEDL-----------
```

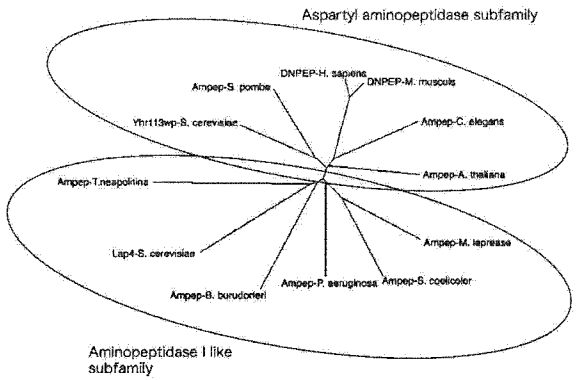

Fig. 4. Dendrogram of the M18 family of metalloproteases. The dendrogram was constructed from the alignments shown in Table 3.

sin were from Sigma (St Louis, MO, USA). We used the following peptides (obtained from Sigma) as substrates: angiotensin I (DRVYIHPFHL); angiotensin II (DRVYIHPF); [Asn1,Val5]angiotensin II (NRVYVHPF); angiotensin II antipeptide (EGVTVHPV); [Sar1,Ala8]angiotensin II (sarcosyl-RVYIHPA); angiotensinogen 1–14 (DRVYIHPFHLL VYS); *N*-acetyl-angiotensinogen 1–14 (Ac-DRVYIHPFH LLVYS); Tyr-bradykinin (YRPPGFSPFR); and Ile-Ser-bradykinin (ISRPPGFSPFR). NH-Mec substrates were obtained from the Peptide Institute (Osaka, Japan). α-Cyano-4-hydroxycinnamic acid was purchased from Aldrich (Milwaukee, WI, USA).

Measurement of aspartyl aminopeptidase activity

Aspartyl aminopeptidase activity was measured using angiotensin I as substrate. The enzyme was mixed with a solution of 50 μM angiotensin I and the reaction solution was incubated at 37 °C for various time periods. The reaction was stopped by the addition of acetic acid, and the reaction solution was then diluted with water to reduce the concentrations of salts. The digested peptide substrate was analyzed using MALDI-TOF MS. The amount of peptide generated was estimated from the ratio of the height of the product peptide peak versus the sum of the heights of the product peptide and substrate peaks.

Purification of yeast aspartyl aminopeptidase

Yeast cells were grown in YPD medium (10 $g \cdot L^{-1}$ yeast extract, 20 $g \cdot L^{-1}$ bactopeptone, and 20 $g \cdot L^{-1}$ glucose) at 29 °C for 72 h with continuous shaking at 160 r.p.m. Cells were then harvested by centrifugation at 2000 *g* for 10 min. Cell extracts were prepared by vortex-mixing a suspension of cells (50%, v/v) with glass beads (0.45 mm) in 50 mM Tris/HCl (pH 7.5)/10% (v/v) glycerol/2 $\mu g \cdot mL^{-1}$ leupeptin (buffer A) on ice. The beads and unbroken cells were removed by centrifugation at 8000 *g* for 20 min. Microsomes and organelles were removed by two successive centrifugations at 40 000 *g* for 40 min and 66 000 *g* for 50 min. Thereafter, high-molecular-mass proteins were precipitated by ultracentrifugation at 200 000 *g* for 5 h. The precipitate was dissolved in buffer A and proteins were fractionated with ammonium sulfate (40–80% saturation). The precipitate was collected by centrifugation (15 000 *g*, 15 min) and dissolved in 50 mM Tris/HCl buffer (pH 7.5) containing 10% glycerol and 0.2 M KCl (buffer B). The solution was passed through a Bio-Gel A-1.5m column (2.5 cm internal diameter × 40 cm; Bio-Rad, Hercules, CA, USA) equilibrated with buffer B. The aspartyl aminopeptidase activity of each fraction was measured. The active fractions were pooled and dialyzed against 25 mM Tris/HCl, pH 7.5 (buffer C). The sample solution was applied to a Mono Q FPLC column (0.5 cm internal diameter × 5 cm; Amersham Bioscience, Uppsala, Sweden) equilibrated with buffer C. After the column had been washed with buffer C (3 × column volume), the enzyme was eluted with a linear gradient (0–0.5 M) of KCl in buffer C. The activities of aspartyl aminopeptidase and leucyl aminopeptidase were measured for each fraction. Fractions that exhibited aspartyl aminopeptidase activity were pooled, and proteins in the combined fractions were precipitated with 80% saturated ammonium sulfate. The precipitate was collected by centrifugation at 15 000 *g* for 30 min and dissolved in 50 mM Tris/HCl (pH 7.5)/0.2 M KCl (buffer D). The solution was passed through a Superose 6 column (1.0 cm internal diameter × 30 cm; Amersham Bioscience) equilibrated with buffer D. Fractions that contained Yhr113wp were stored at −30 °C. Each fraction was analyzed by SDS/PAGE using 15% gels by the methods of Hirano [11]. Yhr113wp was identified by peptide mass finger printing using MALDI-TOF MS.

