Short communication

Tandem affinity purification of exosome and replication factor C complexes from the non-human infectious kinetoplastid parasite *Crithidia fasciculata*

Wakisa Kipandula [a, b], Terry K. Smith [a] and Stuart A. MacNeill [a, *]

[a] Biomedical Sciences Research Complex, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK.

[b] Department of Biomedical Sciences, College of Medicine, University of Malawi, Private Bag 360, Chichiri, Blantyre 3, Malawi.

* Corresponding author:

Email: sam31@st-and.ac.uk

Tel. (+44) 1334 467268

Abstract

Kinetoplastid parasites are responsible for a range of diseases with significant global impact. *Trypanosoma brucei* and *Trypanosoma cruzi* cause human African trypanosomiasis and Chagas disease, respectively, while various *Leishmania* species are responsible for cutaneous, mucocutaneous and visceral leishmaniasis. Understanding the biology of these organisms is key for effective diagnosis, prophylaxis and treatment. The insect parasite *Crithidia fasciculata* offers a safe and low-cost alternative for studies of kinetoplastid biology. *C. fasciculata* does not infect humans, can be cultured to high yields in inexpensive serum-free medium in a standard laboratory, and has a completely sequenced publically available genome. Taking advantage of these features, however, requires the adaptation of existing methods of analysis to *C. fasciculata.* Tandem affinity purification is a widely used method that allows for the rapid purification of intact protein complexes under native conditions. Here we report the application of tandem affinity purification to *C. fasciculata* for the first time, demonstrating the effectiveness of the technique by purifying both the intact exosome and replication factor C complexes. Adding tandem affinity purification to the *C. fasciculata* toolbox significantly enhances the utility of this excellent model system.

Key words: *Crithidia fasciculata*; tandem affinity purification; kinetoplastid; exosome; DNA replication

1. Introduction

Kinetoplastids are a class of protists belonging to the Excavata, which branched away from the Plantae and Opisthokonta extremely early in eukaryotic evolution (1). They form a family distinguished by having only a single flagellum and include three major human pathogenic parasites responsible for chronic severe disease in millions of people worldwide (2). These are *Trypanosoma brucei* which causes human African trypanosomiasis (HAT) and animal trypanosomiasis (sleeping sickness and Nagana, respectively) transmitted by tsetse flies, *T. cruzi* which causes South American trypanosomiasis (Chagas disease) and is transmitted by triatomine bugs, and various *Leishmania* species, which cause cutaneous, mucocutaneous and visceral leishmaniasis, which are transmitted by sandflies (2).

While significant inroads into understanding the biology of these organisms have been made over the last 30 years (3), more rapid progress has been limited by several factors. Most significantly, culturing these human infectious pathogens safely in the laboratory requires dedicated facilities at biosafety level 2 or 3 that are costly to build and maintain. In addition to this, growth of *Trypanosoma* and *Leishmania* species requires relatively expensive serum-containing medium. Together, these concerns have largely restricted kinetoplastid research to resource-rich countries.

The non-human infective parasite *Crithidia fasciculata* (4) offers a very useful model to study biological cellular processes unique to kinetoplastids. *C. fasciculata* is closely related to the pathogenic trypanosomatids (5), but crucially,

this monoxenous non-human infectious organism can be grown safely in a standard laboratory without the need for biosafety level 2 or 3 facilities. A further advantage of *C. fasciculata* is that it can be easily grown to high yields in relatively inexpensive, serum-free liquid medium (complex or defined). Lastly, the complete genome sequence of *C. fasciculata* has been determined by the Beverley group (Washington University School of Medicine) and though unpublished, is publically available in the TriTrypDB database (6), thereby facilitating genome or proteome-wide studies of this organism. Thus *C. fasciculata* offers an excellent, safe, low-cost alternative to working directly with pathogenic trypanosomatids.

Tandem affinity purification offers a rapid and efficient means of isolating tagged proteins from crude cell extracts under native conditions and is very well suited to the isolation and characterisation of multiprotein complexes (7). In this report, we describe the application of tandem affinity purification in *C. fasciculata* for the first time, using two well-defined stable multiprotein complexes, the exosome (8) and replication factor C (9), to demonstrate the efficiency of the process.

The exosome is an essential and well-conserved 3'-5' exonuclease complex that catalyses the processing and/or degradation of a wide range of cellular RNAs (8). The core of the exosome comprises of nine subunits arranged as a six-membered ring, with a three-membered cap. The six proteins forming the ring are distantly related to the *Escherichia coli* RNAse PH protein, while the three forming the cap contain S1 RNA binding domains. In the yeast cytoplasm a tenth RNAse protein (Rrp44) is recruited to this structure to form the 10-subunit cytoplasmic exosome, while in the nucleus an additional RNase (Rrp6) is recruited to form the fully functional 11-subunit nuclear exosome (8).

Stable 11-subunit exosome complexes have been purified from two kinetoplastid organisms: *T. brucei* (10,11) and *L. tarentolae* (12). In both cases, the nine core exosome subunits (forming the ring and cap structures) are present, as is the kinetoplastid RRP6 protein and an additional factor, EAP3, that is distantly related to the yeast exosome cofactor Rrp47. Although both kinetoplastids encode an Rrp44 homologue, this is absent from the purified complexes, consistent with it interacting only weakly with the core (10).

Replication factor C (RFC) is a five-subunit complex that catalyses the ATP-dependent loading of the sliding clamp processivity factor PCNA onto DNA in DNA replication and repair (9). Human and yeast RFC complexes comprise five subunits (one large subunit Rfc1 and four small subunits Rfc2-Rfc5) that together form a heteropentameric structure. The four small subunits of RFC are also components of three alternative RFC-like complexes: Ctf18-RFC, Elg1-RFC and Rad17-RFC, with diverse roles in chromosome cohesion, PCNA unloading and checkpoint control, respectively (9). The last of these three complexes interacts, not with PCNA, but with the PCNA-related 9-1-1 complex. Amongst kinetoplastid organisms, PCNA has been characterised in *T. brucei* (13,14) and components of the 9-1-1 complex in *L. major* (15-18), but RFC remains uncharacterised. Candidate RFC1-RFC5 proteins can be readily identified on the basis of protein sequence similarity, as can a putative RAD17 protein (Table 1).

Here, we apply tandem affinity purification to pull-down exosome and RFC complexes from *C. fasciculata*. Mass spectrometric analysis of purified material identified all nine exosome core and lid proteins plus RRP6, and all five RFC subunits plus RAD17. The results demonstrate the applicability of tandem affinity

purification to *C. fasciculata* and again highlight the potential of this non-infectious organism as a model for the pathogenic trypanosomatids.

