Supplementary Material

Isolation and Proteomic Characterization of Human Parvulin-associating Preribosomal Ribonucleoprotein Complexes

Sally Fujiyama 1, Mitsuaki Yanagida 1, 2, Toshiya Hayano 1, 2, Yutaka Miura 1, Toshiaki Isobe 2, 3 and Nobuhiro Takahashi 1, 2

1 Department of Biotechnology, United Graduate School of Agriculture, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu-shi, Tokyo 183-8509, Japan, and
2 Integrated proteomics system project, Pioneer Research on Genome the Frontier, Ministry of Education, Culture, Sports, Science & Technology of Japan, and
3 Laboratory of Biochemistry, Graduate School of Science, Tokyo Metropolitan University, 1-1 Minamiosawa, Hachiouji, Tokyo

* To whom correspondence should be addressed:

Nobuhiro Takahashi, Ph. D.
Applied Biological Science
Tokyo University of Agriculture and Technology
3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan
Tel:/Fax:.+81-042-367-5709
e-mail: ntakahas@cc.tuat.ac.jp
### Supplementary Table 1. Identification of ribosomal proteins associated with hParvulin.

Each entry contains the identified protein name with NCBI accession No. in parenthesis [e.g. (AF374195) ribosomal protein L6], species of the best matched protein by PMF method (M, mouse; H, human), the theoretical molecular weight (MW) and pl, the number (and the sequence coverage) of matching peptides of the PMF match, the average deviation in mass between measured and calculated values (33, 34). SDS staining band No. corresponding to FIG. 5 is also shown. For putative proteins, the best matching protein after TFAST searching is also indicated.  * , detect in Hela cell.  #, detect in 293 EBNA cell. +, identified also by LC-MS/MS method.

| Protein name | Species identified | SDS band No. | No. of Peptides Matched | Mean Error | Protein MW (Da) / pI | Identified by homology search |
|-------------|------------------|-------------|------------------------|------------|---------------------|-------------------------------|
| **large Subunit** | | | | | | |
| (AK011068) putative#,+ | M | 19 | 14 (38%) | -8.2575 | 47138.1 / 11.06 | ribosomal protein L4 (L1) |
| (BC009655) ribosomal protein L3*#,+ | M | 20 | 15 (37%) | -7.9528 | 46124.4 / 10.22 |
| (AK010267) ribosomal protein P0*#,+ | M | 27 | 12 (46%) | 0.2342 | 34216.7 / 5.91 |
| (AF374195) ribosomal protein L6*#,+ | M | 28,31,32,36 | 9 (30%) | -7.0744 | 33509.9 / 10.69 |
| (AK008489) putative + | M | 29,30 | 9 (35%) | -1.0579 | 34401.0 / 9.78 |
| (M29015) ribosomal protein L7#,+ | M | 32,33 | 14 (44%) | -3.4051 | 31391.9 / 10.83 |
| (M14689) ribosomal protein L7a (surfeit 3 protein)#,+ | M | 32,33,35 | 14 (47%) | -6.1395 | 29976.8 / 10.56 |
| (AL022721) ribosomal protein L10A# | H | 37 | 4 (16%) | 8.0404 | 24831.5 / 9.94 |
| (AK002787) putative + | M | 37 | 9 (46%) | 1.3670 | 24305.7 / 11.54 |
| (AK012277) putative + | M | 38 | 5 (23%) | -3.4016 | 23564.3 / 11.03 |
| (U67771) ribosomal protein L8 + | M | 39 | 6 (21%) | 1.4932 | 28024.9 / 11.04 |
| (BC003358) ribosomal protein L10*# | H | 39 | 11 (38%) | -9.7072 | 24577.1 / 10.11 |
| (X51528) tum- transplantation antigen P198, a mutant of the mouse homolog of rat ribosomal protein L13a# | M | 40 | 11 (45%) | 5.2802 | 23494.4 / 11.02 |
| (AK010674) putative*,#,+ | M | 40 | 17 (64%) | 0.7877 | 24047.1 / 11.57 |
| (AK010440) putative | M | 40 | 4 (25%) | -18.5596 | 23466.2 / 11.48 |
| (AK009044) putative | M | 41 | 6 (35%) | -7.1054 | 21644.7 / 11.79 |
| (AL049597) novel protein similar to ribosomal protein | H | 42 | 5 (17%) | 2.5859 | 21502.5 / 10.21 |
| (AK010983) putative + | M | 42 | 14 (54%) | -6.7830 | 20732.5 / 10.72 |
| (U93863) ribosomal protein L21 + | M | 43 | 8 (44%) | -8.7068 | 18562.0 / 10.49 |
| (X80699) ribosomal protein L26 | M | 44 | 9 (40%) | -2.0535 | 17258.4 / 10.56 |
| **Small Subunit** | | | | | | |
| (AK012284) putative | M | 24, 44 | 6 (20%) | 6.3004 | 28680.9 / 10.85 |
| (Z83368) ribosomal protein S3a + | M | 31 | 15 (50%) | -9.0527 | 29885.1 / 9.75 |
| (X76772) ribosomal protein S3 | M | 32 | 8 (34%) | -4.8712 | 26674.5 / 9.68 |
| (AK002445) putative | M | 34 | 10 (45%) | 0.3286 | 29598.0 / 10.16 |
| (AK009023) putative + | M | 36,37 | 6 (18%) | 5.9890 | 24205.4 / 10.32 |
| (BC000802) Similar to ribosomal protein S9 | H | 41,42 | 13 (46%) | -11.384 | 22591.6 / 10.66 | 99.4% identity to ribosomal protein |
Supplementary Table 2. Identification of trans-acting factors associated with hParvulin.

