METHODS **24,** 218–229 (2001)
doi:10.1006/meth.2001.1183, available online at http://www.idealibrary.com on IDEAL®

The Tandem Affinity Purification (TAP) Method: A General Procedure of Protein Complex Purification

Oscar Puig,[1] Friederike Caspary,[1] Guillaume Rigaut, Berthold Rutz, Emmanuelle Bouveret, Elisabeth Bragado-Nilsson, Matthias Wilm, and Bertrand Séraphin[2]

European Molecular Biology Laboratory Meyerhofstrasse 1, Heidelberg D-69117, Germany

Identification of components present in biological complexes requires their purification to near homogeneity. Methods of purification vary from protein to protein, making it impossible to design a general purification strategy valid for all cases. We have developed the tandem affinity purification (TAP) method as a tool that allows rapid purification under native conditions of complexes, even when expressed at their natural level. Prior knowledge of complex composition or function is not required. The TAP method requires fusion of the TAP tag, either N- or C-terminally, to the target protein of interest. Starting from a relatively small number of cells, active macromolecular complexes can be isolated and used for multiple applications. Variations of the method to specifically purify complexes containing two given components or to subtract undesired complexes can easily be implemented. The TAP method was initially developed in yeast but can be successfully adapted to various organisms. Its simplicity, high yield, and wide applicability make the TAP method a very useful procedure for protein purification and proteome exploration. © 2001 Academic Press

The sequencing of complete genomes of several organisms provides an exceptional opportunity to analyze the different functions governed by their genes. Insights into these complex biological systems can be gained by analysis of gene regulatory networks and by determining the identity, modification, and expression levels of encoded proteins as well as by defining interactions existing among proteins (proteomic analyses). Large-scale two-hybrid screening has been used for this latter purpose (1–3). However, false-positive and false-negative results, the lack of information about stoichiometry, and the limited set of conditions testable make it desirable to use additional strategies to easily detect protein interactions.

Biochemical purification of proteins in combination with mass spectrometry allows identification of interacting partners. This strategy is becoming an important tool to define relations existing among gene products (4, 5). Currently, ~100 fmol of a protein can be detected and identified by mass spectrometry, allowing rapid characterization of any protein present in a complex mixture, provided that the target complex is sufficiently purified in reasonable quantity. Identification of proteins by mass spectrometry is currently facilitated for several organisms by the availability of complete genomic sequences. The current limiting step in protein complex characterization appears therefore to be protein purification rather than protein identification. Each protein has unique properties, which can be exploited for its purification (6). This makes it, however, impossible to design a general purification strategy valid for all cases. A generic purification protocol is therefore desirable to allow routine and possibly automated protein complex purification for proteome analysis. The fusion of tags, peptides, or protein domains to protein targets appeared best suited toward this goal. After comparative testing of several tags, we have recently developed a new tag, the tandem affinity purification (TAP) tag, and we have optimized a procedure, the TAP method, for the native purification of protein complexes (7). This strategy allows for fast purification with high yield of protein complexes under standard conditions. Ultimately, the purified complex can be used for protein

[1] These authors contributed equally to this manuscript.

[2] To whom correspondence should be addressed at his present address at CGM-CNRS, Avenue de la Terrasse, F-91198 Gif sur Yvette, France. Fax: 33 1 69 82 38 77. E-mail: seraphin@cgm.cnrs-gif.fr.

1046-2023/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

identification, functional, or structural studies. Furthermore, variations on the original strategy, including the use of C- or N-terminal tags, the use of a split tag, and/or the use of a subtraction step can easily be developed. These various aspects are described below starting with a presentation of the basic TAP method from gene tagging to protein analysis. We also present several applications of the method and discuss different variations from the original protocol and potential problems. General guidelines useful for various organisms are given; however, as the TAP method was developed with yeast, emphasis is given to applications in this organism. Detailed protocols and latest developments can also be found on our web site (http://www.emblheidelberg.de/ExternalInfo/seraphin/TAP.html).

METHODS

1. Overview of the TAP Method and the TAP Tag

The TAP method involves the fusion of the TAP tag (see below) to the target protein and the introduction of the construct into the host cell or organism. For optimal results, it is preferable to maintain expression of the fusion protein at, or close to, its natural level. Indeed, overexpression of the protein often induces its association with nonnatural partners (heat shock proteins, proteasome; Ref. (8)). Cell extracts are prepared and the fusion protein as well as associated partners is recovered by two specific affinity purification/elution steps. The material recovered can be analyzed in several ways. For protein complex characterization, proteins are concentrated, and eventually fractionated on a denaturing gel, before identification by mass spectrometry. (Alternatively, Edman degradation or Western blot may be used.) Because the various TAP purification steps are performed in a gentle native manner, purified complexes may also be tested for their activities or used in structural analysis.

The TAP tag consists of two IgG binding domains of *Staphylococcus aureus* protein A (ProtA) and a calmodulin binding peptide (CBP) separated by a TEV protease cleavage site. Originally, a C-terminal TAP tag was described (7) (Fig. 1A). We have now also generated an N-terminal TAP tag (Fig. 1A, see below). Note that the relative order of the modules of the TAP tag are inversed in the two tags because the ProtA module needs to be located at the extreme N or C terminus of the fusion protein. Both affinity tags have been selected for highly efficient recovery of proteins present at low concentration. ProtA binds tightly to an IgG matrix, requiring the use of the TEV protease to elute material under native conditions (Fig. 1B). The eluate of this first affinity purification step is then incubated with calmodulin-coated beads in the presence of calcium. After washing, which removes contaminants and the TEV protease remaining after the first affinity selection, the bound material is released under mild conditions with EGTA (Fig. 1B). Optimized conditions have been developed for the generic use of the TAP strategy (see below). The TAP tag is, however, very tolerant to buffer conditions and changes can easily be implemented to optimize recovery of specific complexes.

2. Tagging the Target Protein with the TAP Tag

The choice of the strategy for fusing the TAP tag to the target protein depends on the methods available to introduce recombinant nucleic acids into the corresponding cell or organism. One should also keep in mind that strong overexpression of the target protein is not preferable except if one is interested in producing large amounts of this protein by itself. Indeed, protein overexpression may often lead to the formation of nonspecific and/or nonnatural protein interactions with host proteins (8). This should be avoided if one wants to identify the structure, composition, and/or activity of a complex. The TAP tag has been specifically designed to allow recovery of proteins expressed at their low natural levels. Usually, standard DNA cloning procedures can be used to introduce the N- or C-terminal TAP tag in-frame with the coding region of the protein of interest in an appropriate expression vector. For this purpose, unique restriction sites present upstream and downstream of the N- and C-terminal TAP cassettes are available (Fig. 2A). The recombinant vector can then be transiently or stably introduced into recipient cells or organisms. Optimally, the tagged construct should be used to replace the endogenous wild-type gene. However, depending on the organism analyzed, this is not always possible and often time consuming (e.g., construction of transgenic mice).

