A generic protein purification method for protein complex characterization and proteome exploration

*Guillaume Rigaut, Anna Shevchenko, Berthold Rutz, Matthias Wilm, Matthias Mann[1], and Bertrand Séraphin**

We have developed a generic procedure to purify proteins expressed at their natural level under native conditions using a novel tandem affinity purification (TAP) tag. The TAP tag allows the rapid purification of complexes from a relatively small number of cells without prior knowledge of the complex composition, activity, or function. Combined with mass spectrometry, the TAP strategy allows for the identification of proteins interacting with a given target protein. The TAP method has been tested in yeast but should be applicable to other cells or organisms.

Proteome analysis, in particular using mass spectrometry (MS), requires fast and reliable methods of protein purification. The method most suitable for standardization is affinity purification based on the fusion of a tag, usually a peptide or small protein, to the target protein. However, protein overexpression is not possible for heteromeric complexes of unknown composition and may also lead to the assembly of overexpressed proteins in non-physiological complexes. Protein complex purification therefore requires expression of the target protein at, or close to, its natural expression levels. Thus, a combination of high-affinity tags will be required for purification.

Identification of high-affinity tags

We selected the yeast SmX4p protein[1] to screen for high-affinity tags because one can assay the copurification of the associated U6 snRNA by primer extension, an analysis that is both semiquantitative and dependent on complex integrity. We tested the FLAG tag, two IgG-binding units of protein A of *Staphylococcus aureus* (ProtA), the Strep tag, the His tag, the calmodulin-binding peptide (CBP), and the chitin-binding domain (CBD) (reviewed in refs 2,3). None of the tags impaired protein function, but only the ProtA and CBP tags allowed efficient recovery (roughly 80% and 50%, respectively) of a fusion protein present at low concentration in a complex mixture.

The TAP tag and method

The CBP tag allows for efficient selection and specific release from the affinity column under mild conditions. In contrast, ProtA can only be released from matrix-bound IgG under denaturing conditions at low pH. We therefore inserted a specific TEV protease recognition sequence[4], which has been shown to allow proteolytic release of the bound material under native conditions[5,6], upstream of the ProtA tag. A fusion cassette encoding CBP, a TEV cleavage site, and ProtA (Fig. 1A) was constructed and named the tandem affinity purification (TAP) tag.

The TAP method (Fig. 1B) involves the fusion of the TAP tag to the target protein and the introduction of the construct into the host cell or organism, maintaining the expression of the fusion protein at, or close to, its natural level. The fusion protein and associated components are recovered from cell extracts by affinity selection on an IgG matrix. After washing, the TEV protease is added to release the bound material. The eluate is incubated with calmodulin-coated beads in the presence of calcium. This second affinity step is required to remove the TEV protease as well as traces of contaminants remaining after the first affinity selection. After washing, the bound material is released with EGTA.

