The COP9 complex is conserved between plants and mammals and is related to the 26S proteasome regulatory complex

Ning Wei*, Tomohiko Tsuge*†, Giovanna Serino*, Naoshi Dohmae‡, Koji Takio‡, Minami Matsui† and Xing-Wang Deng*

The COP9 complex, genetically identified in *Arabidopsis* as a repressor of photomorphogenesis, is composed of multiple subunits including COP9, FUS6 (also known as COP11) and the *Arabidopsis* JAB1 homolog 1 (AJH1) ([1–3]; unpublished observations). We have previously demonstrated the existence of the mammalian counterpart of the COP9 complex and purified the complex by conventional biochemical and immunoaffinity procedures [4]. Here, we report the molecular identities of all eight subunits of the mammalian COP9 complex. We show that the COP9 complex is highly conserved between mammals and higher plants, and probably among most multicellular eukaryotes. It is not present in the single-cell eukaryote *Saccharomyces cerevisiae*, however. All of the subunits of the COP9 complex contain structural features that are also present in the components of the proteasome regulatory complex and the translation initiation factor eIF3 complex. Six subunits of the COP9 complex have overall similarity with six distinct non-ATPase regulatory subunits of the 26S proteasome, suggesting that the COP9 complex and the proteasome regulatory complex are closely related in their evolutionary origin. Subunits of the COP9 complex include regulators of the Jun N-terminal kinase (JNK) and c-Jun, a nuclear hormone receptor binding protein and a cell-cycle regulator. This suggests that the COP9 complex is an important cellular regulator modulating multiple signaling pathways.

Addresses: *Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8104, USA. †Laboratory for Photoperception and Signal Transduction, Frontier Research Program and ‡Division of Biomolecular Characterization, RIKEN, Saitama 351-01, Japan.

Correspondence: Ning Wei
E-mail: nwei@peaplant.biology.yale.edu

Received: **5 May 1998**
Revised: **5 June 1998**
Accepted: **5 June 1998**

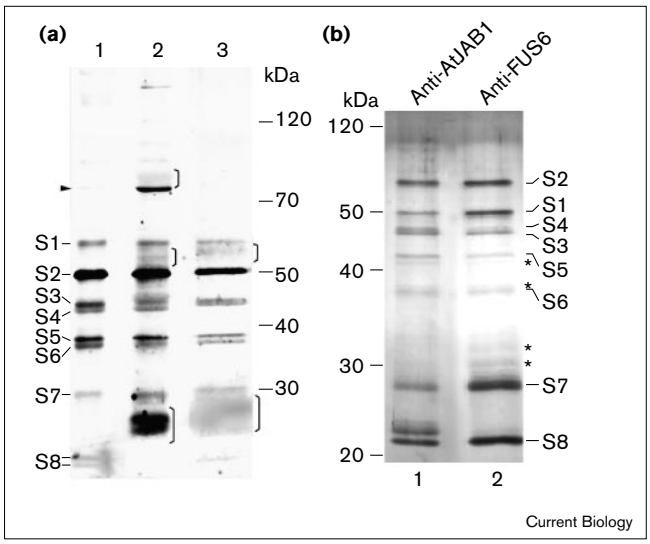
Published: **27 July 1998**

Current Biology 1998, 8:919–922
http://biomednet.com/elecref/0960982200800919

Results and discussion

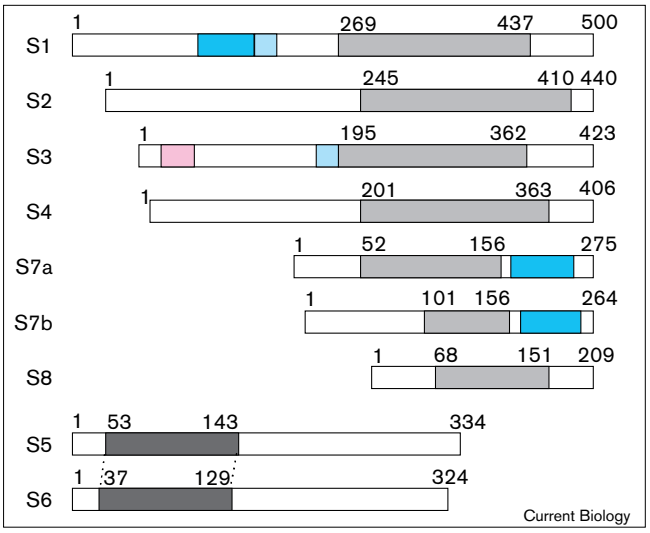
Human COP9 and the Gα pathway suppressor 1 (GPS1) are mammalian homologs of plant COP9 and FUS6, respectively [5,6]. Using an affinity column with anti-GPS1 antibodies, we purified the mammalian COP9 complex from pig spleen. This affinity purified complex had essentially the same protein composition pattern as that purified by a conventional biochemical procedure ([4]; Figure 1a, lanes 1,2). The same set of proteins could also be co-immunoprecipitated using anti-GPS1 antibodies on mouse brain (Figure 1a, lane 3), liver and spleen extracts (data not shown). The subunit composition of the COP9 complex is therefore identical not only between pigs and mice but also among the various organ types examined. We named the