The high efficiency of homologous recombination in yeast bypasses the need to construct a plasmid to fuse the TAP tag to the protein of interest. Polymerase chain reaction (PCR) fragments can indeed be used to integrate the TAP tag directly in the genome (9, 10). We routinely prefer to use the C-terminal TAP tag for this purpose as this maintains expression of the target protein under the control of its natural promoter. However, some proteins undergo loss of function when a peptide is added to its C- terminus. While from our experience this is not very frequent (about 5% of fusions), it is worthwhile to introduce the TAP tag into both haploid and diploid cells in parallel to test this possibility. For

cases where problems are encountered with the C-terminal TAP tag, we have designed a strategy that allows genomic fusion of an N-terminal TAP tag to the protein of interest while maintaining its expression under control of the endogenous promoter (see variations of the TAP method below).

The two plasmids constructed in our laboratory to introduce the C-terminal TAP tag into the yeast genome differ by the presence of either a URA3 or a TRP1 marker from *Kluyveromyces lactis* adjacent to the TAP cassette (Fig. 2A, pBS1479 and pBS1539, respectively). Primers containing a region of similarity to the yeast genome (40–50 nt long) and a constant priming region (Fig. 2B) are synthesized. Primer A hybridizes at the 5′ end of the CBP coding sequence and primer B in the vector downstream of the selection marker. Primer A should be carefully designed such that the last C-terminal residue of the target protein gets fused in-frame to the TAP tag. These primers are used to amplify by PCR the TAP tag from plasmid pBS1479 or pBS1539. The PCR product is extracted with phenol/chloroform/isoamyl alcohol, precipitated, and used to transform haploid and diploid yeast cells (11, 12). Correct integration of the cassette is verified by PCR and/or Southern blot (13, 14). To check for expression of the tagged protein, Western blot is used. Briefly, the cellular pellets corresponding to 1.5 ml of cell culture are vortexed 3 × 30 s with 30 μl siliconized glass beads and 100 μl of SDS–PAGE loading buffer. Samples are boiled, vortexed once more, and loaded directly on an SDS–polyacrylamide gel. Western blots are developed with a peroxidase–antiperoxidase complex (PAP, Sigma P-2026) that detects ProtA. However, one should remember that this strategy might not be sufficiently sensitive if the target protein is expressed at a very low level.

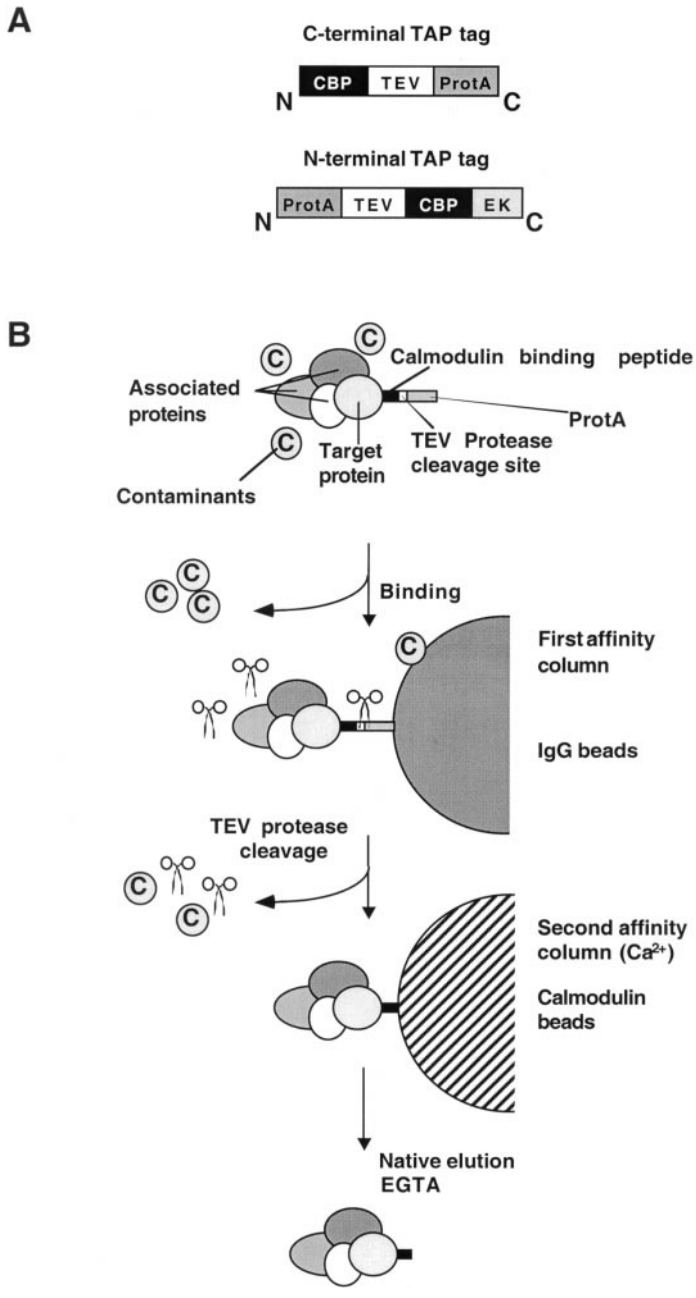

FIG. 1. (A) Schematic representation of the C- and N-terminal TAP tags. (B) Overview of the TAP purification strategy.

3. Extract preparation

Various extraction procedures can be used to prepare extracts from cells or organisms expressing the target protein fused to the TAP tag. The choice of the appropriate extract preparation procedure will depend on the target protein and on prior experience in the field that can be found in the literature. Cell fractionation and/or tissue dissection can facilitate purification by providing a preenrichment step or can be used to assay specifically protein complex composition in various tissues or cell compartments. In general, however, it is advisable to check, by detecting the ProtA moiety of the TAP tag by Western blot, whether extraction is efficient and if the TAP tag is not degraded under these specific conditions.

For yeast, we recommend the following standard procedure that has been extensively used in our laboratory. However, this method is unlikely to be optimal for all proteins and alternative protein extraction methods may be used (see variations in the purification protocol below). Extracts are routinely prepared from 2 liters of yeast cells grown to late log phase (OD_{600} ~2–3). Cell pellets are washed once with water and pelleted again in a 50-ml polypropylene tube (Falcon). The packed cell volume (PCV) is measured and the tube is frozen with liquid nitrogen. Frozen cell pellets may be stored at

−80°C. One PCV of Buffer A (10 mM K-Hepes pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol (DTT), 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 2 mM benzamidine, 1 μM leupeptin, 2 μM pepstatin A, 4 μM chymostatin, 2.6 μM aprotinin) is added to the pellet, which is rapidly thawed and kept at 4°C. All subsequent steps are performed at 4°C with precooled buffer and equipment. Cells are broken by passing them three times through a French press (Sim-Aminco) at a pressure of 8.27 MPa (1200 psi). KCl is adjusted to 0.2 M by addition of 2 M KCl (1/9th volume). The extract is centrifuged at 25,000*g* for 30 min and the supernatant is transferred into a new tube. The extract is then centrifuged at 100,000*g* for 1 h. After this centrifugation step three phases are visible in the tube: a lipidic phase floating on top, a pellet of cellular debris on the bottom, and a middle phase containing the extract. This last phase is recovered and dialyzed against buffer D (20 mM K-Hepes pH 7.9, 50 mM KCl, 0.2 mM EDTA pH 8.0, 0.5 mM DTT, 20 % glycerol, 0.5 mM PMSF, 2 mM benzamidine) for 3 h at 4°C. After dialysis, the extract is frozen and kept at −80°C.

