An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells

Tilmann Bürckstümmer, Keiryn L Bennett, Adrijana Preradovic, Gregor Schütze, Oliver Hantschel, Giulio Superti-Furga & Angela Bauch

Tandem affinity purification (TAP) is a generic two-step affinity purification protocol that enables the isolation of protein complexes under close-to-physiological conditions for subsequent analysis by mass spectrometry. Although TAP was instrumental in elucidating the yeast cellular machinery, in mammalian cells the method suffers from a low overall yield. We designed several dual-affinity tags optimized for use in mammalian cells and compared the efficiency of each tag to the conventional TAP tag. A tag based on protein G and the streptavidin-binding peptide (GS-TAP) resulted in a tenfold increase in protein-complex yield and improved the specificity of the procedure. This allows purification of protein complexes that were hitherto not amenable to TAP and use of less starting material, leading to higher success rates and enabling systematic interaction proteomics projects. Using the well-characterized Ku70-Ku80 protein complex as an example, we identified both core elements as well as new candidate effectors.

Cellular functions are the result of the coordinated action of several proteins in macromolecular assemblies. Protein complex composition varies with time and space to adapt to changing cellular requirements. Therefore, the analysis of protein-complex composition is considered an important step in the genotype-to-phenotype integration process[1–5]. Mass spectrometry–based proteomic tools have proven to be successful in the identification of multicomponent complexes formed under native conditions[6,7]. The TAP procedure is an affinity purification technique originally developed in yeast that enables the purification of protein complexes under close-to-physiological conditions[8]. Protein complex composition is then determined by mass-spectrometric protein identification. TAP is a rapid and reliable technique that can be readily standardized. It has been successfully applied in the analysis of protein-protein interaction networks in prokaryotic and eukaryotic cells[9–13].

The TAP method involves fusion of the TAP tag to the target protein and introduction of the construct into the host cell. The yeast TAP tag (yTAP tag) consists of two immunoglobulin G (IgG)-binding units of protein A from *Staphylococcus aureus*, a cleavage site for the tobacco etch virus (TEV) protease and a calmodulin binding peptide (CBP; **Fig. 1**). The fusion protein and associated components are recovered from cell extracts by two sequential purification steps. In the first step, the protein A moiety of the TAP tag is bound to IgG sepharose, and complexes are released by TEV-protease cleavage. In the second step, the CBP moiety is bound to calmodulin sepharose in the presence of calcium and then eluted with EGTA. Although this technology has been applied successfully in mammalian cells[12,14], it also has limitations. First, and most importantly, the overall yield of the process is very low. As a consequence, TAP requires large initial quantities of cells (typically 5×10^8–1×10^9 cells). Second, the technology has not been applied to highly differentiated cells (for example, neuronal cells). Third, primary cells have not been amenable to the TAP procedure because the availability of cellular material has been limiting.

Here we compare the TAP tag originally developed for yeast with four newly designed TAP fusion protein cassettes encoding alternative affinity binding moieties optimized for use in mammalian systems. Using one of the optimized TAP tags, we obtained overall complex yields that were an order of magnitude higher as compared to the original TAP tag. This increase in affinity purification efficiency permitted the purification and identification of a protein complex (Ku70-Ku80) from 5×10^7 HEK293 cells.

RESULTS

Construction of new TAP tags

We designed a series of new tags and corresponding vectors based on the two-step purification principle proven successful in TAP[8]. The yTAP tag consists of two IgG binding units of protein A from *Staphylococcus aureus* and the CBP unit, separated by a TEV-protease cleavage site[8]. As a first construct, we generated a streamlined version of the conventional yTAP tag by avoiding residues artificially derived from the original cloning junctions and unlikely to contribute to the purification purposes. We designed the resultant protein A–CBP (AC-TAP) tag (**Fig. 1a**) using frequently used human codons (instead of codons that are frequently used in prokaryotes).

The AC-TAP tag as a starting point

We compared the AC-TAP tag to the yTAP tag by fusing each to the N terminus of IKKγ as available data indicated that N-terminal TAP tag fusion proteins of IKKγ are functional[12]. In general, it is advisable to incorporate all available information to decide whether

Research Center for Molecular Medicine (CeMM), Lazarettgasse 19/3, 1090 Vienna, Austria. Correspondence should be addressed to G.S.-F. (gsuperti@cemm.oeaw.ac.at).

RECEIVED 24 MAY; ACCEPTED 26 SEPTEMBER; PUBLISHED ONLINE 22 OCTOBER 2006; CORRECTED ONLINE 6 NOVEMBER 2006 (DETAILS ONLINE); DOI:10.1038/NMETH968