Figure 1



Silver-staining pattern of the purified COP9 complexes. **(a)** Comparison of the silver-staining pattern showing subunit composition of the COP9 complex from different purification procedures. Lane 1, biochemically purified pig spleen COP9 complex [4]; lane 2, immunoaffinity purified complex from pig spleen; lane 3, co-immunoprecipitation from mouse brain extract. The eight mammalian COP9-complex subunits are: S1 (57 kDa), S2 (51 kDa), S3 (47 kDa), S4 (45 kDa), S5 (38 kDa), S6 (36 kDa), S7 (30 kDa) and S8 (22 kDa). The arrowhead denotes a 78 kDa immunoglobulin G (IgG)-binding protein, which is likely to be an artifact of the affinity procedure. The immunoglobulins co-eluted with the complex in lanes 2 and 3 are indicated by brackets. A negative-staining area around the 21–25 kDa region in lane 2 masked the S8 band, but the latter was readily detected by western analysis (data not shown). **(b)** The cauliflower COP9 complex was affinity purified using antibodies against either AJH1 (anti-AtJAB1, lane 1) or FUS6 (lane 2). Analysis of peptide sequences from these proteins revealed that the minor bands (asterisks) were mostly degradation products of the major bands. The additional protein bands above the S8 band in lane 1 were probably co-eluted immunoglobulin light chains. The cauliflower COP9-complex subunits differ in size from their mammalian counterparts.

Figure 2



Summary of the protein sequence features of the COP9-complex subunits. Each subunit is represented by a bar, above which the numbers denote the amino-acid number of the corresponding protein subunit. Light grey, PCI domain (PROSITE profile entry QDOC50250); dark grey, MPN domain (PROSITE profile entry QDOC50249); light blue, TPR domain (PROSITE profile entry QDOC50005); dark blue, coiled-coil structural domain; magenta, leucine-zipper motif. The bars are not drawn precisely to scale.

subunits S1 to S8 as shown in Figure 1a. In parallel, we purified the cauliflower COP9 complex by a similar immunoaffinity procedure, but with two different affinity columns bearing antibodies against two distinct subunits, FUS6 and AJH1. Figure 1b shows that both preparations contain nearly the same set of protein bands, confirming that these proteins, including FUS6 and AJH1, represent integral parts of the plant COP9 complex.

We obtained multiple peptide sequences for each of the subunits. This confirmed that the S8 doublet and S1 proteins were indeed pig homologs of COP9 and GPS1, respectively. It should be noted that GPS1 has been shown to inhibit JNK activity and repress AP-1 activity in mammalian cells [6]. Peptide sequences for the other subunits allowed identification of full-length murine cDNA clones (see Supplementary material available with this article on the internet).

S2 of the mammalian COP9 complex is a 51 kDa protein whose amino-terminal 234 amino-acid region is essentially identical to human Trip15, a protein that was isolated by its ability to interact with thyroid hormone receptor and retinoic acid receptor in a ligand-dependent manner in a yeast two-hybrid assay [7]. S2 has similarity to the 26S proteasome regulatory subunit 9 (p44.5; 25% identity) and to the mammary-tumor-associated protein INT6 (19% identity), also known as the p48 subunit of the human eIF3 complex. The carboxy-terminal half of the protein encompasses a PCI domain (Figure 2), which has been found in many components of the proteasome regulatory complex, the COP9 complex and the eIF3 complex [8]. In fact, six of the eight subunits of the COP9 complex contain a PCI domain, either as a full-length domain or as a half-domain (Figure 2). S3 contains a leucine zipper at the amino terminus, a single tetratricopeptide repeat (TPR) motif in the middle, and a large PCI domain at the carboxy-terminal half of the protein. S4 is also a novel PCI-domain-containing protein of about 45 kDa. Again, S3 and S4 proteins have overall similarity to 26S proteasome regulatory subunits p58 (19% identity) and p55 (22% identity), respectively. We also noted that S1, S2, S3 and S4 have some degree of sequence similarity among themselves.