4. TAP Purification

We perform all the binding and elution steps in 0.8 × 4-cm Poly-Prep columns (Bio-Rad, Hercules, CA) (Fig. 1). One hundred microliters of IgG Sepharose beads (Pharmacia Piscataway, NJ), corresponding to 200 μl of bead suspension, is transferred into the column. The beads are washed with 10 ml IPP150 (10 mM Tris–Cl, pH 8.0, 150 mM NaCl, 0.1% Nonidet (NP-40)). The composition of the extract buffer is adjusted to 10 mM Tris–Cl, pH 8.0, 100 mM NaCl, and 0.1% NP-40

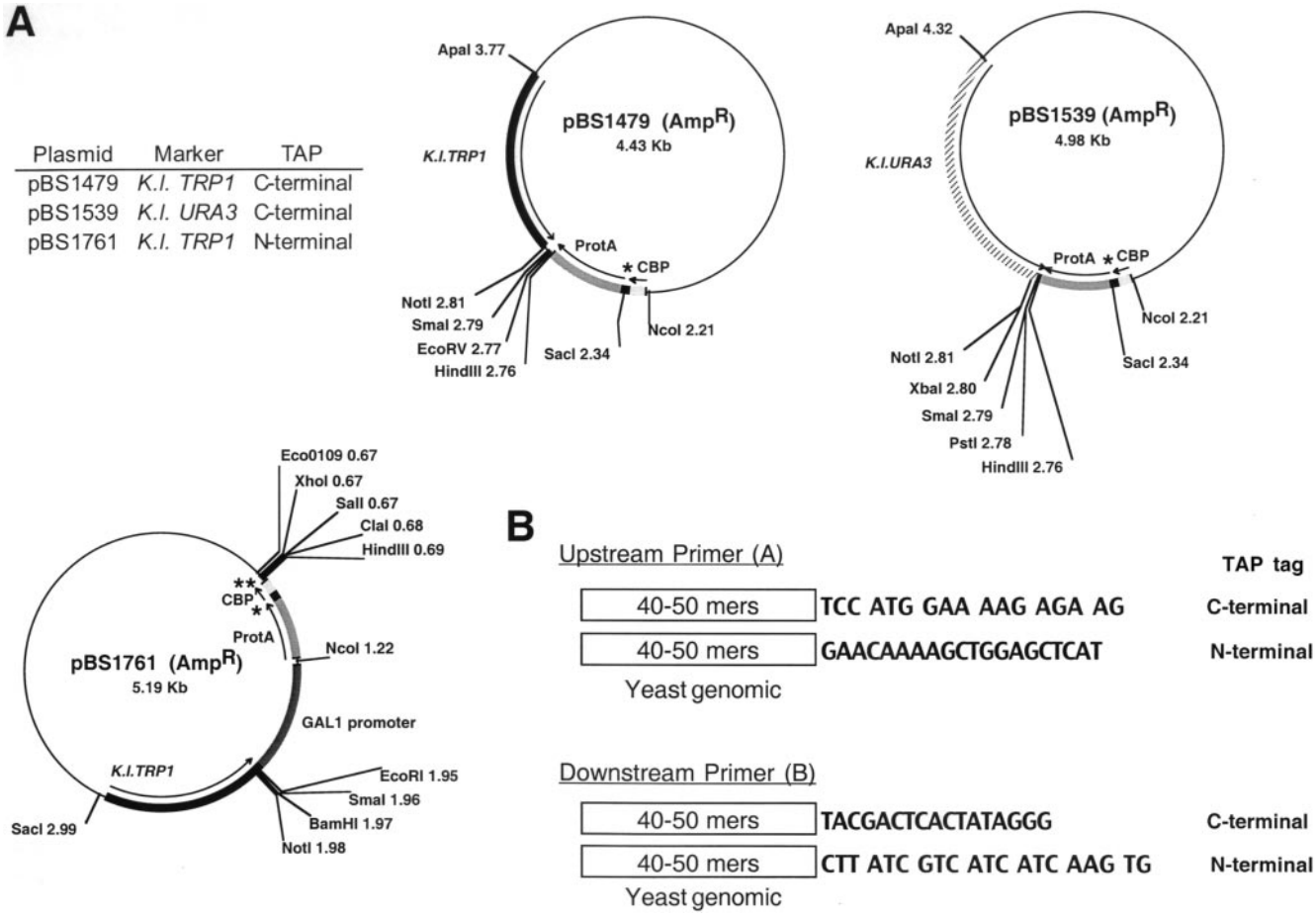

FIG. 2. (A) Plasmid maps including the C- and N-terminal TAP tagging cassettes. Single stars indicate TEV protease cleavage sites; two stars indicate the enterokinase cleavage site. (B) Structure of the oligonucleotides used for tagging. Rectangles represent 40–50 nucleotides of yeast genomic sequence flanking the integration site. Sequences of the constant priming sites are indicated (5′ to 3′). Primers should be carefully designed to fuse the target gene coding sequence in-frame with the TAP tag (upstream primer A for C-terminal tagging or downstream primer B for N terminal tagging). Frames are indicated for the appropriate oligonucleotides (note that this is the noncoding strand for the downstream primer B for N-terminal tagging).

(note that extract already contains 50 mM KCl). Then the extract is transferred into the column containing the washed beads and rotated for 2 h at 4°C.

Elution is done by gravity flow and the beads are washed three times with 10 ml of IPP150 and once with 10 ml of TEV cleavage buffer (IPP150 adjusted to 0.5 mM EDTA and 1 mM DTT). Cleavage is done in the same column by adding 1 ml of TEV cleavage buffer and 100 units of TEV protease (Gibco Ronkonkoma, NY). The beads are rotated for 2 h at 16°C and the eluate is recovered by gravity flow.

One hundred microliters of calmodulin beads (Stratagene, La Jolla, CA), corresponding to 200 μl of bead suspension, is transferred into a column and washed with 10 ml of IPP150 calmodulin binding buffer (10 mM Tris–Cl, pH 8.0, 10 mM 2-mercaptoethanol, 150 mM NaCl, 1 mM magnesium acetate, 1 mM imidazole, 2 mM $CaCl_2$, 0.1 % NP-40).