Testing the TAP strategy identifies a new U1 snRNP subunit

We tested the TAP method by targeting the

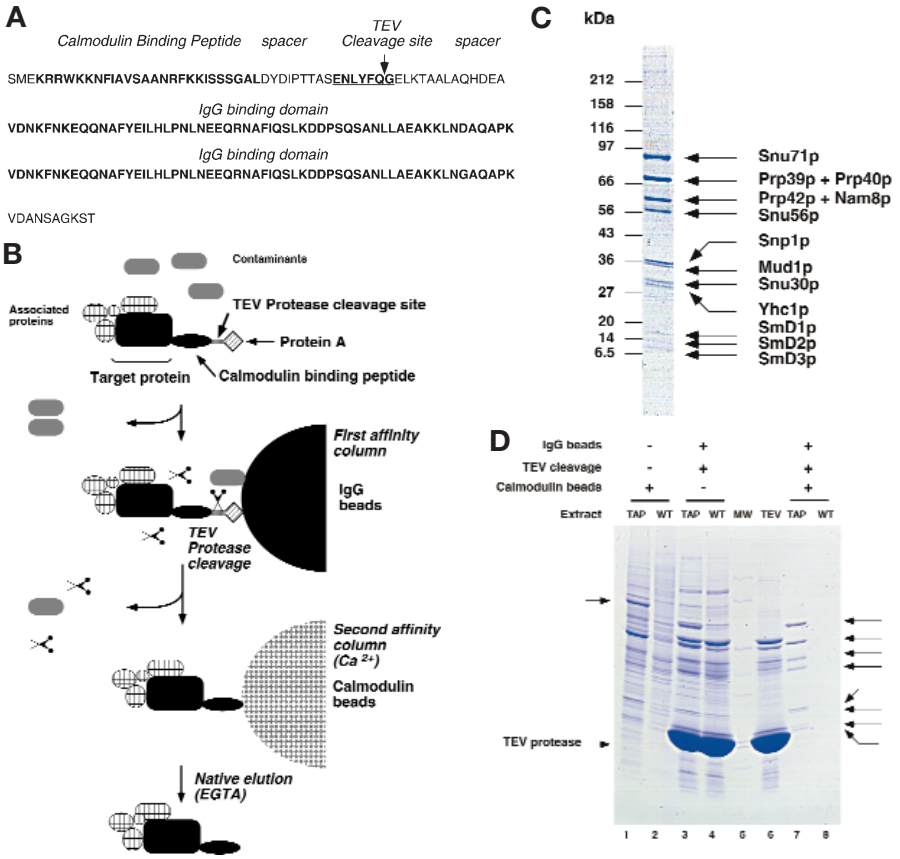

Figure 1. The TAP strategy: rationale and testing. (A) Sequence and structure of the TAP tag. The various domains constituting the TAP tag are indicated. (B) Overview of the TAP procedure. (C) Protein composition of TAP-purified U1 snRNP. Faster migrating bands potentially corresponding to the SmE, SmF, and/or SmG proteins were detected by silver staining but not analyzed. (D) Step-by-step analysis of the TAP strategy. Proteins present in the final TAP fraction (lanes 7 and 8), or present after each of the single affinity purification steps (lanes 1–4), were analyzed. Snu71-TAP (lanes 1, 3, and 7) or wild-type extracts (lanes 2, 4, and 8) were used. Lane 5: molecular weight marker. Lane 6: an amount of TEV protease identical to the amount used to elute proteins bound to IgG beads (lanes 2, 3, 7, and 8). Right arrows indicate the U1 snRNP-specific proteins including the tagged Snu71p after TEV cleavage; the arrow on the left indicates the Snu71p protein fused to the TAP tag before TEV cleavage (see Fig. 1C).

*The authors are at the European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [1]Present address: University of Southern Denmark, Campusvej 55, DK-5230 Odense, Denmark. *Corresponding author (seraphin@embl-heidelberg.de).*

yeast U1 snRNP, a previously characterized multisubunit complex[7] of relatively low abundance[8]. A DNA cassette encoding the TAP tag was integrated into the genome of a haploid cell in frame with the U1 snRNP protein Snu71p so as to maintain its natural level of expression. We routinely recovered sufficient amounts of the yeast U1 snRNP from 2 L of culture to detect 11 protein bands by Coomassie staining (Fig. 1C). The purification was more efficient and the purified fraction cleaner than in previous experiments using specific antibody-mediated affinity purification and 16 L of yeast culture[7].

The identity of the various proteins was determined by MS[9,10]. This confirmed the presence of all known U1 snRNP-specific subunits as well as some Sm proteins. The Prp42p protein, previously identified only in a strain lacking the comigrating Nam8p protein[7], was clearly detected in our purified fraction. Furthermore, one of the proteins, Snu30p/YDL087C, did not correspond to any previously known yeast U1 snRNP subunit. The presence of putative zinc fingers, the sequence of its metazoan homologs, and the fact that Snu30p is encoded by an essential gene (E. Bragado-Nilson and B.S., unpublished; Fortes et al.[11]) strongly suggested that Snu30p is a splicing factor. Purification of the yeast U1 snRNP with a TAP-tagged Snu30p confirmed that Snu30p is a bona fide U1 snRNP subunit.

A two-step procedure is required

To investigate the requirement for both steps of the TAP strategy, we purified the U1 snRNP starting from extracts containing the Snu71-TAP fusion either by a single CBP affinity step or only by ProtA-based selection and TEV protease-mediated release. The profiles of contaminating proteins were determined using nontagged extracts. While several yeast U1 snRNP subunits were detected following a single-step CBP affinity purification, the protein pattern was significantly obscured by a high level of contaminating proteins (Fig. 1D, lanes 1 and 2). A similar observation was made following purification by an IgG affinity step and TEV protease-mediated release (lanes 3 and 4). In this case, however, contaminants corresponded to (abundant) extract proteins and TEV protease (lane 6; a significant amount of TEV protease is required for complete release because of the inefficient cleavage on the solid phase and the limiting levels of substrate). In contrast, no background proteins were detectable in the fully purified fractions (lanes 7 and 8). Therefore, both purification steps are required for highly specific purification with very low background.