| Protein name | Species identified | SDS band No. | No. of Peptides Matched (Peptide cover) | Mean Error | Protein MW (Da/pl) | Yeast Homologue | ORF No. | % identity with yeast homologue (overlap in aa.) | null mutant Phenotype | localization | Known function/Similality/Localization 1) | ref. |
|--------------|--------------------|--------------|----------------|-----------|----------------|----------------|---------|---------------------------------|----------------|-------------|-----------------------------------------|-----|
| (U19891) putative CCAAT binding factor 1 (mCBF1), alternatively spliced transcript mCBF1 + | M | 3 | 25 (28%) | -4.5676 | 113831.9 / 5.45 | Mak21 / Noc1 | YDR060W | 30.6% (984) | lethal | nucleus / nucleolus | 79.7% identity to human CBF2 that stimulates transcription from the hsp70 promoter. Yeast Mak21p is required for 60S ribosomal subunit synthesis, ribosome maturation and transport. Associates with the 66S pre-ribosomes and also with the 90S pre-ribosomes. | 63, 64 |
| (BC007151) Similar to nucleolar protein 1 (120kd) + | M | 4 | 15 (18%) | -5.5914 | 86752.5 / 9.27 | Nop2 / Yna1 | YNL061W | 49.1% (532) | lethal | nucleolus | 73.8% identity to human proliferating cell nucleolar antigen (p120) that is expressed in tumor cells in the early G1 phase of the cell cycle. Yeast Nop2p is RNA methyltransferase that required for 60S ribosomal subunit synthesis and 2'-O-ribose methylation at a specific site in 86.2% identity to human autoantigen NPG-1. Yeast Nop2p is required for export of the 60S ribosomal subunit from the nucleus to the cytoplasm. | 36, 65, 66 |
| (BC003262) Similar to nucleolar GTPase | M | 5 | 12 (21%) | -2.6883 | 83330.2 / 9.27 | Nog2 | YNR053C | 56.2% (450) | lethal | nucleolus | Bop1 plays a key role in the formation of the mature 28S and 5.8S rRNAs and in the biogenesis of the 60S ribosomal subunit. Yeast Erb1p is an essential gene required for maturation of the 25S and 5.8S ribosomal RNAs. | 67, 68, 69 |
| (U77415) Bop1 + | M | 6 | 24 (43%) | -10.0350 | 82546.3 / 5.86 | Erb1 | YMR034C | 41.5% (668) | lethal | nucleolus | | | 36, 37, 70, 71 |
| (X07699) nucleolin + | M | 6,12 | 35 (45%) | -0.5645 | 76723.7 / 4.69 | Nsr1 / She5 | YGR159C | 32.9% (347) | viable | nucleus / nucleolus | Nucleolin is thought to play a role in pre-rRNA transcription and ribosome assembly. Yeast Nup1p is required for efficient pre-rRNA processing at sites A0 to A2. Yeast Nop5p is required for export of the 60S ribosomal subunit from the nucleus to the cytoplasm. | 72, 73 |
| (AK001239) unnamed protein product + | H | 7 | 23 (29%) | -3.9846 | 85766.4 / 9.29 | Nop2 / Nop77 | YPL043W | 27.8% (730) | lethal | nucleolus | Yeast Nop5p is a nucleolar protein required for conversion of 27SA3 pre-rRNA into mature 5.8S and 25S rRNA. Putative essential nucleolar GTP-binding protein. Identified in a two-hybrid screen as a protein that interacts with NOPP44/46, a nucleolar phosphoprotein of Trypanosoma brucei. purifies as one of the proteins associated with the nuclear pore complex. | 74, 69, 75 |
| (AF348208) GTP-binding protein NGB | M | 11,16 | 17 (26%) | -2.1913 | 74113.4 / 9.55 | Nog1 | YPL093W | 47.5% (649) | lethal | nucleus / nucleolus | Yeast Nop5p is involved in the pre-rRNA processing steps that lead to formation of 18 S rRNA; interacts with Nop1p. Component of C/D box snoRNPs, which are necessary for 2'-O-methylation of ribosomal RNAs. | 36, 37, 69, 76 |
| (AF289539) PES1 protein | M | 12 | 24 (38%) | -5.6874 | 67796.1 / 6.40 | Nop7 / Rpl13 | YGR103W | 40.4% (599) | lethal | nucleus / nucleolus | Zebrafish pescadillo is essential for embryonic development. Yeast Nop7p is associated primarily with 66S preribosomes containing either 27S or 25.5S plus 7S pre-ribosomes and contains one copy of the BRC7 motif, which is found in proteins associated with DNA checkpoint pathways. | 36, 37, 69, 76 |