S7a, S7b and S8 do not seem to have close proteasome relatives, neither do they show homology to other proteins in the database. The mouse S7a and S7b proteins have 58% sequence identity between each other, and both have a coiled-coil structure following the PCI domain (Figure 2). As there seems to be only one gene encoding subunit 7 in *Arabidopsis* (D. Chamovitz, personal communication), we believe that S7a and S7b are functional alternative forms encoded by a small gene family in mammals.

S5 is identical to the Jun-activation-domain-binding protein 1 (JAB1), a reported transcription coactivator of the c-Jun oncogene [9]. S5 is highly similar to POH (28% identity), a component of the 26S proteasome and the human homolog of *Schizosaccharomyces pombe* PAD1. S6 is essentially identical to the human Vpr-interacting protein (hVIP; which is involved in the G2–M phase transition of the mammalian cell cycle [10]) and is highly similar (23% identity) to MOV34, which is the p40 subunit of the proteasome regulatory complex. In addition, S5 and S6 also have significant similarity to the p40 and p47 subunits of human eIF3 and are members of the MOV34 family [11], which are characterized by an amino-terminal MPN domain (for MPR1 and PAD1 N termini) [8,11] that is found in a subset of components of the proteasome, COP9 complex and eIF3.

Sequencing of the cauliflower COP9 complex revealed that the plant COP9 complex also contains eight distinct subunits corresponding to those of the mammalian complex (Figure 1b). Homology comparison of the mammalian COP9 complex subunits with the cauliflower COP9 peptide sequences and other relevant sequences is presented in Table 1. Clearly, both the subunit composition of the COP9 complex and sequences of each of the subunits are conserved between mammals and higher plants. Furthermore, putative homologs of the COP9-complex subunits are present in other multicellular eukaryotes such as *Caenorhabditis elegans* and *Drosophila* (Table 1). Therefore, the COP9 complex is widely present among multicellular eukaryotic organisms. Interestingly,

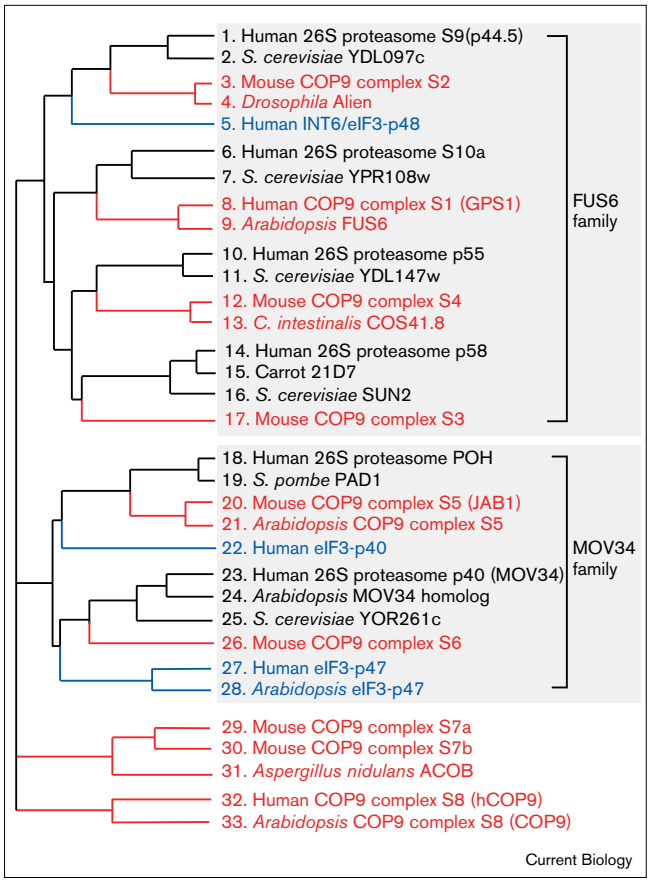
Table 1

Comparison of the mammalian COP9-complex subunits with those of other organisms.