In-gel digestion and peptide mass fingerprinting

The protein band on each SDS/polyacrylamide gel was cut into small pieces which were incubated in 50% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid, and 0.5% (v/v) *N*-ethylmaleimide at 37 °C for 1 h to remove Coomassie Brilliant Blue. Thereafter, the gel pieces were washed with water before 5 μL 0.3 M *N*-ethylmaleimide (pH 8.2) and 1 μL 140 mM 2-mercaptoethanol were added. After a 30-min incubation at 37 °C, 1 μL TPCK-trypsin (50 $\mu g \cdot mL^{-1}$) was added for 16 h at 37 °C to digest the protein. Peptide mass fingerprinting was carried out using MALDI-TOF MS (Tof Spec 2E; Micromass, Manchester, UK). The matrix solution was 60% (v/v) acetonitrile saturated with α-cyano-4-hydroxy-*trans*-cinnamic acid. Peptides produced by in-gel digestion were concentrated with ZipTipC_{18} (Millipore, Bedford, MA, USA). The peptide

and matrix solution (1 μL each) were mixed on a target plate.

Measurement of aminopeptidase activity

Activities of various aminopeptidases were measured using Leu-NH-Mec, Phe-NH-Mec, Met-NH-Mec, Ala-NH-Mec, and Lys-NH-Mec (0.1 mM) as substrates. Fluorescence of 7-methylcoumarin released as the result of enzymatic activity was measured at an excitation/emission wavelength of 390/460 nm using a Labsystems Fluoroskan II (Dainippon Pharmaceutical, Suita, Japan).

Purified yeast aspartyl aminopeptidase and various peptide substrates (50 μM each) were mixed and incubated at 37 °C. Activity with the various peptides was measured by the method used to measure aspartyl aminopeptidase activity with angiotensin I.

Determination of molecular mass of native yeast aspartyl aminopeptidase

The molecular mass of the native yeast aspartyl aminopeptidase was determined by FPLC using a Superose 6 column. The purified enzyme and a standard marker for gel filtration (Bio-Rad) and ferritin (Sigma) were mixed before being separated on a Superose 6 column equilibrated with buffer D. The molecular mass was calculated by comparing the position of the absorbance peak (280 nm) of the standard marker with that of aspartyl aminopeptidase.

Effect of inhibitors on aspartyl aminopeptidase activity

Solutions of various inhibitors of enzymatic activity were mixed with yeast aspartyl aminopeptidase and incubated at 37 °C for 15 min. Enzyme activity was measured at 37 °C for 10 min using 50 μM angiotensin I as substrate.

References

1 Taylor A (1993) Aminopeptidases: structure and function. *FASEB J* **7**, 290–298.
2 Rawlings ND & Barrett AJ (1995) Evolutionary families of metallopeptidases. *Methods Enzymol* **248**, 183–228.
3 Wilk S, Wilk E & Magnusson RP (1998) Purification, characterization, and cloning of a cytosolic aspartyl aminopeptidase. *J Biol Chem* **273**, 15961–15970.
4 Frey J & Rohm KH (1978) Subcellular localization and levels of aminopeptidases and dipeptidase in *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **527**, 31–41.
5 Metz G & Rohm KH (1976) Yeast aminopeptidase I. Chemical composition and catalytic properties. *Biochim Biophys Acta* **429**, 933–949.
6 Rawlings ND, Tolle DP & Barrett AJ (2004) MEROPS: the peptidase database. *Nucleic Acids Res* **32**, D160–D164.
7 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
8 Vallee BL & Auld DS (1990) Zinc coordination, function, and structure of zinc enzymes and other proteins. *Biochemistry* **29**, 5647–5659.
9 Wilk S, Wilk E & Magnusson RP (2002) Identification of histidine residues important in the catalysis and structure of aspartyl aminopeptidase. *Arch Biochem Biophys* **407**, 176–183.
10 Tamura N, Lottspeich F, Baumeister W & Tamura T (1998) The role of tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell* **95**, 637–648.
11 Hirano H (1989) Microsequence analysis of winged bean seed proteins electroblotted from two-dimensional gel. *J Protein Chem* **8**, 115–130.