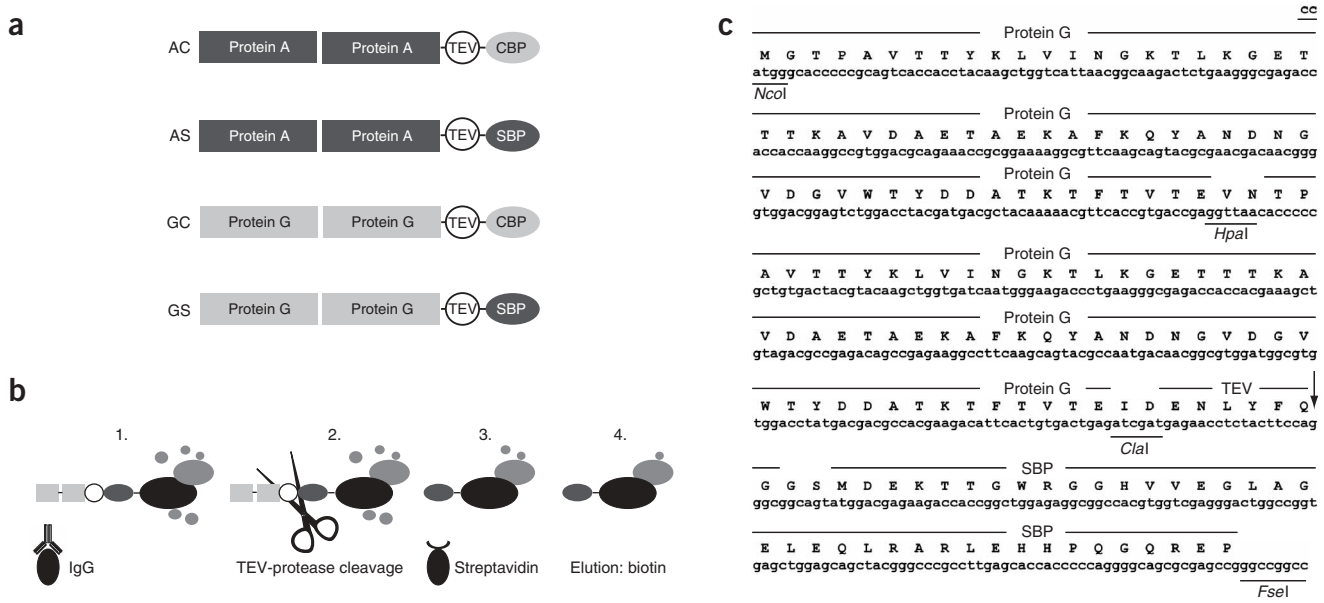

Figure 1 | Schematic representation of TAP. (**a**) Schematics of the four TAP tags that were obtained by permutation of protein A and protein G with CBP and SBP. (**b**) Overview of the TAP protocol. During the first step, TAP-tagged proteins are sequestered by IgG sepharose (1) and released by TEV-protease cleavage (2). TEV protease–cleaved proteins are then bound to streptavidin sepharose (3) and eluted with 1 mM biotin (4). (**c**) Nucleotide and amino-acid sequence of the GS-TAP tag which is comprised of protein G and SBP.

to insert the tag at the N or the C terminus of a protein. In the absence of any information, it is recommended to generate both versions[7,15]. IKKγ is part of the multiprotein IκB kinase (IKK) complex that mediates activation of the transcription factor NF-κB[16]. The regulatory subunit IKKγ is known to copurify with the two catalytic subunits IKKα and IKKβ[12]. We monitored the efficiency of TAP assaying for purification of the bait (IKKγ) and one key interactor, IKKα, by immunoblotting.

We introduced yTAP-IKKγ or AC-TAP–IKKγ into HEK293 cells by retroviral gene transfer. Both proteins were expressed at similar levels (**Fig. 2**); the humanized AC-TAP–IKKγ being only marginally better expressed. The overall purification process was comparable for the two constructs, both at the level of the bait and a key interactor (IKKα; **Fig. 2**). The entire procedure, however, was very inefficient in both cases and only led to the identification of interacting proteins when using large quantities of cells (5×10^8). As the AC-TAP tag performed at least as well as the original yTAP tag, we used the AC-TAP tag as a reference and basis for improvement.

Comparison of the different TAP tag variants

In addition to the AC-TAP tag, we designed three other tags. In two of the constructs, we replaced protein A with two IgG binding units of protein G from *Streptococcus* sp. (**Fig. 1a**). Protein G shows a slightly higher affinity for a broader range of immunoglobulins from different subclasses and species, and has a completely different tertiary structure[17,18] (**Supplementary Fig. 1** online). Moreover, we replaced the CBP moiety with the streptavidin-binding peptide (SBP), which binds to streptavidin with low-nanomolar affinity and can be specifically eluted by biotin[19]. The permutation of protein A and protein G with CBP and SBP produced three additional TAP tags: protein A–SBP (AS), protein G–CBP (GC) and protein G–SBP (GS; **Fig. 1**).

Next we compared the efficiencies of the four TAP tag variants. We stably expressed TAP-tagged fusion proteins with IKKγ in HEK293 cells and subjected them to TAP. Whereas AC-TAP– and GC-TAP–tagged proteins were eluted by boiling in SDS sample buffer, AS-TAP– and GS-TAP–tagged proteins were eluted with 1 mM biotin (AS/Bio and GS/Bio). To monitor the efficiency of the biotin elution, we then boiled the streptavidin sepharose in SDS sample buffer (AS/BB and GS/BB). The TAP-tagged IKKγ fusion proteins were expressed at comparable levels (**Fig. 3a**). Notably, the tags gave different overall purification yields. Although the GC-TAP tag did not result in an improvement over the reference AC-TAP tag, both the AS-TAP and the GS-TAP tags led to a dramatic

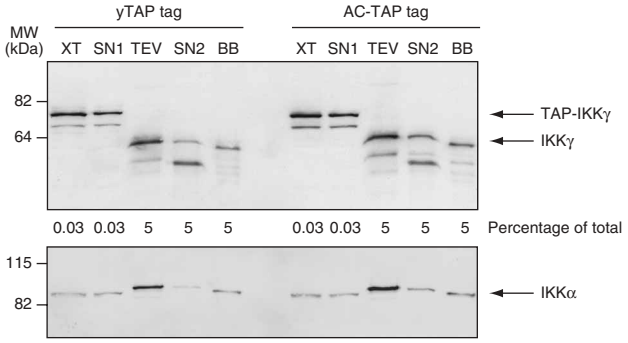

Figure 2 | Comparison of yTAP and AC-TAP. yTAP-IKKγ and AC-TAP–IKKγ were purified by TAP from 5×10^8 cells. Recovery of the bait (IKKγ) and one key interactor (IKKα) was tracked throughout the entire TAP procedure by immunoblotting using IKKγ-specific antiserum and IKKα-specific antiserum. XT, cell extract; SN1, supernatant after IgG binding; TEV, eluate after TEV-protease cleavage; SN2, supernatant after calmodulin binding; BB, boiled beads, that is, the final eluate. Numbers below the lanes indicate which percentage of the total sample was loaded on each lane.

