research articles Journal of proteome research

Sequential Peptide Affinity (SPA) System for the Identification of Mammalian and Bacterial Protein Complexes

Mahel Zeghouf,[†,‡] Joyce Li,[†] Gareth Butland,[†] Anna Borkowska,[†] Veronica Canadien,[§] Dawn Richards,[§] Bryan Beattie,[§] Andrew Emili,[†,||] and Jack F. Greenblatt*[,†,||]

Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ontario, Canada M5G 1L6, Affinium Pharmaceuticals, 100 University Avenue, Toronto, Ontario, Canada M5J 1V6, and Department of Molecular and Medical Genetics, University of Toronto, 1 Kings College Circle, Toronto, Ontario, Canada M5S 1A8

Received October 8, 2003

A vector system is described that combines reliable, very low level, regulated protein expression in human cells with two affinity purification tags (Sequential Peptide Affinity, or SPA, system). By avoiding overproduction of the target protein, this system allows for the efficient purification of natural protein complexes and their identification by mass spectrometry. We also present an adaptation of the SPA system for the efficient purification and identification of protein complexes in *E. coli* and, potentially, other bacteria.

Keywords: SPA system • protein complexes • protein−protein interactions • affinity purification • mass spectrometry • human cell • *E. coli*

Introduction

The "complexity" of an organism may arise partly from combinatorial interactions of its gene products.[1] Indeed, protein−protein interactions play a role in nearly all cellular processes, and protein complexes are so ubiquitous[2] that the biological function of an unknown protein can often be inferred from the functions of the proteins with which it is associated ("guilt by association").

Tandem affinity purification (TAP),[2] which makes use of a tag containing two IgG binding domains of the *Staphylococcus aureus* protein A and a calmodulin binding peptide (CBP) separated by a tobacco etch virus (TEV) protease cleavage site, allows for the generic purification to near homogeneity of yeast protein complexes under native conditions. This method relies on homologous recombination to introduce the TAP tag sequence into chromosomal genes, usually so as to produce a C-terminal tag. The tagged protein is not overproduced and should associate only with its normal protein partners.

Our goal was to adapt the TAP method for use in mammalian and bacterial cells. Because a human full length cDNA collection may soon be completed,[3] large-scale human protein interaction studies will become possible using generic expression vectors and purification procedures. A key objective in our generic human expression system was obtaining expression levels of the tagged protein that were reliable and low enough to avoid unnatural protein associations. To achieve this objective, we used a regulated promoter and an expression cassette bracketed by insulator elements that avoid chromosomal position effects. Despite very low levels of protein expression (about 10^{-6}−10^{-5} of total cell protein, or about $3-5 \times 10^4$ molecules per cell), the use of two affinity tags in our SPA system made purification of the tagged protein to near homogeneity entirely feasible.

We also developed a system for systematically tagging and purifying *E. coli* proteins and identifying *E. coli* protein complexes. In this method, the homologous recombination system of bacteriophage λ was used to place the SPA tag at the C-terminal of a target protein in the *E. coli* genome.[4,5] As in yeast,[2] normal levels of the tagged proteins were produced in the recombinant bacteria so as to avoid unnatural protein associations.

Experimental Section

Plasmids. SPA system mammalian expression vectors were created by sequential cloning of PCR products into the pCI-neo (Promega, Madison, WI) vector as described in the Supporting Information. PCR products were verified by sequencing (ACGT Corporation, Toronto, ON). The pMZ3F, pMZI3F, and pMZS3F vectors contain the mouse Rpb1 promoter, inducible ecdysone promoter and CMV promoter, respectively. pMZI is pMZI3F with a TAP tag[6] instead of a SPA tag (Figure 1). pVgRxR encodes the ecdysone receptor heterodimer and was purchased from Invitrogen (Carlsbad, CA). The plasmid pJL72, which contains a version of the TAP cassette,[6] was created for use in *E. coli* and related bacteria by introduction of a selectable kanamycin marker. This plasmid was further modified by replacement of the TAP tag with the SPA tag sequence from pMZI3F to create pJL148. Details of the cloning protocol are

* To whom correspondence should be addressed: jack.greenblatt@utoronto.ca.

[†] Banting and Best Department of Medical Research, University of Toronto.

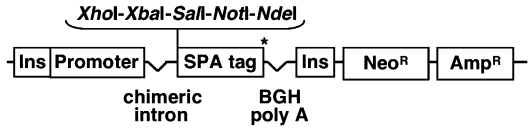
[‡] Present address: Laboratoire d'Enzymologie et Biochimie Structurales (LEBS), CNRS UPR 9063, 1 Avenue de la Terrasse, 91198 Gif sur Yvette, France.