| (AK009799) putative + | M | 13 | 13 (25%) | 0.3731 | 64449.9 / 9.18 | Nop56 / Sik1 | YLR197W | 49.9% (509) | lethal | nucleolus | 88.9% identity with human Nop56. Depletion of yeast counterpart results in inhibition of pre-rRNA processing at sites A0 to A2. Nucleolar protein component of box C/D snoRNPs, which are necessary for 2'-O-methylation of ribosomal RNAs. | 35, 38, 42 |
| (AF053232) SIK similar protein (Nop5/58) + | M | 14 | 11 (26%) | -6.0509 | 53112.0 / 8.66 | Nop5 / 58 | YOR310C | 49.6% (452) | lethal | nucleolus | Yeast Nop5p is involved in the pre-rRNA processing steps that lead to formation of 18 S rRNA; interacts with Nop1p. Component of C/D box snoRNPs, which are necessary for 2'-O-methylation of ribosomal RNAs. | 35, 38, 41, 42, 77, 79 |
| (U79773) NNP-1 | M | 15 | 8 (22%) | -6.1379 | 50013.0 / 8.25 | Rpl1 | YDR087C | 30.2% (278) | lethal | nucleolus | Yeast Rpl1p is recruited at late stages of nucleologenesis. Rpl1p is involved in pre-rRNA processing from 27S to 25S and 5.8S | 37, 80, 81 |
| Accession | Description | Gene Symbol | Identity | Coverage | Log10 P-value | Molar Concentration | Phenotype | Function |
|-----------|-------------|-------------|----------|----------|---------------|-------------------|-----------|----------|
| AK012937 | putative    | +           | 15       | 13 (35%) | -3.5418       | 41385.4 / 9.87    | Cic1 / Nsa3 | YHR052W  | 26.6%    | proteasome/nucleus | 60.1% identity to human PBK1 (A007398) that could be involved in the regulation of proliferation/differentiation and potentially in invasion of trophoblast cells. Yeast Cic1p, the protease substrate recruitment factor, localizes in the nucleus where it is concentrated in a crescent-shaped region that colocalizes with the nucleolar marker Nop1p, which is involved in ribosome biogenesis. |
| AK007463 | unnamed protein product | +          | 18       | 12 (30%) | -17.4856      | 47286.8 / 5.36    | Ynm1       | YOR272W  | 30.6%    | nucleus/nucleolus | Yeast Ytm1p is microtubule-associated protein essential for the G1/S transition, member of WD (WD-40) repeat family. |
| AF239765 | nuclear protein Ytm1p# | +          | 21       | 23 (46%) | -8.8611       | 41495.1 / 10.77   | Rxs1       | YOR294W  | 39.2%    | nucleus/nucleolus | Yeast Ytm1p is microtubule-associated protein essential for the G1/S transition, member of WD (WD-40) repeat family. |
| AK019913 | putative    | +           | 23       | 12 (37%) | -1.6847       | 34703.0 / 10.08   | Ebp2       | YKL172W  | 34.7%    | nucleus/nucleolus | Yeast Ytm1p is microtubule-associated protein essential for the G1/S transition, member of WD (WD-40) repeat family. |
| AB056870 | nucleolar protein mNIFK | +          | 25       | 14 (45%) | -7.9924       | 36265.4 / 10.24   | Nop15      | YNL110C  | 31.7%    | nucleus/nucleolus | Nucleolar RNA-binding protein. It is tempting to speculate that NIFK plays a role in rRNA synthesis and the silencing of rDNA transcription during mitosis is accomplished through the mitosis-specific and phosphorylation-dependent interaction of nNIFK with pKi-67 . |
| Z22593  | fibrillarin# | +           | 26       | 9 (27%)   | -6.0710       | 34439.2 / 10.29   | Nop1      | YDL014W  | 68.3%    | nucleus/nucleolus | mammalian homologue of yeast Nop1p that is a component of C/D-box snRNPs. Depletion results in inhibition of 2'-O-ribose methylation and pre-rRNA processing at sites A0 to A2. |
| AK011387 | putative    | +           | 33       | 12 (38%) | -0.9223       | 27545.9 / 8.63    | Mrt4       | YKL009W  | 38.3%    | unknown | mRNA turnover 4. Mrt4p required for an early step in mRNA decay. Associates with Nop7p / Rrp13p and RNA helicases. |