Three milliliters of IPP150 calmodulin binding buffer and 3 μl of 1 M $CaCl_2$ are added to the 1 ml of eluate recovered after TEV cleavage. This solution is then transferred to the column containing washed calmodulin beads and rotated for 1 h at 4°C. After the beads are washed with 30 ml of IPP150 calmodulin binding buffer, the bound proteins are eluted with 1 ml of IPP150 calmodulin elution buffer (10 mM Tris–Cl, pH 8.0, 10 mM 2-mercaptoethanol, 150 mM NaCl, 1 mM magnesium acetate, 1 mM imidazole, 0.1% NP-40, 2 mM EGTA). Five elution fractions of 200 μl each are collected. The elution peak is usually found in fractions 2 and 3.

It is noteworthy that the two purification steps of the TAP method can be performed in the reverse order. However, one should remember that, in this case, the final purified fraction remains contaminated with TEV protease.

5. Applications of the TAP Strategy: Analysis of Eluted Samples

The material recovered from TAP purification can be analyzed in several ways. A major application of the TAP method is to identify proteins interacting with the target protein. However, the TAP method can also be used to analyze the structure or the activity of the purified complex. Finally, the TAP method can also be used to purify recombinant proteins that are expressed at low level in yeast, bacteria, or other organisms.

5.1. Identification of Proteins Functionally Interacting with the Target Protein

To identify proteins interacting with the target protein it is often desirable to concentrate the eluate fractions before loading them on an analytical gel. We routinely use TCA precipitation (15) or a procedure essentially similar to that described by Wessel and Flügge (16). Proteins present in these fractions are

TABLE 1

Complexes Purified Using the TAP Method

Complex	Function	Protein tagged	Reference
U1 snRNP	Pre-mRNA splicing	Snu71-TAP	(7)
		Luc7p/Snu30p-TAP	(7)
		SmB-TEV-ProtA Nam8p-CBP	This article
		Nam8p-TAP	(23) and this article
U2 snRNP[a]	Pre-mRNA splicing	Lea1p-TAP	(18)
		SmB-TEV-ProtA Lea1p-CBP	(18)
"U6 snRNP"[b]	Pre-mRNA splicing	Lsm8p-TAP	(21)
CBC	Pre-mRNA splicing, nucleocytoplasmic RNA transport	Mud13p-TAP	(7)
BBP-associated	Pre-mRNA splicing, nuclear RNA retention	BBP-TAP	(22)
Mud2p-associated	Pre-mRNA splicing, nuclear RNA retention	Mud2p-TAP	(22)
SF3b	Pre-mRNA splicing	TAP-Rse1p	This article
RNases P/MRP	rRNA and tRNA processing	Pop4-TAP	This article
Dbp5p-associated	Nucleocytoplasmic mRNA transport	Dbp5p-TAP	(19)
Mex67p-associated	Nucleocytoplasmic mRNA transport	Mex67p-TAP	(20)
Mak3/10/31	Protein modification	Mak31-TAP	(7)
Lsm3p-associated	RNA degradation, pre-mRNA splicing	Lsm3p-TAP	(21)
LsmI complex	RNA degradation	Lsm3p-TAP Lsm8p-ProtA	(21)
Xrn1-associated	RNA degradation	Xrn1-TAP	(21)

[a] In this case not all the known components of the complex have been identified.
[b] Contains a mixture of U6, U4/U6, and U4/U6.U5 snRNPs; see Ref. (21).

separated on an exponential gradient SDS–polyacrylamide gel with acrylamide concentration ranging from 4 to 25% (top to bottom). After Coomassie or silver staining, bands are cut out and analyzed by mass spectrometry (4). Alternatively, mass spectrometry analysis could be performed directly on the concentrated eluate without prior fractionation of proteins by gel electrophoresis. However, information about the approximate stoichiometry of the various proteins present in the purified fraction would then not be available. In this vein, it is noteworthy that the TAP strategy was useful in identifying proteins interacting in stable complexes as well as more transiently interacting partners present in nonstoichiometric amounts (Table 1).

An example of the effectiveness of the TAP method for such application is shown in Fig. 3. Following the PCR strategy described above, the TAP tag was fused to the C terminus of the yeast Pop4 protein, a component of the RNase P and RNase MRP holoenzymes. The purified material obtained from 4 liters of yeast culture was fractionated by gel electrophoresis and stained with silver. The protein pattern obtained is similar to that obtained following a conventional biochemical purification of the yeast RNase P (17), except that degradation of the largest subunit (Pop1p) did not appear to occur with the TAP purification. Strikingly, however,

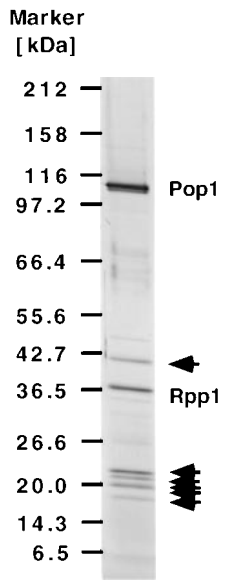

FIG. 3. Example of TAP purification: silver-stained gel depicting proteins recovered following purification of TAP-tagged Pop4p, a subunit of RNases P and MRP. The purification was done from 4 liters of yeast cell culture. The protein pattern can be compared with the pattern obtained following conventional purification of RNase P from 100 litres of cell culture (17). Proteins identified by mass spectrometry to confirm the identity of the complex are indicated. Additional subunits of this complex are indicated by arrowheads.

100 liters of yeast culture and specifically designed purification conditions were required for the conventional biochemical purification of RNase P (17) compared with 4 liters of culture and the standard 1-day purification for the TAP method. Mass spectrometry analysis verified the identity of the purified complex by confirming that two of the purified proteins correspond to Pop1p and Rpp1p. This result demonstrates the effectiveness of the TAP method.

This advantage of the TAP method has already been used to identify a new U2 snRNP-associated protein in yeast (18), a new subunit of the yeast U1 snRNP (7), as well as proteins associated with the Dbp5 RNA helicase, the Mex67 protein, and the Xrn1 protein involved in mRNA transport or RNA degradation (19–21). In addition, up to 24 proteins were identified in the fraction purified with the Lsm3-TAP fusion protein in our analysis of the Lsm proteins (21). The TAP method also revealed the subunit composition of the Mud2p/BPP splicing factor (22) and of the Mak3/Mak10/Mak31 complex involved in protein modification (7). These results indicate that the TAP method can be used to purify complexes involved in many cellular functions (pre-mRNA splicing, RNA transport and degradation, protein modification, etc.) and of various complexities (2–24 subunits, ~100–900 kDa). These results are summarized in Table 1.

The power of the TAP strategy to analyze protein interactions and complex composition can be further used to analyze the effect of mutations on protein association or complex assembly by comparing the patterns of proteins obtained after purification of both wild-type and mutant complexes. It is noteworthy that the TAP tag can be fused directly to the mutant protein, allowing the specific purification of the mutant complex in a wild-type background. An example of such study is provided by the analysis of the protein composition of the yeast U1 snRNP in a strain carrying a mutation of one of its subunits: Luc7p/Snu30p (23). Analysis of the TAP-purified material revealed that several proteins were missing in the mutant background, indicating a dramatic effect of this mutation on the structure of the particle (23).