[§] Affinium Pharmaceuticals, Toronto.

[||] Department of Molecular and Medical Genetics, University of Toronto.

10.1021/pr034084x CCC: $27.50

Published on Web 01/31/2004



Name	Tag	Promoter	Expression	size (kb)
pMZ3F	SPA	Rpb1	ubiquitous	6.6
pMZS3F	SPA	CMV	strong	6.3
pMZI3F	SPA	Ecdysone	regulated	6
pMZI	TAP	Ecdysone	regulated	6.4

```
TCCATGGAAAAGAGAAGATGGAAAAAGAATTTCATAGCCGTCTCA
S  M  E  K  R  R  W  K  K  N  F  I  A  V  S
          CBP
GCAGCCAACCGCTTTAAGAAAATCTCATCCTCCGGGGCACTTGAT
A  A  N  R  F  K  K  I  S  S  S  G  A  L  D

TATGATATTCCAACTACTGCTAGCGAGAATTTGTATTTTCAGGGT
Y  D  I  P  T  T  A  S  E  N  L  Y  F  Q  G
                        TEV site
GAGCTCGACTACAAAGACCATGACGGTGATTATAAAGATCATGAC
E  L  D  Y  K  D  H  D  G  D  Y  K  D  H  D
      3xFLAG
ATCGACTACAAGGATGACGATGACAAGTAG
I  D  Y  K  D  D  D  D  K  *
```

Figure 1. SPA system strategy for mammalian cells. Schematic representation of the mammalian SPA vector and sequence of the SPA tag (see text for details). SPA and TAP vectors used in this study are summarized in the table. Ins, chicken β-globin insulator sequences; *, stop codon; BGH poly A, bovine growth hormone polyadenylation signal; Neo^R, neomycin resistance gene; Amp^R, ampicillin resistance gene.

given in the Supporting Information. Full DNA sequences are available upon request.

Cell Culture, Transfections, and Stable Cell Lines. 293T cells (a kind gift from Dr. David McLennan, University of Toronto, ON) were grown as a monolayer in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 units/ml penicillin and 100 μg/mL streptomycin, at 37 °C, 5% CO_2. EcR-293 cells, which are 293 cells stably transfected with pVgRxR, were obtained from Invitrogen and grown as 293T cells except that 30 μg/mL bleocin (Calbiochem, San Diego, CA) was added to the medium. Transient transfection of 293T cells was performed by the calcium phosphate procedure or using Lipofectamine 2000 (Invitrogen) according to the supplier's instructions. For induction of the ecdysone promoter, cells were cotransfected with pVgRxR in a 1:1 ratio. The precipitates were removed 15 h post transfection and cells were grown for an additional 24 h in a fresh medium with or without 3 μM of Ponasterone A. To establish stable cell lines, EcR-293 cells were transfected in 10-mm dishes using 8 μg of circular plasmid DNA by the calcium phosphate procedure. Clonal cell lines were selected in 300 μg/mL Geneticin (Invitrogen) and screened for SPA-tagged protein expression by western blot using the anti-FLAG M2 antibody (Sigma, St Louis, MO). For purification of protein complexes, stable cell lines were grown in 150-mm dishes to about 60% confluence, and expression of the bait protein was induced using 3 μM of Ponasterone A. After 24 h, cells were washed twice with phosphate-buffered saline then harvested with a cell lifter. The cell pellets were stored at −80 °C until purification.

Preparation and Transformation of *E. coli* Cells with Induced λ Recombination Functions. This was performed as previously described,[5] with the following modifications. Briefly, a 50 mL culture of *E. coli* DY330 was used to induce λ recombination functions. After incubation on ice, cells were washed twice with 1 volume of ice-cold sterile water by centrifugation at 5500 × *g*, then suspended in 1 mL of ice cold sterile water and centrifuged at 10 000 × *g* for 30 s. The cell pellet was then resuspended in a final volume of 300 μL of ice-cold sterile water. Linear, gene specific, PCR products were introduced into *E. coli* by electroporation. Further details are given in the text and in the Supporting Information.

Culturing of Tagged *E. coli* Strains. *E. coli* strains in which a bait protein had been SPA tagged or TAP-tagged were cultured in 1 L of Terrific Broth (TB) at 32 °C until late log phase. Cells were harvested and pellets frozen at −80 °C until purification.