Ku70 and Poly(ADP-Ribose) Polymerase-1 Competitively Regulate β-Catenin and T-Cell Factor-4–Mediated Gene Transactivation: Possible Linkage of DNA Damage Recognition and Wnt Signaling

Masashi Idogawa,[1,3,4,5] Mitsuko Masutani,[2] Miki Shitashige,[1] Kazufumi Honda,[1] Takashi Tokino,[3] Yasuhisa Shinomura,[4] Kohzoh Imai,[4] Setsuo Hirohashi,[1] and Tesshi Yamada[1]

[1]Chemotherapy Division and [2]ADP-Ribosylation in Oncology Project, National Cancer Center Research Institute, Tokyo, Japan; [3]Department of Molecular Biology, Cancer Research Institute; [4]First Department of Internal Medicine; and [5]Department of Biomedical Engineering, Biomedical Research Center, Sapporo Medical University, Sapporo, Japan

Abstract

Formation of the T-cell factor-4 (TCF-4) and β-catenin nuclear complex is considered crucial to embryonic development and colorectal carcinogenesis. We previously reported that poly(ADP-ribose) polymerase-1 (PARP-1) interacts with the TCF-4 and β-catenin complex and enhances its transcriptional activity. However, its biological significance remains unexplained. Using immunoprecipitation and mass spectrometry, we found that two Ku proteins, Ku70 and Ku80, were also associated with the complex. Knockdown of Ku70 by RNA interference increased the amount of β-catenin associated with TCF-4 and enhanced the transcriptional activity. PARP-1 competed with Ku70 for binding to TCF-4. Treatment with bleomycin, a DNA-damaging alkylating agent, induced polyADP-ribosylation of PARP-1 protein and inhibited its interaction with TCF-4. Bleomycin conversely increased the amounts of Ku70 coimmunoprecipitated with TCF-4 and removed β-catenin from TCF-4. We propose a working model in which the transcriptional activity of TCF-4 is regulated by the relative amount of Ku70, PARP-1, and β-catenin proteins binding to TCF-4. Identification of the functional interaction of Ku70 as well as PARP-1 with the TCF-4 and β-catenin transcriptional complex may provide insights into a novel linkage between DNA damage recognition/repair and Wnt signaling. [Cancer Res 2007;67(3):911–8]

Introduction

The Wnt signaling pathway plays important roles in embryogenesis and carcinogenesis (1). Secreted Wnt molecules bind to cell membrane Frizzled receptors and evoke downstream intracellular signaling. The signal is then transmitted to a multiprotein complex consisting of the *APC* gene product, Axin/Axil, and glycogen synthase kinase 3β (GSK3β), a chaperone that supports the phosphorylation of β-catenin by GSK3β (2, 3). Phosphorylated β-catenin protein is subject to rapid degradation via the ubiquitin-proteasome pathway (4). The Wnt signaling inhibits GSK3β and increases the cytoplasmic β-catenin content. Mutation of either the *APC* or β-catenin (*CTNNB1*) gene is frequently seen in colorectal carcinoma and mimics the constitutively active Wnt signaling (5, 6). The excess β-catenin protein acts as a transcriptional coactivator by forming complexes with T-cell factor (TCF)/lymphoid enhancer factor (LEF) family DNA-binding proteins (7). TCF-4 is a member of the TCF/LEF family commonly expressed in colorectal epithelium and cancer cells (8). TCF-4 has been implicated in the maintenance of undifferentiated intestinal crypt epithelial cells because no proliferative compartments have been detected in the intestinal crypts of mice lacking TCF-4 (9). Constitutive transactivation of the target genes of TCF-4 by accumulation of β-catenin protein imposes a crypt progenitor phenotype on intestinal epithelial cells and is considered crucial to the initiation of colorectal carcinogenesis (10).

In our previous study, we found that poly(ADP-ribose) polymerase-1 (PARP-1) interacted with the TCF-4 and β-catenin nuclear complex (11). PARP-1 was originally identified as a nuclear DNA-binding protein that catalyzes the transfer of ADP-ribose from NAD^+ to acceptor proteins (12). PARP-1 is activated by DNA damage and plays an important role in the process of DNA repair and genomic stability (13).

Besides DNA damage recognition and apoptosis, the role of PARP-1 as a regulator of various transcription factors has recently attracted a great deal of attention (14). We have found that PARP-1 is a component and enhancer of the TCF-4 and β-catenin transcriptional complex (11). PARP-1 polyADP-ribosylates its own automodification domain in response to DNA damage (12). PolyADP-ribosylation of PARP-1 inhibits the interaction with TCF-4 and its transcriptional activity (11). However, the biological significance of the interaction between TCF-4 and PARP-1 and its inhibition by polyADP-ribosylation of PARP-1 remains unexplained.