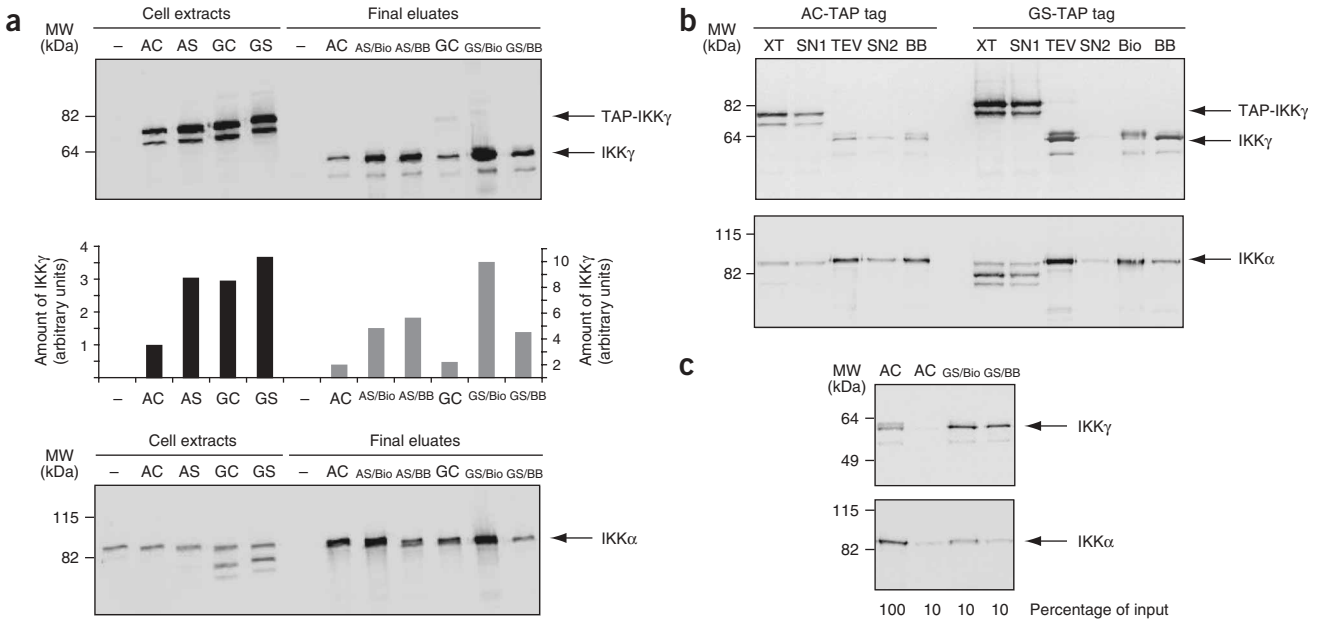

Figure 3 | Comparison of the different TAP tag variants. (**a**) AC-TAP–, AS-TAP–, GC-TAP– and GS-TAP–tagged IKKγ were purified by TAP from 5×10^8 cells and compared with respect to the overall recovery of the bait (IKKγ; top) and one key interactor (IKKα; bottom). Note that the protein G–containing constructs (GC, GS) reacted specifically with the mouse secondary antibody used for the detection of IKKα (compare lower band). IKKγ levels in the top panel were quantified by the LI-COR Odyssey system (middle). (**b**) Purification of AC-TAP– and GS-TAP–tagged IKKγ including all purification steps. Sample nomenclature is the same as in **Figure 2**. (**c**) AC-TAP– and GS-TAP–tagged IKKγ were purified by TAP from 5×10^7 cells (10% of standard starting material) as described in **a**. Final eluates were compared to AC-TAP-tagged IKKγ purified from 5×10^8 cells (100% of standard starting material) by analyzing both recovery of IKKγ and IKKα.

increase in the overall recovery of the bait (**Fig. 3a**). This is striking taking into account that both elution fractions should be added for quantitative comparison with the other tags. When calculating the overall yield for the procedure, it is important to consider that protein G and protein A differ in their ability to be detected by different antibody subtypes from different species (**Supplementary Fig. 1**). We estimate that the difference in recognition between the AC- and the GS-TAP tag amounts to a factor of 2. Using this factor, we calculated the differences in the overall recovery rates as exemplified for the comparison of the AC- with the GS-TAP tag: the difference in recovery (12) is divided by the difference in expression level (3.4) and multiplied by the recognition factor (2). Overall, we estimate an improvement in yield of bait recovery of about tenfold (**Fig. 3a**).

We not only detected the increased yield at the level of the bait (IKKγ), but also at the level of the key interactor IKKα (**Fig. 3a**). There was no major difference in performance between the protein A– and the protein G–containing TAP tags (AC was similar to GC, AS was similar to GS; **Fig. 3a**). We obtained similar results for another bait (MyD88; **Supplementary Fig. 1**). For simplicity, we used the GS-TAP tag for all subsequent analyses.