SPA Purification of Protein Complexes and Protein Identification. Human whole cell extracts were prepared as described previously,[7] except that the dialysis buffer contained 0.1 M potassium acetate, 10 mM Hepes, pH 7.9, 0.1 mM EDTA and 10% glycerol. *E. coli* cell pellets were resuspended in 35 mL of sonication buffer (10 mM Tris-Cl pH 7.9, 100 mM NaCl, 0.2 mM EDTA, 10% glycerol, 0.5 mM DTT) and broken by sonication. Cell debris was removed by centrifugation at 20 000 × *g* for 30 min. SPA purifications were performed essentially as described before for TAP purifications[6] with some modifications. Detailed protocols are given in the Supporting Information. The purified complexes were separated by SDS-PAGE on a 10% gel. After silver staining, protein bands were excised, reduced and alkylated, then subjected to in-gel trypsin hydrolysis. The peptides were purified and analyzed by MALDI-TOF mass spectrometry (MS) using an α-cyano-4-hydroxycinnamic acid matrix (Fluka, Buchs SG, Switzerland) on a Reflex IV instrument (Bruker Daltonics, Billerica, MA) in positive ion reflectron mode.[8]

Immunofluorescence. EcR-293 cells stably expressing RAP74 were grown on poly-L-lysine coated coverslips, induced for 24 h with 5 μM of Ponasterone A and fixed in 4% paraformaldehyde for 15 min at room temperature. Slides were blocked in 3% BSA phosphate-buffered saline and incubated with anti-FLAG M2 antibody (1:160) for 1 h at room temperature. Immunofluorescence was monitored after incubation for 1 h at room temperature with an anti-mouse IgG-Texas red conjugate (Molecular Probes, Eugene, OR**)** (1:400) using a Leica DMR microscope (Wetzlar, Germany). Nuclei were stained with DAPI (4′,6′-diamidino-2-phenylindole) (Roche, Basel, Switzerland) for 4 min before mounting.

***β*-Galactosidase Assays.** Human cells were washed twice with phosphate-buffered saline before addition of 150 μL of lysis buffer (0.1% Triton X-100, 1 M Tris-HCl pH 8.0), then lysed by freezing the cells overnight at −80 °C and thawing at 37 °C. The lysates were centrifuged for 10 min at 10 000 rpm at 4 °C and the supernatants immediately assayed for β-galactosidase as described previously[9] using purified β-galactosidase (Sigma) as a standard. Protein concentrations were determined using the micro BCA reagents (Pierce, Rockford, IL) and bovine serum albumin as a standard.

Results and Discussion

Design and Characterization of the SPA System for Human Cells. We designed a new C-terminal tag, called SPA (Sequential Peptide Affinity), that contains three modified FLAG sequences (3×FLAG) and the calmodulin binding peptide (CBP) separated by a TEV protease cleavage site (Figure 1). We chose to use 3×FLAG in place of the protein A moiety of the TAP tag[6] for three reasons: first, the FLAG epitope has been widely and successfully used for tagging, identifying, and purifying proteins