2. Materials and methods

2.1 Organisms and reagents

C. fasciculata clone HS6 (19) promastigotes were grown at 27°C with gentle agitation in serum-free defined media containing 5 g/L yeast extract, 4 g/L tryptone, 15 g/L sucrose, 4.4 g/L triethanolamine, 0.5% Tween 80 and 10 µg/mL haemin. DNA manipulation procedures were conducted using *E. coli* DH5α. Chemicals and reagents were purchased from Sigma-Aldrich. Enzymes for DNA manipulation were purchased from New England Biolabs, Promega or Bioline. Oligonucleotide primers were synthesised by IDT.

2.2 Construction of expression vectors

Plasmid pNUS-PTPcH was constructed as follows. Sequences encoding the PTP tag were amplified by PCR from plasmid pC-PTP-NEO (20) using primers CfPTP-5ENXVN and CfPTP-3E (see Supplementary Table S1 for sequences), digested with *Eco*RI and cloned into pNUS-SPnH (21), replacing the 102 bp *Eco*RI-*Eco*RI region in this vector that encodes a signal peptide. The newly cloned PTP sequence lacks the *Afl*II, *Eco*RI and *Bam*HI sites internal to the PTP tag sequence in pC-PTP-NEO, includes unique restriction sites for *Nde*I, *Xho*I, *Eco*RV and *Not*I

upstream of the PTP sequence for fusing target sequences to the tag, and contains multiple stop codons at the 3' end of the construct (see Supplementary Information, Figures S1 and S2).

Plasmids pNUS-PTPcH-CfRfc3 and pNUS-PTPcH-CfRrp4 for expression of RFC3-PTP and RRP4-PTP proteins were constructed by amplifying the *RFC3* (Gene ID: CFAC1_300082900) and *RRP4* (Gene ID: CFAC1_110005300) ORFs by PCR from *C. fasciculata* genomic DNA using primers CfRfc3-5Nde and CfRfc3-3Not or CfRrp4-5Nde and CfRrp4-3Not (see Supplementary information, Table S1 for primer sequences), digesting the products with *Nde*I and *Not*I and cloning into pNUS-PTPcH, replacing the 19 bp *Nde*I-*No*tI region of the multiple cloning site. The *RFC3* and *RRP4* ORFs were sequenced to confirm the absence of errors introduced during the cloning procedure.

2.3 Transfection and generation of cell lines

For transfection, *C. fasciculata* HS6 promastigotes were grown to log phase at 27°C and harvested at a density of ~1 x 10^7 cells/mL by centrifugation at 1600 *g* for 5 minutes. The cell pellet (~2 x 10^7 cells/mL) was suspended in 100 µL of Human T-cell Nucleofector™ solution (Lonza) and transferred to a 0.4 cm cuvette containing 15-60 µg of purified pNUS-PTPcH-CfRfc3 or pNUS-PTPcH-CfRrp4 supercoiled plasmid DNA. The mixture was electroporated using an Amaxa Nucleofector instrument (program X-014). Electroporated cells were left on ice for 5 minutes and transferred into a culture flask containing 5 mL of fresh medium and incubated at 27°C to recover. After 24 hours of recovery, the cell culture was

supplemented with 5 mL of fresh medium containing 25 µg/mL hygromycin. Resistant cell lines were subsequently grown with 50 µg/mL of hygromycin and viable clones observed within 5 to 10 days. Resistant cell lines were maintained by supplementing the culture with fresh medium containing 50 µg/mL hygromycin.

2.4 Tandem affinity purification

To confirm expression of PTP fusion proteins, cell lysates were prepared from exponentially-growing cultures and analysed by SDS-PAGE and Western blotting. Briefly, a 1 mL sample (~1 x 10^7 cells) was harvested by centrifugation at 2000 *g* for 5 minutes. The cell pellet was then heated to 95°C for 5 minutes in 2 X SDS-PAGE sample buffer, prior to SDS-PAGE and blotting to a PVDF membrane (Bio-Rad). Blots were probed with PAP reagent (Peroxidase Anti-Peroxidase Soluble Complex antibody, Sigma-Aldrich) at a 1:2000 dilution in PBS containing 0.05% Tween and 5% skimmed milk, followed by goat-anti-rabbit IgG (H+L) DyLight™ 680 conjugated secondary antibody (1:10000) (Thermo Fisher Scientific) according to the manufacturer's instructions and scanned on Odyssey scanner (Li-Cor Biosciences) utilizing the 700 nm channel.

Purification of the target proteins and their interacting partners in the complexes was conducted according to the method described in (20) but with minor modifications as outlined here. Briefly, a 2.5 L culture of transfected *C. fasciculata* was grown in serum-free defined medium supplemented with 50 µg/mL hygromycin until the cell number reached ~ 2 x 10^7 cells/mL. The cells were then

harvested by centrifugation at 800 *g* for 10 minutes and the cell pellet washed three times with 5 mL ice-cold PBS before being resuspended in 1.5 volumes of ice-cold PA-300 buffer comprising 150 mM sucrose, 300 mM potassium chloride, 40 mM potassium L-glutamate, 3 mM $MgCl_2$, 20 mM HEPES-KOH (pH 7.7), 2 mM dithiothreitol, 0.1% Tween 20 and 1X cOmplete™ Mini EDTA-free protease inhibitor cocktail (Roche). Cells were dounced in continuous strokes for 5 minutes on ice in a cold room using a 7 mL dounce homogenizer (Sigma) and subsequently centrifuged at 20,500 *g* for 10 minutes at 4°C. For IgG affinity chromatography, the resultant lysate was filtered straight into a 10 mL Poly-Prep® chromatography column (Bio-Rad) containing 200 µL of IgG Sepharose 6 Fastflow beads (GE Healthcare) pre-equilibrated with PA-300 buffer. The top and the bottom of the column were sealed with Parafilm and the column was rotated for 2 hours at 4°C allowing the PTP tagged protein to bind to the IgG beads. The beads were later washed twice with 10 mL PA-300 before being equilibrated with 8mL TEV buffer (20). To cleave the IgG matrix bound proteins, 20 µL of 10 U/µl AcTEV™ protease (Invitrogen) was diluted in 2 mL TEV buffer and added to the column. After overnight rotation at 4°C, the TEV and column dead-volume were eluted by washing the IgG beads with 4 mL PC-150 buffer (20) containing 1X cOmplete™ Mini EDTA-free protease inhibitor cocktail and 1 mM $CaCl_2$. The mixture was added to a second equilibrated Poly-Prep® column containing 200 µL Anti-ProtC Affinity Matrix beads (Roche) and was rotated for 2 hrs in a cold room to allow the tagged protein to bind to the matrix. After washing the anti-ProtC matrix six times with 10 mL PC-150 buffer, purified proteins were eluted with a 1.8 mL EGTA/EDTA buffer (20) at room temperature. To concentrate the