Subunit	Mammalian protein	Sequence identity: Cauliflower homolog	*C. elegans* homolog	*Drosophila* homolog
S1	GPS1	64% (35)	–	49% (109)
S2	Trip15*	78% (37)	65% (420)	70% (Alien†)
S3	SGN3	68% (37)	–	47% (238)
S4	Novel	60% (45)	41% (129)	65% (257)
S5	JAB1	70% (81)	49%	–
S6	hVIP	50% (95)	–	–
S7a	Novel	38% (61)	–	46% (188)
S7b	Novel	38% (61)	–	–
S8	hCOP9	35% (23)	–	–

Pair-wise sequence comparison using the mammalian proteins as references. Sequences used for *C. elegans* and *Drosophila* are published expressed sequence tags (ESTs) or genomic sequences. The names of homologs, if known, are provided within brackets. In the case of proteins for which only partial sequences are available, the lengths of amino acids used in the sequence comparison are indicated within brackets. The lengths of the peptides from the cauliflower COP9-complex subunits used for sequence comparison are also indicated within the brackets. *Trip15 [7] is the amino-terminal 234 amino-acid fragment of S2 protein (the full-length protein is 440 amino-acids long). †Protein sequence from Genbank; see Figure 3 legend for accession number. A dash indicates that the sequence of the corresponding homolog is not available.

Figure 3



The homology relationships between the COP9-complex subunits and the related components of 19S proteasome regulator and eIF3. The lineage tree was constructed using MULTIALIN [12] and PIMA [13] algorithms. The lineages for proteasome components are in black. The COP9 complex branches are in red, and the eIF3 branches are in blue. The Genbank accession numbers for the sequences used are as follows: 1, AF001212; 2, 2131330; 3, AF071312; 4, U57758; 5, U54562; 6, human open reading frame (ORF) KIAA0101, D14663; 7, 2132285; 8, U20285; 9, L26498; 10, AB003103; 11, Z74195; 12, AF071314; 13, 1764018; 14, D67025; 15, D13434; 16, D78023; 17, AF071313; 18, U86782; 19, P41873; 20, U70736; 21, unpublished; 22, U54559; 23, D50063; 24, U54560; 25, Z75169; 26, AF071315; 27, U94855; 28, U54561; 29, AF071316; 30, AF071317; 31, 604427; 32, U51205; 33, L32874. We refer to the group of proteins including S1, S2, S3, S4 and their corresponding proteasome components and eIF3-p48/INT6 as the 'FUS6 family', after FUS6, the first gene characterized from this family. A number of genes from the FUS6 family have previously been described as 'COP11-like proteins' [5]. All members of the FUS6 family contain a PCI domain, but their sequence similarities extend to the entire length beyond the PCI domain itself. The second group include S5, S6 and their relatives in proteasome and eIF3 complexes. This group has been described as the 'MOV34 family' [11].

no close homologs of any COP9 complex subunits were found in the completely sequenced genome of *S. cerevisiae*, implying that this single-cell eukaryote probably does not possess the COP9 complex.

To further clarify relationships among components of the COP9 complex, the 19S proteasome regulator and the human eIF3 complex, all of the genes for the COP9-complex subunits and their similar genes from proteasome and eIF3 were analyzed by multi-sequence alignment algorithms [12,13] (Figure 3). With the exception of S7 and S8, each subunit of the COP9 complex (Figure 3, red) is most closely related to one of the proteasome non-ATPase regulatory subunits (Figure 3, black). Three COP9-complex subunits, S2, S5 and S6 also have close relatives in the eIF3 complex: p48, p40 and p47, respectively (Figure 3, blue). Both the proteasome and the COP9 complex are highly conserved evolutionarily, and greater than 40% amino-acid identity is usually observed between homologous genes across species. In contrast, the sequence identity between a particular COP9-complex subunit and its corresponding proteasome relative ranges from 19% to 28%. This suggests that the respective COP9-complex subunits from S1 to S6 are paralogs of the six distinct proteasome components, and that these six COP9-complex subunits appear to have diverged, in parallel, from the corresponding proteasome components to about the same distance. This phenomenon is best explained if the COP9 complex and a subcomplex of the 19S proteasome regulatory complex containing six

such subunits, were evolved from a common ancestor; in single-cell eukaryotes such as *S. cerevisiae*, the COP9 complex was lost because of adaptation to the simple life style. Alternatively, the COP9 complex might have simply originated from the proteasome regulatory complex during the emergence of multicellular organisms. It might then have further evolved to serve a new function(s) to meet the demands imposed by the developmental complexity of multicellular organisms. Taken together, our results suggest that the proteasome regulatory complex and the COP9 complex are distinct structural and functional entities, but share similarities in overall architecture and features of macro-molecular assembly.