TAP-purified proteins are functional

The TAP method should allow for direct assay of the activity of purified proteins. To test this possibility, we purified the yeast cap binding complex (CBC)[12] using standard TAP conditions starting from a strain carrying the TAP tag fused to the C terminus of the small subunit of CBC. The largest and smallest proteins detected in the purified fraction by Coomassie staining (Fig. 2A) were identified as the two subunits of yeast CBC. The third protein, present at a substoichiometric level, was identified as Srp1p/Importin α, reflecting a physiological interaction with CBC that may only occur transiently[12] rather than contamination. We conclude that the TAP method is generally applicable, and allows copurification of interacting partners present at substoichiometric levels.

A gel-shift analysis was used[12] to test activity of TAP-purified CBC (Fig. 2B). The purified CBC fraction formed a specific complex with a radiolabeled capped RNA (lanes 4 and 5). This demonstrates TAP-purified complexes can retain activity.

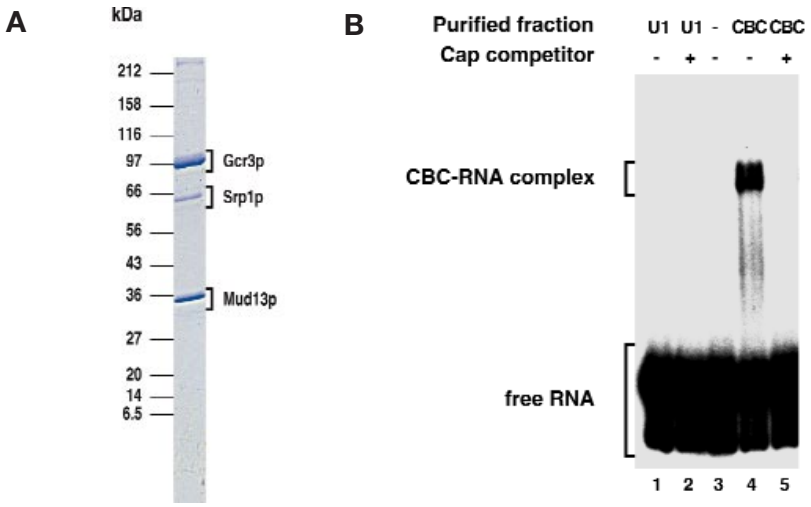

Figure 2. Proteins purified with the TAP procedure are active. (A) TAP purification of yeast CBC and identification of associated proteins. Mud13p and Gcr3p are the small and large subunit of CBC, respectively. (B) TAP-purified CBC is active. A capped labeled RNA probe (lane 3) was mixed with an aliquot of the CBC or U1 snRNP fractions purified with the TAP procedure, and complexes formed were detected by autoradiography following gel electrophoresis (lanes 1 and 4). Competition with unlabeled cap analog demonstrates the specificity of the interaction (lane 5).

Characterization of a new protein complex

To test the generality of the TAP procedure, we undertook the purification of a previously uncharacterized protein. The yeast Mak31p/SmX1p protein, required for maintenance of the yeast killer plasmid, was selected for this purpose. The TAP-tagged Mak31p was purified from 10 L of yeast culture (Fig. 3). Mass spectrometric analysis identified the two copurifying proteins as Mak10p and Mak3p. A doublet of variable abundance in different purifications was identified as the TEV protease. The identification of Mak3p and Mak10p associated in stoichiometric amounts with Mak31p suggests that these three proteins form a stable complex. The fact that these proteins have been assigned related functions reinforces this conclusion[13]. Since Mak3p is a protein *N*-acetyltransferase, the Mak3/10/31 complex is most likely involved in protein modification.

Discussion

The complete TAP purification, including yeast strain construction, can be performed in less than a month. The total costs are low, allowing large-scale applications. Problems could arise if a subunit of the target complex contains a TEV protease cleavage site. However, this will occur very rarely given the high specificity of the TEV protease[4]. EGTA may affect complex stability. This may prevent functional characterization of the activity of the complex but not subunit identification. A C-terminal TAP tag appears to be well tolerated. N-terminal tagging with a TAP tag containing the same units in the reverse order could be used alternatively. The TAP procedure is highly versatile, in that the two affinity modules—protein A associated with a TEV protease cleavage site (TAP-A) and the calmodulin-binding peptide (TAP-C)—do not need to be present on the same passenger protein[14]. This strategy allows for the purification of a specific complex if the two proteins are independently present in several complexes. Additional combinatorial variations, including the use of uncleavable ProtA tag to subtract undesired proteins, could easily be developed.