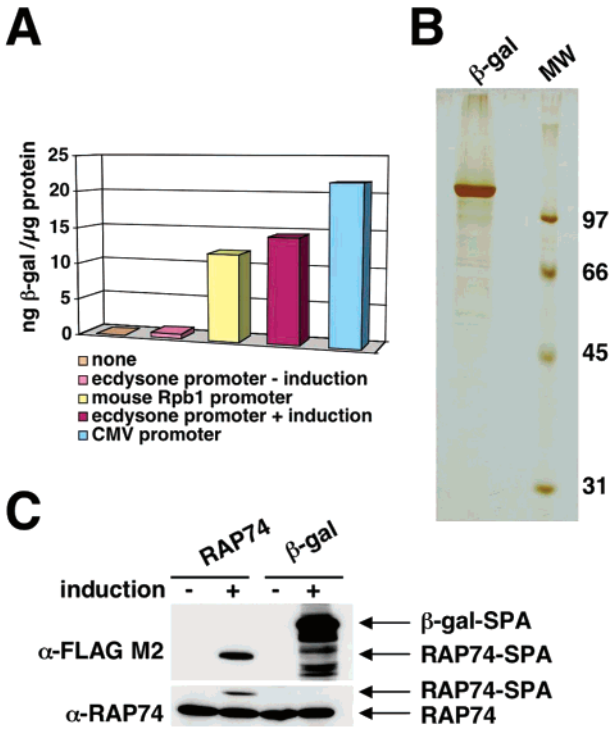


Figure 2. Characterization of SPA vectors with various promoters. **A**, Transient transfections of 293T cells were performed in a 24-well plate using 0.5 μg of the indicated *lacZ* construct and 1.5 μL of Lipofectamine 2000. pVgRXR, which expresses the ecdysone receptor, was cotransfected with the construct containing the ecdysone promoter. β-galactosidase activity was assayed in whole cell extracts as described in the Experimental Section. The ecdysone-inducible promoter was induced with 3 μM of Ponasterone A for 24 h. **B**, SDS-PAGE analysis of tagged β-galactosidase purified from 7×10^7 293T cells after transient cotransfection with the pMZI-*lacZ* construct and pVgRXR (see the Experimental Section for details). MW, molecular weight marker proteins. **C**, Western blot analyses on stable cell lines were performed using anti-FLAG M2 or anti-RAP74 as primary antibody. Whole cell extracts were prepared from EcR−RAP74(SPA) and EcR-β-galactosidase(SPA) stable cell lines established with pMZI3F. Protein induction was with 5 μM of Ponasterone A for 24 h (+), or omitted (−). Each lane contained the same amount of total protein.

in mammalian cells; second, 3×FLAG is much smaller than protein A (22 amino acid residues as opposed to 137), which may lessen its effect on the function and/or stability of some tagged proteins; finally, 3×FLAG is recognized with high affinity by the monoclonal antibody M2, which is known to work well in both western blots and immunoprecipitation[10] (an agarose-conjugate of the anti-FLAG M2 antibody is also commercially available).

We then designed expression vectors that could be manipulated in *E. coli* and allowed expression of a SPA-tagged protein in mammalian cells (Figure 1). Multiple cloning sites facilitated the insertion of target cDNAs upstream of the SPA tags. These elements and the promoter were framed by two chicken β-globin gene insulator sequences that shield a promoter from nearby regulatory elements. These elements have already been used successfully to protect vectors from chromosomal position effects.[11,12]

Three different promoters were characterized, initially in transient transfection assays, using the *E. coli lacZ* gene as a reporter (Figure 2A). The strong CMV promoter is useful for identifying protein complexes containing abundant viral or human proteins and produced β-galactosidase as about 2% of total protein in transfected human 293T cells. The mouse Rpb1 promoter, from the gene encoding the largest subunit of RNA polymerase II (RNAPII), expected to be ubiquitously expressed and relatively weak in a chromosomal context (e.g., in stable cell lines or transgenic animals), was only slightly weaker than the CMV promoter in transfected 293T cells. An ecdysone-inducible promoter presented a low basal level, and the addition of the inducer, Ponasterone A, led to high expression levels in transfected 293T cells expressing the ecdysone receptor (Figure 2A). Most importantly, all stable EcR-293 cell lines established with pMZI3F, which contains the ecdysone inducible promoter and a SPA tag, were tightly regulated by Ponasterone A, as illustrated in Figure 2C for β-galactosidase and the human TFIIF subunit RAP74, and showed comparable expression levels for a given protein (data not shown). β-galactosidase was produced at significantly higher levels than RAP 74 (and other purified human proteins), most likely because of its greater stability. Therefore, pMZI3F is most appropriate for achieving reliable, low expression levels and the vector of choice for expression of SPA-tagged toxic proteins. Because there are about 3×10^5 molecules of TFIIF per cell,[13] we used the data shown in Figure 2C to estimate that about $3-5 \times 10^4$ molecules per cell of RAP74 were produced after induction with 3 μM Ponasterone A.

We expressed β-galactosidase in transient transfection experiments with otherwise identical SPA- and TAP-based vectors to compare the purification efficiencies of the SPA and TAP procedures. Our optimized purification procedures allowed for the purification to near homogeneity of β-galactosidase from four 15-cm plates (Figure 2B). The recovery of β-galactosidase activity after TEV protease cleavage was 90% when using the pMZI3F construct (SPA) and affinity chromatography on anti-FLAG M2 beads, whereas recovery was 70% when using the pMZI construct (TAP) and affinity chromatography on IgG beads. Despite good binding to calmodulin beads (50%) in the subsequent purification steps, final EGTA eluates from the calmodulin beads contained only 10 to 15% of the initial β-galactosidase activity for both the TAP and SPA systems. This somewhat low recovery was caused by inefficient elution from the calmodulin beads, as revealed by eluting the calmodulin beads with 1% SDS and analyzing the eluates by SDS-PAGE. This phenomenon has been observed with other SPA- and TAP-tagged proteins and likely results from nonspecific interactions of CBP with the Sepharose beads. Changing the salt and/or EGTA concentrations in the final elution buffer, as well as adding ethylene glycol, gave no significant improvement (data not shown).