While the TAP strategy is targeted primarily at identification of interacting proteins it should be remembered that it can also be used to identify ligands (nucleic acids, lipids, peptides, etc.) that interact, directly or indirectly, with the target protein. This is demonstrated by the use of primer extension and Northern blotting to detect specifically the U snRNAs copurified with snRNPs. Using appropriate assays, other ligands interacting with purified complexes could similarly be detected.

5.2. Activity Tests of the Purified Complex

Since the TAP purification is performed under gentle native conditions, the purified material can be used for *in vitro* activity tests. Note, however, that this may not be the case if EGTA or other reagents used in the purification interfere with the integrity or activity of the complex. Several macromolecular complexes obtained by TAP purification have been shown to be active. This includes the specific RNA binding activity of yeast CBC (7) and the ability of yeast U1 snRNP to form splicing complexes (O.P. and B.S., unpublished data). Activity can be analyzed directly after elution. However, for some applications a prior dialysis against buffer containing 10–20% polyethylene glycol (PEG) 20,000 greatly improved activity by concentrating the sample in native conditions. The TAP tag could therefore be used to analyze protein complex function at the genomic scale (24).

5.3. Structural Studies of Purified Complexes

The material recovered from TAP purification can also be used for structural studies using electron microscopy, providing that the purified complex is stable, sufficiently large, and concentrated. Successful analysis of the yeast U1 snRNP following this strategy has been achieved (O.P., K. Leonard, and B.S., unpublished data). In this case again, dialysis against buffer containing 10–20% PEG 20,000 was used to concentrate the sample before electron microscopy analysis.

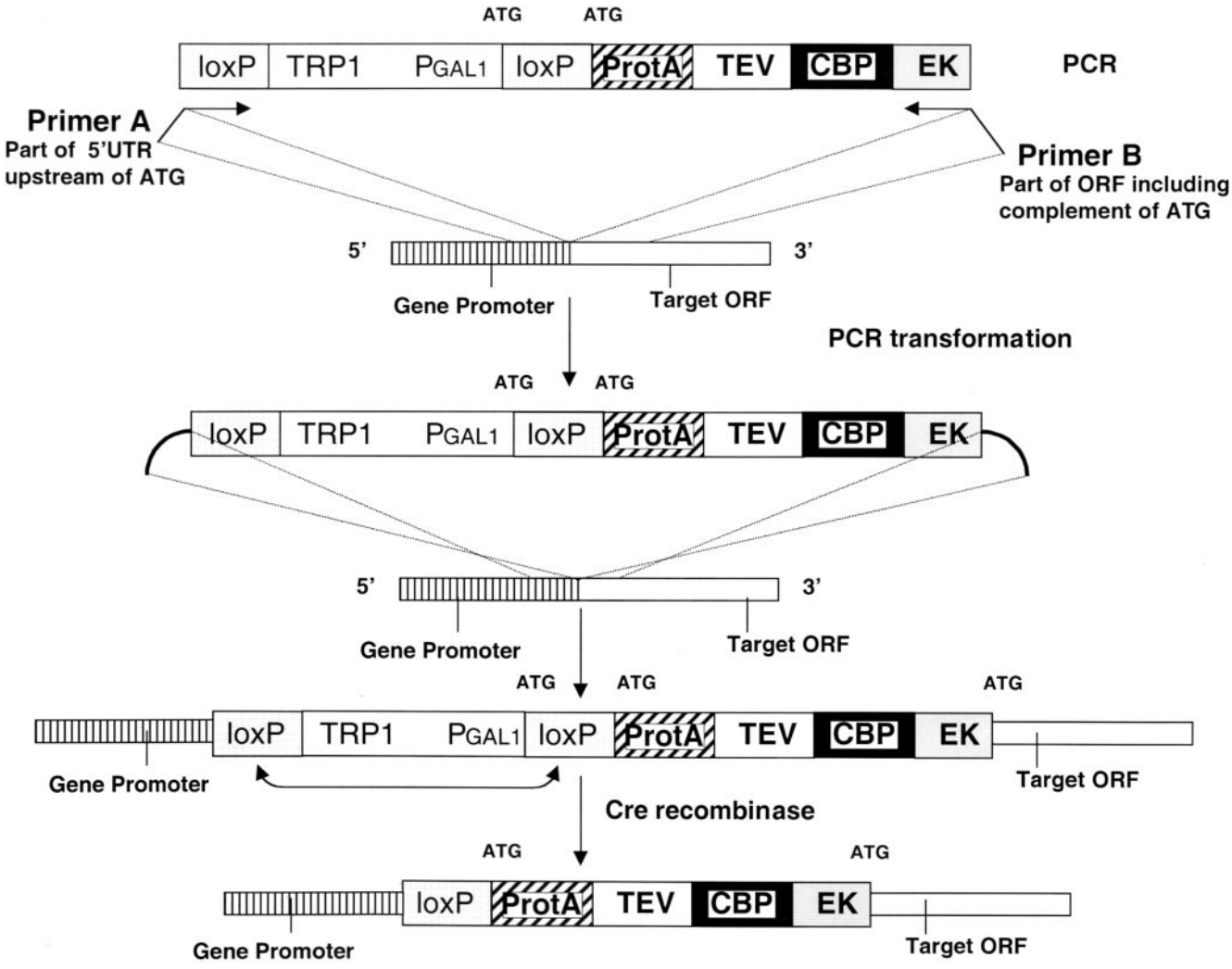

FIG. 4. N-terminal tagging strategy. A PCR fragment, is amplified including the tagging cassette and flanking regions of homology to the target gene. Following transformation into yeast cells, the PCR fragment integrates into the genome, placing the target ORF under the control of the *GAL1* promoter. In the final step, Cre recombinase is used to remove the marker and the *GAL1* promoter, leaving the N-terminal TAP-tagged ORF under the control of its natural promoter.

6. Versatility of the TAP Method

Several variations of the original TAP method have been tested successfully by others and us. These variations increase the number of applications that the TAP method offers and provide examples of the versatility of this procedure. They can be divided into variations affecting the tag system and variations of the purification protocol.

6.1. Variations of the Tag

6.1.1. N-Terminal Tag

Sometimes addition of a C-terminal tag to a protein impairs its function, producing a growth defect or even killing the cells if the protein is essential and the level of activity obtained is insufficient to keep viability.