To dissect the contribution of the individual steps to the overall purification yield, we tracked the bait throughout the purification process. Two steps were consistently more efficient with the GS-TAP tag than with the AC-TAP tag (**Fig. 3b**): (i) more bait was retrieved after TEV-protease cleavage; and (ii) the TEV protease–cleaved bait was quantitatively captured by the streptavidin resin. In summary, these experiments suggest that the GS-TAP tag is superior to the yTAP tag in terms of bait and interactor recovery.

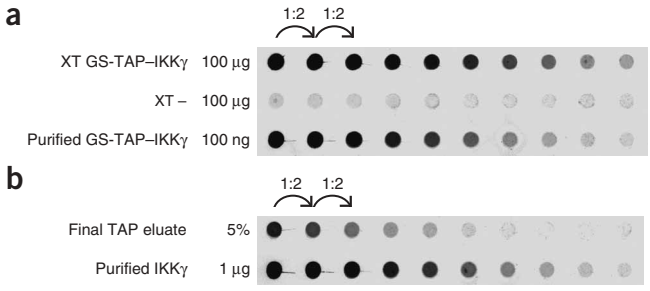

Figure 4 | Quantification of the overall process. (**a**) A twofold serial dilution of 100 µg of cell extract (XT) from either parental HEK293 cells (–) or HEK293 cells expressing GS-TAP–IKKγ was compared to a serial dilution of 100 ng recombinant purified GS-TAP–IKKγ derived from *E. coli*. Spotted proteins were analyzed by immunoblotting (IKKγ). Signal intensities were quantified using the LI-COR Odyssey system and provided the basis for calculating the amount of GS-TAP-IKKγ that is present in the cell extract. (**b**) HEK293 cells expressing GS-TAP–IKKγ were subjected to TAP. Serial dilution of 5% of the final TAP eluate was compared to a serial dilution of 1 µg of recombinant purified IKKγ derived from *E. coli*. Spotted proteins were analyzed as described in **a**.

Reduction of the required biological material for GS-TAP

One of the most cumbersome limitations of the present TAP protocol is the massive amount of cellular material required. Based on the improved efficiency of the GS-TAP tag, we investigated whether it was possible to reduce the quantity of starting material. We purified both AC-TAP and GS-TAP–IKKγ from 5×10^7 cells (10% of the standard starting material). We compared the final eluates of the purifications to the final eluate of AC-TAP–IKKγ recovered from 5×10^8 cells (100% of the standard starting material). Purification of GS-TAP–IKKγ from one tenth of

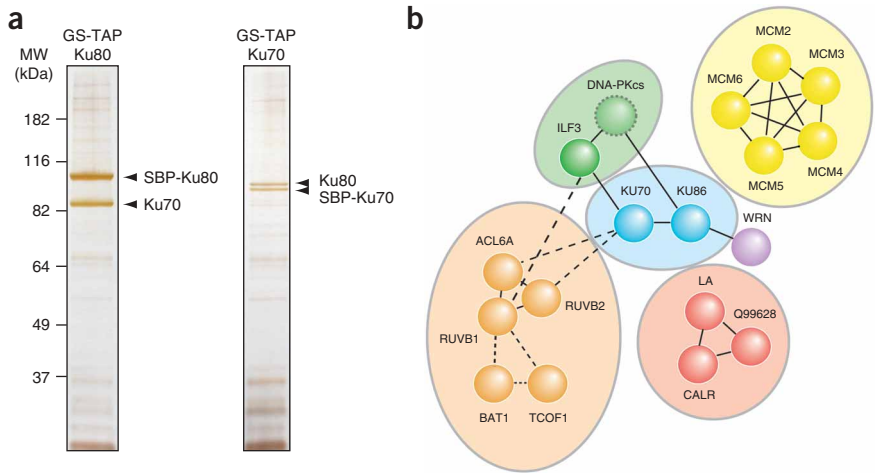

Figure 5 | TAP of the Ku70-Ku80 complex. (**a**) GS-TAP–Ku70 and GS-TAP–Ku80 were purified from 5×10^7 HEK293 cells by TAP. Purified proteins were visualized by silver staining and subsequently identified by LC-MSMS. The final data set contains a total of 39 proteins that were grouped according to biological function (**Table 1**). (**b**) The iHOP database was searched to identify previously reported interactions within the Ku70-Ku80 interactome. Proteins were grouped in one module if there was evidence in the literature that the proteins interact with one another (solid lines). Interaction data stemming from yeast two-hybrid data are indicated with dashed lines.

the usual starting material retrieved as much IKKγ and interacting IKKα as obtained using AC-TAP–IKKγ under standard conditions (**Fig. 3c**).

Quantification of the overall TAP yield

To create a quantitative reference for future users of the protocol and provide a basis for experimental design, we embarked on an approximate quantification of the process. We used purified recombinant proteins as references and compared them to the input and the output of the purification procedure. Notably, the protein A or protein G moiety of the TAP tag binds with high affinity to the F_c portion of virtually all antibodies used for detection. This may lead to an underestimation of the overall yield.

To circumvent this problem, we determined the protein input and output by using two different recombinant standards: GS-TAP–IKKγ and untagged IKKγ, both expressed and purified from *Escherichia coli*. We created a serial dilution of the cell extract from HEK293/GS-TAP–IKKγ (that is, the starting material) and compared it to a serial dilution of recombinant purified GS-TAP-IKKγ. Comparison of the dot blot signal intensities revealed that 100 μg of cell extract contained approximately 169 ng of purified GS-TAP–IKKγ (**Fig. 4a**). In other words, 36 mg of cell extract (obtained from 5×10^7 cells) contain 69 μg (690 pmol) of purified GS-TAP–IKKγ.