In this study, we further explored the protein components of the TCF-4 and β-catenin nuclear complex and identified that Ku70 and Ku80 proteins interact with TCF-4. The Ku autoantigen was originally identified as a nuclear protein recognized by autoantibodies in sera of patients with polymyositis-scleroderma overlap syndrome (15). The Ku autoantigen consists of two subunit proteins of ~70 kDa and 80 to 86 kDa (named Ku70 and Ku80). Ku recognizes DNA double strand breaks and then recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs; ref. 16). The Ku70/Ku80/DNA-PKcs complex mediates nonhomologous end joining and repairs double strand breaks (17). Ku proteins are also involved in other cellular processes such as immunoglobulin gene rearrangement, telomere maintenance, apoptosis, and transcriptional regulation (18). Here, we report that Ku70 is a novel inhibitor of the β-catenin/TCF-4 transcriptional complex.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Tesshi Yamada, Chemotherapy Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. Phone: 81-3-3542-2511, ext. 4270; Fax: 81-3-3547-6045; E-mail: tyamada@gan2.res.ncc.go.jp.

doi:10.1158/0008-5472.CAN-06-2360

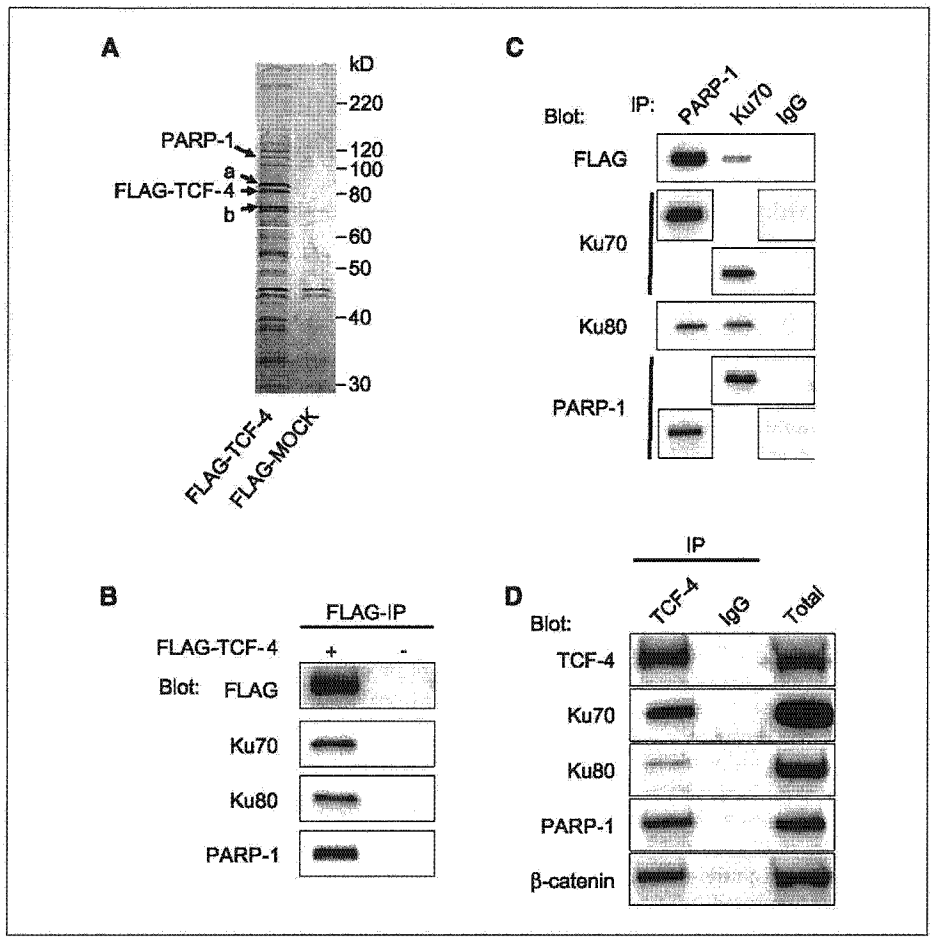

Figure 1. Identification of interaction between Ku and TCF-4. *A*, HEK293 cells were transfected with FLAG–TCF-4 or control FLAG-MOCK. Nuclear extracts were immunoprecipitated with anti-FLAG affinity gel (modified from ref. 11 with permission). *B*, Western blot analysis of the immunoprecipitates (*IP*) of HEK293 cells transfected with FLAG–TCF-4 (+) or control FLAG-MOCK (−). The immunoprecipitates with anti-FLAG affinity beads were blotted with anti-FLAG, anti-Ku70, anti-Ku80, and anti–PARP-1 antibodies. *C*, lysate of HEK293 cells transfected with FLAG–TCF-4 was immunoprecipitated with anti–PARP-1 and anti-Ku70 antibody or normal mouse IgG and blotted with anti-FLAG, anti-Ku70, anti-Ku80, and anti–PARP-1 antibodies. *D*, nuclear extract of HCT116 cells (*Total*) was immunoprecipitated with anti–TCF-4 antibody or normal mouse IgG and blotted with anti–TCF-4, anti-Ku70, anti-Ku80, anti–PARP-1, and anti–β-catenin antibodies.