**RNA helicases**

| Accession | Description | Gene Symbol | Identity | Coverage | Log10 P-value | Molar Concentration | Phenotype | Function |
|-----------|-------------|-------------|----------|----------|---------------|-------------------|-----------|----------|
| AF220365 | nucleolar RNA helicase Id/Gu DDX21 + | DDX21 + | 5        | 23 (38%) | -10.6417      | 88605.3 / 9.17    | Dbp1 / Lplh8 | YPL119C | 31.2%    | viable | nucleolus | RNA helicase Yeast Dsp1p is involved in ribosome assembly and 27 S to 25 S rRNA maturation. |
| BC001132 | Similar to hypothetical protein FLJ120296 + | FLJ120296 + | 7        | 19 (36%) | -11.0116      | 64920.3 / 9.89    | Dsr1      | YLL008W | 42.5%    | viable | nucleolus | Yeast Dsp1p is involved in ribosome assembly and 27 S to 25 S rRNA maturation. |
| AF279891 | dead box protein 15 | DDX24 | 8        | 15 (20%) | -9.4345       | 90933.6 / 7.12    | Pp43      | YGL120C | 65.7%    | viable | nucleus | Pre-mRNA processing factor involved in disassembly of spliceosomes after the release of mature mRNA |
| AF214732 | ATP-dependent RNA helicase (Asp-Glu-Ala-Asp/His)(DDX24) | DDX24 | 10       | 9 (13%)   | -3.7041       | 96472.1 / 9.44    | Prp5 / Rna5 | YBR237W | 23.3%    | viable | nucleus | Yeast Prp5p is pre-mRNA processing RNA helicase of the DEAD box family and mRNA-associated spliceosomal protein. |
| AK019845 | putative    | +           | 10       | 23 (38%) | -2.3262       | 74239.5 / 9.52    | Has1      | YMR290C | 60.4%    | lethal | nucleus | 85.8% identity with human myc-regulated DEAD box protein (DDX18). Yeast Has1p is purified as one of the proteins associated with the nuclear pore complex. |
| X65627 | p68 RNA helicase TNZ2 | DDX24 | 14       | 11 (16%) | -9.0053       | 69320.9 / 9.06    | Dbp2      | YNL112W | 56.3%    | viable | nucleus | Yeast Dbp2p acts before or at decapping, is predominantly cytoplasmic, and associates with polysomes. Dbp2p also plays an important role in rRNA processing. Required for 25 S rRNA production and 60S ribosomal subunits assembly. |
| AK011136 | putative    | +           | 15       | 6 (15%)   | -4.2276       | 61212.6 / 9.26    | Dbp9      | YLR276C | 43.5%    | lethal | nucleolus | 90.9% identity with human putative nucleolar RNA helicase (BC001233). Yeast Dbp9p is required for synthesis of the 27S precessors to mature 25S and 5.8S rRNA. |
**Supplementary Table 3. Identification of hParvulin-associating non-ribosomal proteins of previously unknown function in ribosome biogenesis.**

Each entry contains the identified protein name with NCBI accession No. in parenthesis, the theoretical molecular weight (MW) and pI, the number (and the sequence coverage) of matching peptides of the PMF match, the average deviation in mass between measured and calculated values 33, 34). Origin of the proteins identified is indicated in the second column (M, mouse; H, human). SDS staining band No. corresponding to FIG. 5 is also shown. For putative proteins, the best matching protein after TFAST searching is also indicated with % identity and the sequence length of the match. When yeast homologues are known or identified by TFAST search, protein names and their corresponding ORF names are given. If data are available, cellular localization and function are also indicated. As a control experiment, we identified the proteins present in GST-pulled down fraction, they include Bip and Hsp73 as indicated in Fig 4. Although these proteins were also detected in the hParvulin-associating complexes isolated, they are excluded from Supplementary Table 3. a, listed as the protein constituents of the human nucleolus reported by Andersen et al. (61).

| Protein name | Species identifi ed | SDS band No. | No. of Peptides Matched (Peptide cover) | Mean Error | Protein MW (Da)/pI | Yeast Homologue | Yeast ORF No. | % identity with yeast homologue | deletion mutant phenotype | localization | Known function/Similarity/Localization ref. |
|--------------|--------------------|--------------|--------------------------------------|------------|--------------------|----------------|---------------|---------------------------------|--------------------------|-------------|----------------------------------------|
| (X93167) fibronectin | M | 1 | 19 (19%) | -5.8032 | 160947.5 / 5.19 | - |  |  | nucleus | associate with nuclear matrix | 48, 49 |
| (U63648) p160 myb-binding protein | M | 2 | 31 (34%) | -9.3510 | 151957.2 / 9.05 | - |  |  | nucleolus | localizes predominantly in nucleolus, and has homology with Drosophila lethal 1B and UBF. | 43, 54-55 |
| (AF244547) homologue of targeting protein for Xklp2 (TPX2) | H | 9 | 12 (19%) | 1.8175 | 80989.8 / 9.47 | - |  |  | nucleus | In Xenopus, TPX2 is required for the Ran GTPase- and chromatin-induced microtubule assembly in M phase. | 44-46, 56-58 |
| (AF080863) PITSLRE protein kinase alpha SV9 isoform | H | 15 | 10 (21%) | -14.1581 | 88426.9 / 5.52 | Cdc28 / Cdk1 / Hsl5 / Sm5 | YBR160W | 47.3% (294) | lethal | nucleus/ cytoplasm | Members of the cyclin-dependent kinase superfamily. Yeast Cdc28p is cyclin-dependent protein kinase essential for completion of START and for mitosis, associates with Cks1p and cyclins. | 47, 60, 100 |
| (BC006379) tubulin alpha 1 | H | 17 | 10 (26%) | 0.0521 | 50152.1 / 4.94 | Tub1 | YML085C | 74.7% (446) | lethal | microtubule | constituent of microtubule | 101 |
| (AK010960) putative | M | 17 | 10 (27%) | -7.6338 | 49640.3 / 4.78 | Tub2 | YFL037W | 74.6% (441) | lethal | microtubule | constituent of microtubule | 52, 53, 59 |
| (AK007869) putative | M | 19 | 9 (24%) | -1.2094 | 50061.2 / 6.31 | Elongation factor l-gamma 1/CAM1/TEF3 | YPL048W | 34.2% (404) | viable | - | 99.3% identity with mouse eukaryotic translation elongation factor 1 (EF-1) gamma (BC000384). | 50, 51 |
| (AF050078) growth arrest specific 11 | H | 22 | 9 (16%) | 0.3290 | 56356.0 / 7.71 | - |  |  | - | up-regulated in growth-arrested cells and is a candidate tumor suppressor, is reported to co-localize with microtubules in the cell. | 50, 51 |
| (AK010776) putative | M | 29 | 10 (40%) | -10.5050 | 35364.0 / 10.05 | Rpf2 | YKR081C | 43.6% (305) | lethal | - | Yeast Rpf2p has a PEST motif | - |
| (AY028916) GAJ | H | 43 | 7 (38%) | -9.7968 | 23753.3 / 8.28 | - |  |  | - | 41.6% identity with S.pombe hypothetical coiled-coil protein (SPAC13A11.03) | - |
Supplementary Fig. 1. Identification of Proteins by LC-MS/MS method.