SPA Purification of Native Protein Complexes from Human Cells. We tested the SPA system using cDNAs encoding RAP30 and RAP74, the two subunits of the general transcription factor TFIIF. TFIIF associates with RNAPII[13] and FCP1, a phosphatase that dephosphorylates the carboxy-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII.[14] Clonal EcR-RAP30(SPA) and EcR-RAP74(SPA) stable cell lines were obtained using pMZI3F constructs. As a control, β-galactosidase was also purified from an EcR-β-galactosidase(SPA) stable cell line. Comparable amounts of the same protein complex were purified using tagged RAP30 or tagged RAP74 (Figure 3A, lanes 2 and 3). As expected, mass spectrometry led to identification of RAP74 and RAP30, as well as FCP1 and the Rpb1, Rpb2, and Rpb5 subunits of RNAPII, confirming the utility of this method.

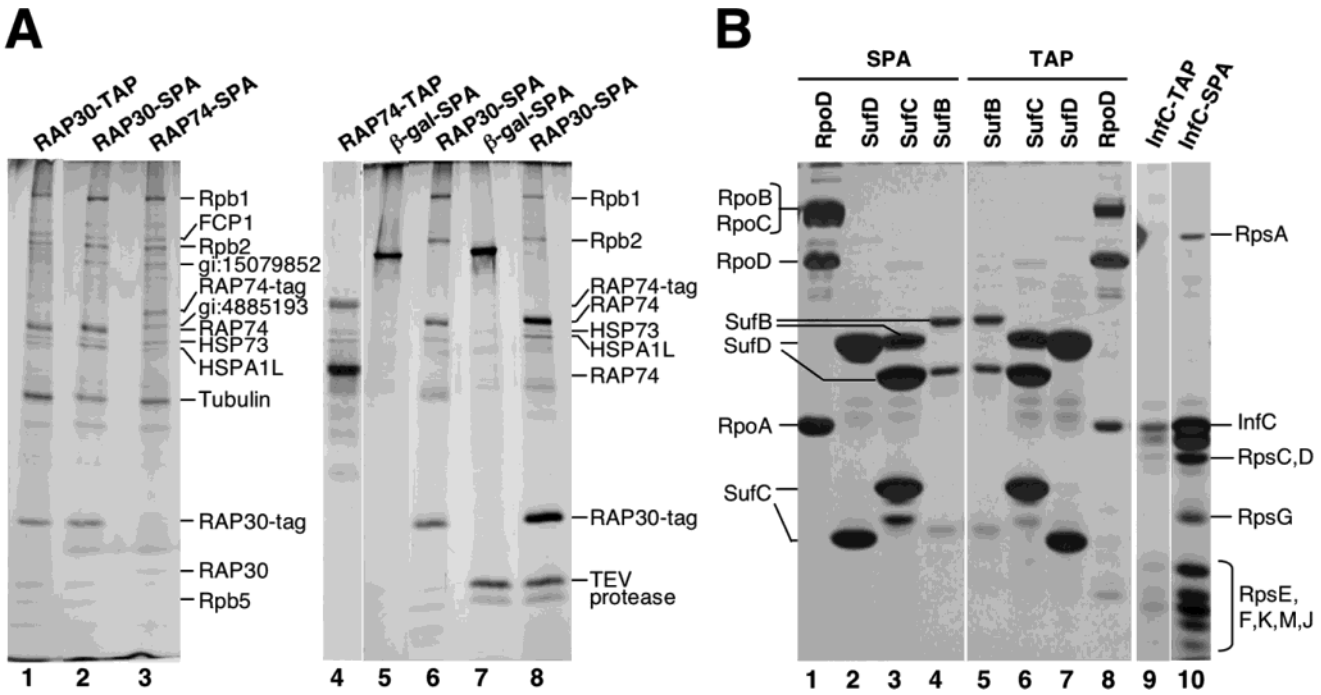


Figure 3. Purification of protein complexes using the SPA system. **A**, SDS−PAGE and silver stain analysis of protein complexes purified from human cells. The tagged proteins are indicated on the top of each lane and were expressed from pMZI3F (SPA tagged proteins) or pMZI (TAP tagged proteins). Purifications were carried out from approximately 7×10^8 cells when using induced stable cell lines (lanes 1−3 and 5−8) and from 7×10^7 cells for the transient transfection experiment (lane 4). Elutions were with EGTA (lanes 1−6) or 1% SDS (lanes 7, 8) In all cases, expression was induced for 24 h using 3 μM of Ponasterone A. **B**, SDS−PAGE analysis of protein complexes purified from *E. coli*. The SPA- or TAP-tagged protein used for each purification is indicated at the top of each lane.