Materials and Methods

Cell culture. The human embryonal kidney cell line HEK293 was obtained from the Riken Cell Bank (Tsukuba, Japan). Hepatoblastoma cell line HepG2 and colorectal cancer cell lines HCT116, DLD-1, and SW480 were purchased from the American Type Culture Collection (Manassas, VA). The Li7 cell line was established from a patient with hepatocellular carcinoma as reported previously (19). PARP-null mouse embryonic fibroblast (MEF) was established from a PARP-1 knockout (*Parp1*$^{-/-}$) mouse (20).

Cells were treated with 5 mmol/L hydroxyurea (Sigma, St. Louis, MO) for 18 h at 37°C. The medium was then removed, and incubation was continued with serum-free medium without or with bleomycin (50 μg/mL; Sigma).

Plasmid constructs. Human TCF-4 cDNA and its truncated forms were subcloned into pFLAG-CMV4 (Sigma). Human Ku70 cDNA and its truncated forms were subcloned into pcDNA3.1/myc-His (Invitrogen, Carlsbad, CA). Human PARP-1 cDNA (kindly provided by Dr. M. Miwa, Nagahama Institute of Bio-Science and Technology, Nagahama, Japan) was subcloned into pcDNA3.1/myc-His. Human β-cateninΔN134 cDNA was subcloned into pCR3.1 (Invitrogen), which lacks a 134-amino-acid sequence at its NH_2 terminus. The composition of all of the constructs in this study was confirmed by restriction endonuclease digestion and sequencing.

Immunoprecipitation. Cells were extracted with lysis buffer [50 mmol/L Tris-HCl (pH 7.4), 150 mmol/L NaCl, 1 mmol/L EDTA, 1% Triton X-100] containing a protease inhibitor cocktail (Sigma). Nuclear extracts were prepared with the CelLytic nuclear extraction kit (Sigma). Immunoprecipitation was done with 50 μL of anti-FLAG M2 affinity gel (Sigma) or anti–PARP-1 monoclonal antibody (BD PharMingen, San Diego, CA), anti-Ku70 (Ab-5) monoclonal antibody (Lab Vision, Fremont, CA), and anti–TCF-4 monoclonal antibody (Upstate, Charlottesville, VA) along with 10 μL of Dynabeads Protein G (Dynal, Oslo, Norway). After being washed with washing buffer [50 mmol/L Tris-HCl (pH 7.4), 150 mmol/L NaCl], immobilized immunocomplexes were eluted from anti-FLAG M2 affinity gel by incubation at 4°C with 150 ng/μL 3×FLAG Peptide (Sigma) or from Dynabeads by boiling in SDS loading buffer. Proteins were fractionated by SDS-PAGE and detected using a negative gel stain MS kit (Wako, Osaka, Japan) or by Western blotting.

Protein identification by mass spectrometry. SDS-PAGE gels were cut into ~1-mm^3 sections, reduced with NH_4HCO_3, and alkylated with iodoacetamide. The gel sections were then washed with acetonitrile, hydrolyzed with modified trypsin (Promega, Madison, WI), and incubated at 37°C overnight. Peptides eluted from the gel sections were spotted onto a steel target plate along with 2,5-dihydroxybenzoic acid (gentisic acid; Sigma) as a matrix. Mass spectra were obtained in the refractor mode by using a Q-star Pulsar-*i* mass spectrometer (Applied Biosystems, Foster City, CA) and analyzed using Mascot software (Matrix Sciences, London, United Kingdom; ref. 21).

Western blot analysis. Anti-FLAG M2 monoclonal antibody was purchased from Sigma; anti-Ku70 (Ab-4) and anti-Ku80 (Ab-2) monoclonal antibodies were from Lab Vision; anti–β-catenin monoclonal antibody was from BD Transduction (Lexington, KY); anti-TCF3/4 monoclonal antibody was from Upstate; and anti-PARP polyclonal antibody was from Trevigen (Gaithersburg, MD). Total cell lysates were extracted at 4°C with radioimmunoprecipitation assay buffer [150 mmol/L NaCl, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mmol/L Tris-HCl (pH 8.0)]. Samples were

fractionated by SDS-PAGE and transferred onto Immobilon-P membranes (Millipore, Billerica, MA), and the blots were detected using an enhanced chemiluminescence method (Amersham, Piscataway, NJ).

Reverse transcription-PCR. Total RNA was prepared with an RNeasy mini kit (Qiagen, Valencia, CA), and 1-μg samples of total RNA were reverse transcribed. cDNA samples from tissues of human sporadic colorectal cancer and the corresponding normal tissues were obtained from Clontech (Palo Alto, CA). The PCR products were analyzed by agarose gel electrophoresis. The sequences of all the PCR primers in this study are available upon request.

Luciferase reporter assay. A pair of luciferase reporter constructs, TOP-FLASH and FOP-FLASH (Upstate), were used to evaluate TCF/LEF transcriptional activity. Cells were transiently transfected in triplicate with one of the luciferase reporters and phRG-TK (Promega) using Lipofect-AMINE 2000 reagent (Invitrogen). Luciferase activity was measured with the Dual-luciferase reporter assay system (Promega) and *Renilla* luciferase activity as an internal control.