eluted proteins, eluates were incubated with 30 µL of StrataClean Resin (Agilent) and pelleted at 5,000 *g* for 1 minute. The beads were then resuspended in 20 µL 4X NuPAGE™ LDS sample buffer (Thermo Fisher Scientific) and boiled at 95°C to release the proteins. 20 µL of the sample was loaded onto a NuPAGE™ 4-12% Bis-Tris pre-cast SDS-PAGE gel and the proteins were resolved by SDS-PAGE before stained with SYPRO® Ruby stain (Thermo Fisher Scientific). The stained gels were visualised using a UV transilluminator. Sections of the gel were then excised, digested by trypsin and the resulting peptides analysed in-house by mass spectrometry.

3. Results and discussion

3.1 Construction of a PTP expression vector

Tandem affinity purification offers a rapid and efficient means of isolating tagged proteins from crude cell extracts under native conditions (7). The standard C-terminal TAP tag comprises – in order – a calmodulin-binding peptide, a TEV protease cleavage site and two protein A domains (7). In *T. brucei*, however, a modified tandem affinity purification cassette called the PTP tag has been reported to increase yields (20). The PTP tag has the calmodulin-binding peptide replaced with the protein C epitope.

In order to express C-terminally PTP tagged proteins in *C. fasciculata*, we modified the previously described *C. fasciculata* shuttle vector pNUS-SPnH (21) by replacing sequences encoding a signal peptide with those encoding the PTP

tag derived from plasmid pC-PTP-NEO (20). As with pNUS-SPnH, expression of a PTP tagged protein from the new plasmid, which we named pNUS-PTPnH, is driven by the *C. fasciculata* phosphoglycerate kinase gene *PGKB* promoter. pNUS-PTPnH was constructed to include unique restriction sites for *Nde*I, *Xho*I, *Eco*RV and *Not*I upstream of the PTP tag sequence for convenient cloning (see Supplementary information, Figures S1 and S2) and confers hygromycin resistance on *C. fasciculata* cells. It was previously shown (21) that plasmid pNUS-H1-GFP (from which pNUS-SPnH was later derived) was maintained episomally in *C. fasciculata* cells. i.e. that chromosomal integration did not occur (21). We did not test pNUS-PTPnH for this property, but expect that this plasmid is also episomally maintained, given its similarity to pNUS-H1-GFP.

3.2 Tandem affinity purification

Exosome complexes have previously been purified from *T. brucei* (11) and *L. tarentolae* (12), allowing the identification of 11 subunits. Putative *C. fasciculata* orthologues of these proteins can be readily identified by sequence database searches (see Table 1 for accession numbers). RFC has not been previously characterised in any kinetoplastid organism but genes encoding candidate subunits of the canonical RFC complex (Rfc1-Rfc5), as well as Rad17 protein (the large subunit of the Rad17-RFC complex) can also be identified (Table 1).

To investigate whether we could isolate the *C. fasciculata* exosome by tandem affinity purification, sequences encoding the putative RRP4 subunit were cloned into pNUS-PTPcH for expression of a RRP4-PTP fusion protein. The resulting

plasmid, pNUS-PTPcH-CfRrp4, was transfected into *C. fasciculata* and transfected cell lines selected using 25 μg/mL hygromycin (see Materials and methods). Expression of the RRP4-PTP fusion was confirmed by immunoblotting using PAP (peroxidase anti-peroxidase soluble complex) antibodies (Figure 1A) and tandem affinity purification performed from a 2.5 litre *C. fasciculata* culture (at ~ 2 x 10^7 cells/mL) essentially as described elsewhere (20), but with minor modifications (see Materials and methods, section 2.4). SDS-PAGE separation followed by MS/MS analysis identified RRP6, EAP1, RRP45, RRP40, RRP4-PTP, RRP41B, CSL4, EAP2, RRP41A and EAP4 with peptide coverage ranging from 11 – 83% and MASCOT scores from 326-1943 (see Figure 1B, Table 1 and Supplementary information for details). We did not find EAP3 subunit in our experiments, possibly due to the weak interaction of EAP3 with the core exosome complex observed previously (10-12), but did identify a number of non-exosomal proteins with significant MASCOT scores, specifically two HSP70-like proteins, several members of a β-fructofuranosidase family and α- and β-tubulin (see Supplementary Table S2). As these proteins were also detected in purified RFC preparations (see below), they were assumed to be contaminants.

For purification of RFC, plasmid pNUS-PTPcH-CfRfc3 was constructed, transfected into *C. fasciculata*, and transfectants selected using hygromycin, as above. Expression of the RFC3-PTP protein was confirmed by immunoblotting (Figure 1A). Tandem affinity purification led to the isolation of all five subunits of the canonical RFC complex with peptide coverage ranging from 21 – 77% and MASCOT scores from 586-1507 (see Figure 1B, Table 1 and Supplementary information for details). In addition, the alternative RFC complex component

Rad17 was also detected (peptide coverage 6%, MASCOT score 174), as were the contaminating proteins described above. Taken together, the results demonstrate the utility of plasmid pNUS-PTPcH for expression of C-terminal PTP fusion proteins suitable for tandem affinity purification from *C. fasciculata*.

3.3 Conclusions

In summary, we have constructed an expression vector that can be used to express PTP-fusion proteins in the non-human infectious kinetoplastid parasite *C. fasciculata* and have shown that these proteins and their interacting partners can be readily purified using standard PTP purification procedures (20). The successful trialling of this method adds significantly to the tools available for studying *C. fasciculata* biology and again underpins the potential usefulness of this organism as a safe, low-cost model for probing kinetoplastid biology.

Acknowledgements

We are indebted to Dr Emmanuel Tetaud (MFP-CNRS-UMR5234, University of Bordeaux) for providing the pNUS-SPnH plasmid and to Prof. Stephen M. Beverley (Washington University School of Medicine, St Louis, USA) for the *C. fasciculata* genome sequence deposited in the TriTrypDB database. This work was supported through the Global Health Implementation programme at the University of St Andrews. We are grateful to the University of St Andrews mass

spectrometry facility for collecting and processing mass spectrometry data and to other members of the TKS and SM groups for their assistance with this project.