Nearly identical results were obtained from a stable line expressing ecdysone-inducible RAP30 containing a TAP tag rather than a SPA tag (Figure 3A, lane 1). Also co-purifying with the TFIIF subunits were two chaperones, HSP73 and HSPA1L, and tubulin, an abundant cellular protein that is a common contaminant when we have purified other tagged proteins, including β-galactosidase (unpublished results). Two proteins of unknown function also co-purified with TFIIF: gi:15079852, which is very similar to the Rho guanine nucleotide exchange factor 1 protein; and gi:4885193, a homologue of the Drosophila dead ringer protein. The relevance of these proteins interactions with TFIIF remains to be established.

Because EGTA elution of the calmodulin column is relatively inefficient, we have subsequently eluted these columns with SDS. Aside from the presence of two low molecular weight proteins, one of which is TEV protease, backgrounds in the SDS elutions were as low as those in the EGTA elutions (Figure 3A, compare lanes 7 and 8 with lanes 5 and 6). Therefore, elution of the calmodulin column with SDS is recommended to maximize the yield and optimize the signal-to-noise ratio.

Establishing a stable line is time-consuming, taking 1−2 months, whereas expressing SPA-tagged proteins in transient transfection experiments is much faster. However, using the latter approach with pMZI-RAP74, we obtained only chaperones along with RAP74 and its degradation products, but no RNAPII or FCP1 (Figure 3A, lane 4). Transient transfection, in our conditions, leads to a high plasmid copy number in the successfully transfected cells and therefore to the overexpression of the tagged protein, which is not the case with stable cell lines. Evidently, over-producing RAP74 during transient transfection swamps out its normal interaction partners and leads to observable interactions only with the chaperones. On the other hand, we were able to purify, following transient transfection, several human proteins specifically associated with the EBNA1 protein of the Epstein−Barr virus.[15] Thus, transient transfection may be successful in certain cases, perhaps especially with viral proteins, but expressing low levels of the tagged protein in stable cell lines is more appropriate for the study of human protein interactions.

Recently, Forler et al. reported an elegant method for protein complex purification:[16] plasmids that express the TAP-tagged human protein are stably transfected into *Drosophila melanogaster* (Dm) S2 cells, the corresponding endogenous protein being suppressed by RNA interference (RNAi) before the purification. Indeed, some complexes could not be purified without the RNAi step suggesting a competition effect from the endogenous protein. Nevertheless, this phenomenon was not observed using our SPA system: the presence of an excess of the endogenous untagged protein (e.g., in the case of RAP74) did not prevent the purification of a protein complex associated with the SPA-tagged protein. This competition effect could be due to the presence of the TAP tag itself, which is about 6 times bigger than the SPA tag and thus more likely to disturb protein−protein interaction events. The use of a heterologous system in the previous study could also explain or amplify this competition effect, since it requires a high homology between the tagged target human and the corresponding endogenous *Drosophila* proteins.

The SPA Tag is also Useful for Subcellular Localization of Protein Complexes. The SPA tag provides 3 FLAG epitopes suitable for immunocytochemistry. Immunofluorescence on a stable cell line expressing $3-5 \times 10^4$ SPA-tagged RAP74 per cell localized RAP74 exclusively to the nucleus, with nucleolar exclusion, as expected for this general RNAPII transcription initiation factor (Figure 4). From this and others experiments (not shown), we estimated that at least 80% of the cells express

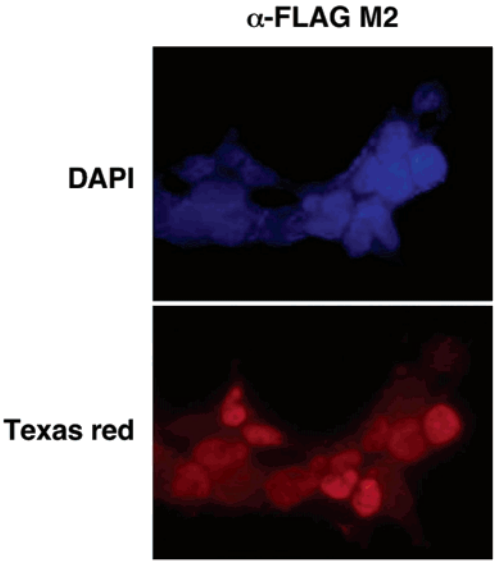


Figure 4. Cellular localization of SPA-tagged RAP74. An immunofluorescence experiment was performed on an EcR-RAP74-(SPA) stable cell line using the anti-FLAG M2 primary antibody as described in the Experimental Section. DAPI staining was used to identify nuclei and Texas red fluorescence to identify RAP74

a SPA-tagged protein after Ponasterone A induction. Therefore, systematic construction of cell lines, or even mice, expressing SPA-tagged versions of the various subunits of a protein complex should not only lead to reliable identification of the subunit composition of the complex but also establish the intracellular localization of all the subunits of that complex.