We compared the quantity of IKKγ in the final TAP eluate to the amount of recombinant purified untagged IKKγ (**Fig. 4b**). Based on this comparison, 5% of the final TAP eluate contained approximately 98 ng of purified IKKγ. This means that 100% of the final TAP eluate contained the equivalent of 2 μg (35 pmol) of IKKγ. In conclusion, the overall yield amounts to approximately 5% of the input material (35 pmol from 690 pmol). We obtained similar results for three other baits in two other cell lines (**Supplementary Fig. 1**).

Purification of GS-TAP–tagged Ku70 and Ku80

Finally, we applied the improved procedure to purify a protein complex that had not yet been isolated by the TAP procedure. We chose the DNA-dependent protein kinase (DNA-PK), which consists of a catalytic subunit ($DNA\text{-}PK_{cs}$) and two cofactors (Ku70 and Ku80), as a model complex. The DNA-PK holoenzyme is of fundamental importance for DNA double-strand-break repair (nonhomologous end-joining) and for VDJ recombination (for review, see ref. 20). We expressed GS-TAP–Ku70 and GS-TAP–Ku80 in HEK293 cells. We assessed the functionality of GS-TAP–Ku70 by monitoring DNA-PK activity associated with Ku70 (**Supplementary Fig. 2** online). We purified GS-TAP–Ku70 and GS-TAP–Ku80 from 5×10^7 cells. We included untransduced HEK293 cells as negative control (**Supplementary Fig. 2**). Ku70 and Ku80 always copurified in equimolar amounts as judged by silver staining (**Fig. 5a**), indicating that Ku70 and Ku80 form a stoichiometric complex. This is in agreement with the observation that the biologically functional unit is the Ku70-Ku80 heterodimer[21].

Next we investigated whether interactors of Ku70 could be identified by subjecting TAP-purified Ku70 to mass-spectrometric analysis. To deconvolute the dataset, we subtracted a number of unrelated purifications. The lists of nonspecific interactors are available in **Supplementary Table 1** online. The final dataset contained several interacting proteins, which we grouped according to biological function (**Table 1**). Indeed, some of the identified proteins have been previously reported to be associated with Ku70-Ku80, such as the Werner syndrome helicase[22] and the transcription factor ILF3/NF90 (ref. 23). In this analysis we did not detect DNA-PK (**Table 1**). This is in agreement with DNA-PK preferentially binding to Ku70-Ku80 in the presence of DNA[21]. In subsequent experiments, however, we identified DNA-PK with a low peptide coverage.

In an attempt to interpret the results of the Ku70-Ku80 purification according to the concept of modularity that has recently been proposed[10], we assessed which of the Ku70-Ku80 interactors have been reported to interact with one another. These interactions would not only cross-validate the data from this study, but would also support the concept that protein complexes can be combined in a modular fashion to fulfill distinct biological functions. Indeed, many proteins could be grouped into subcomplexes based on previously reported interaction data, such as the MCM protein family that has been implicated in DNA replication and chromosome maintenance[24] and the RNA-associated protein complex that contains Lupus La, RoBP1/Q99628 and calreticulin[25]. Moreover, Actin-like 6A/BAF53, RuvB1 and RuvB2 have been shown to form a protein complex that has been implicated in transcriptional activation[26] (**Fig. 5b**). This implies that the core complex (consisting of Ku70 and Ku80) is partitioned to a variety of other complexes, thereby

Table 1 | Specific interactors of Ku70

Biological function	Protein
DNA recognition	KU86_HUMAN (49) KU70_HUMAN (45)
DNA helicases	RUVB1_HUMAN (8) RUVB2_HUMAN (6) WRN_HUMAN (3)
DNA replication	MCM5_HUMAN (7) MCM4_HUMAN (6) MCM3_HUMAN (6) MCM6_HUMAN (4) MCM2_HUMAN (2)
DNA-binding proteins	TIF1B_HUMAN (6) ILF3_HUMAN (4) RCC2_HUMAN (2) TCOF_HUMAN (2)
RNA-binding proteins	LA_HUMAN (11) SF3B2_HUMAN (5) SF3A1_HUMAN (3) SF3A3_HUMAN (2) U2AF2_HUMAN (2)
Signaling	PA2G4_HUMAN (9) MAP4 (7) MARCS_HUMAN (5) DRG1_HUMAN (3) CTR9_HUMAN (2) SAM68_HUMAN (2)
Protein biosynthesis	IF36_HUMAN (12) CDC73_HUMAN (9) CALR_HUMAN (5) Q96DV6_HUMAN (3) RL6_HUMAN (3) RL7A_HUMAN (2)
RNA helicases	DDX46_HUMAN (5) BAT1_HUMAN (4)
Miscellaneous	Q99628_HUMAN (5) IMA3_HUMAN (4) GANAB_HUMAN (4) CALU_HUMAN (2) TXND5_HUMAN (2) ACL6A_HUMAN (2)

Protein names refer to the UniProt Knowledgebase. Numbers in parentheses indicate the number of peptides that were identified for a given protein.

utilizing different modules that are required for specific biological purposes.