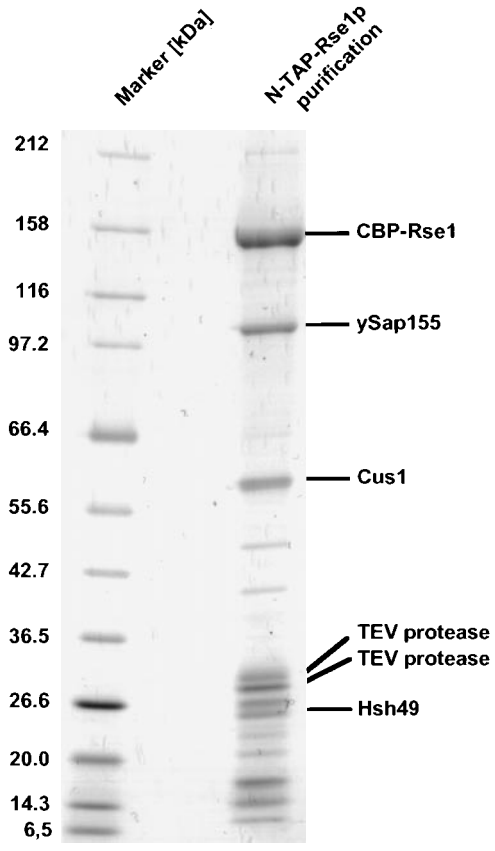

FIG. 5. Purification of proteins associated with Rse1p by using an N-terminal TAP tag. A Coomassie blue-stained gel presenting proteins recovered from purification of the N-terminal TAP-tagged Rse1 is depicted. The left lane represents molecular weight markers. The known components of the yeast SF3b complex identified by mass spectrometry as well as the TEV protease are labeled on the right. Additional bands present in the purified fraction may represent contaminants or new subunits of the yeast SF3b complex and are currently under analysis.

While we have found that this is not frequent (~5% of the protein tested), a remedy has to be found for this problem. A simple alternative solution is to add the tag at the N terminus of the target protein. Indeed, we have observed three proteins for which C-terminal tag fusion was not functional while N-terminal fusion was. We have therefore built an N-terminal TAP tag (Fig. 1). This tag contains the same modules as the C-terminal TAP tag in reverse order. Indeed, as the ProtA module needs to be cleaved off during the first affinity purification step, it must be located at the extreme terminus of the protein. In addition an enterokinase (EK) cleavage site has been introduced downstream of the CBP, allowing the complete removal of tag residues from the tagged protein (Fig. 1A). Convenient restriction sites are present upstream and downstream of the N-terminal tag, allowing the insertion upstream of the target protein by conventional cloning (Fig. 2A).

The PCR-based strategy described above for genomic insertion of the C-terminal TAP tag in yeast cannot be used for the N-terminal TAP tag because of the need to introduce a promoter between the selection marker and the N-terminal TAP tag (Fig. 4). Substitution of the natural promoter by an exogenous one could lead to overexpression of the protein and its accumulation in nonphysiological complexes (8). To avoid this problem we have developed a promoter switching method that allows fusion of the TAP tag to the N terminus of the target protein while maintaining it under the control of its natural promoter. The tagging cassette, cloned in plasmid pBS1761, is composed of the *Kluyveromyces lactis* TRP1 selection marker, the yeast *GAL1* promoter, and the N-TAP tag (Figs. 2 and 4). LoxP sites are introduced N-terminal to the selection marker and between the *GAL1* promoter and the TAP tag, respectively. The cassette is amplified using a primer that anneals upstream of the LoxP site and the selection marker (primer A) and that contains an extension of ~50 nt of the region immediately upstream of (but excluding) the start codon (Figs. 2B and 4). The second primer (primer B) anneals downstream of the cassette and contains an extension including the start codon and the first ~50 nt of the coding sequence (Figs. 2B and 4). Primer B is carefully designed to fuse in-frame the N-TAP tag and the target protein. The PCR is performed as described previously (10) but using 54°C as annealing temperature and an extension time of 2 min 30 s at 72°C. The PCR product is extracted with phenol/chloroform/isoamyl alcohol, precipitated, and transformed into a yeast strain lacking the *GAL1* promoter (YDL401 (25)). The use of this strain is necessary to avoid integration of the PCR fragment at the endogenous *GAL1* promoter. Transformants are selected on Trp^- selective plates (containing 2% galactose, 2% raffinose, 2% sucrose, and

0.05% glucose). It is important to remember that transformants grow slowly on this medium. Transformants are subcloned before being tested for correct integration and expression of the tagged protein. The selected strain contains the TAP-tagged protein under the control of the *GAL1* promoter and should therefore be grown in galactose-containing medium. This strain could, in principle, be used to purify the overexpressed target protein. The *GAL1* promoter together with the selection marker is then removed by taking advantage of the site-specific recombination activity of Cre recombinase. This enzyme will induce recombination between both LoxP sites, popping out the selection marker and the *GAL1* promoter, leaving one LoxP site and the TAP tag inserted in the genome (Fig. 4). We have constructed the LEU2-marked plasmids pBS1776 and pBS1777 (2μ and centromeric, respectively) to express Cre recombinase in yeast cells. One of these vectors is introduced into yeast cells by transformation with selection on galactose-containing medium. Transformants are then grown in rich medium and tested for the loss of the marker associated with the *GAL1* promoter and of the Cre-expressing plasmid. This step is highly efficient. In the resulting strains, the endogenous promoter of the target protein is now directing expression of the N-terminal TAP-tagged fusion.

The applicability of the N-terminal TAP tag is shown in the next example. Rse1p is a component of the SF3b complex associated with the yeast U2 snRNP (18). C-Terminal tagging impaired protein function, resulting in a thermosensitive phenotype and very slow growth at 30°C. To detect interaction partners of this protein, N-terminal tagging of Rse1p was carried out using the method described above. Figure 5 shows the pattern of proteins obtained after purification. Known components of the yeast SF3b complex, like ySAP155, Cus1p, and Hsh49p, are specifically identified among proteins recovered with Rse1p. This result demonstrates that N-terminal TAP tagging allows purification of Rse1p together with other proteins of the SF3b complex, overcoming the problem caused by the C-terminal addition of the TAP tag.

6.1.2. The Split Tag

A second variation of the original TAP method consists of the addition of the two functional halves of the TAP tag to two different proteins of the same complex. Henceforth, the ProtA together with the TEV protease cleavage site is fused to one protein while CBP is fused to a second target (Fig. 6). This strategy allows the purification of protein complexes containing two given proteins even when only a small fraction of the target proteins is associated, e.g., when large fractions remain free or bound to other complexes. Under certain conditions, this strategy allows therefore the purification of complexes even if they do not contain a specific subunit. This strategy proved useful to characterize the yeast U1 and U2 snRNP (18; O.P. and B. S., unpublished result, see Table 1). For the split tag strategy, plasmid pBS1479 or pBS1539 (Fig. 2A) is used as the source of the ProtA/TEV protease cleavage site cassette to tag one of the target proteins while plasmid pBS1512 (18) is used to tag the second protein with only CBP.

6.1.3. The Subtraction Method

This strategy is useful when two (or more) complexes share a common subunit but only one of these complexes is of interest. A protein specific for the undesired complexes is fused to ProtA (without a TEV cleavage site). While both complexes are retained on the first IgG affinity column, the undesired complex cannot be eluted as the ProtA-tagged specific subunit remains attached to the solid support (Fig. 7). Therefore, only the target complex is eluted and purified during the second step. This strategy was successfully used to isolate and characterize a complex that functions in RNA degradation, demonstrating that it is possible to specifically "subtract" a complex from the mixture during the first step of the TAP purification (21). For the substraction strategy, plasmids pBS 1173 and pBS1365 (10) are used to tag the substracted protein with a ProtA cassette lacking a TEV protease cleavage site.