RNA interference. Two short hairpin RNA (shRNA) sequences targeting Ku70 mRNA were designed by B-Bridge (Sunnyvale, CA). Synthesized double-stranded oligonucleotides were cloned into the pSUPER RNA interference vector (OligoEngine, Seattle, WA) carrying the H1 promoter and neomycin resistance gene.

Immunofluorescence microscopy. Cells were grown on poly-L-lysine–coated coverslips (Asahi Technoglass, Funabashi, Japan). After being fixed with 3.7% paraformaldehyde, the cells were incubated with anti-PARP rabbit polyclonal antibody and anti-PARP-1 mouse monoclonal antibody (BD Transduction) overnight at 4°C. Following incubation with Alexa Fluor 488–labeled goat anti-mouse IgG and Alexa Fluor 594–labeled goat anti-rabbit IgG (Molecular Probes, Eugene, OR), the coverslips were inspected with a laser scanning confocal microscope (Bio-Rad, Hercules, CA).

Immunohistochemistry. Ten familial adenomatous polyposis (FAP) patients were selected from the surgical pathology panel of the National Cancer Center Central Hospital. Formalin-fixed and paraffin-embedded intestinal tissues containing adenomas were stained by the avidin-biotin complex method as previously described (22).

Results

Identification of a novel interaction between the TCF-4 and Ku proteins. HEK293 cells were transiently transfected with FLAG-tagged TCF-4 (FLAG–TCF-4) or a control plasmid (FLAG-MOCK). Immunoprecipitation with anti-FLAG antibody and SDS-PAGE revealed that several proteins were selectively coimmunoprecipitated with FLAG–TCF-4, but not with the control (Fig. 1*A*). We had

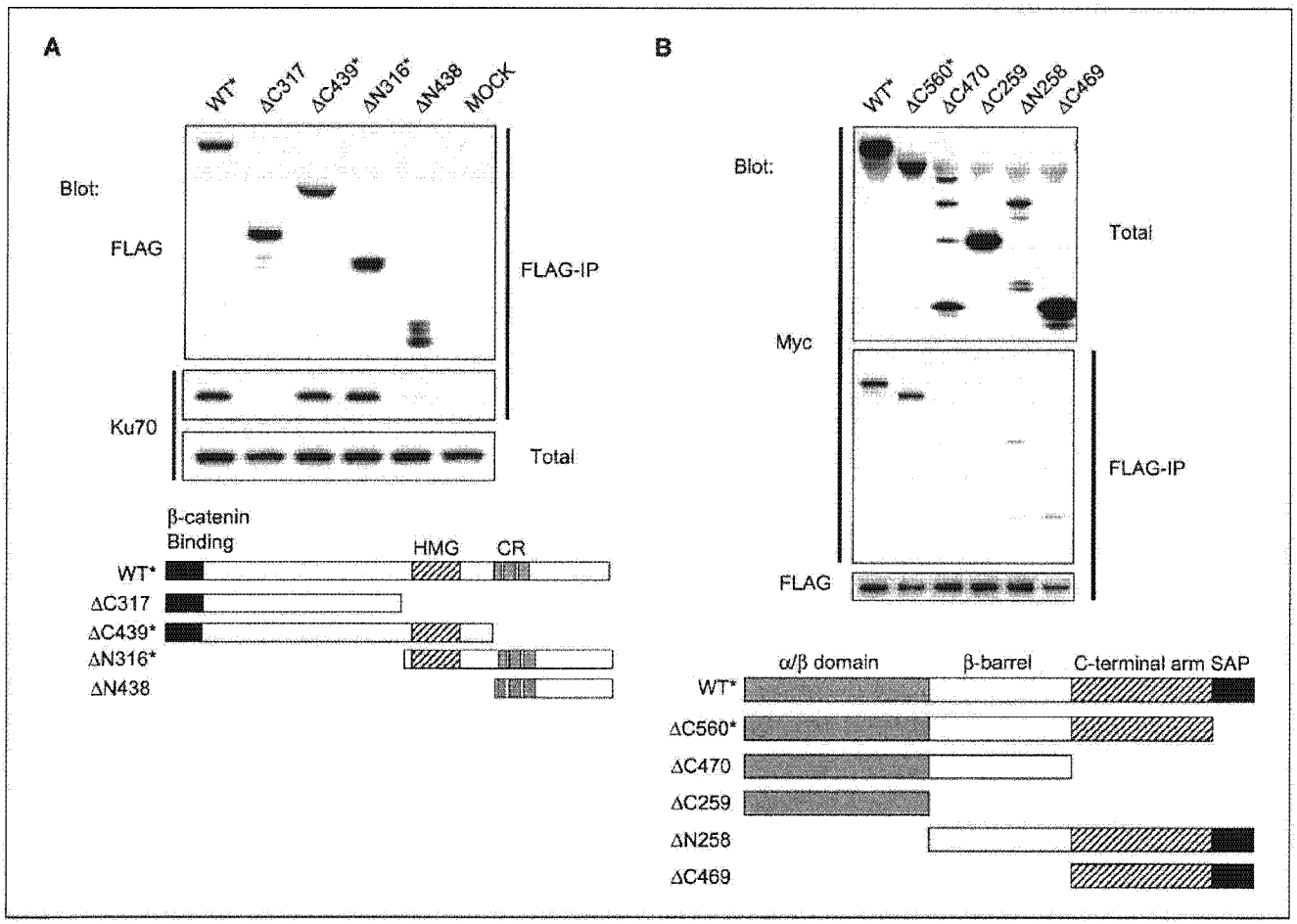

Figure 2. Binding domains necessary for interaction between Ku70 and TCF-4. *A*, full-length or truncated forms of FLAG–TCF-4 were transfected into HEK293 and immunoprecipitated with anti-FLAG affinity beads. The complexes were analyzed by blotting with anti-FLAG and anti-Ku70 antibodies. The full-length and truncated forms of TCF-4 are represented schematically at the