DISCUSSION

Most proteins exert their function as part of a protein complex or cellular machine[1,9–11,27]. The characterization of these machines as building blocks of complex organization units such as pathways, is thought to be critical for the understanding of disease and represents a comprehensive approach toward the identification of new drug targets[28,29]. Several affinity purification procedures have been applied to purify protein complexes from mammalian cells, using for example, the Myc tag, the HA tag, the Flag tag or the LAP tag[15,30]. Nevertheless, a tool that enables efficient and systematic purification of protein complexes from mammalian cells for mass spectrometry analysis has been lacking. Here we present a robust modified TAP procedure. How should the reader judge whether it is worthwhile to embark on this technology? In **Figure 6** we summarize the cornerstone parameters of the new TAP procedure that are based on the quantification presented in **Figure 4**, experience from our laboratory and some approximations.

Depending on the bait under consideration, the number of cells required ranges from 5×10^7 to 1×10^9. For HEK293 cells, 5×10^7 cells are routinely used, and this typically yields 35 mg of total protein. For a standard bait stably expressed after retroviral gene transfer, this may correspond to approximately 700 pmol of TAP-tagged protein. The two subsequent steps (first affinity binding step and elution, and second affinity binding step and elution) show reproducible levels of efficiency. In the first step, one can expect to bind 40% of the bait and retrieve 30% thereof after TEV-protease cleavage. In the second step, approximately 75% of the remaining material is recovered. The biotin elution is the step that shows the greatest bait-dependent variability with an average of 50% recovery. If elution by boiling in SDS sample buffer is chosen, the yield can be increased at the expense of specificity. Overall, the TAP procedure recovers about 5% of the bait present in the lysate. Therefore, the overall amount of purified bait expected is in the picomole range (35 pmol in the example given above). If this is true for the bait, what about the interacting proteins? Depending on the stoichiometry of the interacting proteins (for example, 2:1 to 1:20), copurification of between 70 and 1.5 pmol of the interacting proteins can be envisaged. As a reference, the lower amount is visible by silver stain and is well within the limits of detection of nano liquid chromatography–tandem mass spectrometry (LC-MSMS) in a relatively complex sample.

Are these parameters strongly influenced by the expression levels of the introduced TAP-tagged protein? We believe that moderate transcriptional activity associated with the retroviral gene transfer system warrants that the expression level of the bait is dramatically below levels obtained by transient transfection. In the cases in which we have directly compared the expression levels of the TAP-tagged bait to the expression levels of the endogenous counterpart, we only found a modest overexpression (three- to sixfold, data not shown).

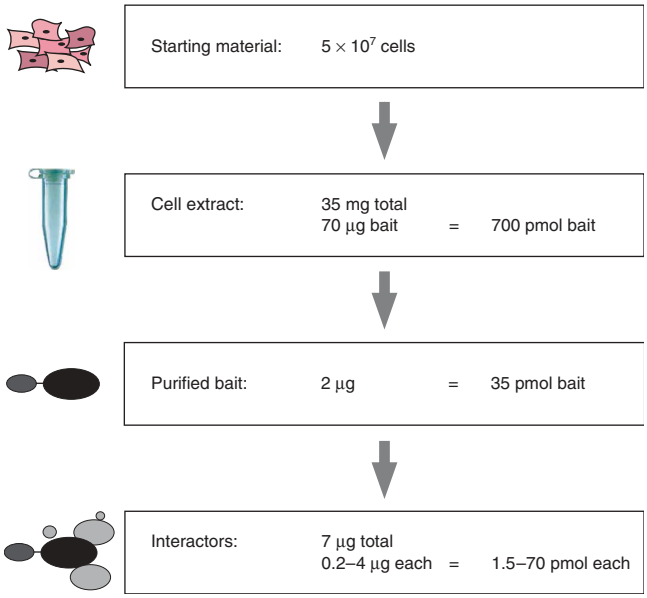

Figure 6 | Quantitative overview of the TAP procedure. A typical TAP purification is done from 5×10^7 cells, yielding 35 mg of total cell extract. Based on the quantification presented in **Figure 4**, the cell extract contains 70 µg (700 pmol) of bait protein. The final eluate contains 2 µg (35 pmol) of bait, which correspond to an overall yield of 5%. Depending on the stoichiometry of the interactions (2:1 to 1:20), 0.2–4 µg (1.5–70 pmol) of interactors will be copurified.

Until now, the application of TAP in mammalian cells was unsatisfactory and did not find wide application. Only a few laboratories have successfully used this technology (see for example, refs. 31–33). In this study we achieved an up to tenfold improvement of the TAP technique with respect to bait recovery. This has two consequences: (i) protein complexes of low abundance can be routinely purified as a result of the improved robustness of the method, and (ii) protein complexes that had previously been successfully purified using the yTAP tag can now be purified from less cellular starting material. An additional advantage of the optimized TAP protocol is the specificity of the final elution step using biotin instead of using SDS sample buffer.

The new GS-TAP tag, in contrast to the previously used yTAP tag, offers the opportunity to perform single-step purification by streptavidin followed by specific elution with biotin or straight boiling (**Supplementary Fig. 3** online). Using the same cell lysate, the experimenter therefore has the choice of obtaining different levels of yield and purity depending on the nature of the protein complex and the type of question being asked.

To demonstrate the application of this method, we focused on the well-characterized DNA-PK protein complex using Ku70 and Ku80 as bait proteins. In both purifications, Ku70 and Ku80 were isolated in equimolar quantities, implying that Ku70 and Ku80 form a stable 1:1 complex. In addition to known interacting proteins such as the Werner syndrome helicase and the transcription factor ILF3/NF90, we also identified several new binding partners. Based on these findings, we propose that Ku70 and Ku80 are involved in a variety of biological processes, as outlined in **Table 1**. The substoichiometric interaction of the Ku70-Ku80 pair with different protein groups is reminiscent of the modular organization of cellular machines observed on a large scale in yeast[10]. Only extensive proteomic analyses, such as purification using interacting proteins of Ku70 as the baits, and combination with techniques allowing temporal and spatial resolution will allow the refinement of this hypothesis and determine which modules may or may not interact in a mutually exclusive fashion and in what order. The challenge will be to understand the logic behind integration of various activities over the Ku70-Ku80 core.