Conflicts of interest: none.

References

1. Simpson, A.G., Stevens, J.R. and Lukes, J. (2006) The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol*, **22**, 168-174.
2. Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gurtler, R.E., McKerrow, J., Reed, S. and Tarleton, R. (2008) Kinetoplastids: related protozoan pathogens, different diseases. *J Clin Invest*, **118**, 1301-1310.
3. Barrett, M.P. and Croft, S.L. (2012) Management of trypanosomiasis and leishmaniasis. *Br Med Bull*, **104**, 175-196.
4. Léger, L. (1902) Sur un flagellé parasite de l'*Anopheles maculipennis*. *CR Seances Soc Biol Fil*, **54**, 354-356.
5. Maslov, D.A., Votypka, J., Yurchenko, V. and Lukes, J. (2013) Diversity and phylogeny of insect trypanosomatids: all that is hidden shall be revealed. *Trends Parasitol*, **29**, 43-52.
6. Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X. *et al.* (2010) TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res*, **38**, D457-462.
7. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M. and Seraphin, B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, **24**, 218-229.
8. Buttner, K., Wenig, K. and Hopfner, K.P. (2006) The exosome: a macromolecular cage for controlled RNA degradation. *Mol Microbiol*, **61**, 1372-1379.

9. Majka, J. and Burgers, P.M. (2004) The PCNA-RFC families of DNA clamps and clamp loaders. *Prog Nucleic Acid Res Mol Biol*, **78**, 227-260.
10. Estevez, A.M., Lehner, B., Sanderson, C.M., Ruppert, T. and Clayton, C. (2003) The roles of intersubunit interactions in exosome stability. *J Biol Chem*, **278**, 34943-34951.
11. Estevez, A.M., Kempf, T. and Clayton, C. (2001) The exosome of *Trypanosoma brucei*. *EMBO J*, **20**, 3831-3839.
12. Cristodero, M., Bottcher, B., Diepholz, M., Scheffzek, K. and Clayton, C. (2008) The *Leishmania tarentolae* exosome: purification and structural analysis by electron microscopy. *Mol Biochem Parasit*, **159**, 24-29.
13. Valenciano, A.L., Ramsey, A.C. and Mackey, Z.B. (2015) Deviating the level of proliferating cell nuclear antigen in *Trypanosoma brucei* elicits distinct mechanisms for inhibiting proliferation and cell cycle progression. *Cell Cycle*, **14**, 674-688.
14. Kaufmann, D., Gassen, A., Maiser, A., Leonhardt, H. and Janzen, C.J. (2012) Regulation and spatial organization of PCNA in *Trypanosoma brucei*. *Biochem Biophys Res Commun*, **419**, 698-702.
15. Damasceno, J.D., Obonaga, R., Santos, E.V., Scott, A., McCulloch, R. and Tosi, L.R. (2016) Functional compartmentalization of Rad9 and Hus1 reveals diverse assembly of the 9-1-1 complex components during the DNA damage response in *Leishmania*. *Mol Microbiol*, **101**, 1054-1068.
16. MacNeill, S.A. (2014) Identification of a candidate Rad1 subunit for the kinetoplastid 9-1-1 (Rad9-Hus1-Rad1) complex. *Biology (Basel)*, **3**, 922-927.
17. Damasceno, J.D., Nunes, V.S. and Tosi, L.R. (2013) LmHus1 is required for the DNA damage response in *Leishmania major* and forms a complex with an unusual Rad9 homologue. *Mol Microbiol*, **90**, 1074-1087.
18. Nunes, V.S., Damasceno, J.D., Freire, R. and Tosi, L.R. (2011) The Hus1 homologue of *Leishmania major* encodes a nuclear protein that participates in DNA damage response. *Mol Biochem Parasit*, **177**, 65-69.

19. Shim, H. and Fairlamb, A.H. (1988) Levels of polyamines, glutathione and glutathione-spermidine conjugates during growth of the insect trypanosomatid *Crithidia fasciculata*. *J Gen Microbiol*, **134**, 807-817.
20. Schimanski, B., Nguyen, T.N. and Gunzl, A. (2005) Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell*, **4**, 1942-1950.
21. Tetaud, E., Lecuix, I., Sheldrake, T., Baltz, T. and Fairlamb, A.H. (2002) A new expression vector for *Crithidia fasciculata* and *Leishmania*. *Mol Biochem Parasit*, **120**, 195-204.

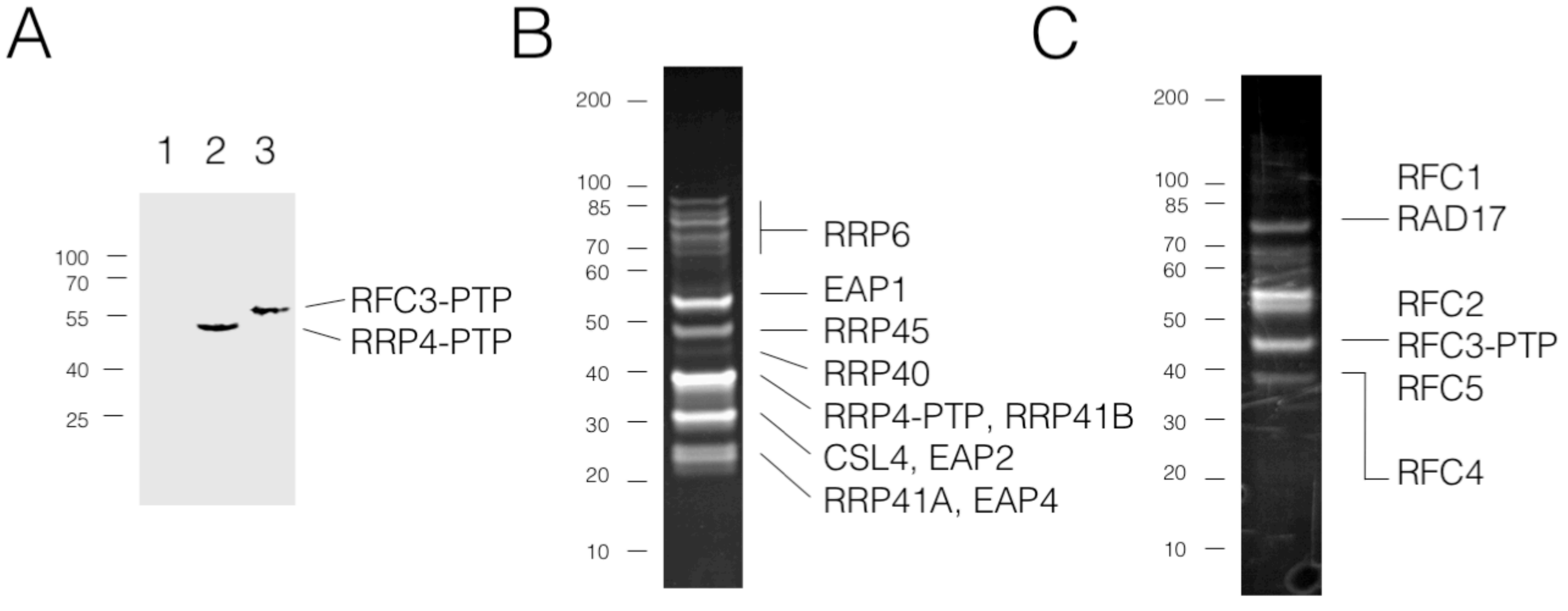
Figure 1. A. Immunoblotting with PAP reagent (Peroxidase Anti-Peroxidase Soluble Complex antibody) of *C. fasciculata* cell lines carrying pNUS-PTPcH (lane 1), pNUS-PTPcH-RRP4 (lane 2) or pNUS-PTPcH-RFC3 (lane 3). **B.** Tandem affinity purification of RRP4-PTP protein pulls down nine other exosome components. **C.** Tandem affinity purification of RFC3-PTP pulls down the canonical RFC subunits RFC1, RFC2, RFC4 and RFC5, as well as the alternative RFC complex subunit RAD17. SDS-PAGE gels were stained with SYPRO® Ruby. Molecular weight markers are indicated to the left of each gel (sizes in kDa).