6.2. Variations in the Purification Protocol

Depending on the aims of each experiment, the TAP purification can be modified to improve the yield or shorten the processing time due to the high tolerance of the TAP tag for buffer conditions. For example, it is possible to skip the final dialysis step in the preparation of extracts before freezing if glycerol is added to a final concentration of 10% right after the last centrifugation step. Extracts can be also prepared with glass beads and a bead beater (26) or other standard methods (27). In our hands, any of the methods used for extract preparation can be scaled up without affecting the results.

When the target protein is not soluble under the conditions described, increasing the salt concentration or adding a nonionic detergent (1% Triton X-100 or 1% NP-40) might help to solubilize it (G. Rigaut, unpublished results; M. Aldea, personal communication). The IgG–ProtA interaction has very high affinity and specificity and tolerates many environments, resisting up to 500 mM NaCl and even low concentrations of SDS (0.1%). Under these conditions a major fraction of the tagged

protein still remains bound to the beads. Interactions with other partners might, however, be disrupted.

7. Limitations and Troubleshooting

A problem intrinsic to the TAP strategy and any tagging method is the possibility that a tag added to a protein might not be sufficiently exposed to allow binding of the protein to the affinity beads or might affect protein function. A second problem is that the tag can affect protein expression levels. As indicated above, changing the location of the tag (N or C terminus) often helps to solve the problem. Another problem that might be encountered is the cleavage of the target protein or of an associated protein by the TEV protease. This is unlikely to be frequent given the high specificity of the TEV protease (28). Indeed, we have not yet encountered this situation with yeast proteins. Consistently, database searches suggest that only a very limited number of cellular proteins are cleaved by the TEV protease. Extract preparation must also maintain the TAP tag structure. We have observed that some enzyme preparations used to break the yeast cell wall, like zymolyase, interfere negatively with the procedure by specifically removing the tag from the protein, most likely through the action of contaminating proteases (data not shown).

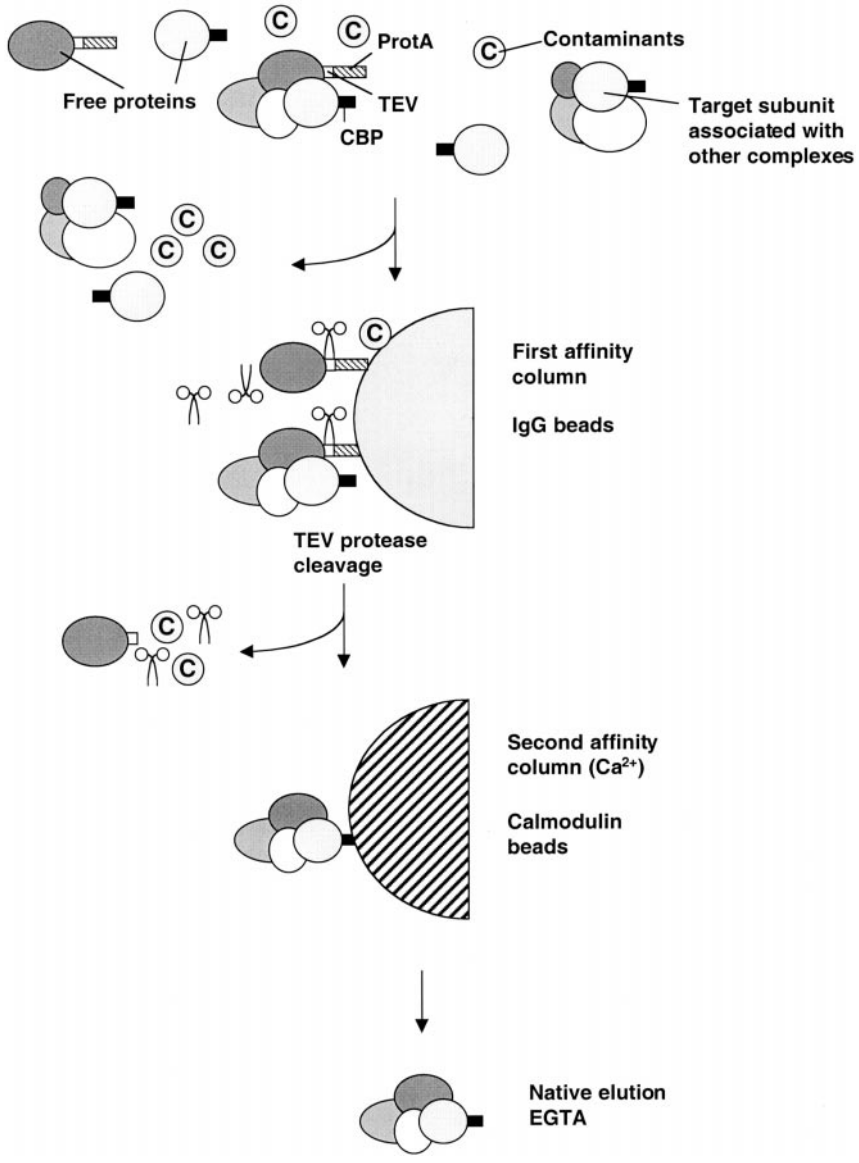

FIG. 6. Schematic representation of the split TAP tag strategy.

It is recommended that independent purifications in parallel be performed, instead of a single large TAP purification if the purification has to be scaled up. For example, five columns each with 200 μl of bead suspension gave better results than a large column with 1 ml of bead suspension. We have noted that the background is strongly augmented when more than 200 μl of bead suspension is used in the same column, even if the size of the column is increased.

A frequent concern is related to the possibility that endogenous calmodulin could bind to the CBP moiety of the TAP tag and prevent interaction with the calmodulin column during the second affinity purification step (Fig. 1B). This can be prevented by adding EGTA during the first affinity purification step which should release any bound calmodulin. While EGTA addition may be useful in some cases, comparative analyses indicate that this step was found to be completely dispensable for purification from yeast extracts.