bottom. *, TCF-4 constructs that bound to the Ku70 protein. *B*, Myc-tagged full-length or truncated forms of pcDNA3.1-Ku70 and full-length FLAG–TCF-4 were cotransfected into HEK293 and immunoprecipitated with anti-FLAG affinity beads. The complexes were analyzed by blotting with anti-Myc and anti-FLAG antibodies. The full-length and truncated forms of Ku70 are represented schematically at the bottom. *, Ku70 constructs that bound to the TCF-4 protein.

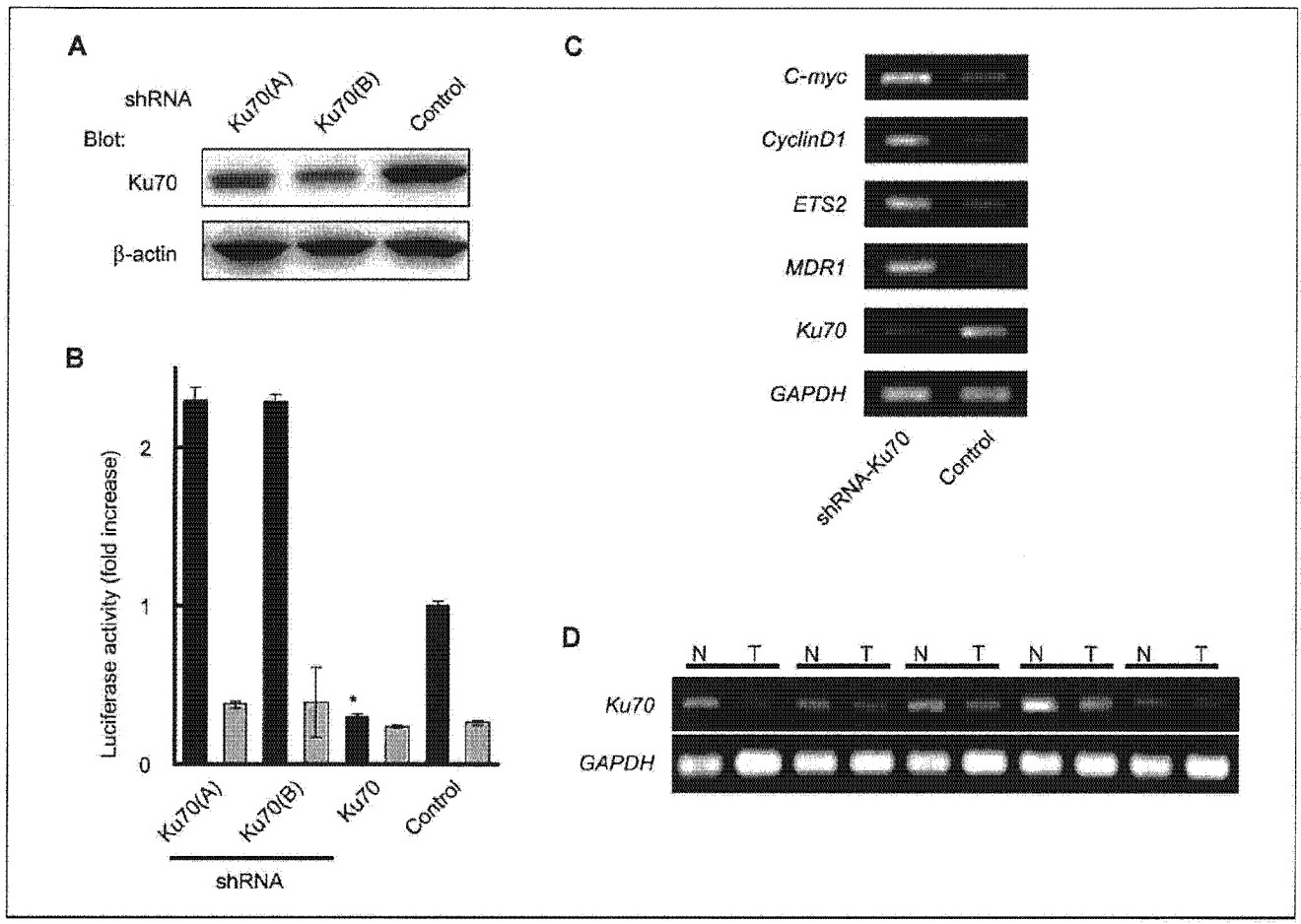

Figure 3. Ku70 suppresses gene transcriptional activity of TCF-4. *A*, Western blot analysis showing the protein level of Ku70 (*top*) and β-actin (loading control, *bottom*) of HCT116 cells transfected with pSUPER-Ku70 [*Ku70(A), Ku70(B)*] or pSUPER-control (*Control*). *B*, HCT116 cells were cotransfected with pSUPER-Ku70(A), pSUPER-Ku70(B), pcDNA3.1-Ku70 (*Ku70*), or control plasmid as well as canonical (TOP-FLASH) or mutant (FOP-FLASH) TCF/LEF luciferase reporter. Forty-eight hours after transfection, the luciferase activity of TOP-FLASH (*black columns*) and FOP-FLASH (*gray columns*) was measured. Activity was adjusted to the TOP-FLASH activity of the control transfectant and expressed as a fold increase. *C*, HCT116 cells were transiently transfected with a mixture of pSUPER-Ku70(A) and pSUPER-Ku70(B) (*shRNA-Ku70*) or empty pSUPER (*Control*). Forty-eight hours after transfection, the expression levels of *c-myc, cyclin D1, ETS2, MDR1, Ku70*, and *GAPDH* mRNA were analyzed by reverse transcription-PCR. *D*, expression of Ku70 and GAPDH mRNA in paired samples of normal intestine (*N*) and cancer (*T*) tissues from five patients with sporadic colorectal cancer.

previously identified one of these proteins as PARP-1 (Fig. 1*A*). Proteins of ~70 kDa (Fig. 1*A*, *b*) and 86 kDa (Fig. 1*A*, *a*) were also constantly coimmunoprecipitated with FLAG-tagged TCF-4 and were subjected to protein identification by mass spectrometry. Peptide mass fingerprinting and tandem mass spectrometry (data not shown) revealed that these proteins were Ku70 (70-kDa thyroid autoantigen/thyroid-lupus autoantigen/G22P1) and Ku80 (X-ray repair, complementing defective, in Chinese hamster, 5/XRCC5).