The composition and subunit identities of the mammalian COP9 complex are nearly identical to the recently described JAB1-containing signalosome from humans [14]. The SGN3 subunit of the signalosome is identical to S3 of the COP9 complex. In addition, peptide sequencing of the signalosome has also yielded sequences of GPS1, human COP9, Trip15 and JAB1 [14]. Therefore, the other subunits described here, S4, S6, S7a and S7b, are also likely to be part of the signalosome. The composition of the COP9 complex and our understanding of its genetic function in *Arabidopsis* suggest that this complex is indispensable in multicellular organisms in mediating the responses to extracellular signals.

Materials and methods

Affinity purification and immunoprecipitation

The affinity column was made by coupling about 3 mg affinity purified anti-GPS1 antibodies to 1.5 ml protein-A–Sepharose beads (Sigma) and 0.3 ml protein-G–Sepharose beads (Sigma). Pig spleen (65 g; Pel-Freez Biologicals) was homogenized in 150 ml extraction buffer followed by 12% polyethylene glycol precipitation and Q–Sepharose ion-exchange chromotagraphy [4]. The COP9-complex-enriched fraction was loaded onto the affinity column. The column was washed with 25 mM Tris-HCl pH 7.2, 0.6 M NaCl and 0.12% Tween-20 before elution with 100 mM glycine pH 2.7. To remove the excessive co-eluted immunoglobulins, the eluent was adjusted to pH 8.8 and 250 mM NaCl, and was slowly run through a Hi-trap protein A column (Pharmacia) to absorb most of the immunoglobulins. All purification procedures were performed with the aid of an automatic FPLC system (Pharmacia) in a 4°C cold room. The silver staining was performed according to the manufacturer's instruction for Silver Stain Plus (BioRad). Affinity purification of the COP9 complex from cauliflower head was performed similarly with anti-COP11 or anti-AJH1 antibodies except that the supernatant fraction of the 8% polyethylene glycol cut was used. The small-scale immunoprecipitation was performed according to a published protocol [3]. The immunoprecipitation buffer was 0.3 × RIPA, and the wash buffer was 50 mM Tris-HCl pH 7.4, 500 mM NaCl, 0.02% NP-40 and 2% glycerol.

Supplementary material

Additional methodological details and a figure showing the predicted amino-acid sequences of the COP9-complex subunits are published with this paper on the internet.

Acknowledgements

We are grateful to T. Masaki for APl protease, S.F. Kwok for the *Arabidopsis* AJH1 antibody, C. Joseph for his assistance in DNA sequencing, and G. Wagner for discussion. We thank A. Galston and S.F. Kwok for critical reading of this manuscript. This work is supported by a grant from the Council for Tobacco Research to N.W., and a NSF Presidential Faculty Fellow award to X-W.D. The protein sequencing analysis is supported by Grant-in-Aid (No. 09251218) of the Plant Organ Plan from the Ministry of Culture, Science and Education of Japan and by a Research for the Future grant of Japan.