The TAP procedure has similar applications as the yeast two-hybrid screen[15]. However, an advantage of the TAP procedure is that, under given conditions, all directly or indirectly interacting components are identified in a single experiment. In addition, the TAP system provides an indication of the approximate stoichiometry of the proteins

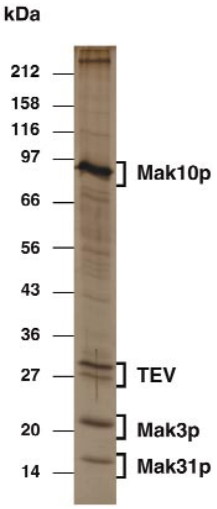

Figure 3. TAP purification of the Mak31p identifies a new complex. A silver-stained gel shows the proteins identified.

present in a given complex and allows for direct biochemical analysis of the purified protein(s). In particular, the activity of mutant complexes can easily be analyzed. The TAP method is not limited to identifying protein–protein interactions, because copurifying ligands can also be characterized. The relative sensitivities and error rates of the TAP strategy and the two-hybrid system remain to be determined. It is noteworthy, however, that all 10 U1 snRNP-specific proteins were identified with the TAP method while no U1 snRNP-specific protein was recovered in two-hybrid screens using two different yeast U1 snRNP proteins as bait[15].

Combined with mass spectrometry methods currently available, the TAP method will be useful to characterize protein complexes and to confirm and/or test for the activity of monomeric or multimeric proteins that have been identified in large-scale nucleic acid sequencing projects. Because it is generic and rapid, the TAP procedure constitutes an important new tool for proteome exploration.

Materials and methods

A detailed experimental description of the TAP method can be found at the laboratory web site[16]. The TAP tag was introduced in strain MGD353-13D as described[17]. Yeast cells were grown at 30°C in YPD medium to OD_{600}=2 and lysed by two passages in a French press (Sim-Aminco) at 8.27 MPa. Extracts were stored frozen at -80°C after dialysis. Purifications were done using standard conditions according to the TAP strategy (Fig. 1B). Proteins in the various fractions were concentrated, fractionated on exponential 7–25% SDS–PAGE gels and identified either by MALDI peptide mapping or nano-electrospray tandem MS[9,10].

Acknowledgments

We thank F. Caspary for testing the method and suggestions on TEV cleavage, G. Stier for TEV protease, and E. Bouveret, F. Caspary, P. Lopez, O. Puig, R. Ramirez-Morales, I. Mattaj, and members of EMBL for support and/or comments on the manuscript. The German Technology Ministry (BMFF) and Glaxo-Wellcome partially supported the M.M. laboratory at EMBL.

1. Séraphin, B. *EMBO J.* **14**, 2089–2098 (1995).
2. Ford, C.F. et al. *Protein Expr. Purif.* **2**, 95–107 (1991).
3. Sheibani, N. *Prep. Biochem. Biotechnol.* **29**, 77–90 (1999).
4. Dougherty, W.G. et al. *Virology* **171**, 356–364 (1989).
5. Senger, B. et al. *EMBO J.* **17**, 2196–2207 (1998).
6. European Patent Application No. 98 115448,7.
7. Gottschalk, A. et al. *RNA* **4,** 374–393 (1998).
8. Riedel, N. et al. *Proc. Natl. Acad. Sci. USA* **83**, 8097–8101 (1986).
9. Shevchenko, A. et al. *Proc. Natl. Acad. Sci. USA* **93**, 14440–14445 (1996).
10. Wilm, M. et al. *Nature* **379**, 466–469 (1996).
11. Fortes P. et al. *Genes & Devel.* **13**, in press (1999).
12. Görlich, D. et al. *Cell* **87**, 21–32 (1996).
13. Wickner, R.B. *Microbiol. Rev.* **60**, 250–265 (1996).
14. Caspary, F. et al. *EMBO J.* **18**, 3463–3474 (1999).
15. Fromont, R.M. et al. *Nat. Genet.* **16**, 277–282 (1997).
16. http://www.embl-heidelberg.de/ExternalInfo/seraphin/TAP.html
17. Puig, O. et al. *Yeast* **14**, 1139–1146 (1998).