SPA Purification of Protein Complexes from *E. coli*. Although a modified two-hybrid system has been used to identify protein−protein interactions in *Helicobacter pylori*,[17] biochemical methods have not yet been used systematically to identify protein complexes in any bacterium. To tag and purify *E. coli* proteins, we constructed cassettes containing SPA tags linked to antibiotic resistance markers for selection in *E. coli*, as described in more detail in the Supporting Information. Such a cassette was amplified by PCR using primer sequences with homologies to the sequences just upstream and just downstream of the stop codon of the targeted gene and then recombined into the chromosome,[5] as illustrated in Figure 5. Tagging was confirmed by western blotting extracts from transformants with M2 antibody against the FLAG epitopes of the SPA tag. Although these recombination events are expected to occasionally disrupt either the function of the operon into which the cassette is inserted or the functioning of the tagged protein, we have successfully tagged more than 90% of 192 essential *E. coli* genes (data not shown). Therefore, this method is expected to be successful in identifying the vast majority of the protein complexes in *E. coli*.

Examples of purifications from *E. coli* extracts of SPA-tagged proteins are shown in Figure 3B. In lane 1 is shown a protein complex containing *E. coli* RNA polymerase (RNAP). When the SPA tag is present on RpoD, the σ^{70} subunit of RNAP, which is required for specific initiation at most promoters, a complex is obtained that contains only the RpoA (α), RpoB (β), RpoC (β'), and RpoD subunits of RNAP, as was expected. Although the small RpoZ (ω) subunit of RNAP ran off the bottom of the SDS gel, it was also identified by LC−MS/MS analysis of an aliquot of the material that was applied to the SDS gel (data not shown).

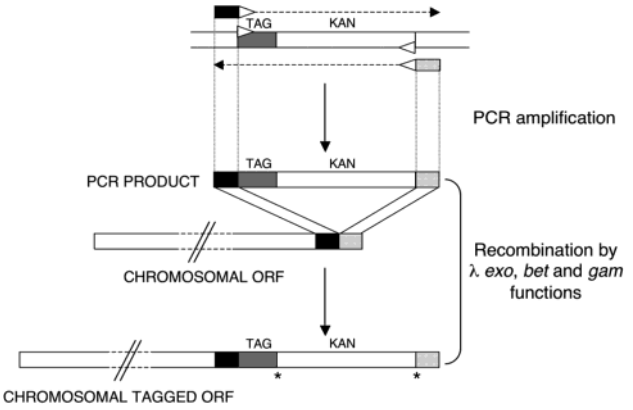


Figure 5. SPA system strategy for bacterial cells. A PCR-based method using a kanamycin-resistance cassette and the bacteriophage λ recombination system for introducing the SPA tag into the *E. coli* genome. *, stop codon.

A different protein complex containing the SufB, SufC, and SufD proteins was obtained after the SPA-tagging and purification of any one of these proteins (Figure 3B, lanes 2, 3, and 4). The adjacent *sufB*, *sufC*, and *sufD* genes, involved in iron-sulfur cluster formation, were recently also found by Outten et al. to be present in the same protein complex.[18,19] Two of these proteins contain ATP-binding cassettes characteristic of membrane transporters in many organisms, and so it may be possible with the SPA method to sometimes isolate membrane protein complexes, as well as soluble proteins.

Both RNAP complexes and the SufB, C, D complex, as well as most other complexes that we have purified (data not shown), could be isolated with virtually identical yields when a TAP tag was used instead of a SPA tag (compare lanes 1−4 with lanes 5−8 in Figure 3B). However, there were cases in which only the SPA tag was successful, perhaps because of its smaller size, and an example is shown in Figure 3B, lanes 9 and 10. The translational initiation factor IF3, encoded by the *infC* gene, co-purified efficiently with proteins of the small ribosomal subunit, with which it associates, only when it was derivatized with a SPA tag and not with a TAP tag. Nevertheless, SPA and TAP tags are equally suitable for most *E. coli* proteins.