References

1. Wei N, Chamovitz DA, Deng X-W: ***Arabidopsis* COP9 is a component of a novel signaling complex mediating light control of development.** *Cell* 1994, **78:**117-124.
2. Chamovitz DA, Wei N, Osterlund MT, von Arnim AG, Staub JM, Matsui M, *et al.*: **The COP9 complex, a novel multisubunit nuclear regulator involved in light control of a plant developmental switch.** *Cell* 1996, **86:**115-121.
3. Staub JM, Wei N, Deng X-W: **Evidence for FUS6 as a component of the nuclear-localized COP9 complex in *Arabidopsis*.** *Plant Cell* 1996, **8:**2047-2056.
4. Wei N, Deng X-W: **Characterization and purification of the mammalian COP9 complex, a conserved nuclear regulator initially identified as a repressor of photomorphogenesis in higher plants.** *Photochem Photobiol* 1998, in press.
5. Chamovitz DA, Deng X-W: **The novel components of the *Arabidopsis* light signaling pathway may define a group of general developmental regulators shared by both animal and plant kingdoms.** *Cell* 1995, **82:**353-354.
6. Spain BH, Bowdish KS, Pacal AR, Fluckiger S, Koo D, Chang C-Y, *et al.*: **Two human cDNAs, including a homolog of *Arabidopsis* FUS6 (COP11), suppress G-protein and mitogen-activated protein kinase-mediated signal transduction in yeast and mammalian cells.** *Mol Cell Biol* 1996, **16:**6698-6706.
7. Lee JW, Choi H-S, Gyuris J, Brent R, Moore DD: **Two classes of protein dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor.** *Mol Endocrinol* 1995, **9:**243-253.
8. Hofmann K, Bucher P: **The PCI domain: a common theme in three multiprotein complexes.** *Trends Biochem Sci* 1998, **23:**204-205.
9. Claret F-X, Hibi M, Dhut S, Toda T, Karin M: **A new group of conserved coactivators that increase the specificity of AP-1 transcription factors**.*Nature* 1996, **383:**453-457.
10. Mahalingam S, Ayyavoo V, Patel M, Kieber-Emmons T, Kao GD, Muschel RJ, Weiner DB: **HIV-1 Vpr interacts with a human 34-kDa mov34 homologue, a cellular factor linked to the G2/M phase transition of the mammalian cell cycle.** *Proc Natl Acad Sci USA* 1998, **95:**3419-3424.
11. Asano K, Vornlocher H-P, Richter-Cook NJ, Merrick WC, Hinnebusch AG, Hershey JWB: **Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits**. *J Biol Chem* 1997, **272:**27042-27052.
12. Corpet F: **Multiple sequence alignment with hierarchical clustering**. *Nucleic Acids Res* 1988, **16:**10881-10890.
13. Smith RF, Smith TF: **Pattern-induced multi-sequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for comparative protein modeling.** *Protein Eng* 1992, **5:**35-41.
14. Seeger M, Kraft R, Ferrell K, Bech-Otschir D, Dumdey R, Schade R, *et al.*: **A novel protein complex involved in signal transduction possessing similarities to 26S proteasome subunits.** *FASEB J* 1998, **12:**469-478.

Supplementary material

The COP9 complex is conserved between plants and mammals and is related to the 26S proteasome regulatory complex

Ning Wei, Tomohiko Tsuge, Giovanna Serino, Naoshi Dohmae, Koji Takio, Minami Matsui and Xing-Wang Deng

Current Biology 1998, **8**:919–922

Material and methods

Protein digestion and peptide sequencing

Subunits of the purified complexes were separated on an 8%–16% gradient SDS–PAGE gel. Coomassie-stained bands were individually excised and treated with 0.2 μg *Achromobacter* protease I at 37°C for 12 h in 0.1 M Tris-HCl pH 9.0, 2 mM EDTA and 0.1% SDS. Peptides were separated on columns of DEAE-5PW (TOSOH, Japan) and Mightysil RP-18 (KANTO Reagents, Tokyo, Japan), connected in series to a model 1100M liquid chromatography system (Hewlett Packard). Peptides were eluted at a flow rate of 0.1 ml/min for 96 min using a linear gradient of 0%–60% of solvent B (solvent A, 0.09% (v/v) trifluoroacetic acid in water; solvent B, 0.075% (v/v) trifluoroacetic acid in 80% (v/v) acetonitrile). Selected peptides were subjected to Edman degradation using a model 477A automated protein sequencer connected to a model 120A PTH-Analyzer (PE Applied Biosystems). The obtained sequences were confirmed by comparing the calculated molecular weight to that measured from the mass spectrometry.

cDNA clones and DNA sequencing

The peptide sequences were used to search public EST database. Multiple EST sequences were used to construct an overlapping contig. For all of the subunits, mouse cDNA clones covering the complete coding region were obtained (Genome Systems, Inc.). We then completely sequenced each of the cDNAs using the fluorescence-based automatic sequencer. Two criteria were used to determine whether the complete coding region with the initiation methionine was within the cDNA clone. First, using the deduced amino-acid sequence length to match the size of the subunit on the SDS–PAGE gel. Second by comparing a homologous sequence from another organism. Usually the coding region has higher homology than the non-coding region.

Figure S1

Current Biology

The deduced amino-acid sequences of the COP9-complex subunits. The peptides generated from the purified pig COP9 complex correspond to the underlined murine or human sequences. The non-identical amino acids between pig peptides and murine or human proteins are indicated in lower case. S1 (GPS1; Genbank accession number U20285) and S8 (hCOP9; Genbank accession number U51205) are published human protein sequences. S2, S3, S4, S6, S7a and S7b are murine sequences (see Materials and methods); the Genbank accession numbers are AF071312, AF071313, AF071314, U70736, AF071315, AF071316 and AF071317, respectively.