CONCLUDING REMARKS

Understanding protein function is a major goal in biology. With the availability of full genome sequences,

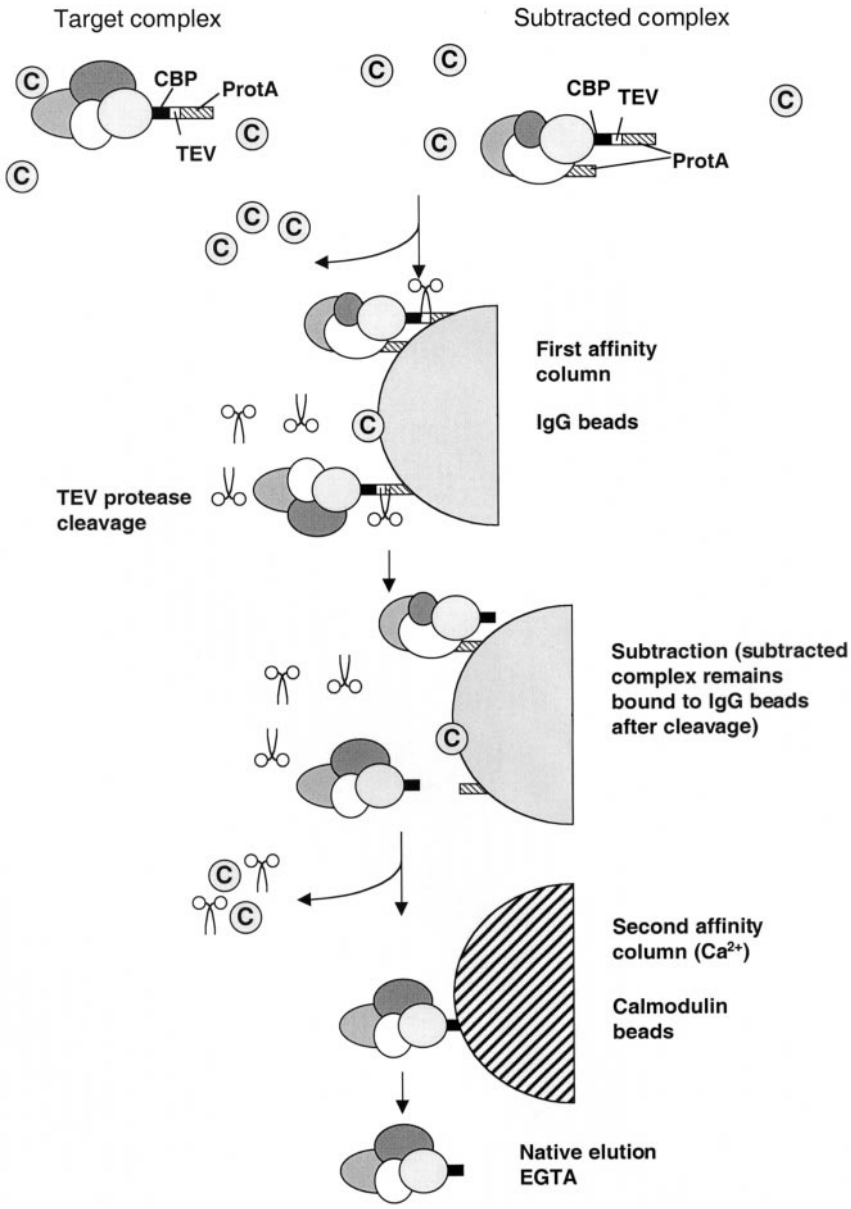

FIG. 7. Schematic representation of the subtraction strategy.

determining which macromolecules interact with a given protein is becoming a limiting step in analyzing protein function. The TAP method has proven to be a very useful tool for the detection of interacting partners of a target protein and for determining protein composition of macromolecular complexes with low levels of false positives and false negatives (Table 1). It has been also used to obtain active complexes for *in vitro* studies.

A feature of the TAP method is the possibility of automation. PCRs, selections of clones, and growth of cultures could be carried out by robots (3, 24) Similarly, because of their generic nature, purification reactions could be automated. In this way large-scale analysis of proteomes is technically feasible; the TAP strategy could therefore become a major tool for proteome exploration. Combined with automated mass spectrometry analysis this would considerably increase the amount of data available on protein interactions. This should ultimately lead to the establishment of protein interaction networks involved in cell function. Integration of data obtained from this and other approaches should help us to unravel how the information contained in the genome leads to the formation and function of living organisms.

REFERENCES

1. Fromont-Racine, M., Rain, J. C., and Legrain, P. (1997) *Nat. Genet.* 16, 277.
2. Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S., and Sakaki, Y. (2000) *Proc. Natl. Acad. Sci. USA* 97, 1143.
3. Uetz, P., Giot, L., Cagney, G., *et al.* (2000) *Nature* 403, 623.
4. Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., and Mann, M. (1996) *Proc. Natl. Acad. Sci. USA* 93, 14440.
5. Blackstock, W. P., and Weir, M. P. (1999) *Trends Biotechnol.* 17, 121.
6. Deutscher, M. P. (1990) *Methods Enzymol.* 182.
7. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. *Nat. Biotechnol.* (1999) 17, 1030.
8. Swaffield, J. C., Melcher, K., and Johnston, S. A. (1996) *Nature* 379, 658.
9. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., and Cullin, C. (1993) *Nucleic Acids Res.* 21, 3329.
10. Puig, O., Rutz, B., Luukkonen, B. G., Kandels-Lewis, S., Bragado-Nilsson, E., and Séraphin, B. (1998) *Yeast.* 14, 1139.
11. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* 153, 163.
12. Soni, R., Carmichael, J. P., and Murray, J. A. (1993) *Curr. Genet.* 24, 455.
13. Ward, A. C. (1992) *Biotechniques.* 13, 350.
14. Seraphin, B., Simon, M., and Faye, G. (1987) *J. Biol. Chem.* 262, 10146.
15. Ozols, J. (1990) *Methods Enzymol.* 182, 587.
16. Wessel, D., Flügge, U. I. (1984) *Anal Biochem.* 138, 141.
17. Chamberlain, J. R., Lee, Y., Lane, W. S., and Engelke, D. R. (1998) *Genes Dev* 12, 1678.
18. Caspary, F., Shevchenko, A., Wilm, M., Séraphin, B. (1999) *EMBO J.* 18, 3463.
19. Schmitt, C., von Kobbe, C., Bachi, A., Pante, N., Rodrigues, J. P., Boscheron, C, Rigaut, G., Wilm, M., Séraphin, B., Carmo-Fonseca, M., and Izaurralde, E. (1999) *EMBO J.* 18, 4332.
20. Stutz, F., Bachi, A., Doerks, T., Braun, I. C., Séraphin, B., Wilm, M., Bork, P., and Izaurralde, E. (2000) *RNA* 6, 638.
21. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000) *EMBO J.* 19, 1661.
22. Rutz, B. (2000) Thesis, Freie Universität Berlin.
23. Fortes, P., Bilbao-Cortes, D., Fornerod, M., Rigaut, G., Raymond, W., Séraphin, B., and Mattaj, I. W. (1999) *Genes Dev.* 13, 2425.
24. McCraith, S. M., Spinelli, S. L., Torres, F. M., Fields, S., Grayhack, E. J., and Phizicky, E. M. (1999) *Science* 286, 1153.
25. Lafontaine, D., and Tollervey, D. (1996) *Nucleic Acids Res.* 24, 3469.
26. Logie, C., and Peterson, C. L. (1999) *Methods Enzymol.* 304, 726.
27. Jazwinski, S. M. (1990) *Methods Enzymol.* 182, 154.
28. Dougherty, W. G., Cary, S. M., and Parks, T. D. (1989) *Virology* 171, 356.