Table 1: Mass spectrometry data			
Protein	Gene ID	Mascot score	Coverage
Exosome complex components			
RRP41A	CFAC1_280032800	988	83%
RRP41B	CFAC1_160013900	326	34%
RRP6	CFAC1_290060300	1943	62%
EAP1	CFAC1_170027500	950	34%
RRP45	CFAC1_240024400	829	31%
RRP40	CFAC1_030007200	1173	70%
RRP4	CFAC1_110005300	1330	48%
CSL4	CFAC1_150031200	573	40%
EAP2	CFAC1_300052700	585	33%
EAP4	CFAC1_280054900	691	40%
RFC and RFC-like complex components			
RFC1	CFAC1_210019200	586	21%
RFC2	CFAC1_260050500	1144	59%
RFC3	CFAC1_300082900	1507	77%
RFC4	CFAC1_230046600	941	70%
RFC5	CFAC1_280077100	997	52%
RAD17	CFAC1_260035800	174	6%

396

397

Table 1: Mass spectrometry data			
Protein	Gene ID	Mascot score	Coverage
Exosome complex components			
RRP41A	CFAC1_280032800	988	83%
RRP41B	CFAC1_160013900	326	34%
RRP6	CFAC1_290060300	1943	62%
EAP1	CFAC1_170027500	950	34%
RRP45	CFAC1_240024400	829	31%
RRP40	CFAC1_030007200	1173	70%
RRP4	CFAC1_110005300	1330	48%
CSL4	CFAC1_150031200	573	40%
EAP2	CFAC1_300052700	585	33%
EAP4	CFAC1_280054900	691	40%
RFC and RFC-like complex components			
RFC1	CFAC1_210019200	586	21%
RFC2	CFAC1_260050500	1144	59%
RFC3	CFAC1_300082900	1507	77%
RFC4	CFAC1_230046600	941	70%
RFC5	CFAC1_280077100	997	52%
RAD17	CFAC1_260035800	174	6%


A
1 2 3
100
70
55
40
25
RFC3-PTP
RRP4-PTP
B
200
100
85
70
60
50
40
30
20
10
RRP6
EAP1
RRP45
RRP40
RRP4-PTP, RRP41B
CSL4, EAP2
RRP41A, EAP4
C
200
100
85
70
60
50
40
30
20
10
RFC1
RAD17
RFC2
RFC3-PTP
RFC5
RFC4

Kipandula W, Smith TK and MacNeill SA

Supplementary information

Supplementary Tables

Supplementary Table S1: Oligonucleotide primers used in this work		
Name	Sequence	Restriction sites
CfPTP-5ENXVN	5′-CGAATTCCATATGCTCGAGGATATCGGCGGCCGCGAAGATCAGGTCGATCCTCGTCTT-3′	*Eco*RI
CfPTP-3E	5′-CCGAATTCCTATTATCAGGTTGACTTCCCCGCGGAGTTCGCGTCTACT-3′	*Eco*RI
CfRfc3-5Nde	5′-GGTGGTGGTGCATATGGCAACTTCGAAGCAGGCAGAG-3′	*Nde*I
CfRfc3-3Not	5′-GGTGGTGGTGGCGGCCGCCAGCGCTGCAGTCGCCGGCCACGGC-3′	*Not*I
CfRrp4-5Nde	5′-GGTGGTGGTGCATATGTCGTCAGGAGTGGTGATTGTC-3′	*Nde*I
CfRrp4-3Not	5′-GGTGGTGGTGGCGGCCGCCCTGGCGGCGGCGCTTCACACCAAT-3′	*Not*I
PNUS-SEQ-F	5′-GACGGCCAGTGCCAAGCTTAC-3′	N/A
PNUS-SEQ-R	5′-ACAAATTCTCGCTAGCAGTAG-3′	N/A
PNUS-PCR-F	5′-ACATTCCGCTCTGTCCACTTCG-3′	N/A
PNUS-PCR-R	5′-GCGAGTCACCGTGGAGCAGCTG-3′	N/A

Supplementary Table S2: List of additional proteins detected in purified exosome and RFC fractions				
Protein	Predicted MW	Gene ID	Mascot score	Coverage
β-fructofuranosidase-like proteins	75 kDa	CFAC1_030010000 CFAC1_030010200	495 237	13% 6%
HSP70-like proteins	72 kDa	CFAC1_3000044700	709	34%
	72 kDa	CFAC1_3000021300	628	27%
α-tubulin	50 kDa	CFAC1_080015700	125	6%
β-tubulin	53 kDa	CFAC1_130028700	270	15%
Flagellar calcium binding protein	23 kDa	CFAC1_120017200	148	17%
Note: β-fructofuranosidase-like and HSP70-like proteins and tubulins were detected in both exosome and RFC preparations.				

Supplementary Figures

Figure S1: Structure of the PTP cassette in pNUS-PTPcH. Abbreviations: PC (protein C), T (TEV cleavage site).