Two examples for the MS/MS spectra obtained by the LC-MS/MS analysis of the lysyl endopeptidase digest of the entire hParvulin-associating complexes are shown. A) A peptide sequence tag from Ytm1p was assembled from the series of C-terminal fragment ions. Peptides identified by the LC-MS/MS analysis are indicated by underlines under the amino acid sequence of Ytm1p. B) A peptide sequence tag from putative protein (AK007491) was assembled from the series of C-terminal fragment ions. Peptides identified by the LC-MS/MS analysis are indicated by underlines under the amino acid sequence of putative protein (AK007491)
Supplementary Results

Isolation of hParvulin-Associating Complexes from the Nuclear Extract of Mouse Fibroblast L929 Cells

To isolate hParvulin-associating proteins, we prepared the nuclear and cytosolic extracts from mouse fibroblast L929 cells. Initially, we tried to isolate hParvulin-associating proteins by eluting with a reduced glutathione from glutathione-Sepharose beads bound to GST-hParvulin after incubating with the nuclear and cytosolic extracts, respectively. While some potential hParvulin-associating proteins were detected following a reduced glutathione elution, the protein pattern was obscured by a high level of contaminating proteins due to nonspecific protein interaction to glutathione-Sepharose beads and GST for both nuclear and cytosolic extracts (data not shown). In order to reduce the contaminating proteins we attempted thrombin-mediated release of hParvulin-associating proteins from the glutathione-Sepharose beads bound to GST-hParvulin. Since the release of hParvulin from glutathione-Sepharose beads bound to GST-hParvulin was nearly complete- without any significant degradation of hParvulin after incubation with thrombin, we adopted the thrombin-mediated release of hParvulin-associating proteins after incubating with the nuclear and cytosolic extracts, respectively (Fig. 3). In a control experiment, although GST and several contaminating proteins were eluted upon thrombin treatment from the glutathione-Sepharose beads bound to GST after incubation with the cell extracts, the number of the contaminating proteins was drastically reduced in the thrombin-released eluate (Fig. 3A lanes 1 and 3). We detected a number of hParvulin-associating proteins upon thrombin treatment of the GST-hParvulin-bound glutathione-Sepharose beads that were incubated with the nuclear extract (Fig. 3A lane 4). In contrast, only a trace of proteins was detected in the
cytoplasmic fraction (Fig. 3A lane 2). Similarly, when human HeLa and 293EBNA cells were used, a number of hParvulin-associating proteins were also detected in their nuclear extracts (Fig. 3B lane 2 and 3C lane 2). Thus, the hParvulin-associating proteins were present only in the nuclear extract. Those results suggest that the proteins isolated by thrombin-mediated release are the nuclear components associated specifically with hParvulin.

**Cellular Localization of hParvulin**

In order to analyze its cellular localization, we constructed two expression vectors; one contains a DNA sequence encoding hParvulin with the N-terminally tagged Flag epitope (Flag-N-hParvulin) (Fig. 1) and another with the C-terminally tagged Flag epitope (hParvulin-C-Flag). Each of the expression vectors was transfected into human 293EBNA cells transiently by use of LipofectAMINE. However, we observed that the efficiency was very low for both Flag-tagged hParvulins if compared with that of Flag-nucleolin or Flag-nucleophosmin observed in separate experiments with the same expression vector and host cell (33). We estimate that the number of the cells transfected with the expression vector for Flag-tagged hParvulins is no more than one fifth of that for Flag-tagged nucleolin or nucleophosmin (data not shown). Although we do not yet know the reasons for the observed inefficiency of the transfection, it is likely that the overexpression of hParvulin causes some toxic effects on the cell. Despite the inefficiency of the transfection, we observed that hParvulin was dispersed throughout the cell, but it was present more in the nucleus including the nucleolus than in the cytoplasm in the 293EBNA cells that were transfected with both expression vectors for Flag-N-hParvulin and hParvulin-C-Flag (only Flag-N-hParvulin is shown in Fig. 4).
Identification of Protein Components in the hParvulin-Associating Complexes