In summary, the procedure described here represents a broadly applicable affinity purification method of protein complexes from mammalian cells and thus meets a long-awaited need in the scientific community. We envision that the robustness of the procedure should result in its wide use in the characterization of molecular machines in a variety of different cell types and experimental conditions. We expect that the method could allow the characterization of complexes from cells that are not easily cultivated in large quantities (for example, neuronal and immune cells) or even primary cells. Finally, large-scale approaches to explore the human proteome and cellular machinery should become more feasible.

METHODS

Vector generation and establishment of stable cell lines. We obtained the yTAP cassette from Euroscarf. The new TAP cassettes were synthesized by Midland Certified. We cloned all cassettes into a Moloney murine leukemia virus–based vector[34]. We introduced an IRES-*GFP* derived from pBMN-I-GFP to monitor infection efficiency[35]. We obtained the cDNAs for IKKγ, Ku70, Ku80, MyD88, RIG-I, Btk and Tbk1 from RZPD, amplified them by PCR and introduced them into the expression vector using the Gateway technology (Invitrogen). We stably transduced HEK293 cells by retroviral gene transfer[12].

Antibodies. We obtained the IKKγ-specific antiserum from Santa Cruz (FL-419). We obtained the IKKα-specific antibody from Imgenex.

Tandem affinity purification. We lysed cells in lysis buffer (50 mM Tris-HCl (pH 7.5), 125 mM NaCl, 5% glycerol, 0.2% NP-40, 1.5 mM $MgCl_2$, 25 mM NaF, 1 mM Na_3VO_4 and protease inhibitors) and cleared the lysate by centrifugation. We incubated the lysate with rabbit-IgG agarose (Sigma) at 4 °C for 2 h. We washed bound proteins with lysis buffer and then with TEV-protease cleavage buffer (10 mM Tris-HCl (pH 7.5), 100 mM NaCl and 0.2% NP-40) and eluted by addition of 40 μg TEV protease (16 °C for 1 h). We incubated the TEV-protease cleavage product with calmodulin sepharose (Amersham) in the presence of 2 mM $CaCl_2$ or with Ultralink Immobilized Streptavidin Plus (Pierce) at 4 °C for 1 h. For the CBP-containing tags, we eluted bound proteins by boiling in SDS sample buffer. For the SBP-containing tags, we eluted bound proteins with 1 mM D-biotin (Sigma). To assess the efficiency of the biotin elution, we boiled the remaining streptavidin beads in SDS sample buffer.

Dot-blot analysis. Preparation of the recombinant proteins that were used as standards is described in **Supplementary Methods** online. We spotted protein samples on a nitrocellulose membrane using the Bio-Dot apparatus (Biorad). We air-dried the nitrocellulose membrane for 10 min and then rehydrated it in phosphate-buffered saline for 2 min. We stained rehydrated membranes with Ponceau Red and analyzed them by immunostaining.

Sample preparation, mass spectrometry, database searches and data analysis. We analyzed TAP samples by one-dimensional SDS-PAGE, silver stained them, excised specific bands and/or regions of interest from the gel and digested them *in situ* with modified porcine trypsin. We analyzed tryptically digested samples by data-dependent nanocapillary reversed-phase LC-MS/MS and identified proteins by automated searching against the International Protein Index protein sequence database. Results from the searches were automatically validated and clustered into protein groups based on the number of shared peptides identified by MS/MS. Details of the sample preparation, mass spectrometric conditions and database searches are available in **Supplementary Methods**.

Data analysis and bioinformatics. We performed all data comparison and filtering using EPICenter (Proxeon Biosystems). In two independent Ku70 TAP experiments, we found 96 protein groups to be common to both purifications (**Supplementary Table 1**). To remove nonspecific interactors from the data set, five unrelated TAPs performed in the same cell line (HEK293) were subtracted. This resulted in the loss of 17 protein groups (**Supplementary Table 1**) with 79 protein groups remaining in the dataset (**Supplementary Table 1**). To further deconvolute the dataset, we also subtracted an additional 15 purifications that used

different affinity reagents and alternate cell lines. This reduced the list by yet another 40 protein groups (**Supplementary Table 1**). The resultant list of 39 proteins is shown in **Table 1** (detailed information is available in **Supplementary Table 1**). For simplicity, generic protein names (as defined in the UniProt Knowledgebase) were given.

To identify previously reported interactions, we manually searched the iHOP database with the 39 proteins that were present in the final data set. We grouped proteins into subcomplexes according to available interaction data.

To provide a more general perspective on nonspecific interactors that are frequently identified by TAP, we analyzed a set of 10 TAP experiments using 6 baits expressed in HEK293 cells. **Supplementary Table 1** online depicts all proteins that were found in more than three purifications. Again, for simplicity, generic protein names are included.

URL. Euroscarf (http://web.uni-frankfurt.de/fb15/mikro/euroscarf/cz_plas.html).