Figure S2: DNA sequence of the *Eco*RI-*Eco*RI insert in pNUS-PTPcH

GAA TTC CAT ATG CTC GAG GAT ATC GGC GGC CGC GAA GAT CAG GTC
GAT CCT CGT CTT ATT GAT GGG AAA TAT GAT ATT CCA ACT ACT GCT
AGC GAG AAT TTG TAT TTT CAG GGT GAG CTC AAA ACC GCG GCT CTT
GCG CAA CAC GAT GAA GCC GTG GAC AAC AAA TTC AAC AAA GAA CAA
CAA AAC GCG TTC TAT GAG ATC TTA CAT TTA CCT AAC TTA AAC GAA
GAA CAA CGA AAC GCC TTC ATC CAA AGT TTA AAA GAT GAC CCA AGC
CAA AGC GCT AAC CTT TTA GCA GAA GCT AAA AAG CTA AAT GAT GCT
CAG GCG CCG AAA GTA GAC AAC AAA TTC AAC AAA GAA CAA CAA AAC
GCG TTC TAT GAG ATC TTA CAT TTA CCT AAC TTA AAC GAA GAA CAA
CGA AAC GCC TTC ATC CAA AGT TTA AAA GAT GAC CCA AGC CAA AGC
GCT AAC CTT TTA GCA GAA GCT AAA AAG CTA AAT GAT GCT CAG GCG
CCG AAA GTA GAC GCG AAC TCC GCG GGG AAG TCA ACC TGA TAA TAG
GAA TTC

*Eco*RI, *Nde*I, *Xho*I, *Eco*RV and *Not*I sites are shown in bold, red, blue, green and purple text respectively. Individual underlined bases indicate sequence changes designed to remove internal *Bam*HI and *Eco*RI sites in plasmid pC-PTP-NEO (Schimanski *et al.*, *Eukaryotic Cell* **4**, 1942-1950, 2005). Stop codons are double underlined.

Mass spectrometry data

Rfc1 (CFAC1_210019200)

```
  1 MSTLSFSTGI EAMTPVNTQS SATAAPRLSE LWADKYKPRS IAEMCYPSYA
 51 NKLKAWLENF TPVGSPGDDP NKHHGVLLSG SPGVGKTTTV YVVARELGRT
101 VIEYNASDFR SRKSLRENVL DLISNRAFAA QATSYTRAVL LMDEVDGCDI
151 GGVGEVVKML FITKIPILCT CNDRWHPKLQ TLVKYVEDMR FSHPPCNIVA
201 NYLCERVLAR EGITLSKPLL QDIIKKSGSD IRNMLNNLQL WCLNRRSLEQ
251 RQLAECAAQA TKDGDAGLFD SAEYFLLQGT SRGERHSIAE MQACYYNSDL
301 IDMFVQENYL HYNPEPVDGR DWMTAVAQAA SSISRADAAQ RIMYYEQNWS
351 VSRFHVLSSS IAPCVYTRGK YETFMTGQQK FFDLQRPVKF PQWLGHNSTA
401 GKNRRLLRCV AMQASHPTRG ISGNQEDVAA DYMPNGWERP LTQPLAEKEK
451 DGIAEVIALM DQYNLMRDDW DLVQTLPHFR HMETPRQQPP VSITTAVKAA
501 FTREFNKTHR FDSFAKTTLK RTDKADEDDG IDEEEGESQK EGAGAKAGTK
551 GRVIADGVTA VTITGSDAAK PKAKTSAARK PRAKKSAANA AAAADDSGET
601 KPARKRAASA STRKPAKPAG KASKAAAGGK ARKRARVESS SESEVEISSD
651 SSSDSSDSE
```

Rfc2 (CFAC1_260050500)

```
  1 MSLSSQPVTK KAKTEAAASP AAAATPWIEK YRPRTLDEVE AQDEAVSALR
 51 ACLKEGANMP HFLFHGPPGT GKTTSILAVA HELFGPDYIR SRVRELNASD
101 DRGINVVREK IKVFAQGAVS SGGSSVTQSD GKVYPVPGFK LIILDEADAL
151 LPDAQGALRR MMEDFSDVTR FCILCNYVSR IIDPIASRCA KYRFKPLVKT
201 ALYNRIQFVA NAEGIELSDA SLQALDSVSG GDLRLAIMHL QSAHKASGSD
251 LTREDFVSVS GSVPADAMQT YVAALVSRRL EDVIAVSRQL VAQGYAAAQV
301 LVQLQRYLVS AECPLNSAQR GRMMLKLCQT ERRLADGGDD YLQLLDMGSS
351 VCAS
```

Rfc3 (CFAC1_300082900)

```
  1 MATSKQAEDA KAGGSHLPWV EKYRPDNLDS VVAHEDILST LRHLMNSGNM
 51 PHLLLYGPPG TGKTTTIKAC AYYLYGKDRV RANVLEMNAS DDRGIDVVRQ
101 QIREFSSTSS IFSMMGPSSS SGGGGNGGSG PLASFKLVIL DEADQMSHDA
151 QAALRRVIEK YTKNVRFCIL CNHINKVIPA LQSRCTRFRF APVKKSAMMP
201 RLKYVAEQEK VKYTTEGLAA AFRLSHGDLR RCMNTMQSSA LSADEITEES
251 VYRVTGNPTP AEVTAIVSDM LSGDFATSWA KVEVAVTQKG ISIADLAREI
301 HPIMMAMDLP QDCKCFLLMK LSDMEYYAAG GAREAAGLGG LLGAFQLVKE
351 AVTQRKPIKA VAGDCSA
```

Rfc4 (CFAC1_230046600)

```
  1 MLCLARDLLL QNTDAATGAD KAGGKDILKD AVLELNASDD RGLDVVREKI
 51 KLFAQTKKTL PKKFFTTGEG AETMEQVVHL HKIVLLDEAD SMTPAAQQAL
101 RRTMELHSST TRFAFACNNS SKIIEPIQSR CAVVRFKKLS DADILRRLVF
```

```
151 VIQQEKVSYT DDGLEALLYL AEGDLRQAMN SLQATHTGYG LVNADNVFKV
201 CDQPHPVLVE NIITACVTKR NIEEAHKEMN RLLNRGYAPA DVIATFFKVV
251 QTNARLFRSE LQQLEVLKVV GETTMRIAEG VGTSLQLAAM LARMIAAVEN
301 NQS
```