For identification of hParvulin-associating proteins isolated by the thrombin-mediated release, individual protein bands were excised and digested with trypsin in the SDS-PAGE gel. The mass-to-charge ratio of the peptides released from the gel was measured by MALDI-TOF mass spectrometry with high-mass accuracy. The complete set of peptide masses from each protein band was compared with the tryptic peptide masses predicted for each protein in a comprehensive nonredundant database. A protein in the eluate released by thrombin-cleavage of GST-hParvulin was considered identified when the spectrum of its measured peptide masses met the previously established criteria for positive identification of proteins using MALDI-TOF mass spectrometry and automated database analysis (33). For this study, we searched the database for proteins with a mass range of 0 to 30kDa or 10 to 100kDa for smaller proteins (apparent Mr >20kDa by SDS-PAGE), and 50 to 200kDa or 100 to 300kDa for larger proteins, with constraint on mouse and human origins. Incomplete tryptic cleavage and peptide modifications that may alter the peptide masses, such as oxidized methionine or carbamidomethyl cysteine, were calculated for the putatively identified protein and compared with the measured masses. The modified peptides identified in the search were added to the list to increase the number of matching peptides and sequence coverage. Protein that revealed the highest scored PMF matching by database search was retrieved as the identified protein.

Supplementary Tables I, II and III show the proteins identified as the hParvulin-associating proteins by applying the criteria described (33). The identification by the in-gel digestion-MALDI-TOF analysis was done for an individual protein band usually on silver-stained 10% polyacrylamide gels. We also used 11% and 15% polyacrylamide gels to
improve protein separation and to confirm the result obtained by the analysis with the
10% polyacrylamide gels. Most of the proteins listed in the supplementary Tables were
identified at least three times in many independent experiments. A typical SDS-PAGE
staining (10% polyacrylamide gel) of the proteins released by thrombin treatment of the
glutathione-Sepharose beads bound to GST-hParvulin that were incubated with the
nuclear extract from L929 cells, is shown with the number of the protein bands and the
protein names identified on the right side in Fig. 5.

So far, we identified sixty-two proteins in the eluate obtained by thrombin-mediated
release (Supplementary Tables 1, 2 and 3). Of these, 15 proteins were ribosomal
proteins; 3 are for the small subunit, and 12 are for the large subunit (Supplementary
Table 1). In addition to the ribosomal proteins, we also identified 23 non-ribosomal
proteins as well as 24 unnamed or putative proteins (Supplementary Table 2 and 3).
Many of the identified non-ribosomal proteins are trans-acting factors that are known to
be involved in ribosome biogenesis, including those such as p68 RNA helicase TNZ2,
Bop1, nucleolin, Nop5/58, and fibrillarin. We also identified many ribosomal proteins and
protein trans-acting factors for ribosome biogenesis as the hParvulin-associating proteins
in the nuclear extracts obtained from human HeLa and 293E cells, respectively
(Supplementary Tables 1 and 2). Thus, the presence of ribosomal proteins and trans-
acting factors for ribosome biogenesis as the hParvulin-associating proteins in the three
different cell strains analyzed, suggested that hParvulin was associated with
 preribosomal ribonucleoprotein (pre-rRNP) complexes from the nuclear extracts of
mammalian cells. In order to ascertain these protein identification based on PMF data, we
also analyzed a lysyl endopeptidase digest of the entire hParvulin-associating complexes
by liquid chromatography electrospray ionization tandem mass spectrometry (LC-MS/MS)
in a data-dependent manner followed by MS/MS spectra searching against NCBI nr using batch processes of Mascot algorithm (manuscript in preparation). By a single operation of the LC-MS/MS, we confirmed the presence of the 15 ribosomal proteins and the 18 trans-acting factors that were shown to be present in the isolated hParvulin-associating complexes by PMF method (supplementary Table I and II). Two examples of the protein identification with MS/MS spectra data are shown in supplementary Fig. 1. Thus, we believe that all of the proteins identified by PMF method are the reliable components of the hParvulin-associating protein complexes.