Requests for materials. materials@cemm.oeaw.ac.at

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank M. Brehme for help with graphical illustrations. We also thank J.-M. Peters (Institute for Molecular Pathology, Vienna), S.P. Jackson (Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge), M. Mann (Max Planck Institute for Biochemistry, Martinsried) and R. Aebersold (Eidgenoessische Technische Hochschule, Zuerich) for critical reading of the manuscript. We thank M. Planyavsky for preparation of the TAP samples for analysis by mass spectrometry and the entire Superti-Furga laboratory for suggestions. T.B. is supported by a research fellowship (BU 2180/1-1) from the German Research Foundation (DFG). CeMM is supported by the Austrian Academy of Sciences. Work described here was funded by the Austrian Proteomics Platform-II of the GenAU Program of the Austrian Ministry of Education and Research, and by a grant from the Austrian National Bank.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Methods* website for details).

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **92**, 291–294 (1998).
2. Papin, J.A., Hunter, T., Palsson, B.O. & Subramaniam, S. Reconstruction of cellular signalling networks and analysis of their properties. *Nat. Rev. Mol. Cell Biol.* **6**, 99–111 (2005).
3. Bauch, A. & Superti-Furga, G. Charting protein complexes, signaling pathways, and networks in the immune system. *Immunol. Rev.* **210**, 187–207 (2006).
4. Hinsby, A.M. *et al.* A wiring of the human nucleolus. *Mol. Cell* **22**, 285–295 (2006).
5. Fox, A.H. & Lamond, A.I. Nuclear processes controlled by molecular machines. *Genome Biol.* **3** REPORTS4016 (2002).
6. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
7. Puig, O. *et al.* The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**, 218–229 (2001).
8. Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032 (1999).
9. Gavin, A.C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
10. Gavin, A.C. *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636 (2006).
11. Ho, Y. *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180–183 (2002).
12. Bouwmeester, T. *et al.* A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat. Cell Biol.* **6**, 97–105 (2004).
13. Butland, G. *et al.* Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* **433**, 531–537 (2005).
14. Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T. & Drewes, G. Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. *J. Biol. Chem.* **279**, 12804–12811 (2004).
15. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **60**, 523–533 (2003).
16. Ghosh, S. & Karin, M. Missing pieces in the NF-kappaB puzzle. *Cell* **109**, S81–S96 (2002).
17. Sjobring, U., Bjorck, L. & Kastern, W. Streptococcal protein G. Gene structure and protein binding properties. *J. Biol. Chem.* **266**, 399–405 (1991).
18. Sauer-Eriksson, A.E., Kleywegt, G.J., Uhlen, M. & Jones, T.A. Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. *Structure* **3**, 265–278 (1995).
19. Keefe, A.D., Wilson, D.S., Seelig, B. & Szostak, J.W. One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expr. Purif.* **23**, 440–446 (2001).
20. Collis, S.J., DeWeese, T.L., Jeggo, P.A. & Parker, A.R. The life and death of DNA-PK. *Oncogene* **24**, 949–961 (2005).
21. Gottlieb, T.M. & Jackson, S.P. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* **72**, 131–142 (1993).
22. Li, B. & Comai, L. Functional interaction between Ku and the werner syndrome protein in DNA end processing. *J. Biol. Chem.* **275**, 28349–28352 (2000).
23. Aoki, Y., Zhao, G., Qiu, D., Shi, L. & Kao, P.N. CsA-sensitive purine-box transcriptional regulator in bronchial epithelial cells contains NF45, NF90, and Ku. *Am. J. Physiol.* **275**, L1164–L1172 (1998).
24. Crevel, G., Ivetic, A., Ohno, K., Yamaguchi, M. & Cotterill, S. Nearest neighbour analysis of MCM protein complexes in *Drosophila melanogaster*. *Nucleic Acids Res.* **29**, 4834–4842 (2001).
25. Fouraux, M.A., Bouvet, P., Verkaart, S., van Venrooij, W.J. & Pruijn, G.J. Nucleolin associates with a subset of the human Ro ribonucleoprotein complexes. *J. Mol. Biol.* **320**, 475–488 (2002).
26. Park, J., Wood, M.A. & Cole, M.D. BAF53 forms distinct nuclear complexes and functions as a critical c-Myc-interacting nuclear cofactor for oncogenic transformation. *Mol. Cell. Biol.* **22**, 1307–1316 (2002).
27. Krogan, N.J. *et al.* Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**, 637–643 (2006).
28. Brown, D. & Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug Discov. Today* **8**, 1067–1077 (2003).
29. Fishman, M.C. & Porter, J.A. Pharmaceuticals: a new grammar for drug discovery. *Nature* **437**, 491–493 (2005).
30. Cheeseman, I.M. & Desai, A. A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci. STKE* **2005**, pl1 (2005).
31. Bertwistle, D., Sugimoto, M. & Sherr, C.J. Physical and functional interactions of the Arf tumor suppressor protein with nucleophosmin/B23. *Mol. Cell. Biol.* **24**, 985–996 (2004).
32. Canton, D.A. *et al.* The pleckstrin homology domain-containing protein CKIP-1 is involved in regulation of cell morphology and the actin cytoskeleton and interaction with actin capping protein. *Mol. Cell. Biol.* **25**, 3519–3534 (2005).
33. Hacker, H. *et al.* Specificity in Toll-like receptor signalling through distinct effector functions of TRAF3 and TRAF6. *Nature* **439**, 204–207 (2006).
34. Naviaux, R.K., Costanzi, E., Haas, M. & Verma, I.M. The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses. *J. Virol.* **70**, 5701–5705 (1996).
35. Kinsella, T.M. & Nolan, G.P. Episomal vectors rapidly and stably produce high-titer recombinant retrovirus. *Hum. Gene Ther.* **7**, 1405–1413 (1996).