Rfc5 (CFAC1_280077100)

```
  1 MLWVDRYRPK TLKEVELYPE LNDVLGRLAK AQDLPHLLFY GPSGSGKKTR
 51 AMAVLHEIYG PSVYSVRLEH KSVQVSDSKV VDIATLSSPH HIDINPSDAG
101 NYDRVIVMQM IREIAQTVPL HTTASSAKAV PYKVVVLNEV DKMSRSAQHA
151 LRRTMEKYMK TCRLVLICNS TSRLIPPLRS RCLGIRVAAH SKDNLALAVQ
201 HVCEGESRPM PSPAFLNSLA LRSDGNLRRG LLMLEASAMT KVDWSGNGAA
251 IPQADWKLFL DEISHDILAE QTPKKLHEVR LKFYDLLAQC ISGETILKTL
301 LDSLLLAVPP KHQAALIQLA ATYDHNMKLG TKPILHLEAF VAGVMKLIKQ
351 Q
```

Rad17 (CFAC1_260035800)

```
   1 MLNEVYAPTT VADLAWSRQK IVALSTLVRS TRSGAQNPRI LLLYGPPGCG
  51 KLESLKVLLR EAPPAAASTT SKSKTPAPAP QVIEPPTTVS VFHTCEASST
 101 AYSQFLQHVL SLCSGQLVGS ALMLTPKDMH GGRDTPSAPS DVQHAHIIKL
 151 YGEPATHVLH RATVAFLRQY EALRLQAIRE EEQQQHQRRY LAKVLASPAS
 201 PSTTLMDHLR RNLIFFVHTT HDSHNDKVDL GSALPAAVLQ SAAVELFHCT
 251 PVTEINLKKR LRHILDTEAR RRANRSAQQR RADVAEATDV DDLFGIAPAL
 301 SGSSAAPRRV AARGGAGSSR GKKGKENAKH APVTALHIPD AADVLDSLAL
 351 DAIAAGSQGD IRQALLQVQW AALVPPGSST AASLVETVAD SSDVVWARLQ
 401 HRRALAQAFA SGSSKADESS LVLSTKSVAP LAEACAAPQQ QDSTVAEDDD
 451 GVVLLISSSS SEFDAPLPLS AAEVTRRQHL PSRSHEAATR KRSRSSENDV
 501 VDVDDVGTTS KAAPPSAQAR ATDMLSLLDS QMNGAGESRA AAAASRGAAK
 551 KLLRAAPVRR DGLAAKNNTD ADDGAAVLPD HRTVLPTTRD EYLGLSHATG
 601 RLLSQKYSVD AVLDILNVPP RKMLDYLTNN QVRYFSDAQL PQYLVCAAAA
 651 SEVDALRTAE FDGGYGGSAA ALRERRQLAD RTTAGESVGN VARLLDVIAL
 701 QTFHRRYLVE QTAVQAPPGF TPQEPPPFLR SAYPRVRDVG STTNTTGPYM
 751 TQRGEAVLEL LAGVSEHEWM EQFLLRLDSA VSASGAITSF SSIGRRRMPP
 801 AASVVSSDAI FSQPALGLTS PSIRLDEVDI LREGLPDLLY RCGCTESVVM
 851 DHYALAPYIV LNLPASSQPS AAAVASQPSP AVTPAGISSA ASGDSADAGG
 901 APLSVARPRR TVFKFAASTP PPPPPTQLHS QPAPLSLRET HAARLSARRP
 951 CTSLQLKILQ RGRDSAAATL RGDHFVLVAT ENIAEEGSMS EKGGVEERPW
1001 MPEGDDIEDD
```

Rrp41A (CFAC1_280032800)

```
  1 MSRQKEYVSP AGLRLDGRRP LEARRMDIAF STLSGCDGSC DITLGRSKVC
 51 ASVFGPRESV HRQEAKHDEV IITCEVAVAA FAGEVRRNPQ RRGRLSEDIS
101 AAVVQVARSV VLLPQYPNSQ IHIYLEVLQQ DGNEKIACIN AACLALVDAN
151 VAMRDAVCCT NVGLLDEHVL VDLTNEELRS QCPVIAAAFT GHDTRNIIWL
201 ETTSRLLPEA AIRLLKAAGQ SAKELFEGTV RGALVEHATQ ILALQS
```

Rrp41B (CFAC1_160013900)

```
  1 MSSALSGQSA VTLSSSSSSS HPTAAATAGA SYTRRDGRTA LEIRGKEMRL
 51 SEMADFDGSS WYAQGQTAVL VTLHGPTLAK NEEYDTCLVR VRVQHAHGLT
101 PSAGGAERAV YEEMKLEMLT RTDALELESL LESTIDAVVM RDRFPRCVLV
151 VDVVVVQDDG SLAAVALNAV MCALLDAGVP CRTTMAAVCV AAVTRAEDAA
201 AGDASRAVGS SLELLLDPTT AEETLGAGNT AAATAAGGEK ARSTVDATMA
251 EKGDLSGAAA AKLSLLRPDA LQGHYRCVST GVFVFSNPAC GGGVLAQLVR
301 RRSGGDSGTG ANTVSVEVYG QMMTLAERAA VVLFDFFRQC NVAE
```

Rrp6 (CFAC1_290060300)

```
  1 MPPKSAEASL PATKAVVSAV FGAVKDYSKL SAQIPADDFE YHLAFAGFRK
 51 HIRDDSVGLV EVMDACCQML PKRRRTNLVA EEDPHSGAVH LAETQRNAVM
101 EAIDSLLENV DSLLDEVKGR KLDAQDQLSV TFGSELAVSA HHDASRGGSS
151 ASNAAGVVRL AHVRRPQLSF ETPVDNSAAP FTPTYRDASG VQHTGVAGEH
201 PFHDAIRAFS VPEAQMMPKA EIPPVPLETC PLSFVDTPDA MQAMVAKLLS
251 ASEIAVDLEH HDFYSYQGFT CLMQISTREE DFIVDCLQLR ASMGALAPVF
301 LNPSILKVFH GAREDVRWLQ KDFALYLVNF FDTGVALQTL HMPYSLAFAV
351 DHFCQVKLNK KYQTADWRVR PLPADMVHYA RQDTHFLLYV YDRLKALLLN
401 SEGRASVGNL LVHVYNESKQ LALQVYAKPN VDPAETYKLA LGRSLGGLTA
451 VQEEVAREIF NWRESAARDV DDSPTAVLHL SSVLAIASKL PTTAKDLLRC
501 CSPATAVLRA NVAHLVELVK KAVASSSEDF ENGVSGSGAG RHGKEEGSRH
551 HNYLDGAAEG SLEWAVYRSR CPTGVHRPMT GTLPSLASVV KTVTPAAVSV
601 SEQAALLSHT MPSSWFSAMS ALSRVLASRQ QHHVELPGAD VRAARQAAAA
651 KSLAGTADAA AAAAVAAEEE TAKAEVESVS SASGEGEQKD EEATGDLPAE
701 ADVAEASSVI ALDKKAFSIK QEYGVGAKSR FKKGEKGGAA KKKK
```