The SPA system could also, in principle, be used for tagging and purifying the natural protein complexes of other bacterial species provided that a plasmid based library exists or can be created for that other bacterium that can be propagated both in *E. coli* and as a low copy plasmid in the other bacterium. The bacteriophage λ recombination system can be used to tag the genes of the other bacterium on the shuttle plasmids in *E. coli*, and then the plasmids can be moved back into the other bacterial species for the purification and identification of protein complexes. Systematic, large-scale projects of this kind should contribute substantially to our understanding of the biology of *E. coli* and other bacteria.

Acknowledgment. This work was supported by grants to JG from the Canadian Institutes of Health Research, the Ontario Challenge Fund, and Genome Canada.

Supporting Information Available: Description of the construction of the mammalian and bacterial SPA vectors; oligonucleotide sequences used in this study; the method for gene-specific tagging in *E. coli*; and detailed protocols for SPA purification from human and *E. coli* cells. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Claverie, J. M. *Science* **2001**, *291*, 1255−1257.
(2) Gavin, A. C.; Bosche, M.; Krause, R.; Grandi, P.; Marzioch, M.; Bauer, A.; Schultz, J.; Rick, J. M.; Michon, A. M.; Cruciat, C. M.; Remor, M.; Hofert, C.; Schelder, M.; Brajenovic, M.; Ruffner, H.; Merino, A.; Klein, K.; Hudak, M.; Dickson, D.; Rudi, T.; Gnau, V.; Bauch, A.; Bastuck, S.; Huhse, B.; Leutwein, C.; Heurtier, M. A.; Copley, R. R.; Edelmann, A.; Querfurth, E.; Rybin, V.; Drewes, G.; Raida, M.; Bouwmeester, T.; Bork, P.; Seraphin, B.; Kuster, B.; Neubauer, G.; Superti-Furga, G. *Nature* **2002**, *415*, 141−147.
(3) Strausberg, R. L.; Feingold, E. A.; Klausner, R. D.; Collins, F. S. *Science* **1999**, *286*, 455−457.
(4) Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6640−6645.
(5) Yu, D.; Ellis, H. M.; Lee, E. C.; Jenkins, N. A.; Copeland, N. G.; Court, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5978−5983.
(6) Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Seraphin, B. *Nat. Biotechnol.* **1999**, *17*, 1030−1032.
(7) Coulombe, B.; Killeen, M.; Liljelund, P.; Honda, B.; Xiao, H.; Ingles, C. J.; Greenblatt, J. *Gene Expr.* **1992**, *2*, 99−110.
(8) Mann, M.; Hojrup, P.; Roepstorff, P. *Biol. Mass Spectrom*. **1993**, *22*, 338−345.
(9) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual, Second Edition*; Cold Spring Harbor: New York, 1989.
(10) Hernan, R.; Heuermann, K.; Brizzard, B. *Biotechniques* **2000**, *28*, 789−793.
(11) Emery, D. W.; Yannaki, E.; Tubb, J.; Stamatoyannopoulos, G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9150−9155.
(12) Chung, J. H.; Bell, A. C.; Felsenfeld, G. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 575−580.
(13) Burton, Z. F.; Killeen, M.; Sopta, M.; Ortolan, L. G.; Greenblatt, J. *Mol. Cell Biol.* **1988**, *8*, 1602−1613.
(14) Archambault, J.; Pan, G.; Dahmus, G. K.; Cartier, M.; Marshall, N.; Zhang, S.; Dahmus, M. E.; Greenblatt, J. *J. Biol. Chem*. **1998**, *273*, 27593−27601.
(15) Holowaty, M. N.; Zeghouf, M.; Wu, H.; Tellam, J.; Athanasopoulos, V.; Greenblatt, J.; Frappier, L. *J. Biol. Chem*. **2003**, *278*, 29987−29994.
(16) Forler, D.; Kocher, T.; Rode, M.; Gentzel, M.; Izaurralde, E.; Wilm, M. *Nat. Biotechnol.* **2003**, *21*, 89−92.
(17) Rain, J. C.; Selig, L.; De Reuse, H.; Battaglia, V.; Reverdy, C.; Simon, S.; Lenzen, G.; Petel, F.; Wojcik, J.; Schachter, V.; Chemama, Y.; Labigne, A.; Legrain, P. *Nature* **2001**, *409*, 211−215.
(18) Takahashi, Y.; Tokumoto, U. *J. Biol. Chem*. **2002**, *277*, 28380−28383.
(19) Outten, F. W.; Wood, M. J.; Munoz, F. M.; Storz, G. *J. Biol. Chem*. **2003**, *278*, 45713−45719.

PR034084X