The protein identification was confirmed by Western blotting with anti-Ku70 and anti-Ku80 antibodies. Ku70, Ku80, and PARP-1 proteins were detected in the immunoprecipitate with anti-FLAG antibody (Fig. 1*B*). Ku70, Ku80, and FLAG-tagged TCF-4 proteins were detected in the immunoprecipitate with anti–PARP-1 antibody (Fig. 1*C, IP: PARP-1*). FLAG-tagged TCF-4 (FLAG), Ku80, and PARP-1 were also detected in the immunoprecipitate with anti-Ku70 antibody (Fig. 1*C, IP: Ku70*) but not with control mouse IgG (Fig. 1*C, IP: IgG*).

Ku70, Ku80, PARP-1, and β-catenin proteins were coimmunoprecipitated with endogenous TCF-4 from a lysate of colorectal cancer HCT116 cells (Fig. 1*D*). Ku70 and Ku80 were also coimmunoprecipitated with PARP-1 (Supplementary Fig. S1), suggesting that Ku70, Ku80, and PARP-1 are native components of the TCF-4 and β-catenin complex.

Binding domains necessary for the interaction between Ku70 and TCF-4. To identify the region of TCF-4 that is essential for its interaction with Ku70, we expressed serially truncated forms of FLAG–TCF-4 and evaluated their binding activity to Ku70 (Fig. 2*A*). Only constructs carrying the high-mobility group (HMG) box [wild-type (WT), ΔC439, and ΔN316] were found to bind to Ku70 (Fig. 2*A*).

The Ku70 protein consists of four domains: the α/β domain, β-barrel, COOH-terminal arm, and scaffold attachment factor (SAP) DNA-binding domain (23). We evaluated the ability of Ku70 serially truncated at the border of each domain to bind to FLAG–TCF-4 (Fig. 2*B*). Only the full-length Ku70 protein (WT) and the Ku70 protein lacking the SAP domain (ΔC560) interacted with TCF-4 (Fig. 2*B*). These results suggest that the three-dimensional structure of Ku70 protein rather than the specific amino acid