Eap1 (CFAC1_170027500)

```
  1 MSVSAASISL AEVRAVQDGV ANDVREDGRT LLQRRPVYIT PRSSPSAVAA
 51 VVGGSDGGDA AGVAQQSYSG SYVEVRASGT VVLAAATPTV VDGCATAAPS
101 PVSAADNADG AKEAAAAAPH DAGRGQLHIT IDAVPHVLDA YAGTVGGRNT
151 HRYRRDYLAF LAATIRAVFG AAQVQVQEQQ GVAEAEVVPE EREGAEDEVG
201 AGTVSSSAVA PAGSGRGDGE TSLASGFPAA DLYIGEGFGF RVHVDVHVLQ
251 CAGGNLFTAI AYAVHAALRS LQLPAVTLHR APGDGAGVSV EVDRSQPYRR
301 PVQWSQLPLL CVLLVSPTGH YVVDPTLREE WALPQQVHVA AGASGQVFYF
351 RYQQLPSRRG NRYQLQEARK ADAEACAAYV APPMALNLLD CWAVLSDAVY
401 VCQAMIHDCE VALQG
```

Rrp45 (CFAC1_240024400)

```
  1 MLLRSAAPVP ADALVQRNVE FARTAWRAGL RPDQREAHQL RMIEIEFPLL
 51 ARDTVQVKCG NTIATASVTC DLVEPMPFRP KHGFFEVHAR QLLHERDPLD
101 QPKAVKQLSM YLTRLLSGSV VETEGLCVIP GRRVWSIAAE VLILNNDGNL
151 HDVAQWAVMA ALQHVRRPEL TIRGDDVVVH PPHERDPVPL SLHHIPLSFT
201 FAVCANPQQV QLAARAAALR RASPVSAGAA GQGSSDNAGE KEDGADASAW
```

```
251 SDDALQIVAD PSLEEAAAAA CTVSVAVNAE GHVCSLEKAD GCDVSLEHLE
301 QCMQVALQLT PPLLTQMQEA MAAHDVKRKA AVRSQFLWAQ KRLGIQAAGG
351 AGASQTQEEQ AAKKSKTE
```

Rrp40 (CFAC1_030007200)

```
  1 MSTHSPTLKS VSELVPLKGH VCLPGEPVLM VQSSAVVAVG GGLRLLAQPS
 51 TATDASQDVA DVFLAEYCAP LQRSSHHLHT HVPRYTVATP ASRRYTPRHA
101 DPVIAVIARK VSQHYYYCYI GGSSLAYLEA IAFDGATKVS RPRLAEGDVV
151 YCYVKPRAAA SYVDGAAASS AAATAAAVSS GGEVELACTA AEVGLPPKDW
201 TSGEAVFGPL LGGRLLTLPL AYVRRLLAPL PATLSGEGPA VKRARVEGGG
251 GEAEEVPASY LLHLLGQRVP FEVAVGMNGL VWVKGLTSEA DATAAARRTV
301 AVSACISEAQ YDATRAEMEA RVESYFPS
```

Rrp4 (CFAC1_110005300)

```
  1 MSSGVVIVGD SICGGERIQK LNTSNDEVYL RGFNTFAGNN PSDIALVHEG
 51 AGEIVAAING HIEVTDRVVS VKGLLPRYQP EIGDVVVGRI LEVTGNKWQV
101 DVNSTQTAIM LLSNVTEPGG MLRRRGRGDE LGMRQLFDQE DLVAAEVQRI
151 SPDGVVSLHT RAAEKYGRIG GFGVLVSVRP SLVKRAKHQF VELAEHHVRL
201 TIGMNGNIWV SRKEETADGT EDKEREAEAR QNVARVANCV KALGVAHIQI
251 HPATIEAAVA ASVEAGFSAF HVSLEKNRDA LLVSVHDAIG VKRRRQ
```

Csl4 (CFAC1_150031200)

```
  1 MPVLVHTGAR VAPGDALFSS AAHVPTGTDA SAATAGDTVS DSDVIPGEGC
 51 VVHYVEVPSE STGDSSRVRR HIVATRQGVA QWDGRLVSVF AAGATGTTAQ
101 LQGASTAVRS AVTGPRPGDT VHVRITRLSR LFAFGEITAV NWQWCSHRSA
151 AGASVSGVFK GVLRLEDIRP FRPTRDQLQP PPPTMAFALG DVVLAEVISQ
201 SDAHQYQLST VGEGFGVVES YVSTAEEHYS GRERVKLQHL PGRRDAMLVP
251 ATGAVVPRWC PLLP
```

Eap2 (CFAC1_300052700)

```
  1 MSLPPNTGSI ELTAFRAHTS QLLARGERLD KRDFTTCRVP TVVREERAAE
 51 APSSSSSGVV QTGINMANSG NLAAVMYTDS YGACMQCTVQ GLLGPPRPDR
101 PAAGRLNIHV EAPFVEQLGG GAATNYKSFQ YIISNGNADL PLRQLEGYIG
151 SVVDGCFDPT QLSIYDGEAC WVLNVTVTLL SFDGGLRAAS LHAVLAALHQ
201 LRLPRTRLPN GDVIESRRVR LSCLPTACTF GFLAGAQVRL LADTTAIEEY
251 VADGLLTIAV SESGEVVGVH QVGRCPLLAQ ALTAAVQQWT EQSASVRKAL
301 YG
```

Eap4 (CFAC1_280054900)

```
  1 MTRLDGRQST EAVRAIHVAT NVLANCHSSA CVEIGQTRVL CGVRPPQQLV
 51 QEYRGTRGRI SCQLHRSSAS SAAATVADNS ADRDMALALE GVAEQAVVLE
101 RIPQLLVEVL IEVLHDDGAV WDAAATALSA ALTAGGVEVY DTFTACSAAV
```

151 RPDGAIVVDL TQEEEAAATA RVVVCGGVSL GGVYYMCHLG ACEAATMAQL

201 VQAATKGMQV R**KALLLEQIR** NQ