**RNA Integrity of the hParvulin-Associating Complex**

Since hParvulin was suggested to bind to the pre-rRNP complex, we addressed questions whether the association of the identified proteins with hParvulin is by direct interactions among these proteins, or by their association to common RNA molecules, or both. In order to clarify among those possibilities, we treated the nuclear extract with RNase A or ethidium bromide before incubating it with the GST-hParvulin-binding glutathione-Sepharose beads. Only a trace of the protein bands was detected as the hParvulin-associating proteins after treating the nuclear extract with RNase A or ethidium bromide (Fig. 6A). In addition, we also treated the hParvulin-associating complex on glutathione-Sepharose beads with RNase A and confirmed that RNase A dissociated most of the proteins associated with hParvulin (Fig. 6B). These results indicate that when RNA was degraded or precipitated before or after mixing with hParvulin, most of the proteins could no longer be associated with hParvulin. These results indicate that almost all protein constituents of the hParvulin-associating complex isolated are associated
through a common RNA molecule. Thus, RNA integrity is necessary for the association of almost all of the proteins present in the hParvulin-associating complex we isolated. These results also suggest that the isolated hParvulin-associating complex is the pre-rRNP complex.

**Presence of Pre-rRNA in the hParvulin-Associating RNP Complexes**

Because all results we obtained suggest that the hParvulin-associating complexes are pre-rRNP complexes, we examined whether the complexes contain pre-rRNA. The identification of pre-rRNA in the isolated hParvulin-associating complexes was achieved by Northern-hybridization analyses with the different probes (Fig. 2). The hParvulin-associating complexes contained almost all rRNA species expected to be produced during rRNA processing in the nucleolus of mammalian cells. Namely, in addition to the matured 18S, 5.8S and 28S rRNAs, we detected 12S, 32S, 34S, 37S, 41S and 45S pre-rRNA species (Fig. 7). Among the pre-rRNA species detected, 32S and 12S pre-rRNAs are concentrated mostly in the isolated hParvulin-associating complexes. 32S pre-rRNA, is identified as the common precursor to 28S and 5.8S rRNA, encompassing the ITS 2 segment, while it reacts negatively with 18S, and 5'ETS probes. 12S pre-rRNA was detected only with ITS2. Thus, RNA species contained in the hParvulin-associating complexes are definitively shown to be rRNAs and pre-rRNAs. These results indicate that the isolated hParvulin-associating complexes are the nuclear precursor of the cytoplasmic ribosomal particles.

**hParvulin Domain that is Involved in Associating to RNP Complexes**

Since the amino-terminal domain of the hParvulin molecule is proposed to be a
putative nucleic acid-associating domain (31), we addressed questions whether the entire hParvulin molecule participates in associating to the RNP complex, or whether any specific domain of the hParvulin molecule is sufficient to bind it. Therefore, we constructed genes for two domain mutants, the amino-terminal domain (GST-ΔC-Par) and the carboxyl-terminal domain (GST-ΔN-Par) (Fig. 1), each of which has a DNA sequence encoding a GST and a short segment with the thrombin cleavage site (TC). The two domain mutants were expressed in a BL21 (DE3) strain, respectively, and were purified as described in Materials and Methods. Each was confirmed to have the expected molecular weight by SDS-PAGE and mass spectrometry analysis (data not shown).

When pulled down from the nuclear extract of L929 cells, the ΔC-Par-associating protein complexes showed a protein profile similar to that for the entire hParvulin on SDS-PAGE gel (Fig. 8A). We could not detect any significant proteins that bound to ΔC-Par from the cytosolic extract as the case of the entire hParvulin molecule. Identification of the protein bands on SDS-PAGE gels by MALDI-TOF analysis and MS-Fit database search showed that all of the protein components of the ΔC-Par-associating complexes so far analyzed are consistent with the protein components identified in the hParvulin-associating complexes (Supplementary Tables 1, 2 and 3). In addition, when the nuclear extract was treated with RNase A before incubation with GST-ΔC-Par, we could not obtain almost all proteins that were pulled down with the entire GST-hParvulin (Fig. 8B). These results indicate that the N-terminal domain of hParvulin is
the principal domain that interacts with the hParvulin-associating pre-rRNP complexes isolated, which indicates specificity of the association between hParvulin and the isolated complexes. On the other hand, the ΔN-Par-associating complexes obtained from the nuclear extract of L929 cells showed a quite different protein profile on SDS-PAGE gel from those of the hParvulin-associating complexes obtained from the nuclear extract (data not shown). Identification of the ΔN-Par-associating proteins indicated that some of the proteins overlapped with those present in the hParvulin-associating proteins; however, the rest of the ΔN-Par-associating proteins are those not found in the hParvulin-associating proteins (manuscript in preparation). Thus, the association of hParvulin with the isolated pre-rRNP complexes occurs mainly at the amino-terminal domain of hParvulin.

**Supplementary References**

63. Lum, L., Sultzman, L., Kaufman, R., Linzer, D. I. H., and Wu, B., (1990) *Mol. Cell. Biol.* 10, 6709-6717

64. Milkereit, P., Gadal, O., Podtelejnikov, A., Trumtel, S., Gas, N., Petfalski, E., Tollervey, D., Mann, M., Hurt, E., and Tschochner, H. (2001) *Cell* 105, 499-509

65. Gustafson, W. C., Taylor, C. W., Valdez, B. C., Henning, D., Phippard, A., Ren, Y., Busch, H., and Durban, E. (1998) *Biochem. J.* 331, 387-393

66. Hong, B., Wu, K., Brockenbrough, J. S., Wu, P., and Aris, J. P., (2001) *Nucleic Acids Res.* 29, 2927-2937

67. Racevskis, J., Dill, A., Stockert, R. and Fineberg, S.A. (1996) *Cell Growth Differ.* 7, 271-280
68. Saveanu, C., Bienvenu, D., Namane, A., Gleizes, P. E., Gas, N., Jacquier, A.,
Fromont-Racine, M., (2001) EMBO J. 20, 6475-6484

69. Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T.,
(2000) J. Cell Biol. 148, 635-652

70. Strezoska, Z., Pestov, D. G., Lau, L. F., (2000) Mol. Cell Biol. 20, 5516-5528

71. Pestov, D. G., Stockelman, M. G., Strezoska, Z., and Lau, L. F., (2001) Nucleic Acids
Res. 29, 3621-3630

72. Lafontaine, D., and Tollervey, D., (1995) Biochem. Cell Biol. 73, 803-12

73. Allain, F. H., Bouvet, P., Dieckmann, and T., Feigon, J., (2000) EMBO J. 19, 6870-6881

74. Berges, T., Petfalski, E., Tollervey, D. and Hurt, E.C., (1994) EMBO J., 13, 3136-3148

75. Park, J. H., Jensen, B. C., Kifer, C. T., and Parsons, M., (2001) J. Cell Sci. 114, 173-185

76. Allende, M. L., Amsterdam, A., Becker, T., Kawakami, K., Gaiano, N., and Hopkins, N.,
(1996) Genes Dev. 10, 3141-3155

77. Lafontaine, D. L. J., and Tollervey, D. (1999) RNA 5, 455-467

78. Wu, P., Brockenbrough, J. S., Metcalfe, A. C., Chen, S., and Aris, J. P. (1998) J. Biol.
Chem. 273, 16453-16463

79. Lafontaine, D. J. L., and Tollervey, D. (1998) Trends. Biochem. Sci. 23, 383-388

80. Savino, T. M., Bastos, R., Jansen, E. and Hernandez-Verdun, D., (1999) J. Cell. Sci.
112, 1889-1900

81. Jansen, E., Meulemans, S. M., Orlans, I. C., and Van de Ven, W. J., (1997) Genomics
42, 336-341.

82. Huch, G., Hohn, H. P., and Denker, H. W., (1998) Placenta 19, 557-567
83. Jager, S., Strayle, J., Heinemeyer, W., and Wolf, D. H. (2001) *EMBO J.* **20**, 4423-4431
84. Kressler, D., Rojo, M., Linder, P., and Cruz, J. (1999) *Nucleic Acids Res.* **27**, 4598-4608
85. Tsuno, A., Miyoshi, K., Tsujii, R., Miyakawa, T., Mizuta, K., (2000) *Mol. Cell. Biol.* **20**, 2066-2074
86. Chatterjee, A., Freeman, J. W., and Busch, H., (1987) *Cancer Res.* **47**, 1123-9
87. Huber, M. D., Dworet, J. H., Shire, K., Frappier, L., and McAlear, M. A. (2000) *J. Biol. Chem.* **275**, 28764-28773
88. Tsujii, R., Miyoshi, K., Tsuno, A., Matsui, Y., Toh-e, A., Miyakawa, T., and Mizuta, K., (2000) *Genes Cells* **5** 543-553
89. Venema, J., and Tollervey, D. (1995) *Yeast* **11**, 1629-1650
90. Zuk, D., Belk, J. P., and Jacobson, A. (1999) *Genetics* **153**, 35-47
91. Ou, Y., Fritzler, M. J., Valdez, B. C., Rattner, J. B. (1999) *Exp. Cell Res.* **247**, 389-398
92. Ripmaster, T. L., Vaughn, G. P., and Woolford, J. L. Jr., (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11131-1135
93. Arenas, J. E., and Abelson, J. N., (1997) *Proc. Natl. Acad. Sci. USA* **94**, 11798-11802
94. Gee, S., Krauss, S. W., Miller, E., Aoyagi, K., Arenas, J., and Conboy, J. G. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 11803-11807
95. Wells, S. E., and Ares, M. Jr., (1994) *Mol. Cell. Biol.* **14**, 6337-6349
96. Grandori, C., Mac, J., Siebelt, F., Ayer, D. E., and Eisenman, R. N., (1996) *EMBO J.* **15**, 4344-4357.
97. Iggo, R. D., Jamieson, D. J., MacNeill, S. A., Southgate, J., McPheat, J. and Lane, D. P. (1991) *Mol. Cell. Biol.* **11**, 1326-1333
98. Bond, A. T., Mangus, D. A., He, F., and Jacobson, A., (2001) *Mol. Cell Biol.* **21**, 7366-7379

99. Daugeron, M. C., Kressler, D., and Linder, P., (2001) *RNA* **7**, 1317-1334

100. Xiang, J., Lahti, J. M., Grenet, J., Easton, J., and Kidd, V. J., (1994) *J. Biol. Chem.* **269**, 15786-15794

101. Wang, D., Villasante, A., Lewis, S. A. and Cowan, N. J. (1986) *J. Cell Biol.* **103**, 1903-1910